Ievlanov, Maksym (Ed.)

Book

Project Management: Industry Specifics

Provided in Cooperation with:
PC TECHNOLOGY CENTER, Kharkiv

Suggested Citation: Ievlanov, Maksym (Ed.) (2024) : Project Management: Industry Specifics, ISBN 978-617-8360-03-0, PC TECHNOLOGY CENTER, Kharkiv. Ukraine, https://doi.org/10.15587/978-617-8360-03-0

This Version is available at:
https://hdl.handle.net/10419/321952

Edited by
Maksym Ievlanov

PROJECT MANAGEMENT: INDUSTRY SPECIFICS

Collective monograph

Technology® Center

2024

UDC 65
P93

Published in 2024
by TECHNOLOGY CENTER PC®
Shatylova dacha str., 4, Kharkiv, Ukraine, 61165

P93 **Authors:**
Edited by **Maksym Ievlanov**
Maksym Ievlanov, Nataliya Vasiltcova, Olga Neumyvakina, Iryna Panforova, Maksym Naumenko, Iryna Hrashchenko, Tetiana Tsalko, Svitlana Nevmerzhytska, Svitlana Krasniuk, Yurii Kulynych, Viktor Myronenko, Andrii Pozdniakov, Yaroslav Ziubryk, Valerii Samsonkin, Ivan Riabushko, Oksana Malanchuk, Anatoliy Tryhuba, Ivan Rogovskii, Liudmyla Titova, Liudmyla Berezova, Mykola Korobko, Georgii Prokudin, Viktoriia Lebid, Oleksii Chupaylenko, Tetiana Khobotnia, Maksym Roi, Mykhailo Holovatiuk
Project management: industry specifics: collective monograph. – Kharkiv: TECHNOLOGY CENTER PC, 2024. – 179 p.

Among the problems existing in the field of project management, it is necessary to highlight the problem of developing new and adapting existing scientific achievements for the purpose of their practical application for project management in various spheres of human activity. The results obtained in solving this problem have not only scientific, but also applied value, because they determine possible places of effective application of scientific and applied products in project management of specific enterprises and organizations.
The monograph contains the results of research on the application of various formal means of solving project management problems in various spheres of human activity. The monograph separately considers the following areas: IT sphere, food industry, transportation sector, and medical sector. This monograph can be useful to researchers, teachers, postgraduate students and higher education students in the field of project management, as well as specialists in applied project management in various spheres of human activity.
Figures 40, Tables 36, References 147 items.

DOI: 10.15587/978-617-8360-03-0
ISBN 978-617-8360-03-0 (on-line)

Cite as: Ievlanov, M. (Ed.) (2024). Project management: industry specifics: collective monograph. Kharkiv: TECHNOLOGY CENTER PC, 178. doi: http://doi.org/10.15587/978-617-8360-03-0

AUTHORS

CHAPTER 1

MAKSYM IEVLANOV
Doctor of Technical Sciences, Professor
Department of Information Control Systems
Kharkiv National University of Radio Electronics
ORCID: https://orcid.org/0000-0002-6703-5166

NATALIYA VASILTCOVA
PhD, Associate Professor
Department of the Information Control Systems
Kharkiv National University of Radio Electronics
ORCID: https://orcid.org/0000-0002-4043-487X

OLGA NEUMYVAKINA
PhD, Senior Researcher, Leading Engineer
Department of Information Control Systems
Kharkiv National University of Radio Electronics
ORCID: https://orcid.org/0000-0001-6936-6543

IRYNA PANFOROVA
PhD, Associate Professor
Department of Information Control Systems
Kharkiv National University of Radio Electronics
ORCID: https://orcid.org/0000-0001-7032-9109

CHAPTER 2

MAKSYM NAUMENKO
PhD, CFO
Charitable Foundation "Ukrainians Together"
ORCID: https://orcid.org/0009-0006-7590-572X

IRYNA HRASHCHENKO
PhD, Associate Professor
Department of Management of Foreign Economic Activity of Enterprises
National Aviation University
ORCID: https://orcid.org/0000-0002-8735-9061

TETIANA TSALKO
PhD, Associate Professor
Department of Management and Smart Innovation
Kyiv National University of Technologies and Design
ORCID: https://orcid.org/0000-0002-4609-8846

SVITLANA NEVMERZHYTSKA
PhD, Associate Professor
Department of Management and Smart Innovation
Kyiv National University of Technologies and Design
ORCID: https://orcid.org/0000-0001-5392-9030

SVITLANA KRASNIUK
Senior Lecturer
Institute of Law and Modern Technologies
Kyiv National University of Technologies and Design
ORCID: https://orcid.org/0000-0002-5987-8681

YURII KULYNYCH
PhD, Associate Professor
Department of Finance
National University of Food Technologies
ORCID: https://orcid.org/0000-0002-9018-0708

CHAPTER 3

VIKTOR MYRONENKO
Doctor of Technical Sciences, Professor, Head of Department
Department of Management of Commercial Activity of Railways
State University of Infrastructure and Technologies
ORCID: https://orcid.org/0000-0002-6088-3867

ANDRII POZDNIAKOV
PhD Student
Department of Management of Commercial Activity of Railways
State University of Infrastructure and Technologies
ORCID: https://orcid.org/0000-0002-3359-586X

YAROSLAV ZIUBRYK
PhD Student
Department of Management of Commercial Activity of Railways
State University of Infrastructure and Technologies
ORCID: https://orcid.org/0009-0003-4078-0611

VALERII SAMSONKIN
Doctor of Technical Sciences, Professor
Department of Transport Technologies and Management of Transportation Processes
State University of Infrastructure and Technologies
ORCID: https://orcid.org/0000-0002-1521-2263

IVAN RIABUSHKO
Assistant
Department of Technology of Machinery Manufacturing and Machine Maintenance
Kharkiv National Automobile and Highway University
ORCID: http://orcid.org/0009-0007-2492-6886

CHAPTER 4

OKSANA MALANCHUK
PhD, Associate Professor
Biophysics Department
Danylo Halytsky Lviv National Medical University
ORCID: https://orcid.org/0000-0001-7518-7824

ANATOLIY TRYHUBA
Doctor of Technical Sciences, Professor, Head of Department
Department of Information Technology
Lviv National Environmental University
ORCID: https://orcid.org/0000-0001-8014-5661

IVAN ROGOVSKII
Doctor of Technical Sciences, Professor
Department of Technical Service and Engineering Management named after M. P. Momotenko
National University of Life and Environmental Sciences of Ukraine
ORCID: https://orcid.org/0000-0002-6957-1616

LIUDMYLA TITOVA
PhD, Associate Professor
Department of Technical Service and Engineering Management named after M. P. Momotenko
National University of Life and Environmental Sciences of Ukraine
ORCID: https://orcid.org/0000-0001-7313-1253

LIUDMYLA BEREZOVA
PhD, Associate Professor
Department of Technical Service and Engineering Management named after M. P. Momotenko
National University of Life and Environmental Sciences of Ukraine
ORCID: https://orcid.org/0000-0002-8443-8442

MYKOLA KOROBKO
PhD, Associate Professor
Department of Technical Service and Engineering Management named after M. P. Momotenko
National University of Life and Environmental Sciences of Ukraine
ORCID: https://orcid.org/0000-0002-1138-7701

CHAPTER 5

GEORGII PROKUDIN
Doctor of Technical Sciences, Professor, Head of Department
Department of International Transport and Customs Control
National Transport University
ORCID: https://orcid.org/0000-0001-9701-8511

VIKTORIIA LEBID
PhD, Associate Professor
Department of International Transport and Customs Control
National Transport University
ORCID: https://orcid.org/0000-0002-1260-3760

OLEKSII CHUPAYLENKO
PhD, Associate Professor
Department of International Transport and Customs Control
National Transport University
ORCID: https://orcid.org/0000-0002-4837-0727

TETIANA KHOBOTNIA
PhD, Associate Professor
Department of International Transport and Customs Control
National Transport University
ORCID: https://orcid.org/0000-0001-7094-6297

MAKSYM ROI
PhD, Associate Professor
Department of International Transport and Customs Control
National Transport University
ORCID: https://orcid.org/0000-0001-5788-4220

MYKHAILO HOLOVATIUK
Assistant
Department of International Transport and Customs Control
National Transport University
ORCID: https://orcid.org/0009-0005-0331-0105

ABSTRACT

Among the problems existing in the field of project management, it is necessary to highlight the problem of developing new and adapting existing scientific achievements for the purpose of their practical application for project management in various spheres of human activity. Solving this problem requires constant consideration of the features and limitations that arise as a result of the impact on the basic models and methods of the specifics of individual industries of human activity. The results obtained in solving this problem have not only scientific, but also applied value, because they determine possible places of effective application of scientific and applied products in project management of specific enterprises and organizations. Therefore, conducting research that highlights the influence of industry specifics on formal models and methods for solving individual project management problems is relevant both from a theoretical and an applied point of view.

The monograph offered to your attention contains the results of research on the application of various formal means of solving project management problems in various spheres of human activity. The monograph separately considers the following areas: IT sphere, food industry, transportation sector, and medical sector.

Among the variety of formal means of solving project management problems, special attention is paid in the monograph to: Data Mining methods; economic and mathematical models of transportation; differential-symbolic apparatus; apparatus of fuzzy production rules.

It should be noted that the main attention of the authors of the monograph was focused on the application of formal models and methods during project planning in general and their individual aspects (Sections 1, 3, 4 and 5). Section 2 is devoted to the application of formal apparatus for the formation, analysis and decision-making in project management. It should also be noted that, unlike other sections of the monograph, the materials of Section 3 are devoted to the use of the formal apparatus as one of the means of generalizing the experience of project implementation. The knowledge obtained in this way can be reused during the initiation and planning of new projects.

KEYWORDS

Project management, Data Mining, economic models, differential-symbolic approach, fuzzy production rules, IT product, food industry, transport networks, medical support of the population, transport services.

CIRCLE OF READERS AND SCOPE OF APPLICATION

The first section presents the results of research on the application of clustering methods to solve the problem of identifying configuration elements of an IT project. These results can be used in the processes of managing the content of IT projects. It is advisable to use these results when planning an IT project as a whole and planning the activities of individual teams of IT project performers. The materials of this section will be useful to scientists in the fields of IT project management, computer science and information technology, IT project managers, system architects and IT product managers.

The second section presents the methodological foundations and scientific and applied solutions for the use of Data Mining and modeling as tools for adaptive project management in the food industry. These results can be used in the processes of forming, analyzing and making decisions on project management at food industry enterprises. It is advisable to use these results to increase the completeness and accuracy of modeling at all levels of project management. The materials of this section will be useful to scientists and specialists in project management at food industry enterprises.

The third section offers the results of the analysis and practical recommendations for improving management practices in the field of railway transport and the integration of national infrastructure into international transport networks. These results can be used to manage projects for the development of the country's transport networks and their integration into international networks. These results are a generalization of the experience of project implementation, which should be taken into account when initiating and planning new projects in the transport sector. The presented materials will be useful to scientists and practitioners in the fields of transport project management, transport engineering and international integration.

The fourth section offers theoretical foundations and computer models of differential-symbolic planning and risk assessment of projects in the medical sector. These results can be used when planning medical projects related to medical support for the population of communities. It is advisable to use these results as one of the possible means of scenario analysis and risk assessment in medical projects. The presented materials will be useful to scientists and practitioners in the fields of medical project management and the application of information technologies in project management.

The fifth section presents theoretical approaches, models and methods for assessing the quality of freight transportation projects. These results can be used to manage the quality of transport projects. It is advisable to use these results for project analysis and selection of proposals for the provision of transport services during the initiation and planning of projects. The presented materials will be useful to scientists and practitioners in the management of transport projects in the field of freight transportation.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The object of this monograph is the scope of implementation and individual project management processes in various fields of human activity.

In the first section, the problem of identifying configuration elements (CE) of an IT product has been solved. To solve the problem of identifying functional CEs, it is proposed to use hierarchical and non-hierarchical clustering methods. As examples of hierarchical clustering methods, the agglomerative clustering method AGNES using the nearest neighbor algorithm and the divisive clustering method DIANA have been proposed. As an example of non-hierarchical clustering methods, the k-means algorithm is proposed. For comparison with these methods, one of the graph-analytic clustering methods has been used, which has been developed to solve the problem of decomposition of the description of the monolithic architecture of a software product into separate microservices. A comparative analysis of the solution progress and the obtained results of solving the problem of identifying functional CEs using all four selected clustering methods has been carried out. It has been established that the best alternative is to use hierarchical clustering methods to solve this problem.

In the second section, methodological foundations and scientific and applied solutions of multi-mode adaptive intelligent data analysis for food industry enterprises have been developed. During the study of the methodological foundations of multi-mode adaptive intelligent data analysis, the effectiveness of using this analysis in the management of competitive food enterprises has been established. The main advantages, difficulties, challenges and problems of using the proposed analysis have been identified. A list of main corporate tasks has been formed, for the solution of which it is advisable to use this type of analysis. A list of main results of using this analysis for an effective and competitive food enterprise in dynamic and crisis conditions has been determined. During the study of scientific and applied solutions of multi-mode adaptive intelligent data analysis, the features, methods and main applied solutions of the following types of analysis have been determined: High Dimensional big data analysis, Ad-Hoc Data Mining, Anomaly & Fraud Detection, Hybrid Data Mining, Crisis Data Mining. The results of the study can be effective for enterprises and companies in which management decisions have been made in unstable or crisis conditions.

The third section analyzes the management aspects of the integration of Ukrainian railways into the international transport infrastructure. In the process of this analysis, the experience of planning and implementing high-speed railway projects in Europe, Asia and other regions has been studied. Particular attention has been paid to the adaptation of world practices to Ukrainian conditions, taking into account the specifics of the national infrastructure, economic and political factors. The features of the project for the integration of Ukrainian railways into the Trans-European Transport Network have been considered using the example of the Lviv railway junction. For this type of project, the stages of its implementation have been described, including technical, organizational and financial aspects, and an analysis of problems and achievements is conducted. The impact of this

type of project on the development of regional infrastructure and the economy has been studied. Using the example of a project for the automatic identification of rolling stock and large-tonnage containers on Ukrainian railways, practical aspects of management processes, including planning, implementation, and monitoring of projects at the system level, have been explored. Examples of successful and problematic projects have been considered, illustrating the opportunities and challenges of the project management process in the context of Ukrainian railway transport.

The fourth section considers the application of the differential-symbolic approach to planning medical projects for supporting the population of communities. This approach involves the use of differential equations to describe the dynamics of projects as a separate system and the use of symbolic expressions to represent individual parameters and their description. Mathematical models for differential-symbolic planning of projects for improving the health of the population of communities and assessing the risks of projects for medical support of the population of communities have been developed. To implement the proposed models, code has been written in the Python programming language using libraries for solving differential equations, optimizing and visualizing results. Based on the use of the developed and implemented models for the given conditions of the project environment, the results of optimizing the configuration of projects for improving the health of the population in the community and assessing the risks of projects for medical support of the population of communities have been obtained.

The fifth section considers theoretical approaches, models and methods for assessing the quality of transportation projects as a means of increasing the efficiency of providing transport services in cargo transportation projects of a project-oriented enterprise. An N-model for making an optimal decision regarding the importance of a set of criteria that determine the quality of transport services, taking into account expert information, has been developed. It allows determining the advantages of one criterion over another based on the theory of the importance of criteria, which can be applied in project process management. A model for ensuring the relationship between the quality indicators of cargo transportation projects and determining the attractiveness of international routes has been developed. The effectiveness of applying the developed methods and models at project-oriented enterprises in the transport industry has been proven by testing them at motor transport enterprises.

This monograph can be useful to researchers, teachers, postgraduate students and higher education students in the field of project management, as well as specialists in applied project management in various spheres of human activity.

DOI: 10.15587/978-617-8360-03-0.CH1

Maksym Ievlanov, Nataliya Vasiltcova,
Olga Neumyvakina, Iryna Panforova

USE OF CLUSTERING METHODS TO SOLVE THE PROBLEM OF IDENTIFYING CONFIGURATION ITEMS IN IT PROJECT

ABSTRACT

The object of the research is the IT project configuration management process.

During the research, the task of identifying the configuration items (CI) of the IT product has been solved. Research in this field is mainly aimed at solving the problem of configuration analysis during the refactoring of a monolithic IT product into separate services or microservices. The question of decomposition methods of the description of the architecture of the developed IT product into separate functional SIs remains practically unexplored.

In the course of the research, it has been proposed to use hierarchical and non-hierarchical clustering methods to solve the problem of identifying functional CIs. As examples of hierarchical clustering methods, the agglomerative clustering method (AGNES using the nearest neighbor algorithm) and the divisive clustering method DIANA have been proposed. The k-means algorithm has been proposed as an example of non-hierarchical clustering methods. In addition to them, it has been suggested to use one of the grapho-analytic clustering methods for comparison, which was developed to solve the decomposition problem of the description of the monolithic architecture of the software product into separate microservices.

The starting data for the research is the description of the architecture of the "Formation and management of the individual plan of the scientific and pedagogical worker of the department" functional task at the level of individual functions. 10 functions of this problem have been considered as CI. Descriptions of 12 entities of the problem database have been used to define these functions. The features of the solution have been considered and the results of the solution to the problem of identifying functional CIs using all four selected clustering methods have been obtained.

A comparative analysis of the process of solving and the obtained results of solving the task of identifying functional CIs using all four selected clustering methods has been carried out. It has been established that the best alternative is to use hierarchical clustering methods to solve this problem. This makes it possible to further consider the task of assigning to individual teams of IT project executors a list of functional CIs that require implementation as a sequence of individual single-criteria optimization tasks.

KEYWORDS

Configuration item, identification, IT product, architecture description, AGNES method, DIANA method, k-means algorithm, Chebyshev distance, Hamming distance, function, data flow diagram, ER diagram.

Modern views recognize the process of configuration management as one of the main processes of project management within the life cycle of IT products of various purposes. Such projects will be referred to as IT projects hereafter. And although different points of view define the appropriateness of the configuration management process, the formulation of the purpose of this process, which is a rule, coincides. Thus, in [1], the configuration management process is one of the integrated change control processes. In contrast to this point of view, in [2] the process of configuration management is one of the processes of technical management of an IT project. But the descriptions of the purpose of this process are more similar to each other. In [1], the purpose of this process is to systematically control configuration changes and maintain the integrity and tracking of the configuration throughout the entire product life cycle. In [2], the purpose of this process is to manage and control system items and configurations within the system life cycle. In addition, configuration management ensures meaningfulness between a product and the configuration definition associated with that product [2]. The variants of decomposition of the configuration management process also correspond. Thus, in [1] the main tasks of this process are: planning and management of the configuration management process; configuration identification; configuration control; configuration state accounting; configuration audit and product release and delivery management. In [2], the main actions of this process are: configuration management planning; configuration definition; implementation of configuration change management; performance of configuration status accounting; configuration evaluation; implementation of release control.

Certainly, the key work of the configuration management process is the work of identifying or defining the configuration. The results of the execution of almost all other works of this process depend on the results of this work. In general, the points of view on the tasks that are performed within the scope of this work coincide. Thus, in [1], the work on configuration identification includes the following tasks:

– determination of items that are subject to control;
– establishment of item identification schemes and their versions;
– establishment of tools and methods that will be used to obtain and manage selected items.

In [2], the configuration definition action includes the following tasks:

– determination of system items and information objects that are configuration objects;
– determination of hierarchy and structure of system information;
– setting identifiers of the system, system item and information object;

– determination of baselines during the life cycle;

– obtaining the agreement of the party acquiring the system to establish a baseline with the supplier.

But these points of view on the process of configuration management, its purpose and content are too theoretical and methodical in nature. A significant reason for this is, in particular, the lack of a clear idea of what exactly to consider controllable items within the configuration management process [1], or configuration objects [2]. Usually, such items are called "configuration items" (CI). These items can be IT products or their baselines, individual system items of such products, information objects, software, a set of hardware or software-hardware complexes. It is recognized that CI can be described by a hierarchical scheme of decomposition of products into separate system items, system items into separate software services, etc. [2]. As a result, it should be assumed that at different stages of the life cycle of an IT product, different descriptions of system items may appear as CI, depending on the current level of presentation of this IT product. An example of such a multi-level representation is proposed in [3].

Based on this assumption, it becomes possible to conduct research on methods of solving the problem of determining the optimal set of CIs at the early stages of the life cycle of IT products. Optimal here should be understood as such a CI set that will require minimal expenditure of time and resources for the implementation of relevant IT projects. Among these IT projects, special attention should be paid to IT projects of creation or development of information systems (IS) management of enterprises and organizations. As shown in [4], most cases of negative impact of local decisions on the overall design and quality of a large software and technical system are eliminated, in particular, by early division of the system into separate items. Solving the task of determining the optimal set of CI provides an assessment of the possibility of implementing an IT project of creating an IS at the stages of project initiation and planning. Such an assessment will make it possible to make more informed decisions about the possibility and feasibility of implementing such projects. Therefore, such studies should be recognized as relevant.

1.1 ANALYSIS OF MODERN RESEARCH IN THE FIELD OF IDENTIFYING CONFIGURATION ITEMS IN IT PRODUCT

Solving the problem of CI identification is of particular importance when designing large systems (the so-called "System of Systems"), which include a significant number of IS management of enterprises and organizations. Thus, in [4] it is shown that negative cases of the influence of local decisions on system-wide decisions arose mainly as a result of incorrect interpretation of requirements or bias of personal experience. This allows to draw a conclusion about the feasibility of using to solve the problem of CI identification during the identification and planning of an IT project for the creation of artificial intelligence methods IS. In such methods, the subjective influence of an individual specialist is minimized. At the same time, the basic description of the IS, which can be

decomposed into separate CIs, should be considered the description of the architecture of the IS being created. It is this description that combines the descriptions of individual requirements for IS and its individual functions [2].

The vast majority of modern research on solving the problem of decomposition of the description of the system architecture into separate CIs is based on the results of theoretical studies considered in [5]. From these results it follows that:

a) most of the existing approaches to the decomposition of a monolithic architecture into a set of individual microservices are applicable only under certain conditions;

b) there are probably no universal approaches to solving this problem.

However, the research results given in [5] practically do not take into account the transfer of the solution of this problem to a higher level of abstraction. Such a transfer is, in particular, the selection of CIs that implement individual IS functions from the description of the architecture of this system as a coherent set of functions.

The first of the above statements is supported by modern research on refactoring the code of a monolithic software application into individual microservices [6, 7]. At the same time, a density-based clustering algorithm is used to solve this problem in [6]. For such a special case of the decomposition problem of the description of the system architecture into separate CIs, the proposed solutions give positive results. However, a serious limitation of the application of these results is the need to observe the assumption of invariance of the functional decomposition of the original description of the software application architecture. This assumption may be violated as a result of a change in an existing or the appearance of a previously forgotten functional requirement for an IT product.

Theoretical studies, aimed at verifying the second of the above statements, made it possible to establish that the solution of most IT service selection problems requires an a priori definition of a set of functionally equivalent IT services [8]. This means that the decomposition problem of the description of the system architecture, created on the basis of a set of functional system requirements, into individual IT services must be solved separately for abstract descriptions of these services. Such abstract descriptions should not depend on the implementation specifics of these services and the non-functional requirements put forward to the services.

This conclusion is confirmed by the results of solving the problem of identification based on the description of the meta-architecture of the designed software system of individual software artifacts that ensure the reliable operation of this system [9]. This work shows that the use of abstractions to describe the architecture of a software system simplifies the solution of the decomposition problem of the description of the system architecture into separate items. However, the selection of such abstractions by functional feature was not considered in [9].

The application of the approaches proposed in [8] to the implementation of the functional decomposition of the description of the architecture of the software system into separate items is shown in [10]. In this work, at the first stage, the description of the architecture of the software system is divided into separate items, taking into account the necessary evolutionary changes.

The second stage of solving the decomposition problem in [10] is considered as the selection of such decomposition options that satisfy the constraints on the cost of the necessary changes. However, this approach to solving the decomposition problem leaves open the issue of assigning individual CIs to IT project teams. The solution to this issue requires the assignment to each individual team or performer of such a subset of CIs that would contain the most similar functional CIs among themselves.

The decomposition of the description of the system architecture into individual microservices using an object-oriented language was studied in [11]. The proposed Silvera language and its compiler make it possible to automate the decomposition of the system-wide description of the architecture into descriptions of individual microservices during the design of the software system. But the use of this solution is limited by the following features [11]:

– focusing exclusively on the decentralized development of microservices;

– the absence of a description of the business logic in the description of the system architecture.

Also, in [3], the issue of the optimal (from the point of view of IT project costs) number of teams of microservices developers that are allocated is neglected.

The analysis of the considered publications allows to formulate the following conclusions:

a) for large IT projects of creation or development of IS for management of enterprises and organizations, the task of CI identification is an important task;

b) solving the problem of CI identification in similar IT projects is usually considered as solving the decomposition problem of the IS architecture description into separate CI;

c) solving this problem requires its division into two sequentially solved sub-problems:

– the subtask of forming a set of options for decomposing the description of the system architecture into separate functional CIs;

– the subtask of selecting from a set of separate functional CIs a subset that will satisfy the selection conditions in the best way.

At the same time, it should be remembered that the peculiarities of solving the problem of forming a set of options for the decomposition of the description of the system architecture into separate functional CIs have been studied very poorly.

These conclusions make it possible to recognize the need for research in the field of finding special methods for solving the subtask of forming a set of options for the decomposition of the description of the system architecture into separate functional CIs. Such methods should be aimed at forming a set of all possible options for decomposition of the IT product architecture description into individual CIs.

The aim of this research is to identify the advantages and disadvantages of using the simplest clustering methods to solve the subtask of forming a set of options for the decomposition of the description of the IS architecture into separate functional CIs. For this purpose, it is proposed to compare the progress and results of solving this subtask using hierarchical and non-hierarchical clustering algorithms. It is expected that the results of the research will allow to draw a conclusion about the expediency of using clustering methods to solve the problem of identifying functional CIs and reasonably recommend the best of the considered algorithms.

To achieve this aim, it is proposed to solve the following research problems:

– to determine the features of hierarchical and non-hierarchical clustering algorithms that will be used in the research;

– to formulate a description of the initial data of the identification task of functional CIs, which will be used to compare different solutions;

– to consider the progress and results of solving the problem of identifying functional CIs using selected representatives of hierarchical and non-hierarchical clustering algorithms;

– to conduct a comparative analysis of the results of the received decisions.

1.2 METHODS OF RESEARCH

Hierarchical clustering methods are usually divided into two large classes: agglomerative and divisive. As an example of agglomerative methods, it is proposed to use the AGNES hierarchical agglomerative clustering method. To solve the subtask of forming a set of options for the decomposition of the description of the system architecture into separate functional CIs, this method can be represented as a sequence of the following stages [12, 13]:

Stage 1. The entire set CI is represented as a set of clusters C, each of which contains one item CI_i, $i=1,\ldots,n$, where n is the number of items in the CI set. Calculate the distance matrix D between the items of the set C.

Stage 2. Select two clusters c_p and c_q, the distance between which will be minimal, and combine them into a new cluster c_r, entering it instead of clusters c_p and c_q into the set of clusters C.

Step 3. Calculate the values of the distance matrix D using the rule:

$$d_{rs} = \alpha_p \times d_{ps} + \alpha_q \times d_{qs} + \beta \times d_{pq} + \gamma \times \left| d_{ps} - d_{qs} \right|, \; r \neq s, \; r \neq p, \; r \neq q, \tag{1.1}$$

where d_{ps} – the distance between the centers of clusters c_p and c_s; d_{qs} – the distance between the centers of clusters c_q and c_s; d_{pq} – the distance between the centers of clusters c_p and c_q; α_p, α_q, β and γ – parameters whose value is determined based on the selected distance calculation algorithm.

Step 4. Repeat Step 2 and Step 3 until one cluster is formed that includes all items of the CI set.

It is proposed to use the nearest-neighbor clustering algorithm as an algorithm for calculating distances. For this algorithm, the parameters in (1.1) acquire the following values: $\alpha_p=0.5$; $\alpha_q=0.5$; $\beta=0$; $\gamma=-0.5$ [12, 13].

As an example of divisive methods, it is proposed to use the DIANA hierarchical divisive clustering method. In order to solve the subtask of forming a set of options for the decomposition of the description of the system architecture into separate functional CIs, this method can be represented as a sequence of the following stages [12–14]:

Stage 1. Form one cluster C1 consisting of all m items of the original set of clustering objects CI.

Step 2. Select the C1 cluster item with the largest average distance value from other cluster items. The average distance value for the i_p item of cluster C1 can be calculated as follows:

$$D_{C1}(i_p) = \frac{1}{m} \times \sum_{q=1}^{m} d(i_p, i_q) \; \forall i_p, i_q \in C1, \; p \neq q, \tag{1.2}$$

where m – the number of items in cluster C1; i_p – the p-th item of cluster C1; i_q – the q-th item of cluster C1; $d(i_p, i_q)$ – the distance between the items i_p and i_q.

Step 3. Remove the item selected in Step 2 from cluster C1 and include it in the new cluster C2.

Stage 4. Among the remaining items of the cluster, find one for which the difference between the average distance to the items remaining in the cluster C1 and the average distance to the items included in the cluster, C2, is positive and maximal.

Step 5. Delete the item selected in Step 4 from cluster C1 and include it in the new cluster C2.

Step 6. Continue to perform Steps 4, 5 until the differences of the average distances become negative, then terminate the method.

The result of this method is the division of the original cluster C1 into two children – C1 and C2. Next, one of the child clusters is selected and divided into two more child clusters using this method. The separation procedure stops in one of the following cases [11–13]:

a) only one item remains in the child cluster;

b) all items of the child cluster have zero difference from each other.

The results of the use of hierarchical classification methods are dendrograms, which describe the result of the decomposition of the IS architecture description into separate clusters of functional CIs. These results allow to proceed to the solution of the sub-problem of selecting from a set of separate functional CIs a subset that will satisfy the selection conditions in the best way. The solution of this subtask is aimed at grouping selected clusters of functional CIs into lists of tasks (backlogs) for IT project implementation teams. At the same time, it is assumed that each team of performers is homogeneous, that is, it consists of specialists with the same skills and level of quality of work performance.

This approach requires a significant investment of time. These costs could be reduced by using non-hierarchical clustering methods, namely partitioning methods. A classic example of these methods is the k-means algorithm [12–14]. The use of this algorithm makes it possible to determine clusters of functional CIs according to the number of executive teams that may be assigned to the implementation of the IT project of creation or development of this IS. In addition, the k-means algorithm is one of the simplest clustering methods and is the basis for a sufficiently large set of non-hierarchical algorithms [12–14].

Conducting research requires comparing the course and results of the proposed clustering methods. Therefore, it is necessary that the above methods are based on the same method of determining the distance between descriptions of functional CIs. For this purpose, the research introduces an assumption according to which descriptions of functional CIs are sets of entities or classes. Therefore, it is suggested to use the methods of determining the distance for objects

whose descriptions contain qualitative features. Among such descriptions, the most frequently used are [15]:

– Hamming distances $dH(x_i, x_j) = \sum_{l=1}^{n} \left| x_{il} - x_{jl} \right|$;

– distances based on the Hauer coefficient;

– distances based on different correlation coefficients (for example, Pearson, Spearman or Kendall).

However, these methods do not sufficiently take into account the peculiarities of the IS CI description. Therefore, the use of these methods in their pure form to solve the given sub-task is ineffective.

In order to establish a better way of determining the distance between descriptions of functional CIs, let's introduce the second assumption into the study. According to this assumption, the description of each individual IS function as a functional CI should be unified and represent a set of the following groups of descriptions [16]:

a) a group of descriptions of entities or classes that characterize a specific functional CI;

b) a group of descriptions of entities or classes of input data flows of a specific functional CI;

c) a group of descriptions of entities or classes of output data flows of a specific functional CI.

This assumption allows to state that the descriptions of two functional CIs should be considered different if they do not match at least in one of the above groups. Based on this, it is recommended to use the Chebyshev distance to determine the distance between descriptions of functional CIs based on the introduced assumptions. This distance is determined by the formula:

$$d_{\infty}(i_p, i_q) = \max_{1 \le l \le m} \left| i_{pl} - i_{ql} \right|, \tag{1.3}$$

where i_p, i_q – objects between which the distance is determined; i_{pl} – the l-th coordinate of the object i_p, $1 \le l \le m$; i_{ql} – the l-th coordinate of the object i_q, $1 \le l \le m$.

In this study, the objects i_p, i_q will be descriptions of the i-th and j-th functional CIs. The coordinates of each specific object will be the groups of function descriptions, input and output flows of the corresponding CIs discussed above. Based on this, the number m will be equal to 3 for the task of decomposition of the description of the IS architecture into separate clusters of functional CIs.

However, to determine the Chebyshev distance, it is necessary that the coordinate values be numerical for all objects. For the sub-task of decomposition of the description of the IS architecture into separate clusters of functional CIs, groups of descriptions, which are selected as coordinates, are sets of character strings. At the same time, the entities or classes, the descriptions of which form these groups, are considered as sets. The items of these sets are individual attributes (and, in the case of classes, also methods). Therefore, it is suggested to introduce additional assumptions:

a) the descriptions of each specific entity or class do not change depending on their inclusion in the descriptions of the function, input or output flows of different CIs;

b) each entity or class description can be matched with an identifier.

These assumptions make it possible to represent groups of descriptions of any functions, input and output flows of CIs as a set, the items of which are binary variables. These variables take the value 1 if the description of the entity or class with the corresponding identifier is present in the description of the function, input or output flow of the functional CI, and 0 otherwise. The proposed representation allows to modify the descriptions of the Hamming distance determination to determine the distance between separate groups as follows:

$$d_{H_m}(i_{pl}, i_{ql}) = \sum_{k=1}^{n} i_{plk} \oplus i_{qlk}, \tag{1.4}$$

where i_{plk} – a binary variable describing the presence of the identifier of the k-th entity or class, which are included in the description of the function, input or output flow i_{pl} of CI i_p; i_{qlk} – a binary variable that describes the presence of the identifier of the k-th entity or class that is included in the description of the function, input or output flow i_{ql} of CI i_q; n – the maximum number of identifiers that participate in the descriptions of the compared functions, input or output flows of CI i_p and i_q; $\oplus$ – operation "sum modulo 2".

The Hamming distance modified in this way will be equal to 0 in the case when the compared descriptions of functions, input or output flows consist of sets of identical entities or classes. In other cases, this distance will be equal to the number of entities or classes that will be present in only one of the two compared descriptions.

The use of the modified Hamming distance allows, in turn, to adapt the Chebyshev distance (3) to solve the subproblem of decomposition of the description of the IS architecture into separate clusters of functional CIs as follows [16]:

$$d_{\infty}(i_p, i_q) = \max_{1 \le l \le m} \sum_{k=1}^{n} i_{plk} \oplus i_{qlk}. \tag{1.5}$$

Using the modified Chebyshev distance (1.5) allows to determine the proximity degree of compared functional CIs based on representations of these CIs in the form of visual models.

1.3 SOLVING THE PROBLEM OF IDENTIFYING CONFIGURATION ITEMS USING THE DECLARED CLUSTERING METHODS

1.3.1 DESCRIPTION OF THE INITIAL DATA OF THE IDENTIFICATION PROBLEM OF CONFIGURATION ITEMS

The starting data for solving the problem of CI identification of an IT product is the CI of the "Formation and maintenance of an individual plan of a scientific and pedagogical employee of the department" functional task. This task was implemented as a separate IT product for the development of the capabilities of the "University" information and analytical system of the Kharkiv

National University of Radio Electronics. Previously, this system implemented the "Distribution of educational load between teachers of the department" functional task, the main source document of which is one of the sections of the "Individual plan of a scientific and pedagogical employee of the department" document.

A detailed description of the names and designations of the functions and data flows forming separate CIs is given in the **Table 1.1** [16, 17]. The abbreviations adopted in **Table 1.1**: IP – individual plan; KPI – key performance indicators. As numerical numbers in the **Table 1.1** shows the numbers of works, input and output flows that were generated by the AllFusion Process Modeler CASE tool during the creation of the data flow diagram.

● **Table 1.1** Description of the configuration items of the "Formation and management of the individual plan of the scientific and pedagogical employee of the department" functional task (based on the data flow diagram)

Work		**Input flow**		**Output flow**	
No.	**Name**	**No.**	**Name**	**No.**	**Name**
1	**2**	**3**	**4**	**5**	**6**
CI1	Conversion of the "Educational work" section	1	The teacher's educational load for the academic year	2	Information from the IP section "Educational work"
CI2	Formation of the "Scientific work" section	2	Information about the teacher	3	Information from the IP section "Scientific work"
		3	Information about planned work		
		5	Information about recommended work		
		8	Information from the IP section "Scientific work"		
		12	Remaining hours		
CI3	Formation of the "Methodical work" section	2	Information about the teacher	4	Information from the IP section "Methodical work"
		3	Information about planned work		
		5	Information about recommended work		
		9	Information from the IP section "Methodical work"		
		12	Remaining hours		
CI4	Formation of the "Organizational Work" section	2	Information about the teacher	5	Information from the IP section "Organizational work"
		3	Information about planned work		
		5	Information about recommended work		
		10	Information from the IP section "Organizational work"		
		12	Remaining hours		

Continuation of Table 1.1

1	2	3	4	5	6
CI5	Formation of the first list of positions and long-term assignments	4 11	Information about positions and long-term assignments Information from the IP section "List of positions and long-term assignments"	6	Information from the IP section "List of positions and long-term assignments"
CI6	Formation of the list of recommended works	5	Information about recommended work	1	Information about recommended work
CI7	Formation and maintenance of normative and reference information about KPI	6	Information about the department's key KPIs	7	Information about the department's key KPIs
CI8	Formation of the teacher's KPI and part of the department's KPI	8	Information from the IP section "Scientific work"	9	Information about the teacher's KPI and parts of the department's KPI
CI9	Forming a summary table for the academic year	9 7 8 10	Information from the IP section "Methodical work" Information from the IP section "Educational work" Information from the IP section "Scientific work" Information from the IP section "Organizational work"	8	Information about the number of hours by IP sections
CI10	Formation of the source document "IP"	9 7 8 10 11	Information from the IP section "Methodical work" Information from the IP section "Educational work" Information from the IP section "Scientific work" Information from the IP section "Organizational work" Information from the IP section "List of positions and long-term assignments"	10	IP

The entity description of the "Formation and maintenance of an individual plan of a scientific and pedagogical employee of the department" functional task was performed in the form of an ER diagram, which is shown in **Fig. 1.1** [16]. The description of entity names and their numerical identifiers is given in the **Table 1.2** [16, 17]. As numerical identifiers in the **Table 1.2** indicate the entity numbers that were generated using the AllFusion ERwin Data Modeler CASE tool during the development of the one shown in **Fig. 1.1** ER diagrams and imported into the AllFusion Process Modeler CASE tool to link with the functional task data flow diagram.

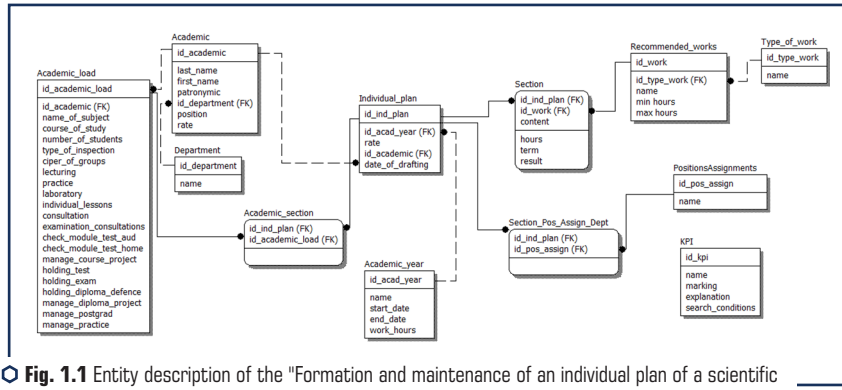


Fig. 1.1 Entity description of the "Formation and maintenance of an individual plan of a scientific and pedagogical employee of the department" functional task in the form of an ER diagram

Table 1.2 Set of descriptions of the entities of the "Formation and maintenance of an individual plan of a scientific and pedagogical employee of the department" functional task

Numerical identifier	Name
1	Academic_load
2	Academic
3	Department
4	Individual_plan
5	Academic_section
6	Academic_year
7	Section
8	Recommended_works
9	Type_of_work
10	Section_Pos_Assign_Dept
11	PositionsAssignments
12	KPI

Supplementing the data flow diagram of the "Formation and maintenance of an individual plan of a scientific and pedagogical employee of the department" functional task with information from ER diagrams allows to fix the corresponding subsets of entities with ER diagrams for each work, input and output flows of the data flow diagram. Therefore, it is proposed to consider each specific work of a data flow diagram with input and output data flows that are associated with this work as functional CIs [16, 17]. The received descriptions of the "Formation and management of the

individual plan of the scientific and pedagogical worker of the department" CI functional task are shown in the **Table 1.3** [17]. Items with the value "1" indicate the facts of the use of the entity with the identifier, which is given in the column of the **Table 1.3**, to describe the operation, input or output data flow with the identifier given in the row of the **Table 1.3**. Items with a value of "0" indicate the opposite facts.

● **Table 1.3** Descriptions of the configuration items of the "Formation and maintenance of an individual plan of a scientific and pedagogical employee of the department" functional task

Description of the CI function												
CI	**Entities IDs**											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
CI1	1	1	1	1	1	1	0	0	0	0	0	0
CI2	1	1	1	1	1	1	1	1	1	0	0	0
CI3	1	1	1	1	1	1	1	1	1	0	0	0
CI4	1	1	1	1	1	1	1	1	1	0	0	0
CI5	0	1	0	1	0	1	1	0	0	1	1	0
CI6	0	0	0	0	0	0	0	1	1	0	0	0
CI7	0	0	0	0	0	0	0	0	0	0	0	1
CI8	1	1	0	1	1	1	1	1	1	0	0	1
CI9	1	1	0	1	1	1	1	1	1	0	0	0
CI10	1	1	0	1	1	1	1	1	1	1	1	0
Description of CI input data flows												
CI	**Entities IDs**											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
CI1	1	1	1	0	0	0	0	0	0	0	0	0
CI2	1	1	1	1	1	1	1	1	1	0	0	0
CI3	1	1	1	1	1	1	1	1	1	0	0	0
CI4	1	1	1	1	1	1	1	1	1	0	0	0
CI5	0	1	0	1	0	1	1	0	0	1	1	0
CI6	0	0	0	0	0	0	0	1	1	0	0	0
CI7	0	0	0	0	0	0	0	0	0	0	0	1
CI8	1	1	0	1	0	1	1	1	1	0	0	0
CI9	1	1	0	1	1	1	1	1	1	0	0	0
CI10	1	1	0	1	1	1	1	1	1	1	1	0

● **Continuation of Table 1.3**

Description of CI output data flows												
CI	**Entities IDs**											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
CI1	1	1	0	1	1	1	0	0	0	0	0	0
CI2	0	1	0	1	0	1	1	1	1	0	0	0
CI3	0	1	0	1	0	1	1	1	1	0	0	0
CI4	0	1	0	1	0	1	1	1	1	0	0	0
CI5	0	1	0	1	0	1	1	0	0	1	1	0
CI6	0	0	0	0	0	0	0	1	1	0	0	0
CI7	0	0	0	0	0	0	0	0	0	0	0	1
CI8	1	1	0	1	1	1	1	0	0	0	0	1
CI9	1	1	0	1	1	1	1	1	0	0	0	0
CI10	1	1	0	1	1	1	1	1	1	1	1	0

1.3.2 DESCRIPTION OF SOLVING THE PROBLEM OF IDENTIFYING CONFIGURATION ITEMS USING THE AGNES AGGLOMERATIVE CLUSTERING METHOD

As a result of the implementation of Stage 1, 10 clusters were formed, each of which contained one functional CI of the researched task. The distance matrix *D* was calculated for these clusters, which is shown in the **Table 1.4** [17]. Distances were calculated according to formula (1.5).

● **Table 1.4** Initial distance matrix *D* (AGNES method, nearest neighbor algorithm)

Clusters	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
C1	0	6	6	6	7	8	7	6	7	9
C2	6	0	0	0	7	7	10	4	3	4
C3	6	0	0	0	7	7	10	4	3	4
C4	6	0	0	0	7	7	10	4	3	4
C5	7	7	7	7	0	8	7	7	6	4
C6	8	7	7	7	8	0	3	7	7	9
C7	7	10	10	10	7	3	0	8	9	11
C8	6	4	4	4	7	7	8	0	2	4
C9	7	3	3	3	6	7	9	2	0	3
C10	9	4	4	4	4	9	11	4	3	0

As a result of the first iteration of Stage 2, a pair of clusters C2 and C3, closest to each other, was selected. The distance between these clusters is equal to 0. The choice was made when viewing the matrix of distances (**Table 1.4**) from left to right and from top to bottom. A new cluster C11 = {C2, C3} was formed [17].

During the first iteration of Stage 3, the distance matrix *D* was recalculated taking into account the presence of the new C11 cluster. The result of the calculation is shown in **Table 1.5** [17].

● **Table 1.5** Distance matrix *D* with cluster C11

Clusters	C1	C11	C4	C5	C6	C7	C8	C9	C10
C1	0	6	6	7	8	7	6	7	9
C11	6	0	0	7	7	10	4	3	4
C4	6	0	0	7	7	10	4	3	4
C5	7	7	7	0	8	7	7	6	4
C6	8	7	7	8	0	3	7	7	9
C7	7	10	10	7	3	0	8	9	11
C8	6	4	4	7	7	8	0	2	4
C9	7	3	3	6	7	9	2	0	3
C10	9	4	4	4	9	11	4	3	0

During the second iteration of Stage 2, a pair of clusters C11 and C4, which are closest to each other, were selected. The distance between these clusters is equal to 0. A new cluster C12 = {C11, C4} was formed.

During the first iteration of Stage 3, the distance matrix *D* was recalculated taking into account the presence of the new cluster C12. The result of the calculation is shown in **Table 1.6** [17].

● **Table 1.6** Distance matrix *D* with cluster C12

Clusters	C1	C12	C5	C6	C7	C8	C9	C10
C1	0	6	7	8	7	6	7	9
C12	6	0	7	7	10	4	3	4
C5	7	7	0	8	7	7	6	4
C6	8	7	8	0	3	7	7	9
C7	7	10	7	3	0	8	9	11
C8	6	4	7	7	8	0	2	4
C9	7	3	6	7	9	2	0	3
C10	9	4	4	9	11	4	3	0

During the third iteration of Stage 2, a pair of clusters C8 and C9, closest to each other, was selected. The distance between these clusters is equal to 2. A new cluster C13 = {C8, C9} was formed.

During the third iteration of Stage 3, the distance matrix *D* was recalculated taking into account the presence of the new C13 cluster. The result of the calculation is shown in **Table 1.7** [17].

● **Table 1.7** Distance matrix *D* with cluster C13

Clusters	C1	C12	C5	C6	C7	C13	C10
C1	0	6	7	8	7	6	9
C12	6	0	7	7	10	3	4
C5	7	7	0	8	7	6	4
C6	8	7	8	0	3	7	9
C7	7	10	7	3	0	8	11
C13	6	3	6	7	8	0	3
C10	9	4	4	9	11	3	0

During the fourth iteration of Stage 2, a pair of clusters C12 and C13 closest to each other was selected. The distance between these clusters is 3. A new cluster C14 = {C12, C13} was formed.

During the fourth iteration of Stage 3, the distance matrix *D* was recalculated, taking into account the presence of the new C14 cluster. The result of the calculation is shown in **Table 1.8** [17].

● **Table 1.8** Distance matrix *D* with cluster C14

Clusters	C1	C14	C5	C6	C7	C10
C1	0	6	7	8	7	9
C14	6	0	6	7	8	3
C5	7	6	0	8	7	4
C6	8	7	8	0	3	9
C7	7	8	7	3	0	11
C10	9	3	4	9	11	0

During the fifth iteration of Stage 2, a pair of clusters C14 and C10 closest to each other was selected. The distance between these clusters is 3. A new cluster C15 = {C14, C10} was formed.

During the fifth iteration of Stage 3, the distance matrix *D* was recalculated taking into account the presence of the new C15 cluster. The result of the calculation is shown in **Table 1.9** [17].

● **Table 1.9** Distance matrix *D* with cluster C15

Clusters	C1	C15	C5	C6	C7
C1	0	6	7	8	7
C15	6	0	4	7	8
C5	7	4	0	8	7
C6	8	7	8	0	3
C7	7	8	7	3	0

During the execution of the sixth iteration of Stage 2, a pair of clusters C6 and C7, which are closest to each other, were selected. The distance between these clusters is 3. A new cluster C16 = {C6, C7} was formed.

During the execution of the sixth iteration of Step 3, the distance matrix *D* was recalculated taking into account the presence of the new C16 cluster. The result of the calculation is shown in **Table 1.10** [17].

● **Table 1.10** Distance matrix *D* with cluster C16

Clusters	C1	C15	C5	C16
C1	0	6	7	7
C15	6	0	4	7
C5	7	4	0	7
C16	7	7	7	0

During the execution of the seventh iteration of Stage 2, a pair of clusters C15 and C5, closest to each other, was selected. The distance between these clusters is 4. A new cluster C17 = {C15, C5} was formed.

During the execution of the seventh iteration of Stage 3, the distance matrix *D* was recalculated, taking into account the presence of the new C17 cluster. The result of the calculation is shown in **Table 1.11** [17].

● **Table 1.11** Distance matrix *D* with cluster C17

Clusters	C1	C17	C16
C1	0	7	7
C17	7	0	7
C16	7	7	0

During the execution of the eighth iteration of Stage 2, a pair of clusters C1 and C17, closest to each other, was selected. The distance between these clusters is 7. A new cluster C18 = {C1, C17} was formed.

During the eighth iteration of Stage 3, the distance matrix *D* was recalculated taking into account the presence of a new cluster C18. The result of the calculation is shown in **Table 1.12** [17].

Table 1.12 Distance matrix *D* with cluster C18

Clusters	C18	C16
C18	0	7
C16	7	0

During the ninth iteration of Stage 2, cluster C19 was formed on the basis of clusters C18 and C16, which includes all the original clusters. This completes the execution of the nearest neighbor algorithm.

The result of applying the AGNES agglomerative clustering method (using the nearest neighbor algorithm) is a dendrogram that looks like the one shown in **Fig. 1.2** [17].

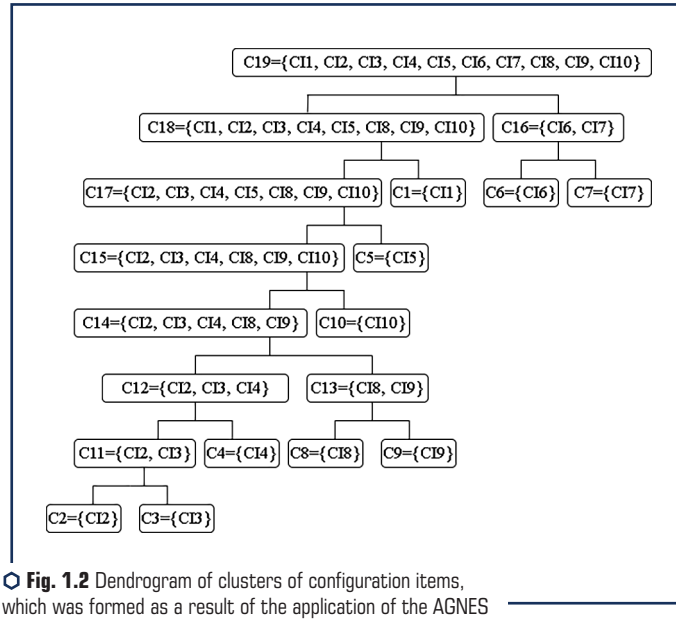


Fig. 1.2 Dendrogram of clusters of configuration items, which was formed as a result of the application of the AGNES agglomerative clustering method

1.3.3 DESCRIPTION OF THE SOLUTION TO THE PROBLEM OF IDENTIFYING CONFIGURATION ITEMS USING THE DIANA DIVISIVE CLUSTERING METHOD

Let's consider the features of using the DIANA divisive clustering method on the example of using this method to decompose the initial cluster C1, which includes all ten CIs defined in **Table 1.1** and **Table 1.3**. During Stage 1 and Stage 2, the value of the modified Chebyshev distance (1.5) was calculated for each pair of items of the initial cluster C1. These values are given in **Table 1.13** [16].

Table 1.13 Values of modified Chebyshev distances for each pair of items of the initial cluster C1

CI	CI1	CI2	CI3	CI4	CI5	CI6	CI7	CI8	CI9	CI10
CI1	0	6	6	6	7	8	7	6	7	9
CI2	6	0	0	0	7	7	10	4	3	4
CI3	6	0	0	0	7	7	10	4	3	4
CI4	6	0	0	0	7	7	10	4	3	4
CI5	7	7	7	7	0	8	7	7	6	4
CI6	8	7	7	7	8	0	3	7	7	9
CI7	7	10	10	10	7	3	0	8	9	11
CI8	6	4	4	4	7	7	8	0	2	4
CI9	7	3	3	3	6	7	9	2	0	3
CI10	9	4	4	4	4	9	11	4	3	0

During Stage 2, the value of the average modified Chebyshev distance was determined for each item of the initial cluster C1. These values are given in **Table 1.14** [16].

Table 1.14 Values of average modified Chebyshev distances for each item of cluster C1

CI	CI1	CI2	CI3	CI4	CI5	CI6	CI7	CI8	CI9	CI10
Average distance	6.2	4.1	4.1	4.1	6	6.3	7.5	4.6	4.3	5.2

As a result of Stage 2, item CI7 was selected. As a result of Stage 3, this item was excluded from the original cluster C1 and included in the formed child cluster C2.

During Step 4, new mean modified Chebyshev distances were determined for each item remaining to be considered in cluster C1, as well as the difference between these distances. The performance results are shown in **Table 1.15** [16].

As a result of Stage 5, the item CI6 was transferred to the daughter cluster C2. After that, the calculations of Stage 4 were repeated for items CI1, CI2, CI3, CI4, CI5, CI8, CI9 and C10, which remained in cluster C1. The results of these calculations are shown in the **Table 1.16** [16].

● **Table 1.15** Values of the difference between the average Chebyshev distances for each item of the cluster C1

Chebyshev distances										Average by items C1	Averages by items C2	Difference
CI	CI1	CI2	CI3	CI4	CI5	CI6	CI8	CI9	CI10			
CI1	0	6	6	6	7	8	6	7	9	6.11	7	–0.89
CI2	6	0	0	0	7	7	4	3	4	3.44	10	–6.56
CI3	6	0	0	0	7	7	4	3	4	3.44	10	–6.56
CI4	6	0	0	0	7	7	4	3	4	3.44	10	–6.56
CI5	7	7	7	7	0	8	7	6	4	5.89	7	–1.11
CI6	8	7	7	7	8	0	7	7	9	6.67	3	3.67
CI8	6	4	4	4	7	7	0	2	4	4.22	8	–3.78
CI9	7	3	3	3	6	7	2	0	3	3.78	9	–5.22
CI10	9	4	4	4	4	9	4	3	0	4.56	11	–6.44

● **Table 1.16** Values of the difference of the average Chebyshev distances for each item that remained in the cluster C1

Chebyshev distances									Average by items C1	Averages by items C2	Difference
CI	CI1	CI2	CI3	CI4	CI5	CI8	CI9	CI10			
CI1	0	6	6	6	7	6	7	9	5.875	7.5	–1.625
CI2	6	0	0	0	7	4	3	4	3	8.5	–5.5
CI3	6	0	0	0	7	4	3	4	3	8.5	–5.5
CI4	6	0	0	0	7	4	3	4	3	8.5	–5.5
CI5	7	7	7	7	0	7	6	4	5.625	7.5	–1.875
CI8	6	4	4	4	7	0	2	4	3.875	7.5	–3.625
CI9	7	3	3	3	6	2	0	3	3.375	8	–4.625
CI10	9	4	4	4	4	4	3	0	4	10	–6

Since all the differences in the **Table 1.8** are negative, then a new child class C3 was formed. This class includes items CI1, CI2, CI3, CI4, CI5, CI8, CI9 and CI10.

As a result of performing Step 6 of the method, it was established that the stopping conditions for clusters C2 and C3 are not fulfilled. Each of these clusters is not a singleton, and the distances between the items of each of these clusters are not equal to 0. This concludes the first iteration of the DIANA method. For the second iteration of the method, the proposed cluster is C2.

As a result of iterative execution of the DIANA method for the given problem, a dendrogram of clusters was formed, which is shown in **Fig. 1.3** [16].

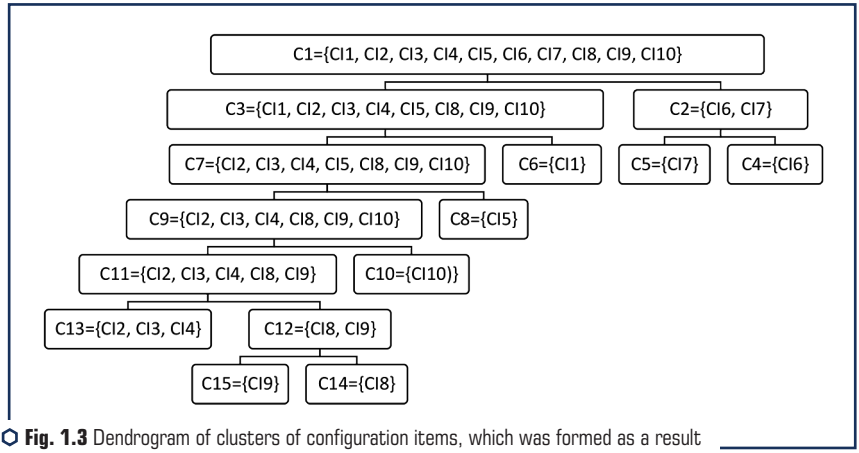


Fig. 1.3 Dendrogram of clusters of configuration items, which was formed as a result of applying the DIANA divisive clustering method

1.3.4 DESCRIPTION OF SOLVING THE PROBLEM OF IDENTIFYING CONFIGURATION ITEMS USING THE K-MEANS CLUSTERING ALGORITHM

Before proceeding with the task of analyzing the configuration of an IT product using the k-means algorithm, let's assume that the number of teams performing this IT project is three. Therefore, as centers for the initial division of the description of the CI of the functional problem considered in 1.3.1 into clusters, vectors describing the items of CI1, CI3 and CI7 are proposed. The choice of CI data is driven by the following considerations:

a) CI1 describes the function of forming the "Educational Work" section, which is an important initial item for drawing up the final individual plan of the employee;

b) CI3 describes the function "Formation of the "Methodical work" section ", which is an example of a number of typical functions for the formation of separate sections of the individual plan of the employee;

c) CI7 describes the "Formation and maintenance of normative reference information about KPI" function, which is the most isolated from the description of the final individual plan of the employee.

The value of the accuracy coefficient δ in the k-means algorithm is 0.1.

In the course of preliminary calculations, a matrix of values of modified Chebyshev distances (1.5) between individual CIs of the functional problem is formed. This matrix is shown in **Table 1.17**.

Based on the data in the **Table 1.17**, the matrix of CI functional task membership to the clusters of the initial partition (partition matrix) is calculated. This matrix is shown in **Table 1.18**.

Since the distances between CI5 and the centers of the initial partition were the same, the membership of CI5 to the cluster with the center CI1 was determined by the first comparison from left to right during the analysis of the distance matrix (**Table 1.17**).

Table 1.17 Matrix of values of modified Chebyshev distances between individual configurational items of the functional problem (k-means algorithm)

Configuration items	CI1	CI2	CI3	CI4	CI5	CI6	CI7	CI8	CI9	CI10
CI1	0	6	6	6	7	8	7	6	7	9
CI2	6	0	0	0	7	7	10	4	3	4
CI3	6	0	0	0	7	7	10	4	3	4
CI4	6	0	0	0	7	7	10	4	3	4
CI5	7	7	7	7	0	8	7	7	6	4
CI6	8	7	7	7	8	0	3	7	7	9
CI7	7	10	10	10	7	3	0	8	9	11
CI8	6	4	4	4	7	7	8	0	2	4
CI9	7	3	3	3	6	7	9	2	0	3
CI10	9	4	4	4	4	9	11	4	3	0

Table 1.18 Partition matrix at the zeroth iteration of the k-means algorithm for a functional problem

Configuration items	CI1	CI3	CI7
CI1	1	0	0
CI2	0	1	0
CI3	0	1	0
CI4	0	1	0
CI5	1	0	0
CI6	0	0	1
CI7	0	0	1
CI8	0	1	0
CI9	0	1	0
CI10	0	1	0

Next, cluster centers were determined on the first iteration of the k-means algorithm. The calculation was carried out according to the formula [12–14]:

$$c_l^{(i)} = \sum_{j=1}^{d} u_{ij}^{(l-1)} \times m_j \Big/ \sum_{j=1}^{d} u_{ij}^{(l-1)}, 1 \le i \le c, \tag{1.6}$$

where $c_l^{(i)}$ – the designation of the center of the i-th cluster on the l-th iteration of the algorithm; $u_{ij}^{(l-1)}$ – designation of the item of the partition n matrix for the j-th CI and the i-th cluster at the (l–1)-th iteration of the algorithm; m_j – designation of the vector describing the j-th CI; c – the number of clusters selected on the (l–1)-th iteration of the algorithm. For this problem $c=3$.

During the calculations, new cluster centers were discovered, which were described as conventional items CI11, CI12 and CI13. The vectors of their descriptions are shown in **Table 1.19**.

● **Table 1.19** Vector descriptions of items of conditional configuration CI11, CI12 and CI13

Description of CI function												
CI	**Entities IDs**											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
CI11	0.5	1	0.5	1	0.5	1	0.5	0	0	0.5	0.5	0
CI12	1	1	0.5	1	1	1	1	1	1	0.17	0.17	0.17
CI13	0	0	0	0	0	0	0	0.5	0.5	0	0	0.5
Description of CI input data flows												
CI	**Entities IDs**											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
CI11	0.5	1	0.5	0.5	0	0.5	0.5	0	0	0.5	0.5	0
CI12	1	1	0.5	1	0.83	1	1	1	1	0.17	0.17	0
CI13	0	0	0	0	0	0	0	0.5	0.5	0	0	0.5
Description of CI output data flows												
CI	**Entities IDs**											
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
CI11	0.5	1	0	1	0.5	1	0.5	0	0	0.5	0.5	0
CI12	0.5	1	0	1	0.5	1	1	0.83	0.67	0.17	0.17	0.17
CI13	0	0	0	0	0	0	0	0.5	0.5	0	0	0.5

Then, during the first iteration of the k-means algorithm, the distance matrix was recalculated according to formula (1.5) taking into account the newly introduced cluster centers in the form of conditional items CI11, CI12 and CI13. To perform the modulo 2 summation operation, the fractional values of the coordinates of the cluster center vectors were rounded to the nearest whole number (0 or 1). The results of calculating the distances to the new cluster centers are shown in the **Table 1.20**.

Based on the values obtained from the **Table 1.18**, the partition matrix was updated. The result of the update is shown in **Table 1.21**.

To check the stopping condition of the k-means algorithm, the partition matrix at the first iteration (**Table 1.21**) was subtracted from the partition matrix at the zero iteration (**Table 1.18**) of the k-means algorithm. The result is a zero partition matrix. This means that the stopping condition of the k-means algorithm is fulfilled.

● **Table 1.20** Matrix of modified Chebyshev distances between configuration items and new centers of clusters CI11, CI12 and CI13 at the first iteration of the algorithm

Configuration items	CI11	CI12	CI13
CI1	5	6	9
CI2	6	2	8
CI3	6	2	8
CI4	6	2	8
CI5	3	7	9
CI6	11	7	1
CI7	10	10	2
CI8	6	3	8
CI9	6	1	8
CI10	4	3	9

● **Table 1.21** Partition matrix on the first iteration of the k-means algorithm for a functional problem

Configuration items	CI11	CI12	CI13
CI1	1	0	0
CI2	0	1	0
CI3	0	1	0
CI4	0	1	0
CI5	1	0	0
CI6	0	0	1
CI7	0	0	1
CI8	0	1	0
CI9	0	1	0
CI10	0	1	0

Thus, the solution to the problem of clustering using the k-means algorithm for CI of the considered functional problem will be three clusters consisting of the following items:

a) cluster 1 (C1), the items of which are CI1 and CI5 with the center in the conditional item CI11;

b) cluster 2 (C2), the items of which are CI2, CI3, CI4, CI8, CI9 and CI10 with the center in the conditional item CI12;

c) cluster 3 (C3), the items of which are CI6 and CI7 with the center in the conventional item CI13.

1.3.5 DESCRIPTION OF SOLVING THE PROBLEM OF IDENTIFYING CONFIGURATION ITEMS USING THE GRAPHOANALYTIC METHOD

Let's consider the solution of the problem of CI identification according to the conditions outlined in 1.3.1, by the method outlined in [10]. According to this method, it is proposed to transform the description of the studied system into a set of the following descriptions:

– description of data structures defined state variables;

– descriptions of operations performed on state variables.

At the same time, interactions between operations and state variables are divided into two main groups [10]:

– read operation of the state variable;

– recording of the value in the state variable by the operation.

Then in [10] it is proposed to transform the set of these descriptions into an undirected graph. The vertices of this graph represent state variables and operations, and the arcs represent interactions between operations and state variables. At the same time, each arc has one of the following weights:

– weight equal to 1, if reading is performed by the state variable operation;

– weight equal to 2, if the value operation is written to the state variable.

System decomposition is performed on the formed graph by selecting individual clusters. Each such cluster is a part of the graph that is connected to other parts by the minimum possible number of arcs with minimum weights. The rationale for this method of cluster selection is detailed in [10].

It should be noted that the method described in [10] is focused on the use of UML diagrams that describe the software system being created. In our case, the description of the functional problem is a DFD. Therefore, as a description of operations, let's use descriptions of works from the **Table 1.1**. As a description of the state variables, let's use the descriptions of the input and output flows from the **Table 1.1**. At the same time, let's accept the following assumption: the output flows of any work from the **Table 1.1** are the input streams of other works from the **Table 1.1** if the names of these streams match. The results of selection of operations and state variables of the "Formation and maintenance of an individual plan of a scientific and pedagogical employee of the department" functional task are shown in the **Table 1.22** [16]. The description of interactions between operations and state variables is given in the **Table 1.23** [16].

The result of graph construction according to the method described in [10] is shown in **Fig. 1.4** [16]. Its weight is indicated on each edge. At the same time, it is considered that the weight of the edge that describes both reading and writing is equal to 3 ($(r=1)+(w=2)=3$).

Most of the vertices of the constructed graph are strongly connected. This means that the studied description of the architecture of the functional task is highly monolithic. Therefore, during the selection of clusters on the constructed graph, the following will be formed:

– relatively small clusters that include one operation;

– one large cluster, which includes a large number of operations that will have to be implemented as a monolithic software module.

● **Table 1.22** Description of operations and state variables of the functional task (based on the data flow diagram)

Designation	Name
Operations	
O1	Conversion of the "Educational work" section
O2	Formation of the "Scientific work" section
O3	Formation of the "Methodical work" section
O4	Formation of the "Organizational work" section
O5	Formation of the first list of positions and long-term assignments
O6	Formation of the list of recommended works
O7	Formation and maintenance of normative and reference information about KPI
O8	Formation of the teacher's KPI and part of the department's KPI
O9	Formation of a summary table for the academic year
O10	Formation of the source document "IP"
State variables	
V1	Teaching load for the academic year
V2	Information about the teacher
V3	Information about planned work
V4	Information about positions and long-term assignments
V5	Information about recommended work
V6	Information about the department's key KPIs
V7	Information from the "Educational work" IP section
V8	Information from the "Scientific work" IP section
V9	Information from the "Methodical work" IP section
V10	Information from the "Organizational work" IP section
V11	Information from the "List of positions and long-term assignments" IP section
V12	Remaining hours
V13	Information about the number of hours by IP sections
V14	Information about the teacher's KPI and parts of the department's KPI
V15	IP

Table 1.23 Description of interactions between operations and state variables

Operations	State variables														
	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15
01	*r*	–	–	–	–	–	*w*	–	–	–	–	–	–	–	–
02	–	*r*	*r*	–	*r*	–	–	*r*, *w*	–	–	–	*r*	–	–	–
03	–	*r*	*r*	–	*r*	–	–	–	*r*, *w*	–	–	*r*	–	–	–
04	–	*r*	*r*	–	*r*	–	–	–	–	*r*, *w*	–	*r*	–	–	–
05	–	–	–	*r*		–	–	–	–	–	*r*, *w*	–	–	–	–
06	–	–	–	–	*r*, *w*	–	–	–	–	–	–	–	–	–	–
07	–	–	–	–	–	*r*, *w*	–	–	–	–	–	–	–	–	–
08	–	–	–	–	–	–	–	*r*	–	–	–	–	–	*w*	–
09	–	–	–	–	–	–	*r*	*r*	*r*	*r*	–	–	*w*	–	–
010	–	–	–	–	–	–	*r*	*r*	*r*	*r*	*r*	–	–	–	*w*

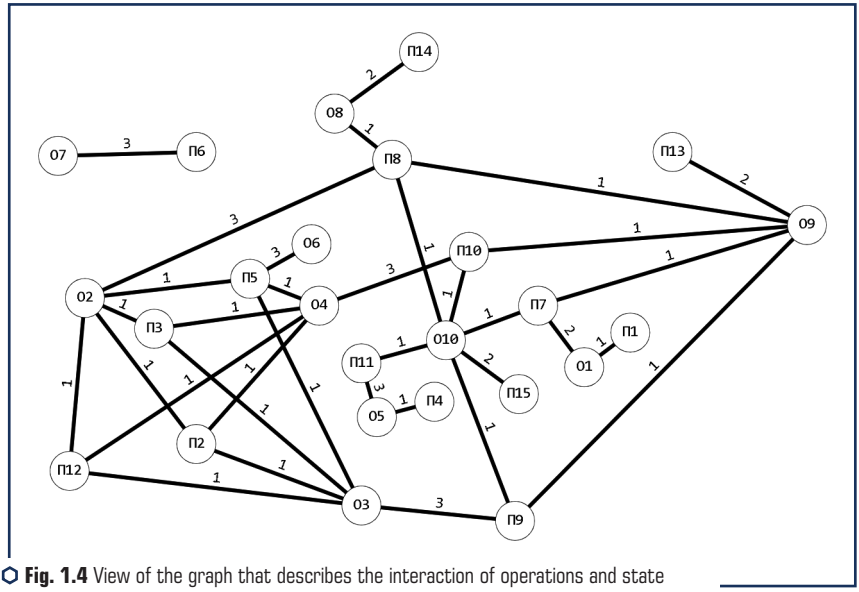


Fig. 1.4 View of the graph that describes the interaction of operations and state variables of the functional task

Solutions for selecting clusters on the constructed graph (**Fig. 1.4**) are given in **Table 1.24** [16].

The method described in [10] requires minimizing the number of intersecting arcs. However, the fulfillment of this condition (see solution 1 in **Table 1.24**) leads to issuing the task of creating

a large monolithic software module (cluster C2). Such a decision is inefficient for the further division of work between IT project implementation teams.

● **Table 1.24** Description of the results of solving the problem by the method described in [10]

Cluster designation	Designation of the vertices included in the cluster	Designation of functional task CIs that correspond to operations
1	**2**	**3**
Solution 1 (the number of cut arcs of minimum weight is 0)		
C1	O7, V6	CI7
C2	O1, O2, O3, O4, O5, O6, O8, O9, O10, V1, V2, V3, V4, V5, V7, V8, V9, V10, V11, V12, V13, V14, V15	CI1, CI2, CI3, CI4, CI5, CI6, CI8, CI9, CI10
Solution 2 (the number of cut arcs of minimum weight is 1)		
C1	O7, V6	CI7
C2	O8, V14	CI8
C3	O1, O2, O3, O4, O5, O6, O9, O10, V1, V2, V3, V4, V5, V7, V8, V9, V10, V11, V12, V13, V15	CI1, CI2, CI3, CI4, CI5, CI6, CI9, CI10
Solution 3 (the number of cut arcs of minimum weight is 1)		
C1	O7, V6	CI7
C2	O5, V4, V11	CI5
C3	O1, O2, O3, O4, O6, O8, O9, O10, V1, V2, V3, V5, V7, V8, V9, V10, V12, V13, V14, V15	CI1, CI2, CI3, CI4, CI6, CI8, CI9, CI10
Solution 4 (the number of cut arcs of minimum weight is 2)		
C1	O7, V6	CI7
C2	O8, V14	CI8
C3	O5, V4, V11	CI5
C4	O1, O2, O3, O4, O6, O9, O10, V1, V2, V3, V5, V7, V8, V9, V10, V12, V13, V15	CI1, CI2, CI3, CI4, CI6, CI9, CI10
Solution 5 (the number of cut arcs of minimum weight is 2)		
C1	O7, V6	CI7
C2	O1, V1, V7	CI1
C3	O2, O3, O4, O5, O6, O8, O9, O10, V2, V3, V4, V5, V8, V9, V10, V11, V12, V13, V14, V15	CI2, CI3, CI4, CI5, CI6, CI8, CI9, CI10
Solution 6 (the number of cut arcs of minimum weight is 3)		
C1	O7, V6	CI7
C2	O8, V14	CI8
C3	O1, V1, V7	CI1
C4	O2, O3, O4, O5, O6, O9, O10, V2, V3, V4, V5, V8, V9, V10, V11, V12, V13, V15	CI2, CI3, CI4, CI5, CI6, CI9, CI10

● **Continuation of Table 1.24**

1	2	3
Solution 7 (the number of cut arcs of minimum weight is 3)		
C1	O7, V6	CI7
C2	O5, V4, V11	CI5
C3	O1, V1, V7	CI1
C4	O2, O3, O4, O6, O8, O9, O10, V2, V3, V5, V8, V9, V10, V12, V13, V14, V15	CI2, CI3, CI4, CI6, CI8, CI9, CI10
Solution 8 (the number of cut arcs of minimum weight is 4)		
C1	O7, V6	CI7
C2	O8, V14	CI8
C3	O5, V4, V11	CI5
C4	O1, V1, V7	CI1
C5	O2, O3, O4, O6, O9, O10, V2, V3, V5, V8, V9, V10, V12, V13, V15	CI2, CI3, CI4, CI6, CI9, CI10
Solution 9 (the number of cut arcs of minimum weight is 4)		
C1	O7, V6	CI7
C2	O6, V5	CI6
C3	O1, O2, O3, O4, O5, O8, O9, O10, V1, V2, V3, V4, V7, V8, V9, V10, V11, V12, V13, V14, V15	CI1, CI2, CI3, CI4, CI5, CI8, CI9, CI10

In [10], the stopping condition for solving the clustering problem on the graphical description of the system architecture is not specified. So, the **Table 1.24** shows those division options for which the number of intersecting graph arcs does not exceed 4.

1.4 COMPARATIVE ANALYSIS OF THE OBTAINED RESULTS

Shown in **Fig. 1.2, 1.3** dendrograms obtained as a result of solving the problem by the AGNES agglomerative clustering method and DIANA divisive clustering method, respectively, coincide almost completely. An exception is clusters C2, C3, C4, and C11, selected during the construction of the dendrogram shown in **Fig. 1.2**. The appearance of these clusters is due to the fact that the AGNES agglomerative clustering method (using the nearest neighbor algorithm) is unable to recognize CIs which descriptions completely match. This shortcoming is the reason why the first two iterations of Stage 2 of the AGNES method were aimed at merging clusters that contained CIs with completely identical descriptions. In practice, this shortcoming of the AGNES method can lead to IT project executors being allocated separate sprints to implement CIs with overlapping

descriptions (in this case, CI2, CI3 and CI4) during the planning of their work. This will lead to an unjustified overestimation of labor costs and time spent on the implementation of these CIs.

To eliminate the mentioned shortcoming, it is proposed to adjust the AGNES agglomerative clustering method by supplementing it with operations to adjust the set of initial clusters formed at Stage 1. As a result of the addition, the modified AGNES agglomerative clustering method will be represented by a sequence of the following stages [16]:

Stage 1. The entire set CI is represented as a set of clusters C, each of which contains one item CI_i, $i=1,\ldots,n$, where n is the number of items in the set CI. Calculate the distance matrix D between the items of the set C.

Step 2. Calculate the distance matrix D between the items of the set of clusters C and adjust the set C by combining in a new cluster each pair of clusters c_p and c_q for which $d_{pq}=0$, then exclude the clusters c_p and c_q from further consideration and do not display them on the final dendrogram.

Stage 3. Select two clusters c_p and c_q, the distance between which will be minimal, and combine them into a new cluster c_r, entering it instead of clusters c_p and c_q into the set of clusters C.

Step 4. Calculate the values of the distance matrix D using rule (1.1).

Step 5. Repeat Step 2, Step 3, and Step 4 until one cluster is formed that includes all items of the CI set.

This modification makes it possible to eliminate the above-mentioned shortcoming and does not lead to a serious increase in the computational complexity of the AGNES method.

It should be noted that the AGNES method considered in the study (using the nearest neighbor algorithm) is less complex in terms of computational complexity than the DIANA divisive classification method. In particular, in order to obtain a similar result, the AGNES method does not require quite complex calculations regarding the calculation of average distances and the search for items transferred from the parent cluster to the newly created child cluster. Therefore, it is more appropriate to use agglomerative clustering methods in order to solve the problem of CI identification of an IT product (provided that the above-mentioned shortcoming is eliminated).

Let's compare the clusters formed as a result of the application of the k-means algorithm with the general parts of the dendrograms shown in **Fig. 1.2, 1.3**. As a representation of such a common part, let's use the dendrogram shown in **Fig. 1.3**. According to the results of this comparison, it should be noted:

– cluster C3 obtained as a result of applying the k-means algorithm coincides with cluster C2 highlighted on the dendrogram;

–cluster C2, obtained as a result of applying the k-means algorithm, coincides with the C9 cluster selected on the dendrogram;

– cluster C1, obtained as a result of applying the k-means algorithm, has no direct analogues on the dendrogram.

The absence of a direct analogue for the C1 cluster allows to claim that this cluster is artificial. To verify this statement, let's analyze the distance of individual CIs from the conditional centers of CI11, CI12, and CI13 clusters determined in the first iteration. For this purpose, let's use the data from **Table 1.20**. The results of the analysis are given in **Table 1.25**.

● **Table 1.25** The distance of individual configuration items from the conditional centers CI11, CI12 and CI13 of the clusters determined in the first iteration of the k-means algorithm

Chebyshev modified distance to the center	A list of CIs that are at the appropriate distance to the center
Conditional center CI11	
3	CI5 ∈ C1
4	CI10 ∈ C2
5	CI1 ∈ C1
6	(CI2, CI3, CI4, CI8, CI9) ∈ C2
10	CI7 ∈ C3
11	CI6 ∈ C3
Conditional center in CI12	
1	CI9 ∈ C2
2	(CI2, CI3, CI4) ∈ C2
3	(CI8, CI10) ∈ C2
6	CI1 ∈ C1
7	CI5 ∈ C1, CI6 ∈ C3
10	CI7 ∈ C3
Conditional center in CI13	
1	CI6 ∈ C3
2	CI7 ∈ C3
8	(CI2, CI3, CI4, CI8, CI9) ∈ C2
9	(CI1, CI5) ∈ C1, CI10 ∈ C2

From the **Table 1.25**, it can be seen that the initial decision to select a cluster centered on CI1 was rather erroneous. This is evidenced by the fact that CI10, which is part of cluster C2, is closer to the conventional center of cluster C1 than CI1, which is part of this cluster.

It would be appropriate to assign cluster centers at the beginning of the k-means algorithm, based on the results of the analysis of the matrix of distances between CIs. The purpose of such an analysis is to select the desired combination of CIs that are as far apart as possible. However, such a combination of clusters, identified at the beginning of the k-means algorithm, does not allow obtaining a set of clusters balanced by the main characteristics of the IT project (indicators of labor intensity, time, cost, and quality of project work). At the same time, the dendrogram of clusters formed as a result of the application of the DIANA method makes it possible to select CI clusters that are maximally balanced in terms of such characteristics in the future. Therefore, the use of the k-means algorithm to solve the problem of CI identification in the IT project is less appropriate than hierarchical clustering methods.

A comparison of the results of solving the problem of CI identification in the IT project using the methods of hierarchical clustering AGNES and DIANA and the k-means algorithm with the results obtained using the graphoanalytic method is given in the **Table 1.26**.

● **Table 1.26** Comparison of the results of solving the task of identifying functional configuration items using the graph-analytic method, the DIANA method, and the k-means algorithm

Designation of the cluster by the graphoanalytical method	Designation of CIs of the functional problem, which are included in the cluster according to the graphoanalytic method	Designation of a similar cluster on the dendrogram (Fig. 1.3)	Designation of a similar cluster based on the results of the k-means algorithm
1	**2**	**3**	**4**
Solution 1 (the number of cut arcs of minimum weight is 0)			
C1	CI7	C5	There is no complete analogue, the closest is C3
C2	CI1, CI2, CI3, CI4, CI5, CI6, CI8, CI9, CI10	There is no complete analogue, the closest is C3	There is no complete analogue, the closest is C2
Solution 2 (the number of cut arcs of minimum weight is 1)			
C1	CI7	C5	There is no complete analogue, the closest is C3
C2	CI8	C14	There is no analogue
C3	CI1, CI2, CI3, CI4, CI5, CI6, CI9, CI10	There is no complete analogue, the closest is C3	There is no complete analogue, the closest is C2
Solution 3 (the number of cut arcs of minimum weight is 1)			
C1	CI7	C5	There is no complete analogue, the closest is C3
C2	CI5	C8	There is no complete analogue, the closest is C1
C3	CI1, CI2, CI3, CI4, CI6, CI8, CI9, CI10	There is no complete analogue, the closest is C3	There is no complete analogue, the closest is C2
Solution 4 (the number of cut arcs of minimum weight is 2)			
C1	CI7	C5	There is no complete analogue, the closest is C3
C2	CI8	C14	There is no analogue
C3	CI5	C8	There is no complete analogue, the closest is C3
C4	CI1, CI2, CI3, CI4, CI6, CI9, CI10	There is no complete analogue, the closest are C3, C7, C9	There is no complete analogue, the closest is C2
Solution 5 (the number of cut arcs of minimum weight is 2)			
C1	CI7	C5	There is no complete analogue, the closest is C3
C2	CI1	C6	There is no complete analogue, the closest is C1
C3	CI2, CI3, CI4, CI5, CI6, CI8, CI9, CI10	There is no complete analogue, the closest is C7	There is no complete analogue, the closest is C2

● **Continuation of Table 1.26**

1	2	3	4
Solution 6 (the number of cut arcs of minimum weight is 3)			
C1	CI7	C5	There is no complete analogue, the closest is C3
C2	CI8	C14	There is no analogue
C3	CI1	C6	There is no complete analogue, the closest is C1
C4	CI2, CI3, CI4, CI5, CI6, CI9, CI10	There is no complete analogue, the closest is C7	There is no complete analogue, the closest is C2
Solution 7 (the number of cut arcs of minimum weight is 3)			
C1	CI7	C5	There is no complete analogue, the closest is C3
C2	CI5	C8	There is no complete analogue, the closest is C1
C3	CI1	C6	There is no complete analogue, the closest is C1
C4	CI2, CI3, CI4, CI6, CI8, CI9, CI10	There is no complete analogue, the closest is C9	There is no complete analogue, the closest is C2
Solution 8 (the number of cut arcs of minimum weight is 4)			
C1	CI7	C5	There is no complete analogue, the closest is C3
C2	CI8	C14	There is no analogue
C3	CI5	C8	There is no complete analogue, the closest is C3
C4	CI1	C6	There is no complete analogue, the closest is C1
C5	CI2, CI3, CI4, CI6, CI9, CI10	There is no complete analogue, the closest is C9	There is no complete analogue, the closest is C2
Solution 9 (the number of cut arcs of minimum weight is 4)			
C1	CI7	C5	There is no complete analogue, the closest is C3
C2	CI6	C4	There is no complete analogue, the closest is C3
C3	CI1, CI2, CI3, CI4, CI5, CI8, CI9, CI10	C3	There is no complete analogue, the closest is C2

The following conclusions can be drawn based on the data from the **Table 1.26**:

a) when selecting clusters consisting of one CI, hierarchical clustering methods and the method described in [10] give the same results;

b) when selecting clusters consisting of several CIs, the method described in [10] is less accurate than hierarchical clustering methods;

c) with an increase in the number of connections between CIs, which must be cut when decomposing the graph into separate clusters, hierarchical clustering methods and the method outlined in [10] generally give similar results (solution 9 in **Table 1.26** completely coincides with clusters C3, C4 and C5, highlighted in **Fig. 1.3**);

d) the results obtained using the k-means method and algorithm described in [10] practically do not coincide with each other.

It should be taken into account that the increase in the number of connections between CIs, which must be cut when decomposing the graph into separate clusters, leads to a further increase in problems when integrating the system from separate CIs. Therefore, the methods of hierarchical clustering should be recognized as the best for solving the problem of CI identification in the IT project of creating a new IT product, in particular – IS. In case of re-planning of an ongoing IT project due to changes, the method described in [10] may be more convenient.

CONCLUSIONS

It has been proposed to use methods of hierarchical and non-hierarchical clustering to investigate the peculiarities of solving the problem of CI identification. As examples of hierarchical clustering methods, it has been proposed to use the AGNES agglomerative clustering method (using the nearest neighbor algorithm) and the DIANA divisive clustering method. As an example of non-hierarchical clustering methods, it has been proposed to use the k-means algorithm. In addition, it has been proposed to use the grapho-analytical clustering method considered in [10] to compare the course and results of solving the problem of CI identification. This method was recommended in [10] to solve the decomposition problem of the architecture of a monolithic software system into separate services.

It has been proposed to use the description of the architecture of the "Formation and maintenance of an individual plan of a scientific and pedagogical employee of the department" functional task as the initial data in the study. This description is quite suitable for presenting the description of the system architecture at the level of individual functional CIs. In this case, descriptions of individual functions of this task act as such CIs. These descriptions, in turn, consist of descriptions of the corresponding functions, as well as the input and output data streams of these functions. The use of the proposed description of the architecture makes it possible to establish differences in the course and results of the application of the researched clustering methods to solve the problem of identifying functional CIs.

The process of solving the problem of identifying functional CIs using the selected clustering methods has been considered. Based on the results of this review, it can be concluded that the simplest of the selected methods is the AGNES agglomerative clustering method (using the nearest neighbor algorithm). However, this statement requires further research, since it is not known how the procedure for solving the above-mentioned problem will change in the case of using other

methods and algorithms (for example, Ward's method). In general, it should be noted that with a small amount of initial data, the computational complexity of the considered clustering methods is approximately the same and does not allow to reasonably choose the best of these methods.

A comparison of the results of solving the task of identifying functional CIs using the selected clustering methods allows to establish that the results of the application of the DIANA method should be considered the most accurate. Using the method of divisive clustering makes it possible to single out in a separate cluster those functional Cis, which descriptions completely match each other. This selection allows to avoid the mistake of assigning the most similar functional CIs to different teams of IT project executors in the future. A modification of the AGNES method has been developed to provide a similar possibility of presenting the results of solving the clustering problem.

It should be noted that the use of hierarchical clustering methods to solve the problem should be considered a better alternative. Such a point of view makes it possible to further consider the task of assigning a list of functional CIs to individual teams of IT project executors that require implementation as a sequence of individual single-criteria optimization tasks. An attempt to establish such lists of functional CIs based on the a priori known number of executive teams (using the k-means algorithm) leads to the formation of individual lists, the similarity of items of which will be sufficiently artificial. However, to verify this conclusion, it is necessary to conduct further research using more complex and modern clustering algorithms.

REFERENCES

1. Bourque, P., Fairley, R. E. (Eds.) (2014). Guide to the Software Engineering Body of Knowledge. Version 3.0. IEEE Computer Society.
2. ISO/IEC/IEEE International Standard – Systems and software engineering – System life cycle processes: ISO/IEC/IEEE 15288:2015 (2015). IEEE. https://doi.org/10.1109/ieeestd.2015.7106435
3. Levykin, V. M., Ievlanov, M. V., Kernosov, M. A. (2014). Pattern planning of requirements to the informative systems: design and application: monograph. Kharkiv: The "Kompaniya "Smit LTD".
4. Cadavid, H., Andrikopoulos, V., Avgeriou, P., Broekema, P. C. (2022). System and software architecting harmonization practices in ultra-large-scale systems of systems: A confirmatory case study. Information and Software Technology, 150, 106984. https://doi.org/10.1016/j.infsof.2022.106984
5. Fritzsch, J., Bogner, J., Zimmermann, A., Wagner, S. (2019). From monolith to microservices: A classification of refactoring approaches. 1st International Workshop on Software Engineering Aspects of Continuous Development and New Paradigms of Software Production and Deployment, DEVOPS 2018, 128–141. https://doi.org/10.1007/978-3-030-06019-0_10
6. Sellami, Kh., Saied, M. A., Ouni, A. (2022). A Hierarchical DBSCAN Method for Extracting Microservices from Monolithic Applications. 2022 ACM International Conference on Eva-

luation and Assessment in Software Engineering, EASE 2022, 201–210. https://doi.org/10.1145/3530019.3530040
7. Krause, A., Zirkelbach, C., Hasselbring, W., Lenga, S., Kroger, D. (2020). Microservice Decomposition via Static and Dynamic Analysis of the Monolith. 2020 IEEE International Conference on Software Architecture Companion, ICSA-C 2020, 9–16. https://doi.org/10.1109/icsa-c50368.2020.00011
8. Reiff-Marganiec, S., Tilly, M. (Eds.) (2012). Handbook of Research on Service-Oriented Systems and Non-Functional Properties: Future Directions. IGI Global. https://doi.org/10.4018/978-1-61350-432-1
9. Shahin, R. (2021). Towards Assurance-Driven Architectural Decomposition of Software Systems. 40th International Conference on Computer Safety, Reliability and Security, SAFECOMP 2021 held in conjunction with Workshops on DECSoS, MAPSOD, DepDevOps, USDAI and WAISE 2021, 187–196. https://doi.org/10.1007/978-3-030-83906-2_15
10. Faitelson, D., Heinrich, R., Tyszberowicz, Sh. (2017). From monolith to microservices: Supporting software architecture evolution by functional decomposition. 5th International Conference on Model-Driven Engineering and Software Development, MODELSWARD 2017, 435–442. https://doi.org/10.5220/0006206204350442
11. Suljkanović, A., Milosavljević, B., Inđić, V., Dejanović, I. (2022). Developing Microservice-Based Applications Using the Silvera Domain-Specific Language. Applied Sciences, 12 (13), 6679. https://doi.org/10.3390/app12136679
12. Han, J., Kamber, M., Pei, J. (2012). Data Mining. Concepts and Techniques. Waltham: Morgan Kaufmann Publishers. https://doi.org/10.1016/c2009-0-61819-5
13. Barseghyan, A. A., Kupriyanov, M. S. Kholod, I. I., Tess, M. D., Elizarov, S. I. (2009). Analiz dannykh i protcessov. Saint-Petersburg: BHV-Petersburg, 512.
14. Kaufman, L., Rousseeuw, P. J. (2005). Finding Groups in Data. Introduction to Cluster Analysis. John Wiley & Sons, Inc.
15. Wierzchoń, S., Kłopotek, M. (2018). Modern Algorithms of Cluster Analysis. Cham: Springer. https://doi.org/10.1007/978-3-319-69308-8
16. Ievlanov, M., Vasiltcova, N., Neumyvakina, O., Panforova, I. (2022). Development of a method for solving the problem of IT product configuration analysis. Eastern-European Journal of Enterprise Technologies, 6 (2 (120)), 6–19. https://doi.org/10.15587/1729-4061.2022.269133
17. Vasiltcova, N., Panforova, I. (2022). Research on the use of hierarchical clustering methods when solving the task of IT product configuration analysis. Management Information Systems and Devices, 178, 37–49. Available at: https://www.ewdtest.com/asu/wp-content/uploads/2024/01/ASUiPA_178_37_49.pdf

DOI: 10.15587/978-617-8360-03-0.CH2

Maksym Naumenko, Iryna Hrashchenko, Tetiana Tsalko,
Svitlana Nevmerzhytska, Svitlana Krasniuk, Yurii Kulynych

CHAPTER 2

INNOVATIVE TECHNOLOGICAL MODES OF DATA MINING AND MODELLING FOR ADAPTIVE PROJECT MANAGEMENT OF FOOD INDUSTRY COMPETITIVE ENTERPRISES IN CRISIS CONDITIONS

ABSTRACT

Developed in this research scientific and practical applied project solutions regarding Data Mining for enterprises and companies (on the example of food industry) involve the application of advanced cybernetic computing methods/algorithms, technological modes and scenarios (for integration, pre-processing, machine learning, testing and in-depth comprehensive interpretation of the results) of analysis and analytics of large structured and semi-structured data sets for training high-quality descriptive, predictive and even prescriptive models. The proposed by authors multi-mode adaptive Data Mining synergistically combines in parallel and sequential scenarios:

– methods of preliminary EDA,
– statistical analysis methods,
– business intelligence methods,
– classical machine learning algorithms and architectures,
– advanced methods of testing and verification of the obtained results,
– methods of interdisciplinary empirical expert interpretation of results,
– knowledge engineering formats/techniques – for discovery/detection previously unknown, hidden and potentially useful patterns, relationships and trends (for innovative project management).

The main methodological and technological goal of this developed methodology of multi-mode adaptive Data Mining for food industry enterprises is to increase the completeness (support) and accuracy of business and technical-technological modeling on all levels of project management of food industry enterprises: strategic, tactical and operational.

By optimally configuring hyperparameters, parameters, algorithms/methods and architecture of multi-target and multidimensional explicit and implicit descriptive and predicative models, using high-performance hybrid parallel soft computing for machine learning – the improved methodology of multimode Data Mining (proposed by the authors) allows to find/detect/mine for new, useful, hidden corporate knowledge from previously collected, extracted, integrated Data Lakes, stimulating the overall efficiency, sustainability, and therefore competitiveness, of food industry enterprises at various organizational scales (from individual, craft productions to integrated international holdings) and in various food product groups and niches.

In more detail, the purposes of this research are revealed in two meaningful modules:

1. The first part of the detailed goals and objectives of this research relate to the effective use of Data Mining (and modeling) in the competitive management of enterprises and companies in modern economy, namely:

– research and verification of the effectiveness of the basic/main three types of Data Mining in the management of a competitive enterprise;

– detection of basic/main difficulties and challenges of Data Mining technology in the management of a competitive enterprise;

– research and generation of a list of basic/main expedient functional applied corporate tasks for the application of the improved concept of Data Mining;

– determination of the list of basic/main results of using the proposed Data Mining concept and methodology for an effective and competitive enterprise in dynamic and crisis conditions;

– finding the basic/main advantages of using the proposed Data Mining concept and methodology for an effective and competitive enterprise in dynamic and crisis conditions;

– research of the basic/main technological problems of using the proposed concept and methodology of Data Mining for an effective and competitive enterprise in dynamic and crisis conditions;

– detection of the basic/main ethical problems of using the proposed Data Mining concept and methodology for an effective and competitive enterprise in dynamic and crisis conditions;

– research and search for basic/main perspectives of intelligent data analysis in the management of a competitive enterprise or company.

2. The second and main part of the detailed goals and objectives of this publication relate to the effective use of Data Mining (and modeling) in the competitive management of enterprises and companies in the food industry, namely:

– determination of features and methods of analysis and analytics of High Dimensional big data of at enterprises of the food industry;

– research of features and development of methodological and technological techniques for effective mode of OnLine Data Mining at food industry enterprises;

– research of specifics and development of recommendations regarding the effective mode of Ad-Hoc Data Mining at food industry enterprises;

– research of the specifics and development of applied recommendations regarding the effective mode of Anomaly & Fraud Detection of technological data of food industry enterprises;

– identification of directions and development of recommendations for effective use of Hybrid Data Mining at food industry enterprises;

– detection of features and development of a complex of scientific and practical recommendations regarding the effective regime of Crisis Data Mining at food industry enterprises in dynamic and unstable external conditions;

– identification of directions and development of recommendations for future trends in the effective use of Data Mining at food industry enterprises.

It can not be argued that in modern conditions (pre-crisis, crisis and post-crisis conditions of both regional food industries and the global world; globalization and simultaneous very narrow specialization of the food industry sectors; the need to take into account a huge amount of stream and packet information from various sources and various formats; the need for a quick adaptive optimal management response/adaptation in response to rapid changes in the global or regional market situation; unstable and difficult to predict dynamics of external influences: international, national, sectoral, local direct regulatory and indirect public regulation of the food industry) – deployment of the multi-mode adaptive Data Mining methodology proposed by the authors – will result in enterprises, companies and organizations/institutions of the food industry gaining additional competitive advantages at the state, regional, branch and corporate management levels.

KEYWORDS

Food industry enterprise, data mining, machine learning, big data, food industry project management, efficiency and competitiveness.

Stable, effective and competitive management involves a cycle of: strategic/tactical/operational planning; then high-quality, productive and timely implementation; monitoring/controlling/auditing results and adaptive feedback. In a business environment that rapidly and deterministically (and sometimes stochastically) changes in different directions with different rates of dynamics (the food industry in developing countries during the period of multimodal global and regional crises is characterized by this) [1], the above-mentioned stable and effective management is of crucial importance for enterprises seeking to achieve and maintain competitive advantage.

Optimal management of competition involves a set of innovative strategies, practices and algorithms aimed at increasing the total efficiency of the enterprise/company compared to its competitors. In other words, effective competitiveness management can be defined as a systematic approach to identifying, developing and implementing strategies, operational measures and tactics that increase the potential/ability of an enterprise to outperform its competitors [2]. This also involves constant monitoring of competitive factors and the environment, strategic planning and effective operational-tactical use of all types of resources (in particular, corporate knowledge) to achieve the goals of enterprises/companies/organizations, especially in difficult crisis conditions (in particular, the post-covid 19 consequences for the global and regional economies) [3].

It should be noted that in the current era of digital technologies Industry 4.0 and Industry 5.0 – ALL enterprises and companies generate huge amounts of structured, semi-structured and unstructured data from various sources (such as: internet social networks/media and network activity of users; sensors, controllers and robots; corporate systems ERP, CRM, MES, WMS, EAM, HRM) at all levels of detail and time horizons of management [4, 5]. The emergence of large structured, semi-structured and unstructured Big Data has radically changed operational, tactical

and strategic management, efficiency, and therefore the development trends of modern competitive enterprises [6]. It is the intelligent use of all big data (in different formats, in different modes, qualitative and not yet qualitative) that significantly increases the integrated efficiency, stability and robustness of enterprises, which is especially relevant in multimodal crisis periods. That is, the analysis and analytics of big data will support the adoption of optimal and timely management decisions at all levels of enterprise management, will contribute to the timely reengineering of technological and business processes at the enterprise, will be an important element of the mechanism for detecting incomprehensible anomalies and detecting potential threats in the internal or external environment of the company [7].

Taking into account the above, it is necessary to emphasize that in the context of a stable, competitive market position of an enterprise or company, it is worth emphasizing adaptive innovative management – this is a dynamic approach to decision-making and resource management that combines training in monitoring and evaluation to adjust strategies in response to changing conditions [8]. Adaptive management involves iterative cycles of planning, implementation, monitoring and evaluation, where decisions are constantly adjusted based on new information and feedback. In adaptive management, these ideas allow enterprises, companies and organizations to improve resource allocation, optimize interventions and increase resilience to uncertainty and change [9]. It is Data Mining that plays a crucial role in adaptive management, using machine learning of all types on large accumulated data sets of all formats to obtain new, useful and implicit information, identify hidden patterns and patterns, and therefore to more effectively support data-driven decision making based on empirical precedents, heuristics.

The immediate classical concept of Data Mining is a process project of discovery/detection of hidden patterns/regularities, i.e. formalized new knowledge from accumulated data sets/heuristics [10]. However, the modern, proposed concept/paradigm of Data Mining (as a subset of Data Science) provides powerful intellectual techniques, methods/algorithms and scenarios for searching and formalizing valuable new, previously unknown, useful information (insights) from multidimensional large sets of structured batch and stream data [11]. That is, modern Data Mining (taking into account: the problem/curse of data dimensionality, on-line data mining, ad-hoc data mining, hybrid data mining, anomaly & fraud detection, crisis data mining) is a relevant and mandatory technology/tool in modern innovative adaptive management of the enterprise (food industry in particular [12]).

Crisis management involves systematic planning, coordination, and implementation of strategies for prevention, preparation, response, and recovery after emergencies or disasters. It must be noted that Data Mining plays an important, often key role in accurate monitoring, effective prevention, thorough preparation and optimal response to crises of all types and levels. That is, the specially adapted and configured Data Mining technology is end-to-end for the settlement of pre-crisis, crisis and post-crisis situations of the enterprise (especially in food industry).

As part of crisis management, adaptive Data Mining effectively scenario-configured and parametrically configured within special technological modes (online Data Mining, ad-hoc Data Mining, hybrid Data Mining, crisis Data Mining) is a relevant and important factor in ensuring not only the

stability/sustainability of the enterprise (food industry in particular [13]), but also, even, ensuring its competitiveness in multimodal crisis conditions.

In crisis management, this knowledge can inform decision-makers, improve proactive awareness of crisis factors/phenomena, optimally allocate limited resources, and effectively reduce risks. Such data mining tasks as hierarchical and non-hierarchical clustering, binary and non-binary classification, search for association rules (in particular, unexpected rules), construction/training/learning (**Fig. 2.1** shows the result of such training progress) regression predicative EXPLICIT models ("white explicit models" – for example, regression trees or "gray explicit models" – for example, logit regression equations) or even IMPLICIT models – in the form of configured and trained ANN (**Fig. 2.2** shows the example of configured and trained ANN for classification of employees of an enterprise of food industry) – will contribute to enterprises and food industry companies to increase their readiness, response efficiency and effectiveness of proactive measures/influences for recovery/stabilization in various scenarios and stages of crisis phenomena.

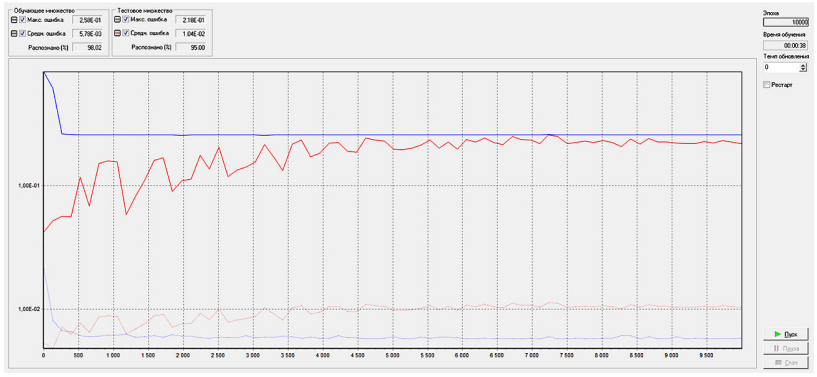


Fig. 2.1 The result of the progress of ANN training for the classification model of employees of an enterprise of food industry
Note: developed by the authors

Taking into account the above, the task of researching an effective and optimal concept of using innovative Data Mining modes in the effective adaptive management of enterprises/companies, with the aim of improving their competitiveness (especially in multimodal crisis conditions), becomes particularly relevant. Therefore, this publication aims to present the author's achievements and scientific and practical results (supported by theoretical studies, thematic industry research, accumulated industry heuristics and the author's empirical experience) regarding a specialized paradigm, concept, methodology and a set of operational and tactical measures for the optimal use of innovative Data modes Mining (and their proposed options/settings) for more effective and adaptive, anti-crisis management of the food industry enterprise.

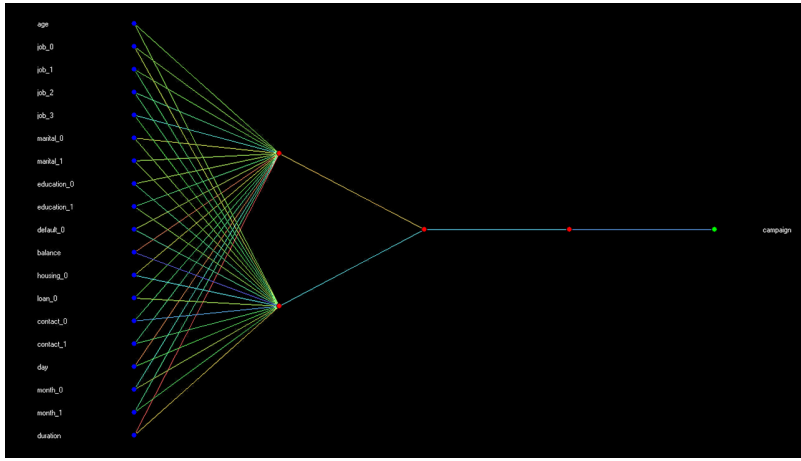


Fig. 2.2 The graph of configured and trained ANN Multilayer Perceptron for the classification model of employees of an enterprise of food industry (using Multilayer Perceptron architecture (3 hidden layers) with back-propagation, Sigmoid activation function, machine learning speed = 0.1 and moment of inertia 0.9)
Note: developed by the authors

The provisions of the MODERN Data Mining (and Machine Learning) theory in unstable crisis conditions are revealed in the publications of such scientists as: V. Derbentsev [14], A. Matviychuk [15, 16], H. Velykoivanenko [17], etc.

But, taking into account the significant and systemic specificities/peculiarities of the food industry, the authors studied and analyzed the following relevant scientific articles in detail. Applying data mining techniques and analytic hierarchy process to the food industry was researched in [18]. Data mining and optimization issues in the food industry were analyzed in [19]. In [20] was paid attention to Data mining application for customer segmentation based on loyalty. Global food production and distribution analysis using data mining and unsupervised learning were developed in [21]. Application of Data mining in food trade network was analyzed in [22]. Mining logistics data to assure the quality in a sustainable food supply chain was developed in [23]. A framework for modeling efficient demand forecasting using data mining in supply chain of food products export industry was proposed in [24]. Predicting consumer preference for fast-food franchises: a data mining approach was described in [25]. A case study of customers grouping using data mining techniques in the food distribution industry was described in [26]. Data mining on time series: an illustration using fast-food restaurant franchise data was investigated in [27]. Consumers' behavior in the food and beverage industry through data mining was researched in [28]. Alternative data mining/machine learning methods for the analytical evaluation of food quality and authenticity

were reviewed in [29]. However, relevant and unresolved issues are the features/specificities of the paradigm, concept, methodology, and set of operational-tactical measures for optimal setting and deployment of special Data Mining Modes when making effective decisions regarding competitive and sustainable management, especially in conditions of multimodal crisis phenomena (in particular, for food industry enterprises).

The aim of the research: research and development of an effective and optimal concept, methodology, technological techniques and modes, scenarios for the use of Data Mining (and modeling) in the adaptive management of food industry enterprises, with the aim of improving their competitiveness (especially in dynamic and unstable external conditions, and even, in crisis conditions).

In more detail, the purpose of this study is revealed in two meaningful modules:

1. The first part of the detailed goals and objectives of this publication relate to the effective use of Data Mining (and modeling) in the competitive management of enterprises and companies in various sectors of the economy, namely:

– research and verification of the effectiveness of the basic/main three types of Data Mining in the management of a competitive enterprise;

– detection of basic/main difficulties and challenges of Data Mining technology in the management of a competitive enterprise;

– research and generation of a list of basic/main expedient functional applied corporate tasks for the application of the improved concept of Data Mining;

– determination of the list of basic/main results of using the proposed Data Mining concept and methodology for an effective and competitive enterprise in dynamic and crisis conditions;

– finding the basic/main advantages of using the proposed Data Mining concept and methodology for an effective and competitive enterprise in dynamic and crisis conditions;

– study of the basic/main technological problems of using the proposed concept and methodology of Data Mining for an effective and competitive enterprise in dynamic and crisis conditions;

– detection of the basic/main ethical problems of using the proposed DATA MINING concept and methodology for an effective and competitive enterprise in dynamic and crisis conditions;

– research and search for basic/main perspectives of intelligent data analysis in the management of a competitive enterprise or company.

2. The second and main part of the detailed goals and objectives of this publication relate to the effective use of Data Mining (and modeling) in the competitive management of enterprises and companies in the food industry, namely:

– determination of features and methods of analysis and analytics of big data of high Dimensionality at enterprises of the food industry;

– study of features and development of methodological and technological techniques for effective mode of OnLine Data Mining at food industry enterprises;

– study of specifics and development of recommendations regarding the effective mode of Ad-Hoc Data Mining at food industry enterprises;

– study of the specifics and development of applied recommendations regarding the effective mode of Anomaly & Fraud Detection of technological data of food industry enterprises;

– identification of directions and development of recommendations for effective use of Hybrid Data Mining at food industry enterprises;

– detection of features and development of a complex of scientific and practical measures regarding the effective regime of Crisis Data Mining at food industry enterprises in dynamic and unstable external conditions;

– identification of directions and development of recommendations for future development, trends in the effective use of Data Mining at food industry enterprises.

In this research, a thorough and systematic review of specialized specialized scientific literature, industry reports and practical examples and author's experience of Data Mining in enterprise management (food industry in particular) was applied. Classical methods of analysis and synthesis, deduction and induction are used in combination with the author's heuristics and empirical insights. This methodological approach provides a detailed, systematic and effective study of the role, functionality, technological mechanisms, methods/algorithms, modes, advantages and caveats of successful Data Mining projects of technological and business data in the effective management of a competitive enterprise in the food industry.

2.1 INNOVATIVE AND EFFECTIVE APPLICATIONS OF DATA MINING IN COMPETITIVE ENTERPRISE MANAGEMENT

The results of the author's empirical observations, accumulated heuristics and conducted research indicate that it is intelligent data analysis that is a transformational tool for effective enterprise management in modern dynamic (and often crisis) conditions, which potentially provides significant advantages in terms of efficiency, adoption of optimal business and technological solutions and increasing competitive advantages. However, successful implementation of data mining requires addressing issues related to data quality, complexity, scalability, and ethical concerns. By responsibly using intelligent data analysis and taking into account the future achievements of intelligent technologies and appropriate specialized hardware, enterprises and companies can use the full potential of Data Mining not only to ensure stability and sustainability, but also to stimulate qualitative transformation and quantitative growth.

Proven Data Mining technologies in the management of a competitive enterprise provide for three applied areas [30]:

1. Descriptive Analytics: involves transformation, generalization of accumulated, historical data – to understand the situation and current state.

2. Predictive Analytics: involves using accumulated historical data to predict future events and trends.

3. Prescriptive Analytics: generates recommendations for decisions and actions, based on prebuilt descriptive and predicative models.

The main difficulties and challenges of Data Mining technology in the management of a competitive enterprise are identified [31]:

1. Data quality and availability (inconsistent and incomplete data can hinder quality intelligence; data privacy and security issues limit access to sensitive information, etc.).

2. Computational complexity (machine learning on big data requires significant computing resources; real-time analysis requires efficient algorithms and infrastructure, etc.).

3. Interdisciplinary integration (combining expertise from different industries, levels of management, regions are critical, but difficult).

The following functional applied corporate sectors are proposed for the application of the improved *Data Mining* concept [32]:

1. Customer Relationship Management (CRM).

Data mining in CRM helps businesses understand and predict customer behavior to improve customer satisfaction and retention. Key techniques include: clustering to group customers based on behavioral or demographic similarities to enable targeted marketing strategies; classification to predict customer churn or propensity to buy to inform retention strategies; defining association rules to define relationships between products to facilitate market basket analysis and cross-selling.

2. Financial Analysis and Fraud Detection. Data mining is critical in financial analysis for risk assessment and fraud detection, including anomaly detection (identifying unusual patterns that indicate fraudulent activity), predictive modeling (assessing credit risk by analyzing historical data to predict the likelihood of credit default).

3. Supply Chain Management (SCM). Data mining optimizes supply chain processes, including demand forecasting, inventory management and supplier evaluation, including through: time series analysis (forecasting future demand to improve inventory planning); clustering (classification of suppliers based on performance indicators to assist in supplier selection and management), etc.

4. Human resource management (HRM). Data mining supports HRM in talent acquisition, performance evaluation and employee retention (predictive analytics to identify potential candidates likely to succeed based on historical hiring data; text mining to analyze employee feedback to assess workplace satisfaction and identify areas, which need to be improved, etc.).

The following list of possible results of using the proposed Data Mining concept and methodology for an effective and competitive enterprise in dynamic and crisis conditions is defined:

1. Making informed decisions. Data mining provides a strong foundation for strategic decisions by uncovering hidden patterns and correlations in data, leading to more accurate and effective decision making.

2. Efficiency and cost reduction. Automated data analysis reduces time and resource consumption compared to traditional methods. Improved forecasting and inventory management reduce operational costs.

3. Improved and enhanced customer experience. Highly personalized marketing and improved adaptive customer service (retail and wholesale) based on data analysis contribute to increasing the level of customer satisfaction and loyalty.

The following basic advantages of using the proposed concept and methodology of Data Mining for an efficient and competitive enterprise in dynamic and crisis conditions are clarified:

1. Additional competitive advantage. Businesses that use data mining can gain a competitive advantage by identifying market trends and customer preferences over their competitors.

2. Improved scalability. Data mining techniques can be scaled to handle large volumes of data, making them suitable for businesses of all sizes.

The following basic technological problems of using the proposed Data Mining concept and methodology for an efficient and competitive enterprise in dynamic and crisis conditions were detected:

1. Quality of input data. The effectiveness of data mining is highly dependent on the quality of the data. Incomplete, noisy, or biased data can lead to inaccurate models and misleading insights. Suggested: Implementation of thorough data cleaning and pre-processing techniques results in higher data quality and more reliable results.

2. Complexity and experience. Data mining requires specialized knowledge in statistics, machine learning, and domain expertise. The complexity of algorithms and the need for qualified specialists can be an obstacle for some enterprises. It is suggested that additional investment in training and hiring skilled data professionals is critical to successful data analytics initiatives.

3. Scalability issues. As data volumes grow, ensuring that data mining algorithms scale effectively becomes a challenge. High computing demands may require advanced infrastructure and resources. It is suggested that the use of cloud computing and distributed processing systems such as Hadoop and Spark can help solve scalability problems.

The following basic ethical problems of using the proposed Data Mining concept and methodology for an effective and competitive enterprise in dynamic and crisis conditions were also detected:

1. Privacy & security. Data mining, as a rule, involves the most detailed and comprehensive analysis of all collected personal data, which has long been a concern of civil society regarding the privacy & security of personal data. Data-driven companies must ensure compliance with all regulations for the protection of personal and sensitive data, at a minimum, such as GDPR, and implement proactive and more comprehensive security measures against the leakage of this data. It should be emphasized that even more clear and transparent methodologies for the collection, integration, pre-processing, analysis and analytics, interpretation of the above-mentioned data, in parallel with obtaining the interactive informed consent of users, simultaneously with the implementation of more reliable security protocols regarding the leakage of this data – are already vital for maintaining credibility and relevance in the Data Mining industry.

2. Bias/fairness. Methods and technologies of data extraction, integration and pre-processing (for further ML) can, for example, inadvertently form input biases, biases/distortions in training samples, which will lead to unfair or discriminatory or incorrect ML results. It is proposed to use a complex of methods and technologies of mathematical statistics and comprehensive human expert analysis to identify and mitigate potential bias, regular audit and descriptive and predicative ML models, in particular, systematic testing of the reliability and completeness of predictive models.

3. Transparency/Accountability. The use of Data Mining (and Data Science) in the process of making management decisions at all levels should be unambiguous, clear and transparent, with detailed documentation of prerequisites, methodologies and justified responsibility for results. It is believed that the regulatory establishment/enforcement/audit of the use of normative ethical rules (rather than just principles) and systematic management of personal data will help to ensure the responsible, non-harmful use of such big data.

Below are the results of research into the future trends & prospects of data mining in the management of a competitive enterprise or company:

1. Integration of artificial intelligence (intelligent, knowledge-oriented technologies) within the framework of modern management. The integration of AI with Data Mining (and Data Science) will lead to significant qualitative and quantitative changes in the management of a modern enterprise. On the one hand, modern DM modes will ensure that AI knowledge bases are filled with new, relevant, hidden regularities/patterns; on the other hand, AI will help improve the accuracy and completeness of semi-supervised ML. Synergies between AI and DM will drive hybrid interdisciplinary innovation and thus further efficiency.

2. Data Mining (and even Data Science) in real time, 24/7/365, not only of batch, but also of streaming data. With the increasing prevalence of streaming structured, semi-structured and unstructured data in real time, companies are moving to Data Mining (and later to Data Science) in the mode of integration, processing, analysis, analytics, testing and interpreting the results – in real time. This approach allows competitive companies to instantly make optimal decisions based on current data, increasing responsiveness and flexibility. Real-time Data Mining (and even Data Science) is especially valuable in areas such as detecting errors/malfunctions in process equipment of conveyors or process flow production lines (for example, in the food industry), etc.

3. Data mining in IoT (especially in the Internet of Industrial Things). The proliferation of devices, equipment, equipment with IoT technology generates additional huge volumes of streaming data. Data Mining will manage to extract new, hidden, valuable additional useful information (regularities/patterns) from this data, contributing to improvements in food industry areas such as optimal predictive maintenance, intelligent adaptive continuous manufacturing, adaptive dispatching, etc.

4. AutoML and AutoDM. There is only one way to know which algorithm/method or ensemble of algorithms, which settings of their hyperparameters and local parameters will help to obtain the best model in each particular case, is to try all algorithms, their combinations, and all combinations of their parameters. But, if the ML engineer also takes into account all possible variants of data normalization and variants of objective functions, a combinatorial "explosion", "curse of dimensionality" will definitely occur. So, trying to try all of this by hand is impractical, so ML tool vendors have put a lot of effort into developing support systems of the AutoML class (so-called automated ML).

The best such tools combine settings and selection of parameter variants, selection of algorithms/methods and data normalization variants. Hyperparametric fitting of a better model (or models) is often performed in the later stages of an ML project. Hence, autoML is used to reduce human interaction and automate all tasks to solve real-world ML (and therefore DM)

problems/problems. This functionality involves automating the entire iterative process from raw data to ML model testing.

This will significantly reduce the time and effort required to analyze and analyze big data and make it more accessible to companies of all sizes. Remember, even though AutoML doesn't require human interaction, that doesn't mean it's completely superior to it.

5. Expanded and improved methods of exploratory intelligent visualization of multidimensional structured and semi-structured data in EDA mode. Expanded and improved methods of visualization of the results of analysis and data analytics will help and make it easier for decision-makers to interpret the results of data mining in depth and interdisciplinary, and then act on them in a reasoned way. Moreover, pictographs, parallel coordinates and other methods of intelligence visualization of multi-dimensional data will help to find/notice hints, trends, ideas that will become the basis of further data research, their analysis and subsequent analytics.

2.2 INNOVATINE TECHNOLOGICAL MODES OF DATA MINING AND MODELLING FOR FOOD INDUSTRY ENTERPRISES IN CRISIS CONDITIONS

Data Mining in conditions of high dimensionality of data of enterprises and corporations of the food industry

Definition: High dimensionality refers to data sets with a large number of features (variables). Such data sets are common in very complex, interdisciplinary industries/productions/technologies of the food industry: baby food production, organic food production, diet (e.g. diabetic) and special (e.g. lactose-free, gluten-free) food production, and other products of deep complex technological processing.

Difficulties and problems associated with the high dimensionality of the company's input data:

1. Curse of Dimensionality (increased sparsity of data points in high-dimensional space; difficulty in recognizing meaningful patterns due to noise and irrelevant features).

2. Computational complexity (higher requirements for memory and processing power; longer computing time for model analysis and training).

3. Overfitting (increased risk of models capturing informational noise along with basic, deterministic patterns; reduced ability to generalize to new data in new external conditions).

4. Complexities of intelligence visualization within the framework of the EDA mode (below in **Fig. 2.3**, there is the attempt to perform discovery visualization (without prior dimensionality reduction) of the high dimensional financial characteristics of customers of network retail (aimed for the sale of food products).

As is clear from **Fig. 2.3**, high dimensionality creates significant analytical challenges, but it must be effectively managed using dimensionality reduction methods (PCA, t-SNE, autoencoders), deterministic feature selection (Filter Methods, Wrapper Methods, Embedded Methods) and regularization (L1 Regularization (Lasso) or L2 Regularization (Ridge)). These approaches improve model performance, reduce computational burden, and improve interpretation.

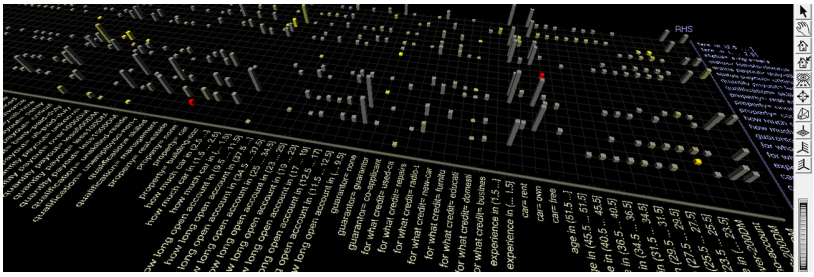


Fig. 2.3 The example of high dimensional EDA – discovery visualization (without prior dimension reduction) of the high dimensional financial characteristics of customers of chain retail (aimed for the sale of food products)
Note: developed by the authors

Online Data Mining (and later – Data Science) for food industry enterprises means the concept, methodology and technologies of analysis and analytics of big data, which are constantly generated, processed and updated in real time. This approach is different from traditional methods of batch processing and analysis, so it is crucial for those tasks and management functions of the food industry enterprise that require immediate analysis of streaming data and operational response. That is why, below, the actual results of the author's research on the principles, methodology, effective application, challenges and future directions of online Data Mining at food industry enterprises will be presented, emphasizing its role in dynamic and time-sensitive applications in the food industry.

In today's data-driven food industry, the need for real-time data analysis has become increasingly important. Online Data Mining addresses this requirement by enabling the continuous extraction of insights from all types and formats of streaming data. Let's consider the basic principles of Online Data Mining:

1. Incremental Processing: Data is processed incrementally (using data partitioning), ensuring that the system can handle high-velocity data streams without the need for complete retraining.

2. Continuous Learning: Online Data Mining involves algorithms that can learn and adapt continuously as new data arrives.

3. Adaptability: Algorithms must be adaptable to changing data clustering, general patterns, distributions, etc.

4. 24/7/365 Real-Time Online Analysis and Analytics: The primary goal is to provide real-time ideas & insights and appropriate decision-making capabilities.

Let's outline the basic methodologies of online Data Mining, because online Data Mining employs various methodologies, including:

– Online Streaming Algorithms: such algorithms designed to process data in a single pass or limited number of passes, such as the Hoeffding Tree, online K-means, and incremental PCA;

– Sliding Windows: techniques that use a fixed-size window of the most recent data points to update models continuously;

– Concept Drift Detection: methods for detecting and adapting to changes in data distribution over time, ensuring model relevance and accuracy;

– Incremental MLearning: machine learning methods that update models incrementally, such as online versions of support vector machines (or even RVM) and shallow neural networks.

The following are the main advantages of online data mining:

– Adaptability: capable of adapting to evolving data patterns and trends;

– Scalability: capable of handling big volumes of multidimensional data generated at high velocity;

– Resource efficiency: requires less computational resources compared to reprocessing entire datasets;

– Timeliness: provides immediate insights/ideas, enabling prompt decision-making and action.

Challenges and cautions of Online Data Mining:

1. Complexity: improvement, adaptation, support of effective algorithms/methods of online pre-processing of data, online analysis and online analytics is a rather non-trivial and complex task, which must also take into account the specifics of the country, region, specific enterprise and the specifics of the current managerial functional task/problem.

2. Data Quality: Ensuring the quality and consistency of streaming data will typically be a non-trivial and resource-intensive task.

3. Scalability: effective and timely management of the scalability of computing systems for online integration, data preprocessing, their analysis and analytics (especially predictive analytics) of high-speed data streams is a complex task that requires a highly experienced interdisciplinary team of human experts.

4. Latency: the need to check and test the results of online Data Mining – can cause a delay in making informed management decisions in real time.

5. Drift and evolution of the concept/paradigm encoded in the input data: the complexity of continuous adaptation and retraining due to temporal dynamics in: the distribution of data, their cluster structure, the number and values of local extremes and other changes in the fitness function or production function of the food industry enterprise.

Prospects and trends of online Data Mining at food industry enterprises:

1. Integration with Edge Computing: Using ancillary/peripheral computing to integrate and pre-process data as close as possible to its source, reducing latency and/or resource lag, and therefore maximizing throughput.

2. Advanced Machine Learning: the use of advanced machine learning techniques, such as deep machine learning using Deep Neural Networks, to enhance the capabilities of online unstructured data science [33].

3. Distributed Processing: use of innovative distributed computing infrastructures for effective script scaling of online Data Mining processes.

4. Explainable AI: development of alternative scenario methods and hybrid technologies for providing in-depth and extended interpretation and transparency of constructed models in the online Data Mining process.

The very clear example of this trend is the decision tree configured and built by the authors (**Fig. 2.4**), for the task "Estimation of Obesity Levels Based on Eating Habits and Physical Condition", which as the result helps to interpret and understand the relationship between the predicted target variable (the weight category of the consumer/patient – divided into 7 degrees: from underweight – to obesity of the 3rd degree) and other input categorical and quantitative factors (gender, age, height, weight, family history of obesity, indicators: Do you usually eat vegetables in your meals? How many main meals do you have daily? Do you eat any food between meals? Do you smoke? Do you monitor the calories you eat daily? How often do you have physical activity? How much time do you use technological devices such as cell phone, videogames, computer and others? How often do you drink alcohol? Which transportation do you usually use?).

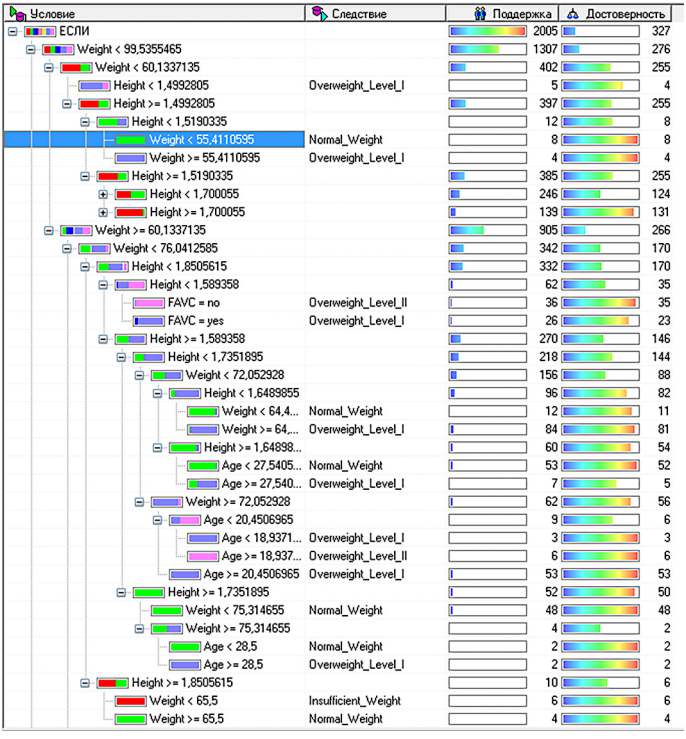


Fig. 2.4. The example of configured and constructed explainable decision tree using the C4.5 algorithm for the "Estimation of Obesity Levels Based On Eating Habits and Physical Condition" task, which helps to interpret and visually understand the relationship between the target variable and other input categorical and quantitative factors
Note: developed by the authors

Another example of the verbalization of such an implicit model (black box) obtained from a shallow ANN configured and trained by the authors with three hidden layers is given in **Fig. 2.5**.

```
Input database fields (features) (input symptoms):
          REGION
          FRESH
          MILK
          GROCERY
          FROZEN
          DETERGENTS
          DELICASSEN

Output database fields (output syndrome):
          CHANNEL

Preprocessing of database field values for transmission to network training
REGION=(REGION-2)/1
          FRESH=(FRESH-56077)/56074
          MILK=(MILK-36776,5)/36721,5
          GROCERY=(GROCERY-46391,5)/46388,5
          FROZEN=(FROZEN-30447)/30422
          DETERGENTS=(DETERGENTS-20415)/20412
          DELICASSEN=(DELICASSEN-23973)/23970

Functional converters
          Sigmoid1(A)=A/(0,1+|A|)
          Sigmoid2(A)=A/(0,1+|A|)
          Sigmoid3(A)=A/(0,1+|A|)

Syndromes of 1 level:
          Syndrome1_1=Sigmoid1( -0,07807371*REGION-0,2318511*-
FRESH-0,1125404*MILK+0,1231333*GROCERY+0,04119471*FROZEN-0,3295803*DETERGENTS+0,21418
85*DELICASSEN-0,1244661 )
----------------------------
          Syndrome1_10=Sigmoid1( 0,04964642*REGION+0,006197082*-
FRESH-0,01756757*MILK-0,05744172*GROCERY-0,09775436*FROZEN-0,03407726*DETERGENTS+0,01
003292*DELICASSEN+0,05423961 )
.....
Syndromes of 3 level:
          Syndrome3_1=Sigmoid3( -0,05175032*Syndrome2_1-0,08382219*Syn-
drome2_2-0,06745245*Syn-
drome2_3-0,08310731*Syndrome2_4+0,01224828*Syndrome2_5+0,01331895*Syndrome2_6-0,05745
047*Syndrome2_7-0,0655295*Syndrome2_8+0,03557626*Syndrome2_9+0,003902948*Syndrome2_10
-0,09182979 )
----------------------------
          Syndrome3_10=Sigmoid3( -0,05422281*Syndrome2_1-0,252343*Syn-
drome2_2-0,02300238*Syn-
drome2_3+0,0822472*Syndrome2_4-0,660238*Syndrome2_5+0,1182159*Syndrome2_6+0,04464602*
Syndrome2_7+0,03296235*Syndrome2_8-0,006042616*Syndrome2_9-0,2184449*Syndrome2_10+0,0
1921556 )

Output syndromes:
          CHANNEL_1=0,005829105*Syndrome3_1+0,3342025*Syn-
drome3_2-0,007643878*Syndrome3_3-0,06917655*Syndrome3_4-0,09772637*Syndrome3_5+0,2204
48*Syndrome3_6+0,2064538*Syndrome3_7-0,1307883*Syndrome3_8-0,1251397*Syndrome3_9-0,41
09597*Syndrome3_10-0,02931657
          CHANNEL_2=-0,04215042*Syndrome3_1-0,2311292*Syn-
drome3_2-0,07857881*Syndrome3_3-0,05260604*Syndrome3_4-0,05044578*Syndrome3_5-0,20043
5*Syndrome3_6-0,1850976*Syndrome3_7+0,289685*Syndrome3_8+0,1089799*Syndrome3_9+0,3452
721*Syndrome3_10-0,06909672

Final post-processing of output syndromes:
          CHANNEL belongs to the range corresponding to the syndrome with the maximum
value of the binary reliability score CHANNEL_1,...,CHANNEL_2
```

⬡ **Fig. 2.5.** Fragment of the experimental attempt of verbalization of the IMPLICIT model ("black box"), generated from the configured and trained shallow ANN with three hidden layers and sigmoid activation function (this model is trained for binary classification of retail distribution channel of 6 main groups of food products for a holding/concern from food industry)
Note: developed by the authors

5. Ethical Considerations: further improvement and expansion of regulation/formalization of regulation (as at the interstate level – GDPR, national, industry and even corporate levels) of possible ethical issues related to data privacy and potential bias when using the results of online Data Mining.

Online Data Mining is primarily recommended for use in solving such problems of food industry enterprises as:

– detection of staff abuses or errors in real time;

– adaptive algorithmic recipe management;

– adaptive algorithmic management of the schedule and modes of operation of technological equipment (especially relevant – considering the current state of energy supply in Ukraine);

– medical and biological monitoring;

– detection of anomalies in raw materials in real time;

– monitoring of the company's information and communication network;

– proactive forecasting and timely detection of malfunctions of technological equipment;

– adaptive planning and dispatching of technical maintenance and all types of repairs (ad-hoc, current, major repairs, modernization) of equipment;

– total comprehensive quality control of the food industry enterprise in real time.

Thus, as a conclusion on this matter, it can be stated that online Data Mining is a key approach to discover new, hidden patterns/knowledge in real-time 24/7/365 from streams of structured, semi-structured and unstructured big data that are constantly generated at food industry enterprises. Using stream (apach kafka & apach sparkl etc.) and parallel (map reduce) technologies of machine learning organization, ensemble machine learning (bagging, stacking and boosting etc.), specialized RE-learning strategies (Learning rate decay, Transfer learning, Training from scratch, Dropout etc.), methods of detecting drift/trend of the concept (for example, the displacement vector of the center of the cluster) – same online Data Mining provides timely, scalable and operational analytical support for the management of food industry enterprises. As the development and simultaneous cost reduction of peripheral distributed computing technologies, AutoML, distributed preprocessing of big data (not only ETL or ELT, but also Change Data Capture, Data replication, Data virtualization, Stream Data Integration) continues, the future of intelligent data analysis is online contains significant potential for further improvement of the process of making effective and competitive management decisions in real time by food industry enterprises.

Ad-hoc Data Mining refers to the exploratory and flexible application of data mining techniques to address specific, often unplanned, analytical queries/problems. Unlike traditional, classical approaches to Data Mining, which adhere to a predefined, deterministic set of algorithmic procedures, Ad-hoc Data Mining is characterized by its high adaptability, flexibility and ability to quickly respond to immediate analytical needs.

Ad-hoc Data Mining involves the spontaneous and situational use of those Data Mining methods/algorithms and technologies that allow data exploration and the search for patterns/patterns WITHOUT a pre-set, fixed, a priori analytical structure/prerequisites/parameters/constraints. This approach is especially valuable in dynamic environments/tasks where the types of analytical tasks,

modes of data processing and their analytics, types of patterns, forecasting time horizons, quantification of classification tasks, etc., are rapidly changing. That is, ad-hoc Data Mining allows enterprises to adapt to new information conditions and requirements (internal and external) very quickly and in a timely manner, which is especially important in times of crisis. The example of performed ad-hoc binary classification (using the kNN method) of promising or non-promising places of commercial fishing of crabs is shown in **Fig. 2.6**, and the fragment of generated PMML code of this model is presented in **Fig. 2.7**.

Let's detail the principles of Ad-Hoc Data Mining:

– User-Driven: typically initiated by end-users or analysts who require immediate answers to specific questions;

– Flexibility: Ad-hoc Data Mining allows for the adjustment of analytical approaches based on the specific context and detailed requirements of the current query;

– Iterative Process: the process is iterative, involving repeated cycles of data exploration, analysis, refinement and deployment;

– Exploratory Nature: it emphasizes exploration and discovery, often used to generate unexpected hypotheses or uncover hidden patterns in data.

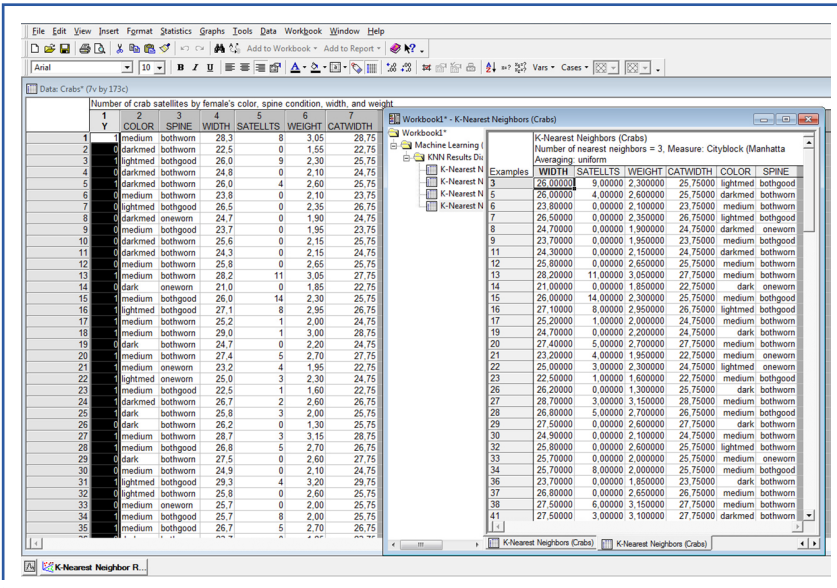


Fig. 2.6 The example of the ad-hoc binary classification (using the kNN method) of promising/unpromising places for commercial fishing of crabs (after a series of experiments, the following optimal settings of this ad-hoc binary classification model were determined: kNN classification algorithm, $k=3$, distance measure = Cityblock, averaging = uniform, sampling = 0.75)
Note: developed by the authors

```
<?xml version="1.0" encoding="windows-1251" ?>
<PMML version="2.0">
<DataDictionary numberOfFields="7">
          <DataField name="Y" optype="continuous">
          </DataField>
          <DataField name="WIDTH" optype="continuous"/>
          <DataField name="SATELLTS" optype="continuous"/>
          <DataField name="WEIGHT" optype="continuous"/>
          <DataField name="CATWIDTH" optype="continuous"/>
          <DataField name="COLOR" optype="categorical">
                    <Category label="dark" value="5"/>
                    <Category label="darkmed" value="4"/>
                    <Category label="lightmed" value="2"/>
                    <Category label="medium" value="3"/>
</DataField>
          <DataField name="SPINE" optype="categorical">
                    <Category label="bothgood" value="1"/>
                    <Category label="bothworn" value="3"/>
                    <Category label="oneworn" value="2"/>
</DataField>
</DataDictionary>
<KNearestNeighborModel modelName="STATISTICA K-Nearest Neighbor" knn_type="Regres-
sion" noOfNearestNeighbors="3" metric="cityblock" weighted="no">
          <KNNSchema>
                    <KNNField name="Y" usageType="predicted">
                    </KNNField>                    <KNNField name="WIDTH"
shift="-1,68000000000000e+000" scale="8,00000000000000e-002" />
                    <KNNField name="SATELLTS" shift="0,00000000000000e+000"
scale="6,66666666666667e-002" />
                    <KNNField name="WEIGHT" shift="-3,00000000000000e-001"
scale="2,50000000000000e-001" />
                    <KNNField name="CATWIDTH" shift="-3,25000000000000e+000"
scale="1,42857142857143e-001" />
                    <KNNField name="COLOR" shift="0,00000000000000e+000"
scale="1,00000000000000e+000" >
                              <Category label="dark" value="5"/>
<Category label="darkmed" value="4"/>                    <Category
label="lightmed" value="2"/>
<Category label="medium" value="3"/>              </KNNField>
<KNNField name="SPINE" shift="0,00000000000000e+000" scale="1,00000000000000e+000" >
                              <Category label="bothgood" value="1"/>
<Category label="bothworn" value="3"/>                    <Category label="oneworn"
value="2"/>
</KNNField>         </KNNSchema>
<KNearestNeighborExamples noOfExamples="129">
<Example id="1">
<AttributeInstance name="Y" value="1,00000000000000e+000"/>
<AttributeInstance name="WIDTH" value="2,60000000000000e+001"/>
<AttributeInstance name="SATELLTS" value="9,00000000000000e+000"/>
<AttributeInstance name="WEIGHT" value="2,30000000000000e+000"/>
<AttributeInstance name="CATWIDTH" value="2,57500000000000e+001"/>
<AttributeInstance name="COLOR" value="2,00000000000000e+000"/>
<AttributeInstance name="SPINE" value="1,00000000000000e+000"/>
</Example>
.....
<Example id="129">
<AttributeInstance name="Y" value="0,00000000000000e+000"/>
<AttributeInstance name="WIDTH" value="2,45000000000000e+001"/>
<AttributeInstance name="SATELLTS" value="0,00000000000000e+000"/>
<AttributeInstance name="WEIGHT" value="2,00000000000000e+000"/>
<AttributeInstance name="CATWIDTH" value="2,47500000000000e+001"/>
<AttributeInstance name="COLOR" value="3,00000000000000e+000"/>
<AttributeInstance name="SPINE" value="2,00000000000000e+000"/>
</Example>
</KNearestNeighborExamples>
</KNearestNeighborModel>
</PMML>
```

Fig. 2.7 Fragment of generated PMML code of configured and trained classification model by using kNN method ($k=3$) (this model is built for binary classification of promising/unpromising places for industrial fishing/harvesting of crabs from data set with 129 records)
Note: developed by the authors

Ad-hoc Data Mining employs a variety of methodologies, including but not limited to: Query-Based Analysis, Visual Data Exploration, Descriptive Statistics, classic models of Machine Learning, Text Mining, Web Mining, Process Mining, SNA, etc.

Ad-hoc Data Mining is applied across various levels and functions of food enterprise management, including:

– Business Intelligence (quickly addressing business queries, such as sales trends, customer behavior, and market analysis, etc.);

– Biological safety and Healthcare staff (investigating patient data for immediate insights into treatment outcomes, disease patterns, and healthcare utilization, etc.);

– Finance and Risk Management (analyzing financial transactions and market data to identify fraud, assess risks, and optimize investment strategies, etc.);

– BtB and BtG E-Commerce (exploring customer purchase data to uncover buying patterns, product preferences, and inventory needs, etc.);

– Information Security and Cyber Security of food enterprise telecommunications (examining call records and network data to detect anomalies, optimize network performance, and improve customer service, etc.).

Advantages of ad-hoc Data Mining: Timeliness (ensures faster understanding of the problem and context, which allows the management of food industry enterprises to respond more quickly to new, unexpected problems and obstacles); Customization (adjusts and adapts Data Mining to specific current urgent needs, providing more targeted, relevant and narrowly focused results); sometimes Resource Efficiency (in some simple cases and situations – reduces the need for careful preliminary planning and long-term allocation of resources – which is typical for traditional corporate Data Mining projects).

Challenges and cautions of Ad-hoc Data Mining for food enterprise management:

– Data Quality: ensuring the accuracy and reliability of data used in ad-hoc analyzes can be challenging, especially with unstructured, incomplete or noisy data;

– Scalability: Ad-hoc approaches may struggle to scale efficiently when dealing with very large semi-structured and unstructured datasets or complex queries from Data Lakes;

– Consistency: maintaining consistency and reproducibility of results can be difficult without standardized procedures;

– Skill Requirements: qualified and experienced interdisciplinary engineers and analysts with the ability to quickly adapt are needed to effectively apply and deploy various methods/algorithms of big data preprocessing, analysis and analytics (including hybrid and soft computation methods [34]).

Prospects and trends of Ad-hoc Data Mining at food industry enterprises:

– Enhanced Tool Integration: development of more sophisticated and customized tools that integrate ad-hoc Data Mining capabilities with user-friendly interfaces and advanced analytical function;

– Automated Ad-hoc Analysis: leveraging artificial intelligence to automate parts of the ad-hoc analysis process, improving speed and reducing the manual effort required;

– Collaborative Platforms: creating platforms that facilitate collaboration among analysts, allowing for shared insights and more comprehensive ad-hoc analyses;

– Real-Time Data Processing and Analysis: advancing real-time data processing and Analysis technologies to support immediate ad-hoc queries and analysis on streaming data.

Ad-hoc Data Mining is a crucial approach in the modern data-driven landscape, offering the flexibility and speed needed to address immediate analytical needs (for example, during unexpected situations in the logistics activities of the enterprise, complications in matters of currency import of important raw materials and equipment, in possible export VAT refund issues, unexpected changes in state and industry standards/norms, tax legislation, etc.). While ad-hoc Data Mining presents unique challenges, its advantages in providing timely, customized insights make it an invaluable tool across various domains of the food industry. Continued advancements in technology and methodologies with relevance of crisis management will further enhance the effectiveness and applicability of ad-hoc Data Mining for enterprises and companies in the food industry.

Mode of detection of anomalies in the intellectual analysis of technological data of food industry enterprises

Anomaly detection in structured, semi-structured, and even unstructured data mining is the process of identifying rare elements, events, or observations that are significantly different from the majority of data. Hence, the detection of anomalies, also known as the detection of outliers, is an important aspect of intelligent data analysis, and focuses on the detection of cases that differ significantly from the accepted/defined norm. These anomalies can detect and even predict critical consequences, such as: financial or logistical fraud, insider abuse in corporate management, intrusion into the company's information network; malfunction of the technological conveyor/equipment/equipment, etc. – which makes their detection (and better proactive forecasting) very important in the conditions of the specifics of the food industry. Thus, the main objective of the author's study of the Anomaly and Fraud Detection regime is to provide a comprehensive understanding of how anomaly detection affects different areas of activity and management levels of food industry enterprises, by identifying unusual data sets, their patterns/regularities and even, rare models (which may indicate relatively regular, but NOT massive events, facts, scenarios (with low statistical support)).

Components of the proposed concept of detecting anomalies in data:

1. Identifying Anomalies: Anomalies are data points that do not conform to expected patterns. They can be classified into point anomalies, contextual anomalies, and collective anomalies.

2. Types of data: Anomaly detection can be applied to different types of data, according to different classifications (numeric, categorical, time series and spatial data; structured, semi-structured and unstructured data; batch and streaming data, etc.).

3. The context of detecting anomalies in data: the context and specifics of the country, region, sector and segment of the food industry, macroeconomic conditions and the conditions of a specific enterprise/factory – greatly influences the choice of the Data Mining mode, the Data Mining method/algorithm, the optimal configuration of its hyperparameters and local parameters of the previously selected algorithm/method of Data Mining.

Accurate and complete detection of anomalies in the data of food industry enterprises involves the iterative, multi-stage use in synchronous and asynchronous modes of the proposed subset of methods/algorithms with different paradigms:

1. Exploratory Data Analysis (EDA), namely discovery visualization: use these methods at the first stage of the task of detecting anomalies in the data of food industry enterprises. To summarize the main characteristics of a dataset, often using visual methods, to understand its structure, detect patterns, anomalies, and test hypotheses. EDA is a crucial step in the data analysis process that provides a thorough understanding of the dataset, guiding further statistical analysis and modeling efforts. The authors recommend using parallel coordinates, pictographs (but NOT the Chernov method). Below, in **Fig. 2.8** given the performed example of the above-mentioned intelligence discovery visualization of data about emissions of harmful substances (CO_2, CH_4, N_2O) within the environmental management of a food industry enterprise – from this visualization, it is possible to notice the abnormal, jump-like change in emissions.

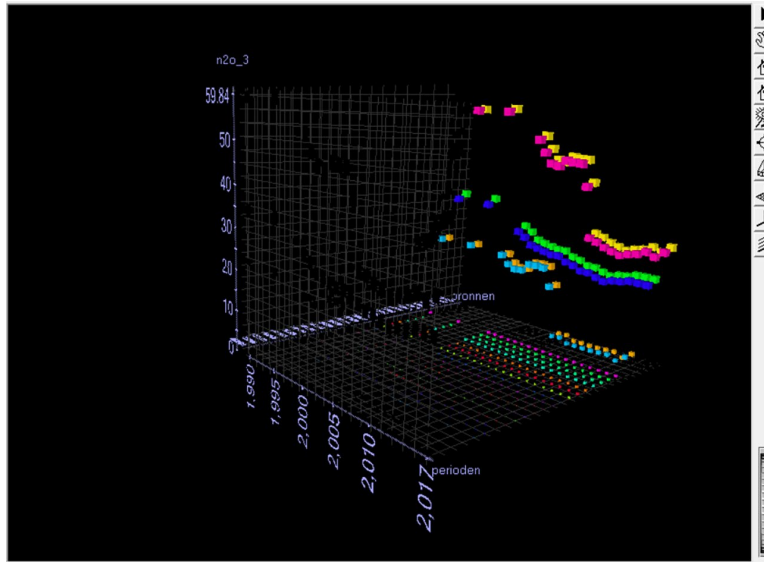


Fig. 2.8. The example of intelligence discovery visualization of data about emissions of harmful substances (CO_2, CH_4, N_2O) within the environmental management of a food industry enterprise
Note: developed by the authors

2. Statistical methods: at the second stage of anomaly detection – use several statistical algorithms simultaneously using the probability of specific data points (categories), in particular: Z-score, Grubbs test, Rosner test and the use of distribution-based threshold values.

3. Machine learning: the authors recommend combining different types of machine learning to detect anomalies in the data of food industry enterprises:

3.1. First you should use unsupervised machine learning: it predicts and detects anomalies in unfamiliar accumulated data (in a new subject area for the analyst/manager, a new food industry market, in a completely new task/situation) – that is why these data are usually NOT pre-labeled, and for this, it is recommended to use (specially configured):

3.1.1. Clustering methods: it is recommended by the authors to apply at the first stage – Hierarchical cluster analysis on a small random sample of big data set – to find the optimal parameter *k*. Using this already known optimal *k* value on second stage – for K-means or K-medians algorithm for the entire Big Data Set. Please, see **Fig. 2.9** with the illustration of the performed such preliminary hierarchical cluster analysis. It shows different variants of a priori *k* parameter: depending on the desired level of management: either $k=2$ for a higher/strategic level of management, or $k=4$ for middle/tactical level of management.

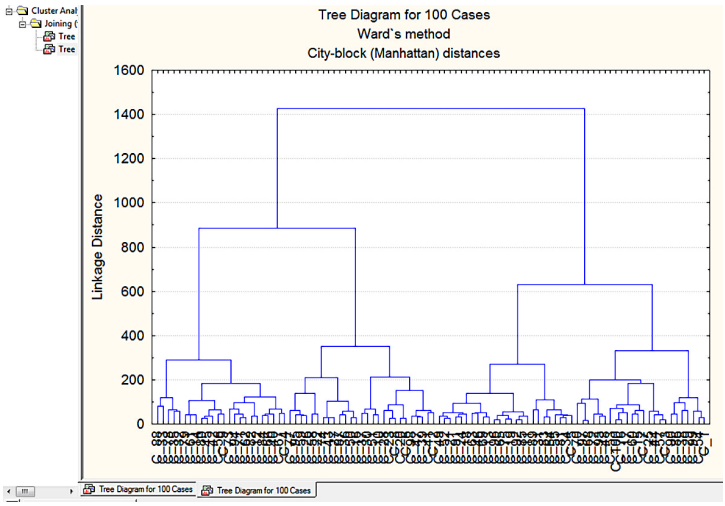


Fig. 2.9 The preliminary hierarchical cluster analysis trained on the random small sample (100 data vectors) from BIG input data set. A data set with the results of consumer blind testing of four different samples of products of a food industry enterprise was used
Note: developed by the authors

3.1.2. Algorithms for finding patterns in the form of rules – with the option of finding UNexpected rules – see author's example below, in **Fig. 2.10**.

3.1.3. Special architectures of Artificial Neural Networks in particular SOM, Autoencoders (as an example – see **Fig. 2.11** with the results of the author's application of the algorithm

of unsupervised machine learning of artificial neural network SOM Kohhonen for detecting anomalies in multidimensional data).

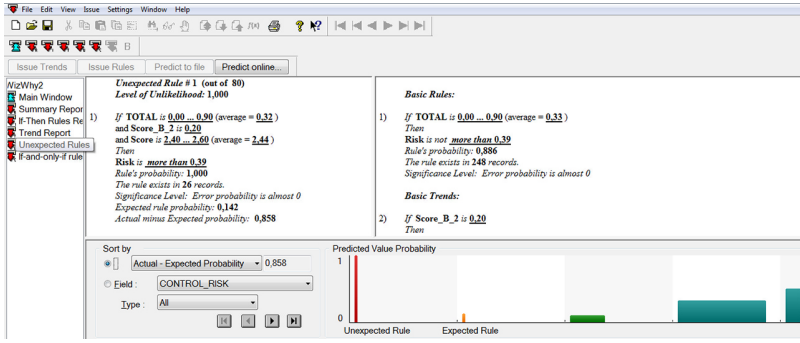


Fig. 2.10 The example of finding UNEXPECTED patterns in the form of rules – using modified Apriori method. The results of the financial audit of enterprises (food industry in particular) were used as input data
Note: developed by the authors

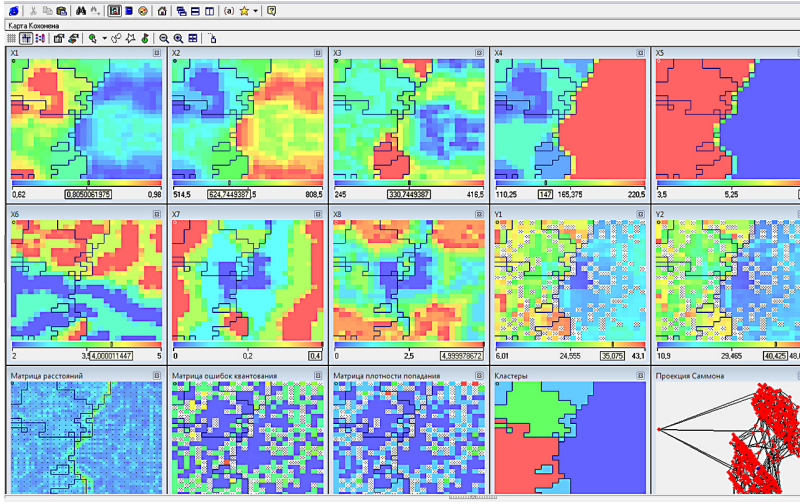


Fig. 2.11 The visualizations of the results of the Anomaly Detection Analysis (for energy audit purposes) using the dataset with projects of energy efficiency for food industry enterprises (mode and purpose of analysis: Anomaly and Fraud Detection visualization; algorithm: SOM Kohhonen; rectangular cells; neighborhood function: step function)
Note: developed by the authors

3.2. In the future, with the accumulation of initial experience and expertise regarding the current management task/situation, it is advisable to apply supervised machine learning: this type of ML uses labeled data to train models that distinguish between normal and abnormal cases. It is recommended that the authors use classification methods/algorithms. At the same time, it is recommended to first apply binary classification settings (typical/template operation/transaction/item or atypical/suspicious), and on the next iteration – ternary, and then use soft computational approaches).

3.3. With the deployment of the project/production of a food industry enterprise, Big Data will be accumulated, the analysis and analytics of which require the use of semi-supervised machine learning. semi-supervised machine learning combines marked/labeled data (of course, not a large amount) and large volumes of unlabeled data, followed by the application of Data Mining methods/algorithms to solve classification and forecasting problems. It is semi-supervised machine learning that will allow enterprises and companies in the food industry to increase the accuracy and efficiency of their Big Data analysis and analytics.

4. Ensemble methods: Combine multiple models to improve reliability and accuracy, such as a combination of statistical and machine learning methods.

It should be noted that effective anomaly detection methods/algorithms are based on two basic approaches to calculating anomalies in data:

– an approach based on distance in a multidimensional feature space: identifies anomalies based on distance measurements between data points, for example, Euclidean distance, Mahalanobis distance and NN, kNN, CBR methods, etc.;

– density-based approach: detects anomalies by examining the density of data points using methods such as local outlier factor (LOF) and isolation forest, etc.

Some practical applications of the anomaly detection mode in the practice of the food industry:

1) detection of fraud, abuses and errors in commercial activity: detection of intentional fraudulent and unintentional (erroneous) transactions, operations and influences in the sphere of procurement, in the sphere of sales activities of the food industry enterprise/holding;

2) cyber security (and, even, physical security monitoring): detection of network intrusions, malicious software and unauthorized access in cyber security, on the fly recognition of unusual actions, movements and movements of employees, third parties, etc.;

3) health care: monitoring of all types and formats of data about employees (and visitors) for early detection of diseases and abnormal conditions, infections;

4) controlling production processes and operations: detection of defects and violations in production processes, technological operations to ensure proactive control of quality and efficiency;

5) detection of fraud, abuse and errors in financial activities: detection of unusual trading patterns and market anomalies, fraudulent or erroneous financial transactions/operations in the field of financial management of the food industry enterprise/holding;

6) environmental monitoring: detection of unusual environmental conditions and/or detection of unusual conditions, processes and manifestations in the internal technological environment of a food industry enterprise.

The authors highlight the following challenges and problems when applying the anomaly detection regime in the practice of all levels of management of food industry enterprises:

1. High dimensionality of all types of data. High dimensionality refers to datasets with a large number of features (variables). The data anomaly detection mode should be activated secondarily, i.e. only AFTER performing the normal mode of detecting widespread, accurate and complete patterns/patterns in the data. But in the conditions of modern big data, even this primary mode of Data Mining is complicated due to "the curse of data dimensionality".

2. Imbalanced and underrepresentative input data: Anomalies occur quite rarely, which is why unbalanced and underrepresentative input data sets make them difficult to detect.

3. Evolution of anomaly patterns: anomalies can evolve and change over time, requiring adaptive methods to detect them in real time.

4. Noise and outliers: distinguishing real anomalies and threats from stochastic informational noise and outliers in the form of mechanical or technical errors in the data can be quite difficult, requiring reliable detection algorithms and, even, a wider selection of human expertise (which makes it difficult to detect anomalies on large data in AutoML mode).

5. Time-consuming and labor-intensive, availability and cost of a comprehensive interpretation of identified potential anomalies in the data of a food industry enterprise. In the above-mentioned conditions of large data of high dimensions, sometimes unbalanced and not sufficiently representative, in the conditions of the dynamics of changing patterns of anomalies, the presence of information noise – the evaluation, interpretation and formalization of detected anomalies requires more and more time and experience of human expertise of preliminary results of detecting anomalies by Data Mining methods. The experience of the authors of this study shows that among the previously detected anomalies in the data, only 3–5 % pass this stage of human interdisciplinary examination and will be included in the enterprise's corporate knowledge base.

Future directions and prospects for improving the anomaly detection regime for food industry enterprises:

1. Detection of anomalies in real time: deployment of real-time streaming data processing and analysis architectures (Apach Kafka and Apache Spark etc.) to detect anomalies as they occur, "on the fly".

2. Deep learning: extensive use of architectures and methods of deep machine learning (Deep ANN) on large unstructured data – to increase the accuracy and speed of pre-processing (feature extraction). Analysis and analysis of complex patterns of unstructured data. It provides for automatic recognition and subsequent classification of big data video, audio, and photo formats for the purpose of detecting and identifying shortages/defects of food products, monitoring the functioning of equipment and facilities, and safety purposes of the food industry enterprise.

3. Explained artificial intelligence: development of methods to improve the interpretation and transparency of detected anomaly models (even further attempts to explain/verbalize the "black boxes" of trained shallow artificial neural networks (Example of verbalization of such an implicit model (black box), obtained from configured and trained by the authors of shallow An ANN with three hidden layers is given above, in **Fig. 2.5**).

4. Adaptive algorithms: Comprehensive use of adaptive algorithms that can learn and evolve with changing data patterns.

5. Integration with big data: integration of anomaly detection methods with big data platforms for efficient processing of large and diverse data sets on the fly.

Summarizing the results of the author's research on effective methods and tactical techniques for detecting anomalies in business and technological data of food enterprises, it should be emphasized that the mode of detecting anomalies in Data Mining is a critical area that allows identifying rare but important events in all functional areas of management of food enterprises industry. Using a range of statistical, machine learning and ensemble ML techniques, it is the anomaly detection mode that uncovers unexpected new, hidden and valuable information/knowledge that can help prevent fraud, abuse, errors, negligence, improve security, improve production results and ensure more thorough controls product quality of food industry enterprises. Constant progress in the field of real-time analysis and analytics, the involvement of deep learning architectures and adaptive algorithms with elements of soft-computations – will further expand the possibilities and effect of the application of the anomaly detection mode in the future in the food industry.

Hybrid Methods and Algorithms of Data Mining to support decision-making in complex, interdisciplinary, complex, multidimensional tasks of enterprise management (on the example of the food industry)

Hybrid Data Mining methods/algorithms combine numerous separate autonomous computing methods (from such areas as: complex multidimensional DB queries, Statistical Analysis, ML, Mathematical Programming, Soft Computations) to increase the accuracy, efficiency, completeness and stability of the results of intelligent in-depth data analysis.

The main tasks of Data Mining hybridization: hybrid clustering, hybrid classification and hybrid predictive modeling.

Hybrid Data Mining methods/algorithms involve the combination of two or more algorithms/methods to create a more efficient, adaptive analytical tool in order to overcome the limitations of individual stand-alone algorithms/methods by synergistically using their strengths.

Hybrid approaches are particularly useful in highly complex, interdisciplinary industries/productions/technologies of the food industry (baby food production, organic food production, diet food production, etc.).

Today, in the era of big data and in anticipation of multimodal crisis phenomena, the greatest potential and hopes are placed on a hybrid approach to processing, analytics and data analysis, which in these dynamic and unstable conditions will be a powerful tool for solving complex problems that cannot be overcome by "pure" methods. classic BI approaches and methods. The implementation and universality of such hybrid approaches is limited by the requirements for the uniformity of data types/formats and the specific variability of decision-making at different levels of management in different functional departments of the food industry enterprise. That is why, for the effective use of hybrid Data Mining scenarios, a unified/standardized presentation of data, metadata and knowledge is an important condition.

Below, the results of the analysis of the three main tasks of Hybrid Data Mining will be presented:

1. Hybrid Regression – most often implemented through technologies of ensemble ML. Ensemble methods combine multiple learning algorithms to achieve better predictive performance than any single algorithm alone. Key ensemble techniques include:

– Bagging (Bootstrap Aggregating) Combines the predictions of multiple models trained on different subsets of the data to reduce variance and improve accuracy;

– Boosting: sequentially applies weak classifiers to the data, adjusting their weights based on the accuracy of previous classifiers to reduce bias and variance;

– Random Forests: an ensemble of decision trees where each tree is trained on a random subset of features, enhancing model robustness and accuracy.

2. Hybrid clustering algorithms integrate different clustering techniques to enhance the quality and interpretability of clusters. Common approaches include:

– Hierarchical-K-means Clustering: combines the hierarchical and K-means clustering methods to first identify broad clusters and then refine them into more precise sub-clusters;

– Density-Based and Partitioning Methods: merges density-based clustering (e.g., DBSCAN) with partitioning methods (e.g., K-means) to identify clusters of varying shapes and densities.

3. Hybrid Classification algorithms integrate multiple classification techniques to improve prediction accuracy and generalization. Examples include:

– Neural Network and Decision Tree Hybrid: combines the high accuracy of neural networks with the interpretability of decision trees;

– Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) Hybrid: utilizes the robustness of SVM for boundary detection and the simplicity of KNN for local classification.

Currently, there is no precisely defined classification of hybrid information technologies in the scientific literature. Most often, options for the classification of hybrid systems are considered according to the following criteria: "homogeneity – heterogeneity", "hardness – softness", hybrid systems of the first order and hybrid systems of the second order, "equality – hierarchy".

It is worth noting that within Soft Computations, three aspects are improved (compared to Non-hybrid technologies): uncertainty management, retraining, adaptation.

Below, in **Fig. 2.12** – the given fragment of the base of fuzzy production rules for adaptive control of the auxiliary climate system for the small grain elevator/storage.

Using the above-mentioned fragment of the fuzzy production rules base for adaptive control of the grain elevator's auxiliary climate control system, the authors conducted a comparative simulation/modeling of the fuzzy technological control system and the classical one with hard technological rules – **Fig. 2.13**.

As it is possible to see from the figure, it can be claimed that the climate control system based on fuzzy logic adapts to changes more quickly and fluctuations in its operation are smaller. That is, the use of control systems based on fuzzy logic makes it possible not only to increase the level of flexibility and interpretability of control algorithms, but also to increase the resource of controlled systems and reduce their energy consumption.

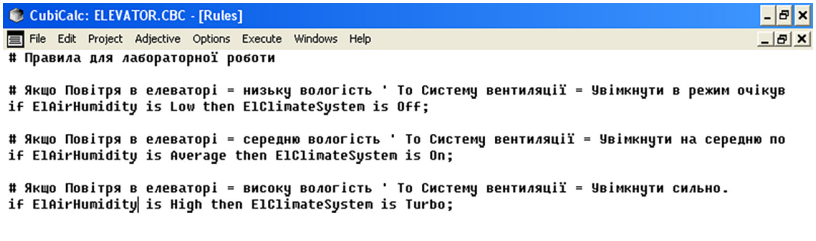


Fig. 2.12 The fragment of the fuzzy base of production rules for adaptive control of the auxiliary climate system for the small grain elevator
Note: developed by the authors

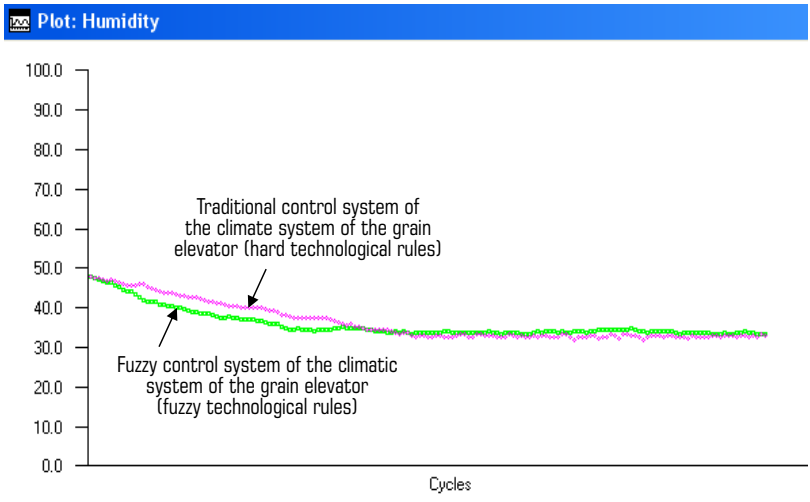


Fig. 2.13 Comparative simulation/modeling of the fuzzy technological climate control system and the classical one with hard technological rules (on the example of small grain elevator/storage)
Note: developed by the authors

Advantages and benefits of using Hybrid Data Mining:

– Robustness: hybrid approaches can handle a variety of data distributions and noise levels, making them more robust;

– Improved Accuracy: by combining different algorithms, hybrid methods often achieve higher predictive accuracy than individual methods;

– Scalability: many hybrid methods are designed to handle large datasets efficiently, making them suitable for big data applications;

– Flexibility: hybrid methods can be tailored to specific problems by selecting and combining appropriate algorithms.

Challenges and difficulties in using Hybrid Data Mining:

– Computational Cost: combining multiple algorithms may increase computational requirements;

– Complexity: hybrid methods can be more complex to implement and understand than single-method approaches;

– Integration: ensuring seamless integration of different algorithms and maintaining consistency can be difficult;

– Parameter Tuning: hybrid methods often require careful tuning of multiple parameters, which can be challenging.

Let's consider promising areas of Hybrid Data Mining:

– Automated Hybrid Systems: development of automated systems that can dynamically select and combine algorithms based on the data characteristics and problem requirements;

– Adaptive Hybrid Methods: creation of adaptive hybrid methods that can learn and evolve based on feedback and changing data patterns;

– Interdisciplinary Approaches: leveraging advances from other fields such as biology, physics, and social sciences to inspire novel hybrid Data Mining techniques;

– Interdisciplinary Approaches to hybridization: using algorithmic advances from other fields, such as biology, physics, and social sciences, to create new hybrid data analysis and modeling techniques (e.g., gravity search, ant colony method, etc.).

Hybrid intelligent systems, depending on the architecture, are divided into three main types:

– sequential hybrid systems (sequentially process data using different algorithms/methods). For example: the analytics subsystem of a food industry enterprise integrates and pre-processes input from users (for example, attracting posts and reactions on social networks with TextMining algorithms in Sentiment Analysis mode) before retraining a regression model for forecasting demand for new types of products;

– parallel hybrid systems (several intelligent methods/algorithms are simultaneously applied to the same data). For example: within the analytics subsystem of a food industry enterprise for optimal planning, dispatching and synchronization of loading and operation modes of logistics (transport and warehouse) equipment – can use parallel and competitive different (classical and alternative) methods of mathematical programming and different sets of optimization parameters of these methods with for the purpose of further comparison of the results of solving such a complex optimization problem (below, in **Fig. 2.14** – the author's example of solving such an optimization problem using a rather innovative (for the classic food industry management) method of genetic algorithms is shown);

– hierarchical hybrid intelligent systems (presuppose cybernetic organization of intelligent system components in a hierarchical structure, where higher-level components control lower-level components). For example: a distributed artificial intelligence system for 24/7/365 diagnostics of very different technological equipment and equipment of a food industry enterprise (to identify various types of current malfunctions and predict future ones).

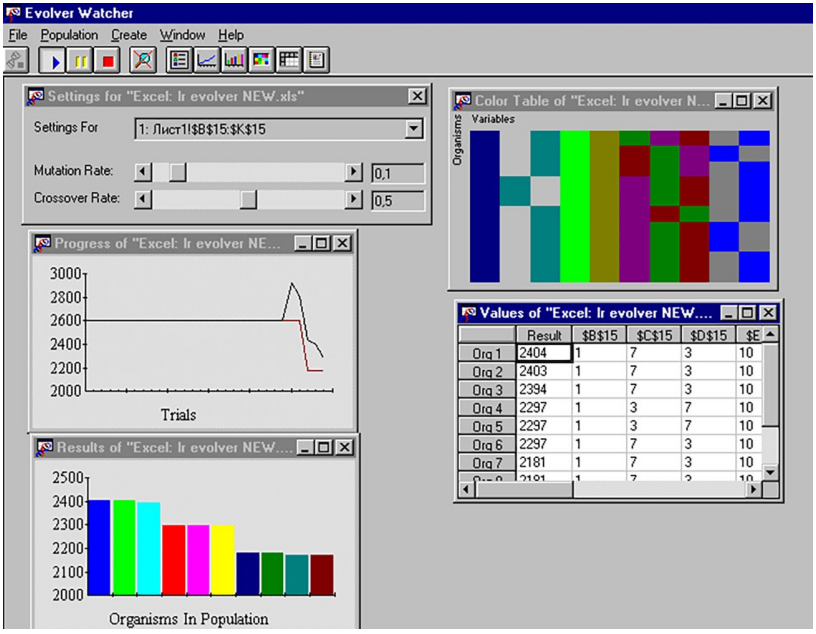


Fig. 2.14 The use of genetic algorithm to solve the complex problem of optimal planning, dispatching and synchronization of loading and operation modes of logistics (transport and warehouse) equipment on the example of a food industry enterprise (population size=10; mutation=0.1; crossover=0.5; stopping conditions=1000 trials)
Note: developed by the authors

Hybrid technologies, Data Mining algorithms/methods and their use scenarios represent an innovative approach to increasing the accuracy, completeness, robustness (i.e. integrated reliability and quality) of Data Mining (also and Data Science) of big data of food industry enterprises. By synergistically combining the strengths, advantages, or radically/absolutely different, or similar, related methods/algorithms of DM, overcoming/overcoming their certain shortcomings and limitations – it is hybrid technologies that will help solve complex problems of enterprises and companies (especially in the food industry) – much more effectively.

Future advances in automated and self-adaptive hybrid intelligent technologies and systems will provide additional potential for sustainability and competitiveness in various functional areas and management levels of food industry enterprises and companies (especially in the context of multimodal crises).

Data Mining for crisis management (in particular for enterprises and companies of the food industry) involves systematic analysis of all available internal and external data sets to identify early warning signs, data anomalies, monitor states, activities and events in real time 24/7/365

for further reactive generation and evaluation of the effectiveness of response strategies to identified crisis situations. In other words, crisis data mining is the process of finding, extracting and formalizing new, hidden and useful patterns/patterns from large data sets to aid in optimal crisis management and adaptive response. This process increases the resilience of the organization, helps reduce risks and supports informed decision-making during crises.

That is why, the urgent tasks of crisis data mining are a comprehensive study of the specifics of the methodology, effective tasks and challenges/cautions related to the intelligent analysis of crisis data, emphasizing the importance of crisis data mining for improving the decision-making process during various types of crises, including natural disasters, pandemics and socio-political disturbances, military actions.

The complexity and unpredictability of crises requires reliable and thorough approaches based on large (structured, semi-structured and even unstructured) data for effective detection of anomalies and threats.

Let's outline below the applied functional tasks/directions of using crisis datamining:

1. Natural disasters (earthquake prediction through seismic data analysis to predict potential earthquakes; flood monitoring through satellite imagery and weather data analysis to predict and monitor floods, etc.).

2. Pandemics (detection of disease outbreaks by searching for early signs of disease outbreaks in social networks and medical records; distribution of medical resources, i.e. optimization of the distribution of medical goods based on predictive models of the development of pandemics, etc.).

3. Socio-political crises (prediction of the development of socio-cultural and/or political conflicts through monitoring of social networks and news to assess and forecast the zones, causes and epicenter of the conflict; displacement tracking, i.e. using Data Mining to track the spontaneously displaced population and provide them with preventive assistance, etc.).

As such, crisis intelligence is a critical tool in modern crisis management, offering the potential to significantly improve response times and outcomes, and reduce potential harm. Continuous progress in data collection, their pre-stream processing, machine learning (including advanced anti-crisis analytics in the formed Data Lakes) and interdisciplinary interpretation are essential to overcome current challenges and maximize the effectiveness of crisis data mining.

Below, the results of the author's case studies will be briefly presented regarding the identified directions of future development and trends of Data Mining in particular (and sometimes Data Science in general) in the dynamic, unstable external conditions of food industry enterprises

Data mining involves the use of advanced computing techniques to extract meaningful information from complex data sets. As the volume and complexity of data continues to grow, the need for more sophisticated analysis tools becomes increasingly critical. In this subsection, the future directions and potential consequences of intelligent data analysis are considered, so the author highlights the following new trends and perspectives of Data Mining:

1. Automated machine learning (AutoML). Automated machine learning aims to automate the end-to-end process of applying machine learning to real-world problems. Future developments

in AutoML are expected to: increase accessibility, i.e. lower the barrier for non-experts by simplifying the process of creating, configuring and deploying machine learning models; efficiency gains: Streamline model selection, hyperparameter tuning, and feature development processes, reducing time and computational resources required.

2. Exploratory and in-depth real-time analysis and analytics. Real-time analytics involves continuous analysis of streaming data to provide instant insights and facilitate rapid decision-making. Future advances are likely to include: real-time decision-making – integrating real-time analytics with business operations to facilitate immediate response to emerging trends and anomalies; improve predictive maintenance: Use real-time data to predict equipment failures and optimize maintenance schedules in industrial applications.

3. Expanded and improved cross-industry, cross-task interpretation of Data Mining results. As machine learning models become more complex, ensuring their interpretability is critical. Future developments in model interpretation are expected to: increase transparency: develop methods to make complex models more understandable and transparent to stakeholders; increasing trust: Increase trust in and acceptance of AI systems by providing clear explanations of model solutions.

4. Integration of diverse and multi-format data sources of varying quality with the formation of specialized Data Lakes. Combining data from different sources can provide more comprehensive information. Future directions for such NON-homogeneous data integration include: data fusion: developing advanced data fusion techniques to seamlessly integrate structured and unstructured data from different domains; interdisciplinary collaboration: fostering collaboration across disciplines to leverage diverse data sets and generate holistic understanding.

5. Ethical considerations regarding total Data Mining. The ethical implications of data mining are becoming increasingly important. Key future areas include: Bias Mitigation: Developing methods to detect and mitigate biases in data and models to ensure equity and fairness; data privacy: improving data privacy protection mechanisms to protect sensitive information while enabling data analysis.

6. Potential systemic and total impact of innovative Data Mining. Advances in data mining are expected to have a profound impact on all areas of the food industry: ERP, CRM, MES, WMS, EAM, HRM; monitoring and proactive actions to ensure sanitary, biological and food safety; in the field of financial management – improved fraud/abuse/error detection and scenario-based long-term risk management (especially insurance risk management for food industry [35, 36]); investment management of food industry enterprises using more accurate and adaptive predicative models; in the field of production management – more accurate and systematic technological forecasting and better optimization of complex production processes and their components (see **Fig. 2.15** the author's example of the trained shallow ANN for forecasting product output (under the influence of external stochastic factors) depending on the number/ volume of 4 components: HCl, NH_3, H_2O and the amount of chemical reaction catalyst); proactive TQM in real time; optimal management of technical maintenance, current and capital repairs of equipment and facilities of food industry enterprises; etc.

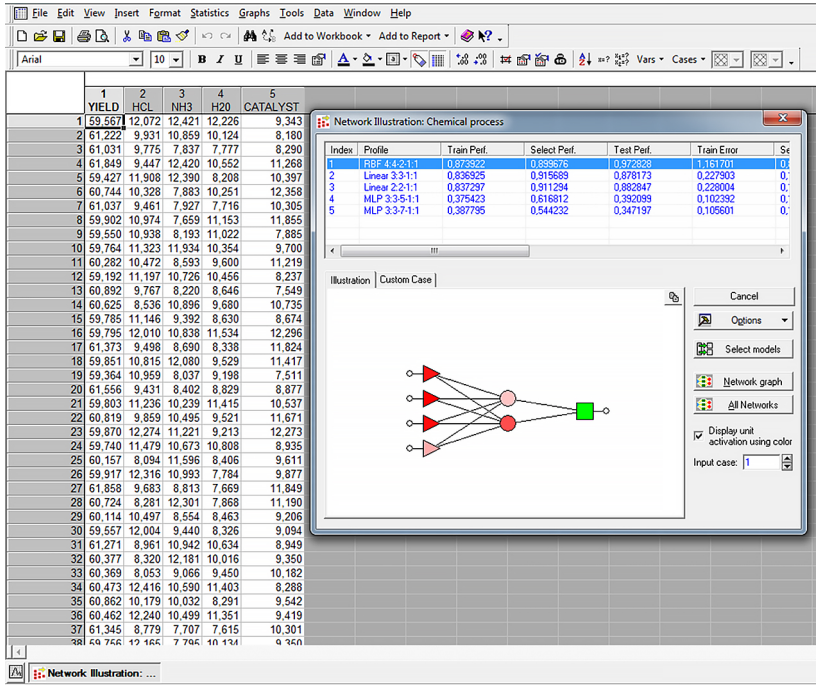


Fig. 2.15 The example of the configured and trained shallow ANN for forecasting the output of ultra-processed food product (under the conditions of the action of the complex of external stochastic factors) depending on the amount/volume of 4 components and the amount of catalyst for the chemical and technological reaction
Note: developed by the authors

Thus, the future prospects of Data Mining in particular (and sometimes Data Science in general) thanks to progress in all types of machine learning; technologies for effective organization of data integration, their processing, analysis and analytics in real time 24/7/365; complex in-depth expert interpretation of results; integration of data, knowledge and relevant ethical regulatory mechanisms – are relevant (for enterprises and food industry companies in particular). These areas of development and improvement of Data Mining in particular (and sometimes Data Science in general) should qualitatively improve the effectiveness of extracting valuable, new and hidden regularities/patterns/insights from large multidimensional low-quality data of all formats, will contribute to the adoption of operational, informed proactive decisions at all levels management, in all functional areas and all sectors of the food industry. Further continuation of research and interdisciplinary, inter-project cooperation will be necessary to realize the full potential of Data Mining

in particular (and sometimes Data Science in general) to increase the integrated sustainability, efficiency, competitiveness of enterprises and companies in the food industry, especially in multimodal crisis external conditions.

CONCLUSIONS

Proposed in this publication scientific and practical applied solutions regarding Data Mining for enterprises and companies (on the example of food industry) involve the application of advanced cybernetic computing methods/algorithms, technological modes and scenarios (for integration, pre-processing, machine learning, testing and in-depth comprehensive interpretation of the results) of analysis and analytics of large structured and semi-structured data sets for training high-quality descriptive, predictive and even prescriptive models.

The proposed by authors multi-mode adaptive Data Mining synergistically combines in parallel and sequential scenarios: methods of preliminary EDA, statistical analysis methods, business intelligence methods, classical machine learning algorithms and architectures, advanced methods of testing and verification of the obtained results, methods of interdisciplinary empirical expert interpretation of results, knowledge engineering formats/techniques – for discovery/detection previously unknown, hidden and potentially useful patterns, relationships and trends.

The main methodological and technological goal of this developed methodology of multi-mode adaptive Data Mining for food industry enterprises is to increase the completeness (support) and accuracy of business and technical-technological modeling on all levels of management of food industry enterprises: strategic, tactical and operational.

By optimally configuring hyperparameters, parameters, algorithms/methods and architecture of multi-target and multidimensional explicit and implicit descriptive and predicative models, using high-performance hybrid parallel soft computing for machine learning – the improved methodology of multimode Data Mining (proposed by the authors) allows to find/detect/mine for new, useful, hidden corporate knowledge from previously collected, extracted, integrated Data Lakes, stimulating the overall efficiency, sustainability, and therefore competitiveness, of food industry enterprises at various organizational scales (from individual, craft productions to integrated international holdings) and in various food product groups and niches.

But, summarizing the above, it is worth emphasizing, that even a very experienced team of specialists/experts will not give an unequivocal effective recommendation on the first attempt, which Data Mining mode/algorithm/scenario, in which configuration of global hyperparameters and method-oriented parameters – will work most effectively (in a specific sector of the food industry, a specific region, for a specific enterprise at this time) without a set of additional studies, a set of experiments and a subsequent series of tests/trials.

What is why, the decision the problems (previously detected by the authors in BI practice of food industry) related to: multidimensionality of input big data; their verifiability and representativeness;

internal inconsistency, duplication and damage to data integrity; hidden presence of various types of data biases; informational noise and interferences; data outliers; missing and/or corrupted data) – are important determinant factors of the proposed concept/methodology of effective multi-mode Big Data Mining for food industry enterprises (for example, in production of vegetable proteins and synthetic meat, in production healthy snacking, in reducing harm of ultra-processed foods, in foods upcycling, for foods traceability, for farm-to-table and local sourcing initiatives, for circular economy practices in food waste reduction, for connected and transparent food supply chains, for net-zero & scope 3 emissions audit, for monitoring of crackdown on greenwashing etc.).

The Online Data Mining mode at food industry enterprises involves the concept, methodology and technology of online rapid data analysis and analytics in real time 24/7/365. This approach differs from traditional Data Mining modes, therefore it is relevant and crucial for those tasks and functions of food industry enterprise management that require maximum immediate (but often approximate) analysis of streaming data and operational response, emphasizing the relevant role of such mode in dynamic and time-sensitive applications in the food industry (in particular, in continuous production chemical and technological processes, production conveyors, etc.).

That is why the authors in this study especially paid considerable attention to the special mode of Anomaly & Fraud Detection Data Mining within the framework of proactive adaptive management in the food industry, which involves 24/7/365 monitoring/analysis/analytics of large data to identify anomalous/frauds/threatening patterns and trends, that help become the basis for proactive strategies, tactics and responses. That is this (Anomaly & Fraud Detection) Data Mining mode involves monitoring and detecting outliers, anomalies in streaming and batch, structured/semi-structured/unstructured big data regarding commercial, financial, technological, logistics, HR, marketing functions of food industry enterprise/corporation/holding in order to recognize maximum possible threats in real time mode. Using this mode, food industry enterprises can increase their robustness, sustainability by improving the efficiency of proactive decision-making in pre-crisis and crisis periods.

In addition, within the framework of this study, the methodology of effective Hybrid Data Mining is proposed (taking into account the specifics of enterprises and companies of the food industry), because it has been proven that it is the hybrid methods/algorithms of Data Mining indeed represent a powerful approach to increasing the accuracy, support, reliability and representativeness of Data Mining models trained from all types and formats of big data for a food industry enterprise. I.e. by combining the strengths of different methods, hybrid data mining methods can solve most complex enterprise/company problems more effectively than single-method/algorithm-oriented approaches. Future advances in automated and multilevel hybrid systems hold the potential for further expanding the capabilities and applied applications of Data Mining in complicated&complex functional areas and product sectors of food industry enterprises (in particular, especially in the modern high-tech food production: farm optimization with precision farming, regenerative & agro-ecological farming, automation in food storage warehouses and supermarkets etc.).

It is also worth highlighting the improved mode of Ad-hoc Data Mining for enterprises and companies of the food industry, which is a relevant and important mode in today's dynamic and

stochastic (and therefore difficult to predict) macroeconomic and industry environment. It is this mode that offers the flexibility and speed needed to address the urgent and critical analytical needs of a food industry enterprise/company in today's environment.

Using the above-mentioned advances in Crisis Data Mining mode and, thanks to deep interdisciplinary expertise, enterprises and companies in the food industry will be able to use the full potential of Data Mining to achieve vertical and horizontal adaptive resilience and sustainability, even in times of crisis.

It is worth noting that by effectively using the fireproofed modes of Data Mining of a food industry enterprise, it is possible to reach a qualitatively top new level in Descriptive, Predictive and Prescriptive Analysis and Analytics of management system with help of such additional innovated technologies, as: automated machine learning (AutoML), streaming data analytics in real time, improved and in-depth cross-industry interpretation of results, innovative pre-processing and integration of various big data sources (in particular, using not only ETL or ELT, but also such technologies of integration and preliminary preparation of input data as: Change Data Capture, Data replication, Data virtualization, Stream Data Integration), new methods and scenarios of ensemble semi-supervised machine learning, multilevel and ternary hybridization of Data Science methods, distributed and multi-agent AI.

Additionally, the authors emphasize that Big Data Mining in the food industry is currently very promising in the context of R&D in the field of genetics, which involves the use of advanced ML methods/algorithms and ML technologies for the extraction/discovery/search/mining of new, useful, hidden genetic patterns from large sets of data related to food genetics (for example, analysis and analytics of the genetic composition of: agricultural crops, in livestock industry, in fishing industry etc.). Using the above proposed advanced innovative computational AI modes, researchers and industry professionals can extract valuable insights from genetic data, leading to innovations in crop and livestock production, food safety, and personalized nutrition. It is the above-proposed improved modes, scenarios and Data Mining technologies that should be widely used for the food industry, in particular, for: improving crop production, livestock breeding, food safety and quality, nutritional personalized genomics, effective&productive organic products etc. Therefore, data mining in the food industry, especially in the field of genetics, has a huge potential to improve productivity of food production, safety and quality, and of course, to reduce industry risks [37, 38] and, eventually, to improve the financial stability of food industry enterprises/companies.

It can not be argued, that in modern global macroeconomic standings (pre-crisis, crisis and post-crisis conditions of both regional food industries and the global world food industry; globalization and simultaneous very narrow specialization, often personalization of the food industry selected sectors; the need to take into account a huge amount of stream and packet information of various formats from different sources; the need for a quick adaptive optimal management actions in response to rapid changes in the global or regional market situation; unstable and difficult to predict dynamics of external influences: international, national, sectoral, local direct regulatory and indirect (from civil society and public organizations) regulations of the food industry standards) – deployment of the multi-mode adaptive Data Mining methodology proposed by the authors – will result

in enterprises, companies and organizations of the food industry shall gained additional competitive advantages at the national state, regional, branch and corporate management levels.

Moreover, the proposed scientific and practical applied solutions, approaches, technologies, modes, configurations, settings (regarding High Dimensionality of Input Data, online Data Mining, ad-hoc Data Mining, hybrid Data Mining, crisis Data Mining, Data Mining for Anomaly & Fraud Detection) will be also particularly effective for enterprises and companies in countries and regions, in that industries (i.e., not only in the food industry) – where managerial decision-making requires complex expert analysis, is associated with significant capital risky investments, there are many branches of scenarios and nodes of risky decision-making, difficult-to-forecast negative effects of stochastic and dynamic external factors are possible, there are open uncertainties situations (probabilities and fuzziness), specific industry/sectoral/product risks.

REFERENCES

1. Ostapenko, T., Onopriienko, D., Hrashchenko, I., Palyvoda, O., Krasniuk, S., Danilova, E. (2022). Research of impact of nanoeconomics on the national economic system development. Innovative development of national economies. Kharkiv: PC TECHNOLOGY CENTER, 46–70. doi: https://doi.org/10.15587/978-617-7319-64-0.ch2
2. Palyvoda, O., Karpenko, O., Bondarenko, O., Bonyar, S., Bikfalvi, A. (2018). Influence of network organizational structures on innovation activity of industrial enterprises. Problems and Perspectives in Management, 16 (3), 174–188. https://doi.org/10.21511/ppm.16(3).2018.14
3. Illiashenko, S., Bilovodska, O., Tsalko, T., Tomchuk, O., Nevmerzhytska, S., Buhas, N. (2022). Opportunities, Threats and Risks of Implementation the Innovative Business Management Technologies in the Post-Pandemic Period COVID-19. WSEAS transactions on business and economics, 19, 1215–1229. https://doi.org/10.37394/23207.2022.19.107
4. Krasnyuk, M., Nevmerzhytska, S., Tsalko, T. (2024). Processing, analysis & analytics of big data for the innovative management. Grail of Science, 38, 75–83. https://doi.org/10.36074/grail-of-science.12.04.2024.011
5. Krasnyuk, M., Elishys, D. (2024). Perspectives and problems of big data analysis analytics for effective marketing of tourism industry. Science and Technology Today, 4 (32), 833–857. https://doi.org/10.52058/2786-6025-2024-4(32)-833-857
6. Krasnyuk, M. T., Hrashchenko, I. S., Kustarovskiy, O. D., Krasniuk, S. O. (2018). Methodology of effective application of Big Data and Data Mining technologies as an important anti-crisis component of the complex policy of logistic business optimization. Economies' Horizons, 3 (6), 121–136. https://doi.org/10.31499/2616-5236.3(6).2018.156317
7. Tsalko, T., Nevmerzhytska, S., Didenko, Ye., Kharchenko, T., Bondarenko, S. (2020). Optimization of goods implementation on the basis of development of business process re-engineering. Journal of Management Information and Decision Sciences, 23 (2). Available at: https://

www.abacademies.org/articles/optimization-of-goods-implementation-on-the-basis-of-development-of-business-process-reengineering-9174.html

8. Krasnyuk, M., Hrashchenko, I., Goncharenko, S., Krasniuk, S., Kulynych, Y. (2023). Intelligent management of an innovative oil and gas producing company under conditions of the modern system crisis. Access Journal – Access to Science, Business, Innovation in the Digital Economy, 4 (3), 352–374. https://doi.org/10.46656/access.2023.4.3(2)
9. Turlakova, S., Bohdan, B. (2023). Artificial intelligence tools for managing the behavior of economic agents at micro level. Neuro-Fuzzy Modeling Techniques in Economics, 12, 3–39. https://doi.org/10.33111/nfmte.2023.003
10. Krasnyuk, M. T., Hrashchenko, I. S., Kustarovskiy, O. D., Krasniuk, S. O. (2018). Methodology of effective application of Big Data and Data Mining technologies as an important anti-crisis component of the complex policy of logistic business optimization. Economies' Horizons, 3 (6), 121–136. https://doi.org/10.31499/2616-5236.3(6).2018.156317
11. Y. Krasnyuk, M., Kulynych, Y., Krasniuk, S. (2022). Knowledge discovery and data mining of structured and unstructured business data: problems and prospects of implementation and adaptation in crisis conditions. Grail of Science, 12-13, 63–70. https://doi.org/10.36074/grail-of-science.29.04.2022.006
12. Tereshchenko, S., Kulinich, T., Matienko, V., Tymchyna, Y., Nevmerzhytska, S., Ievseitseva, O. (2023). Financial and marketing innovative management of ecological agribusiness. Financial and Credit Activity Problems of Theory and Practice, 6 (53), 487–500. https://doi.org/10.55643/fcaptp.6.53.2023.4255
13. Ishchejkin, T., Liulka, V., Dovbush, V., Zaritska, N., Puzyrova, P., Tsalko, T. et al. (2022). Information subsystem of agri-food enterprise management in the context of digitalization: the problem of digital maturity. Journal of Hygienic Engineering and Design, 38, 243–252. Available at: https://keypublishing.org/jhed/jhed-volumes/jhed-volume-38-fpp-18-tymur-ishchejkin-viktoriia-liulka-vita-dovbush-nadiia-zaritska-polina-puzyrova-tetiana-tsalko-svitlana-nevmerzhytska-yuliia-rusina-olena-nyshenko-svitlana-bebko/
14. Derbentsev, V. D., Bezkorovainyi, V. S., Matviychuk, A. V., Hrabariev, A. V., Hostryk, A. M. (2023). A comparative study of deep learning models for sentiment analysis of social media texts. CEUR Workshop Proceedings, 3465, 168–188.
15. Derbentsev, V., Kibalnyk, L., Radzihovska, Yu. (2019). Modelling multifractal properties of cryptocurrency market. Periodicals of Engineering and Natural Sciences (PEN), 7 (2), 690–701. https://doi.org/10.21533/pen.v7i2.559
16. Derbentsev, V., Bezkorovainyi, V., Akhmedov, R. (2020). Machine learning approach of analysis of emotional polarity of electronic social media. Neuro-Fuzzy Modeling Techniques in Economics, 9, 95–137. https://doi.org/10.33111/nfmte.2020.095
17. Derbentsev, V., Velykoivanenko, H., Datsenko, N. (2019). Machine learning approach for forecasting cryptocurrencies time series. Neuro-Fuzzy Modeling Techniques in Economics, 8, 65–93. https://doi.org/10.33111/nfmte.2019.065

18. Carneiro, F., Miguéis, V. (2021). Applying Data Mining Techniques and Analytic Hierarchy Process to the Food Industry: Estimating Customer Lifetime Value. Proceedings of the International Conference on Industrial Engineering and Operations Management. Kotler, 2000, 266–277. https://doi.org/10.46254/sa02.20210125
19. Vlontzos, G., Pardalos, P. M. (2017). Data mining and optimisation issues in the food industry. International Journal of Sustainable Agricultural Management and Informatics, 3 (1), 44–64. https://doi.org/10.1504/ijsami.2017.082921
20. Hajiha Hajiha, A., Radfar, R., Malayeri, S. S. (2011). Data mining application for customer segmentation based on loyalty: An iranian food industry case study. 2011 IEEE International Conference on Industrial Engineering and Engineering Management, 504–508. https://doi.org/10.1109/ieem.2011.6117968
21. Shekhar, H., Sharma, A. (2023). Global Food Production and Distribution Analysis using Data Mining and Unsupervised Learning. Recent Advances in Food, Nutrition & Agriculture, 14 (1), 57–70. https://doi.org/10.2174/2772574x14666230126095121
22. Massaro, A., Dipierro, G., Saponaro, A., Galiano, A. (2020). Data Mining Applied in Food Trade Network. International Journal of Artificial Intelligence & Applications, 11 (2), 15–35. https://doi.org/10.5121/ijaia.2020.11202
23. Ting, S. L., Tse, Y. K., Ho, G. T. S., Chung, S. H., Pang, G. (2014). Mining logistics data to assure the quality in a sustainable food supply chain: A case in the red wine industry. International Journal of Production Economics, 152, 200–209. https://doi.org/10.1016/j.ijpe.2013.12.010
24. Holimchayachotikul, P., Phanruangrong, N. (2010). A Framework for Modeling Efficient Demand Forecasting Using Data Mining in Supply Chain of Food Products Export Industry. Proceedings of the 6th CIRP-Sponsored International Conference on Digital Enterprise Technology. Berlin Heidelberg: Springer, 1387–1397. https://doi.org/10.1007/978-3-642-10430-5_106
25. Hayashi, Y., Hsieh, M.-H., Setiono, R. (2009). Predicting consumer preference for fast-food franchises: a data mining approach. Journal of the Operational Research Society, 60 (9), 1221–1229. https://doi.org/10.1057/palgrave.jors.2602646
26. Taghi Livari, R., Zarrin Ghalam, N. (2021). Customers grouping using data mining techniques in the food distribution industry (a case study). SRPH Journal of Applied management and Agile Organisation, 3 (1), 1–8. https://doi.org/10.47176/sjamao.3.1.1
27. Liu, L.-M., Bhattacharyya, S., Sclove, S. L., Chen, R., Lattyak, W. J. (2001). Data mining on time series: an illustration using fast-food restaurant franchise data. Computational Statistics & Data Analysis, 37 (4), 455–476. https://doi.org/10.1016/s0167-9473(01)00014-7
28. Liao, C.-W. (2009). Consumers' behavior in the food and beverage industry through data mining. Journal of Information and Optimization Sciences, 30 (4), 855–868. https://doi.org/10.1080/02522667.2009.10699915
29. Jiménez-Carvelo, A. M., González-Casado, A., Bagur-González, M. G., Cuadros-Rodríguez, L. (2019). Alternative data mining/machine learning methods for the analytical evaluation of

food quality and authenticity – A review. Food Research International, 122, 25–39. https://doi.org/10.1016/j.foodres.2019.03.063

30. Krasnyuk, M., Krasniuk, S. (2020). Comparative characteristics of machine learning for predicative financial modelling. Scientific bulletin ΛΌГОΣ, 55–57. https://doi.org/10.36074/26.06.2020.v1.21
31. Krasnyuk, M., Tkalenko, A., Krasniuk, S. (2021). Results of analysis of machine learning practice for training effective model of bankruptcy forecasting in emerging markets. Scientific bulletin ΛΌГОΣ. https://doi.org/10.36074/logos-09.04.2021.v1.07
32. Krasnyuk, M., Krasniuk, S. (2021). Modern practice of machine learning in the aviation transport industry. Scientific bulletin ΛΌГОΣ. https://doi.org/10.36074/logos-30.04.2021.v1.63
33. Krasnyuk, M., Krasniuk, S., Goncharenko, S., Roienko, L., Denysenko, V., Liubymova, N. (2023). Features, problems and prospects of the application of deep machine learning in linguistics. Bulletin of Science and Education, 11 (17), 19–34. Available at: http://perspectives.pp.ua/index.php/vno/article/view/7746/7791
34. Krasnyuk, M., Hrashchenko, I., Goncharenko, S., Krasniuk, S. (2022). Hybrid application of decision trees, fuzzy logic and production rules for supporting investment decision making (on the example of an oil and gas producing company). Access Journal – Access to Science, Business, Innovation in the Digital Economy, 3 (3), 278–291. https://doi.org/10.46656/access.2022.3.3(7)
35. Shirinyan, L., Arych, M. (2019). Impact of the insurance costs on the competitiveness of food industry enterprises of Ukraine in the context of the food market security. Ukrainian Food Journal, 8 (2), 368–385. https://doi.org/10.24263/2304-974x-2019-8-2-15
36. Shirinyan, L., Arych, M., Rohanova, H. (2021). Influence of Insurance on Competitiveness of Food Enterprises in Ukraine. Management Theory and Studies for Rural Business and Infrastructure Development, 43(1), 5–12. https://doi.org/10.15544/mts.2021.01
37. Arych, M., Levon, M., Arych, H., Kulynych, Y. (2021). Genetic testing in insurance: implications for non-EU insurance markets as a part of the European integration process. On-line Journal Modelling the New Europe, 36, 25–52. https://doi.org/10.24193/ojmne.2021.36.02
38. Arych, M., Kuievda, I., Dvořák, M., Hinke, J. (2023). The farming costs (including insurance) of the agricultural holdings in the European Union. Journal of International Studies, 16 (1), 191–205. https://doi.org/10.14254/2071-8330.2023/16-1/13

DOI: 10.15587/978-617-8360-03-0.CH3

Viktor Myronenko, Andrii Pozdniakov,
Yaroslav Ziubryk, Valerii Samsonkin, Ivan Riabushko

CHAPTER 3

PROJECT MANAGEMENT OF UKRAINE'S INTEGRATION INTO THE TRANS-EUROPEAN TRANSPORT NETWORK

ABSTRACT

The chapter of the monograph "Project management of Ukraine's integration into the Trans-European transport network" is devoted to a comprehensive study of the management aspects of the integration of Ukrainian railways into the international transport infrastructure. It consists of three parts that analyze in detail the planning, implementation and specifics of management practices in the context of high-speed railways and integration projects.

The first part, "Planning of high-speed railway projects – global experience and Ukrainian perspectives", focuses on the global experience of planning and implementation of high-speed railway projects, in particular on examples from Europe, Asia and other regions. It analyzes various management models, technological innovations and economic aspects that contribute to the successful implementation of such projects. Special attention is paid to the adaptation of world practices to Ukrainian conditions, taking into account the specifics of the national infrastructure, economic and political factors.

The second part, "The project of the integration of Ukrainian railways into the Trans-European transport network (TEN-T) on the example of the development of the Lviv railway node", considers a specific case of the integration of Ukrainian railways into the Trans-European transport network. The stages of project implementation are described on the example of the Lviv railway node, including technical, organizational and financial aspects. An important role is played by the analysis of problems and achievements, as well as the impact of the project on the development of regional infrastructure and the economy.

The third part, "Peculiarities of the practical implementation of system-level project management in railway transport of Ukraine", focuses on the specifics of managing large projects in the railway transport system of Ukraine, using the example of the SAIRS-UZ project. It examines the practical aspects of management processes, including planning, implementation and monitoring of projects at the system level. Examples of successful and problematic projects illustrating the opportunities and challenges of the management process in the context of Ukrainian railway transport are considered separately.

The chapter of the monograph is addressed to teachers, researchers, graduate students, students, as well as practitioners in the field of project management, transport engineering and international integration. It offers in-depth analysis and practical recommendations to improve

management practices in the field of rail transport and the integration of national infrastructure into international transport networks.

KEYWORDS

Project management; railway transport; high-speed railways; integration; Trans-European transport network; Lviv railway node; infrastructure; rolling stock; vehicles; technical devices; international standards; planning; implementation of projects; transport integration; Ukrainian perspectives; economic aspects; technological innovations; system level; practical aspects; management processes; regular monitoring; problems and achievements; infrastructure development.

A project is defined as a temporary enterprise designed to create a unique product, service or result. A project has a defined beginning and end, as well as certain time, resource, and quality constraints. According to the PMBOK standard, a project is a unique activity that has a beginning and an end, and is also aimed at achieving specific goals [1].

Project management is the process of applying knowledge, skills, tools, and techniques to project activities to meet or exceed stakeholder expectations. This process includes the following phases: initiation, planning, implementation, monitoring and control, as well as project completion [1]. Project management is aimed at balancing the triangle of constraints: scope of work, time and cost. A fourth component is often added to this – quality.

The scope of work determines what exactly should be done in the project. This is a description of all the tasks and works necessary to achieve the goals of the project. Determining the scope of work is critically important to avoid changes in the project implementation process that may lead to an increase in cost or time delay [2].

Time in the context of project management includes planning, allocating, and controlling the completion of work on time. Effective time management allows to minimize the risks of delays and ensure timely implementation of the project [2]. For this, various methods and techniques are used, in particular, Gantt charts, critical path method (CPM) and program evaluation and analysis technique (PERT).

Project cost management includes the processes involved in planning, estimating, budgeting, financing, managing and controlling costs so that the project is completed within the approved budget. This includes estimating the cost of resources to be used in the project, as well as cost control during its implementation [3].

Quality in the project is defined as a measure of how well the final result meets the established requirements and expectations of the customer and interested parties. Quality management includes the processes of planning, assurance and quality control, which guarantees compliance of project results with standards and specifications [3].

Sectoral features of project management Each industry has its own project management features. For example, in construction, an important role is played by risk management and control over the

performance of work at all stages of the project [4]. In the IT industry, special attention is paid to flexible project management methodologies, such as Agile and Scrum, which allow quick adaptation to changing requirements [5]. Resource management and logistics are important in the manufacturing industry [6].

Thus, project management is a complex and multifaceted process that includes planning, organization, execution and control of various aspects of activity. Understanding the basic concepts and their features in various fields is key to successful project implementation.

3.1 PLANNING OF HIGH-SPEED RAILWAY PROJECTS – WORLD EXPERIENCE AND UKRAINIAN PERSPECTIVES

3.1.1 EXPERIENCE OF A SUCCESSFUL PROJECT IN THE TRANSPORT INDUSTRY

Successful project in the field of high-speed railways (China) – "Construction of high-speed railway Beijing-Shanghai" project.

High-speed railways (HSRs) represent an advanced sector of transport infrastructure that provides fast, convenient and economical transportation of passengers over long distances. The high-speed railway project between Beijing and Shanghai is a vivid example of the successful implementation of a large-scale infrastructure project. The project not only demonstrated efficiency in the management of major infrastructure initiatives, but also served as an impetus for the further development of high-speed rail transport in China and beyond.

The project became one of the largest infrastructure projects of the 21st century. The 1,318 km long line connects the two largest economic centers of China – Beijing and Shanghai. Construction began in 2008, and the official launch of the line took place in June 2011. The project included not only the construction of new tracks, but also the modernization of existing facilities, which made it possible to increase the speed of trains to 300 km/h. This provided a significant reduction in travel time between cities from 10 hours to approximately 4 hours [7].

Designing and planning for the implementation of the Beijing-Shanghai high-speed railway required a comprehensive approach, therefore it included the following stages and areas of activity:

1. Assessment of needs and opportunities: at the initial stages of the project, a comprehensive analysis of traffic flows, demand for transportation and potential economic benefits was conducted. This made it possible to determine the optimal speed and frequency of train movement [8].

2. Geodetic and environmental studies: to ensure the stability and safety of the construction, detailed geodetic studies were carried out, including an assessment of possible environmental impacts. This made it possible to reduce risks and take into account all factors that could affect the implementation of the project [9].

3. Financing and budgeting: investments in the project amounted to more than 33 billion USD. Both public and private investments, as well as international creditors, were involved to provide financing [10].

4. Project management: project management was carried out through an integrated structure that included the national railway corporation (China Railway Corporation) and local authorities. Modern project monitoring and management systems were used to control quality and deadlines [7].

The implementation of the project involved several key stages:

1. Construction of infrastructure: construction of new tracks, bridges, tunnels and railway stations. This included the construction of more than 1,000 km of new tracks, 30 bridges, 200 tunnels and the modernization of existing facilities [8].

2. Implementation of technologies: the project used the latest technologies to ensure high speed and traffic safety. This included automated control systems, wireless communication technologies for monitoring the state of trains and infrastructure, as well as advanced signaling and control systems [9].

3. Testing and start-up: after the completion of the construction, a series of test runs were carried out to check all the systems under real operating conditions. The testing included checking the speed, safety and comfort of the trains [8].

4. Launch and operation: the official launch of the line took place on June 30, 2011. The first trains began to run between Beijing and Shanghai, providing a high level of comfort and schedule accuracy [7].

The Beijing-Shanghai high-speed railway project has achieved significant results:

1. Reduction of travel time: the travel time between Beijing and Shanghai was reduced to 4 hours, which significantly increased the convenience and efficiency of transportation [8].

2. Economic benefits: the project contributed to the significant growth of economic ties between northern and southern China, increasing the volume of passenger transportation and business activity in the regions [7].

3. Technological achievements: the introduction of the latest technologies in railway construction and management has become a model for further projects both in China and abroad [9].

4. Environmental benefits: high-speed trains have a lower environmental impact compared to road and air transport, which has contributed to the reduction of CO_2 emissions and reduced road traffic [10].

The Beijing-Shanghai high-speed railway project demonstrates the successful implementation of large-scale infrastructure projects through effective planning, management and implementation of innovative technologies. The success of this project is of great importance for understanding how complex projects can contribute to economic development and improve the quality of transport services. The experience of building the Beijing-Shanghai high-speed railway can serve as an important lesson for other countries seeking to modernize their transport infrastructure.

3.1.2 PLANNING OF A LARGE-SCALE PROJECT IN THE TRANSPORT INDUSTRY OF UKRAINE

Taking into account the goals of the National Transport Strategy of Ukraine for the period until 2030 (NTS-2030) regarding the formation of the transport market and increasing the share of non-state transport operators in railway transport to 25 % by 2025 and to 40 % by 2030 [11]

the need to implement large-scale transport projects to improve the convenience and mobility of passengers and the wider use of multimodal transport services, becomes even more urgent and requires the emergence of a separate operator of the infrastructure of the new high-speed railway (HSR) of Ukraine [12].

NTS-2030 provides for the gradual introduction of high-speed rail connections (up to 400 km/h) between the main centers of Ukraine on separate tracks with a width of 1435 mm and their use for mixed passenger and cargo transportation (for accelerated delivery). goods with high added value), as well as joining the national HSR network to the Trans-European TEN-T network [13].

As it is known, even in China with its powerful passenger flows and long distances, not all high-speed railway lines are profitable, so high-value goods that need urgent delivery are also forced to be transported there [14]. In Ukraine, moreover, the use of very expensive VSHZ lines only for passenger transportation will be clearly ineffective and unprofitable. Our preliminary estimates, made even before the full-scale Russian war against Ukraine, show that the payback period for the investment in the infrastructure and rolling stock of HSR, if it is used only for passenger transportation, will be hundreds of years, even with maximum passenger traffic. On the other hand, with a relatively small passenger flow in Ukraine, HSR will have enough free capacity for freight transportation [13].

Therefore, there is no alternative to the joint use of the future HSR of Ukraine for passenger and cargo transportation as part of international multimodal transport systems. The proposed **large-scale project** on the balance of freight and passenger transportation on the HSR according to the criteria of maximum profitability for the infrastructure operator, taking into account the interests of other market participants, is **important** for the development of high-speed multimodal transportation in Ukraine [15].

A future transportation system that includes high-speed rail will be an extremely expensive piece of infrastructure (compared to conventional rail). It will not be effective if it does not provide for the integrated use of infrastructure, firstly, for mass passenger transportation, as well as for cargo transportation (in certain market niches), and secondly, as a system-forming component in other types of economic activity (such as tourism, hotel business or residential construction in the area of HSR attraction). The implementation of such a project and the future functioning of such a transport system can be described only with the help of appropriate complex mathematical models [14].

Designing and planning for the **implementation** of the high-speed railway project of Ukraine requires a comprehensive approach, which includes a technical and economic justification for the business model of the future high-speed railway system in Ukraine. It is advisable to use the economic-mathematical model of the rational ratio of passenger and freight transportation on one line and the location of stations on it. The model proposed in the **large-scale project** [16] is based on taking into account the costs of construction and operation of the railway, the fee for using the railway infrastructure, the maximum and average operational speed of trains, other operational variables and taking into account the subsidization of unprofitable passenger transport.

In the object model, the balance of cargo and passenger transportation on the HSR includes taking into account two restrictions arising from the need to satisfy the demand for transportation [14]:

1. The need to transport all willing passengers.

2. The need to transport the maximum possible amount of cargo within the available capacity.

One of the problems of organizing the joint use of the VSHZ infrastructure for freight and passenger transportation is the different speed, as well as the different acceleration/deceleration dynamics of freight and passenger trains, which is reflected in the model with the proposed solution.

When trains with different speeds run on the line, the phenomenon of "removal" of trains of one speed category by trains of another speed category occurs, which leads to a loss of line capacity [14]. The proposed basic capacity model (3.1) reflects the influence on the capacity of the line on which trains of two speed categories run:

$$N_q = \frac{24 - t_{reg}}{j} - \left[1 + \frac{2}{j}\left(\frac{1}{v_q} - \frac{1}{v_a}\right)l\right]N_a, \tag{3.1}$$

where N_q – total number of freight trains per day, train; t_{reg} – regulated time of interruptions in movement during the day (for example, "windows" for carrying out scheduled works in track and energy management), hours; j – time interval between trains, h; l – length of the limiting run of the line (the distance between the stations at which the average time of the train is the longest), km; v_a – average speed of passenger trains, km/h; v_q – average speed of freight trains, km/h; N_a – total number of passenger trains per day, train.

The application of the model in the project is used for the maximum required number of passenger trains on the line, provided that the rest of its capacity is used for freight trains.

It is known from the world practice of managing railways, including Ukrainian ones, that passenger transportation is mostly unprofitable. In this case, subsidies are needed. In domestic practice, this is the so-called "cross-subsidization" of passenger transportation losses due to revenues from cargo transportation and other sources. In the future, this is unacceptable, both from the point of view of the railway transport legislation of the European Union, and from the point of view of the laws of the market economy [15].

Therefore, the project model offers an analytical tool for substantiating the rational ratio of passenger and cargo transportation on a single infrastructure, taking into account the balance of interests of infrastructure operators and transport enterprises, taking into account that the infrastructure operator cannot bear losses from the provision of infrastructure for the public needs of passenger and cargo transportation.

One of the criteria for the joint use of the HSR infrastructure for freight and passenger transportation in the model is proposed aggregate revenue from running along the line of freight and passenger trains, which have different sources and indicators of profitability (the cost of a ticket in a passenger train and the tariff for freight transportation of the corresponding category of cargo).

The mathematical model for calculating revenue at the maximum speed of passenger trains v_a^{max} has the following general form (3.2):

$$\begin{Bmatrix} d_a N_a + d_q N_q \Rightarrow \max; \\ d_a \succ 0, N_a \succ N_{a\min}, N_a \prec N_{a\max}; \\ d_q \succ 0 \end{Bmatrix}, \quad (3.2)$$

where d_a – train revenue rate for passenger transportation, (for example, euro/train-km); d_q – train revenue rate for container transportation, euro/train-km.

Even under "ideal" conditions, when there is the necessary freight flow to use the full capacity of freight trains in the amount N_q after passing the required number of passenger trains N_a, the total revenue $\sum D$ from transportation decreases with the increase in the number of passenger trains. That is, the unprofitability and the need to subsidize passenger transportation has another confirmation here. Passenger transportation is a "social order" (state obligations), so in all countries it is subsidized in one way or another by the customers of these transportations, if they are unprofitable for the carrier [16].

An important factor in making a decision to implement a **large-scale project** for the HSR implementation in Ukraine is the amount of subsidizing the unearned income in the event of an increase in the number of passenger trains, which will be less profitable for the infrastructure operator than freight ones.

The size of the subsidy during the HSR operation S_{OP} can be determined taking into account such factors as: D_{max} – the maximum income on the HSR line under "ideal" conditions; D_{act} – the actual income on the HSR line, taking into account the "removal" of profitable freight trains by less profitable passenger trains, which occupy a larger capacity of the line [4].

The amount of subsidies is affected by the distance between stations l, both directly and in the form of more complex dependencies, for example, its effect on train speed. The average value of this distance, for a certain length of the line, affects the number of stations that need to be built on the line, and therefore, the size of capital investments, their payback period and the efficiency of using the HSR infrastructure [14].

The proposed economic-mathematical model of the project makes it possible to reflect the decrease in the need for subsidies S_{OP} when the number of stations K_{st} on the line increases, as its operation becomes more flexible when the size and structure of the train flow changes. Increasing train flow makes better use of capacity. On the other hand, the construction of a larger number of stations requires higher capital costs e_C, which is also evident from the results of the calculation based on the model. Instead, the objective function reflecting the total costs $S_{OP} + e_C$ has a clear minimum (in our case at $K_{st} = 6$), which corresponds to the optimal number of stations on the line under the given conditions of its construction and operation. Such a number of stations and the total length of the line make it possible to determine the average length of the run and reasonably place these stations on the line as close as possible to the points of formation and destination of passenger flows [14].

The HSR project of Ukraine provides for obtaining the following results:

1. **Highest profitability**: the profitability of a railway infrastructure operator depends on various factors, such as the average length of a run, the number and speed of freight and passenger trains. The highest profitability is observed with the minimum length of the run, the minimum number of passenger trains and the minimum difference in speed between passenger and freight trains [14].

1.1. An economic-mathematical model of the rational ratio of passenger and freight transportation on one line and the location of stations on it was used for the technological and economic substantiation of the business model.

1.2. The model takes into account the costs of construction and operation of the railway, fees for the use of railway infrastructure, the maximum and average operating speed of trains, other operational variables and taking into account the subsidization of unprofitable passenger transport [16].

2. **Rational ratio of passenger traffic and freight transportation**: rational ratio of passenger traffic and freight transportation on railways, taking into account the balance of interests of the infrastructure operator and transportation operators. The authors of [16] emphasize the need to change approaches to subsidizing passenger transportation, in particular, focusing on the lost revenue from freight transportation in the case of joint use of high-speed railways for passenger and freight transportation

3. **The price of transport services**: the parameters of the business model, including indicators of the technology and economy of passenger and cargo transportation, have been determined. These parameters are used to analyze the technological aspects of the business model, such as the optimal distance between stations, the required number of trains (wagons/containers in trains) for passenger and freight transportation. This allows to calculate transport costs and the price of transport services to achieve a balance of interests of infrastructure and transport operators [14].

3.2 THE PROJECT OF INTEGRATION OF UKRAINIAN RAILWAYS INTO THE TRANS-EUROPEAN TRANSPORT NETWORK (TEN-T) ON THE EXAMPLE OF THE LVIV RAILWAY NODE DEVELOPMENT

There is still a widespread opinion in Ukraine that integration into the EU is possible without changing the railway track width standard, because this process is complex and requires significant investments. However, the latest decisions of the European Union bodies in the field of railway transport indicate that the transition to the 1435 mm gauge is mandatory for all participating countries in the medium term (until 2050). Therefore, in order to join the EU, Ukraine must be ready for the global reconstruction of the railway network.

In 2022, the European Commission made changes to the plans for the development of Trans-European TEN-T networks and included Ukrainian railways in them.

The general strategy of the EU in the field of railway transport indicates that the European Union is unlikely to support the idea of building transshipment terminals on the border with Ukraine as the main strategy for the integration of Ukrainian railways into the EU transport network in the event of Ukraine's accession to the EU.

In this context, it is logical to consider the development of railway transport in Ukraine not in isolation, but in view of its integration into the European transport system. The most common gauge in the world is 1435 mm (4 English feet and 8.5 inches), which is why it is also called "normal gauge" [17, 18].

A 1520 mm wide track is laid to our western border, while in Europe a 1435 mm track is used. Ukraine is not the first country to face the problem of compatibility of two technical standards.

The experience of Spain, Portugal and the Baltic states proves that the operation of internal railway networks of excellent track width leads to the isolation of the railway and its reduced role. For Ukraine, this is unacceptable, as it will lead to an excessive load on highways that are not adapted to it, as well as to negative effects on the environment [19, 20].

Transportation of goods in a connection where different track widths are used is mainly carried out using transshipment technologies. This is the pumping of liquid cargoes (liquefied gas, oil products, chemicals); transshipment from car to car of bulk, bulk and container loads, etc. Overloading technologies require significant expenditure of time, labor and energy resources, and can have a negative impact on the environment. There are problems with the safety of cargo and rolling stock. In addition, in case of overloading of dangerous goods, there is a potential threat of man-made disasters.

The transportation of passengers is accompanied by the replacement of trolleys at the points of change of carriages when passing track nodes of different standards. This requires transfer points specially equipped with expensive equipment, causes significant technological delays for trains and inconvenience for passengers.

Transportation using traditional technologies, involving transshipment operations, causes damage to railway transport due to damage to cargo and rolling stock, and leads to significant time and labor costs.

Therefore, to transfer railways to a different track standard is not just to change wagons and "re-sewn" the tracks to a different standard width, but also to solve a complex technical, technological, and organizational problem.

Transferring the entire network of Ukrainian railways to the "European" track is a very difficult task. It is expensive, because it is necessary not only to "re-stitch" the tracks, but also to completely replace the rolling stock.

In addition, such a large-scale reconstruction can cause significant disruptions in rail traffic and affect the economy of the country as a whole. It is also important to take into account that a significant part of Ukraine's freight flows was directed to the east, where a track width of 1520 mm is used. Therefore, the transition to the European standard will require additional solutions to maintain the efficiency of transportation in the western direction.

Considering all these factors, perhaps a more realistic approach will be the introduction of a phased transition and the use of combined solutions, such as a combined track or trolley change

systems, which will allow gradual adaptation to the new standard without significant negative consequences for the transport system and the economy of Ukraine.

At the same time, it is quite realistic and expedient to reasonably define and develop cross-border corridors with the European track – as well as use the existing infrastructure of the 1435 mm track, which has not been operated for a long time, to build new sections of the European track to large cities in the regions bordering the EU. At the same time, it is necessary to revive the infrastructure of the European track, which already exists on the territory of Ukraine, but is hardly used, in the shortest possible time.

The introduction of intermodal transport between Ukraine and the EU will contribute to the reduction of logistics costs, the reduction of risks due to the use of safer transport, the reduction of cargo losses and damage, the acceleration of capital turnover, the improvement of the efficiency of the use of the wagon fleet and the creation of favorable conditions for users of railway transport.

Therefore, it is more expedient to stimulate the development of capacity for multimodal transportation throughout the territory of Ukraine and EU countries and at the nodes of tracks 1435/1520 and to encourage the introduction of new transportation and transshipment technologies [19].

Analysis of the methods of organizing cargo transportation in international connections with the European Union showed that the following options are possible:

– reloading of goods, in particular containers, from the rolling stock of the 1520 mm gauge to the rolling stock of the 1435 mm gauge;

– replacement of carts at points where wagons are changed during the transition between tracks of different standards;

– use of special rolling stock with sliding wheel pairs;

– extension of the 1520 mm track from the borders of Ukraine to the territory of Europe;

– extension of the 1435 mm track from the borders of Europe to the territory of Ukraine;

– use of a combined track of 1435/1520 mm.

So, it is possible to several possible alternative solutions to the problem of integrating Ukrainian railways into the Trans-European transport network (TEN-T), which requires the use of certain technologies, design and construction solutions. It is obvious that none of the alternatives is ideal, not every technology or solution can be economically expedient or implemented in specific conditions, but complex and phased application of various solutions is possible, which will give the best results on the path of European integration of Ukrainian railways [21].

3.2.1 THE PROJECT OF INTEGRATION OF UKRAINIAN RAILWAYS INTO THE TRANS-EUROPEAN TRANSPORT NETWORK (TEN-T) ON THE EXAMPLE OF THE LVIV RAILWAY NODE DEVELOPMENT

The Lviv railway node (**Fig. 3.1**) is unique: it was historically built at the intersection of the main railway lines. The largest sorting node at the node of track networks 1435 and 1520 [22].

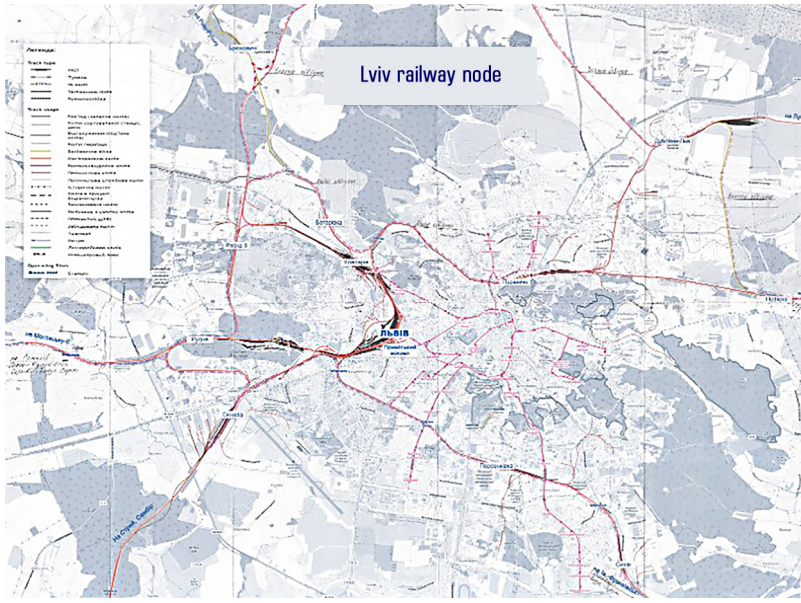


Fig. 3.1 Scheme of the Lviv node in the structure of the city of Lviv

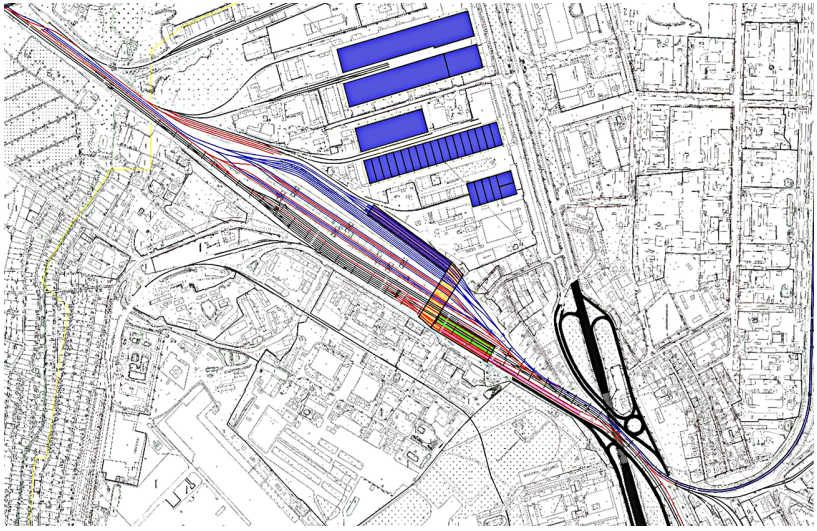
The Lviv railway node has 12 railway stations, 10 of which are located within the Lviv MTG. This includes 2 sorting stations, 3 cargo stations and 4 linear stations:

– Lviv Station (sorting station);
– Klepariv Station (sorting);
– Pidzamche (cargo);
– Sknyliv (cargo);
– Persenkivka (cargo);
– Sykhiv;
– Rudne;
– Riasna-2;
– Dubliany-Lviv;
– Briukhovychi.

As of today, the western cross-border railway crossings of Ukraine provide only 50 % of transport needs for export-import transportation.

The flow across the border of Poland is insignificant due to capacity limitations at the border [22].

It is proposed to consider as the first stage of the project the construction of a narrow track (1435 mm) in the direction of Mostyska II – Sknyliv (**Fig. 3.2**).



For reference:
The Mostyska II – Lviv section is part of the "Cretan" international transport corridor No. 3 (Berlin-Wroclaw-Lviv-Kyiv). Currently, track 1520 runs from Ukraine to Przemyśl station (Poland), Eurotrack runs from Przemyśl to Mostyska I station (Ukraine).
Sknyliv railway station is located in Lviv. It is located 5.5 km from Lviv railway station, 3 km from Lviv International Airport and 1 km from the bus station.

Fig. 3.2 Proposal for the reconstruction of the Sknyliv station passenger and freight terminals of 1520 and 1435 mm tracks and the transshipment front

Designing and planning for the implementation of this project requires a comprehensive approach, which includes:

1. Topo-geodetic searches, pre-project proposals and environmental studies. To ensure the stability and safety of the construction, detailed geodetic studies were carried out, including an assessment of possible environmental impacts. This made it possible to reduce risks and take into account all factors that could affect the implementation of the project.

2. Technical and economic substantiation and design and estimate documentation.

At the initial stages of the project, a comprehensive analysis of traffic flows, demand for transportation and potential economic benefits was conducted.

3. Analysis of methods of organizing cargo transportation in international communication with the European Union.

4. Analysis of the approach scheme of the Lviv railway node (**Fig. 3.3**):

– Krasne – Lviv: main passage, two-track railway line, electrified on alternating current.

– Lviv – Stryi: main passage, two-track railway line, electrified on direct current.

– Lviv – Mostyska-II: two-track railway line, electrified on direct current.
– Lviv – Sambir: single-track railway line, electrified on direct current.
– Lviv – Khodoriv: single-track railway line, not electrified.
– Lviv – Sapizhanka: single-track railway line, not electrified.
– Lviv – Rava-Ruska: single-track railway line, not electrified.

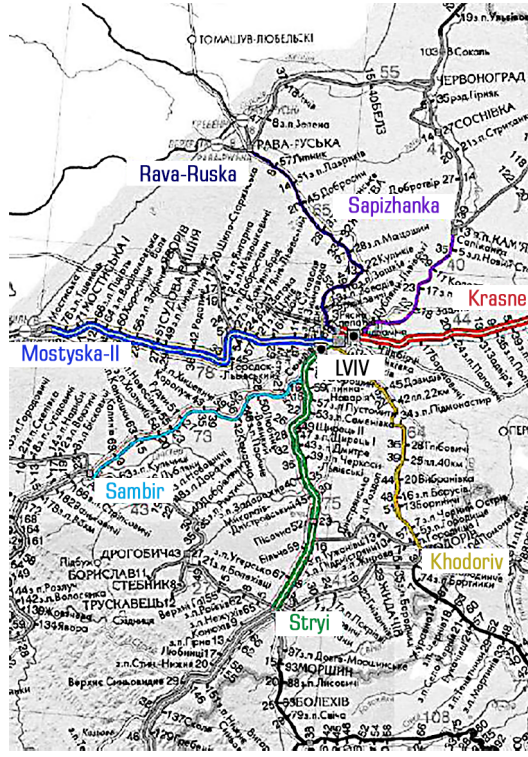


Fig. 3.3 Scheme of approaches to the Lviv railway node

5. Analysis of the structure of freight train flows after the beginning of the war (**Fig. 3.4**).

The flow across the border of Poland is not significant due to the limited capacity of the node at the border.

6. Financing and budgeting: the project had significant financial requirements.

Funding will be provided by the United States Agency for International Development (USAID). Projected investments amount to more than 225 million USD [23].

7. Project management: project management will be carried out through an integrated structure that includes: The United States Agency for International Development (USAID), JSC Ukrzaliznytsia and local authorities. Modern monitoring and project management systems will be used to control quality and deadlines.

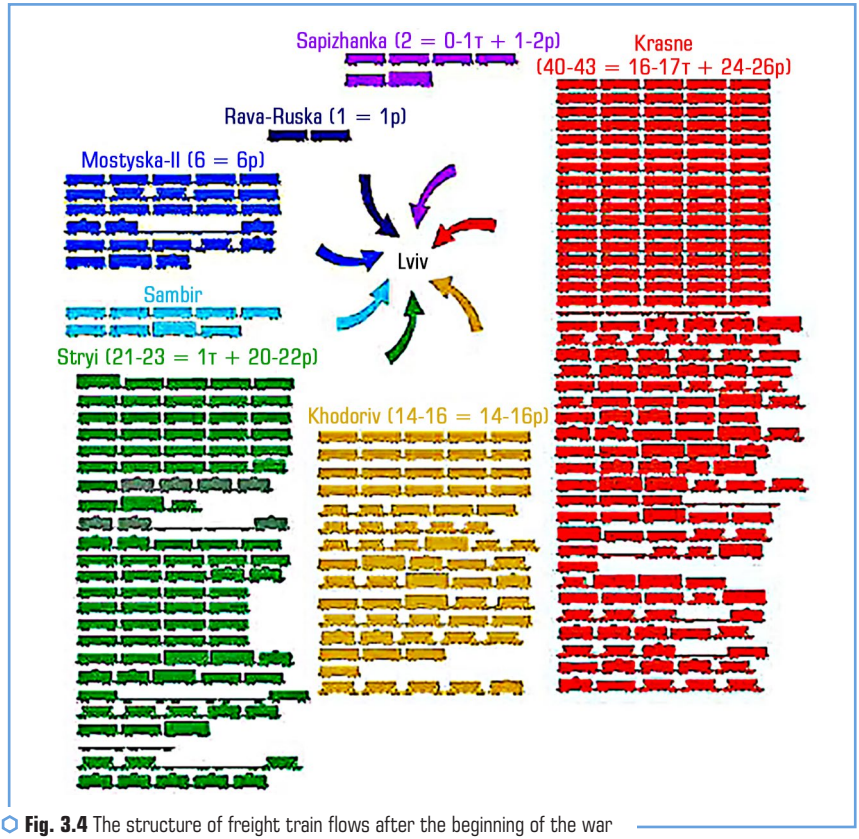


Fig. 3.4 The structure of freight train flows after the beginning of the war

The implementation of the project involves several key stages:

1. Construction of infrastructure: in the general complex of works, it is necessary to lay 69.8 km of combined track of 1435/1520 mm, build 3.1 km of track of 1435 mm, carry out 58.2 km of expansion of the earthworks site, carry out comprehensive rehabilitation of 9.5 km of the existing combined track and reconstruction of 8 stations, perform reconstruction and modernization of the electricity supply infrastructure, build the infrastructure of the station of the Sknyliv station, track 1435 mm.

2. Implementation of technologies: the project will use the latest technologies to ensure high speed and traffic safety. This includes automated control systems, wireless communication technologies for monitoring the condition of trains and infrastructure, as well as advanced signaling and control systems.

3. Testing and start-up: after the completion of the construction, a series of test runs will be carried out to check all the systems under real operating conditions. Testing includes checking the speed, safety and comfort of the trains.

Results:

– provision of a minimum speed limit of 160 km/h for passenger transport on the main and extended support networks, and 100 km/h for freight;

– first and last mile connections through multimodal passenger nodes in all EU cities that are connected to the network and have a population of over 100,000 inhabitants;

– strengthening air-rail connections for all EU airports in the network and those serving more than 4 million passengers; promotion of air-rail multimodal travel on such routes;

– the maximum waiting time at the border for freight trains is fifteen minutes;

– the possibility of transporting trucks by trains of the network;

– increased resistance to natural and anthropogenic disasters;

– implementation of European train control system (ETCS), including signaling and speed control (ERTMS), as well as a communication system (GSM-R).

The project of the integration of Ukrainian railways into the Trans-European transport network (TEN-T) on the example of the development of the Lviv railway node demonstrates the success of the implementation of large-scale infrastructure projects through effective planning, management and implementation of innovative technologies. The success of this project is of great importance for understanding how complex projects can contribute to economic development and improve the quality of transport services. The experience of implementing this project can serve as an important impetus for the further integration of Ukrainian railways into the Trans-European Transport Network (TEN-T).

The country will be able to switch to the 1435 mm gauge by 2040–2050 in various ways. It will be able to build an alternative railway network on the main transport routes – by completely decommissioning the 1520 mm wide tracks, or by replacing the infrastructure in stages. However, the long-term strategy must necessarily include the transition to a single European track gauge standard.

3.3 PECULIARITIES OF PRACTICAL IMPLEMENTATION OF SYSTEM-LEVEL PROJECT MANAGEMENT IN RAILWAY TRANSPORT OF UKRAINE

This chapter will provide examples of the implementation of projects in the railway transport of Ukraine during the first decade of the 21st century. The coordinator and main executor of these

projects was the State Scientific and Research Center of Railway Transport of Ukraine (SSRCRTU), which was established in December 2001 as part of the State Administration of Railway Transport of Ukraine, the legal successor of which became JSC "Ukrainian Railway". In 2016–2020, the Scientific Research and Design and Technology Institute was established on the basis of SSRCRTU as a branch of JSC "Ukrainian Railway".

There were several projects: System of automatic identification of rolling stock and containers, Financial and economic information system, Psychological support of locomotive crews, etc. All of them were implemented according to the same principles. But this work presents the implementation of only one project.

3.3.1 DEVELOPMENT OF BASIC PRINCIPLES OF PROJECT MANAGEMENT

In the early years of the 21st century, the project approach was just beginning to be applied. It was then that the basic principles of project management for research, consulting, construction and some other organizations were developed. The project meant the creation of a new, as a rule, single non-repetitive product (product or technology). Although it should be noted that the project approach could be used for any company when implementing innovative programs and projects. These are the basic principles used in the SSRCRTU:

1) project management includes definition of its goals, formation of structure, planning and organization of works, coordination of actions of executors;

2) in terms of form, the structure of project management can correspond to a brigade (cross-functional) or divisional structure, in which a certain division (department) is created for a specific project and not always, but within the project's implementation period. Advantages of project management: high flexibility, reduction in the number of management personnel. The project cross-functional structure was typical for research enterprises of the military-industrial complex;

3) each project goes through four phases during its existence:

a) phase 1 – proving its **attractiveness**. Define the mission of the project, which should show all the participants of the project and its external environment that each of them will get from its successful implementation; having developed and presented first the concept, and then the business plan of the project;

b) phase 2 – **development**, when the project management team concentrates its efforts on creating the most effective project implementation plan. In the planning process:

– all the works of the project are presented in the form of a structural hierarchical decomposition of the works, which allows even the project to be broken down into components available for inspection;

– draw up a calendar plan for the project, which is optimized based on the availability of resources and the sequence of activities;

– develop schedules of the project's resource needs;

– form the project budget;

– develop the organizational structure and project team;

– develop communication mechanisms and procedures for making changes to plans;

– etc.

The result of the planning is the consolidated project plan approved by the customer, which is the guiding document of the project team;

c) phase 3 – **implementation** or carrying out of the main works. At this stage, monitor the progress of work, monitor changes occurring both within the project and in its environment, and make appropriate changes to the project plans;

d) phase 4 – **completion.** The main tasks of the completion phase: deal with all "tails" of the project; to employ personnel who were temporarily involved in working on the project; analyze the experience, identify positive and negative features of implementation.

3.3.2 SYSTEM OF AUTOMATIC IDENTIFICATION OF ROLLING STOCK AND LARGE-TONNAGE CONTAINERS ON THE RAILWAYS OF UKRAINE (SAIRS UZ)

This is an international project of the CIS and Baltic States (Commonwealth), which have a single standard of railway track width of 1520 mm, in the development of which Ukraine participated independently, including by using the national technical base.

The management structure of the SAIRS-UZ project was built according to the divisional principle: a special department of the Chief Designer of the SAIRS was created at the SSRCRTU with 9 specialists headed by the deputy director of the SSRCRTU.

An operational headquarters under the deputy general director of Ukrzaliznytsia (UZ) was created. The progress of the project was reported twice a year to the UZ Council.

Phase 1

The purpose of the project: SAIRS-UZ is intended for automatic identification of freight cars, locomotives and large-tonnage containers in real time for the formation of information in the hierarchy of the Automated Freight Transportation Control System (AFT CS UZ) to solve the following tasks:

– operational control of locomotive, wagon and container fleets;

– numbered accounting and monitoring of the location and condition of locomotives, wagons and containers in real time;

– operational search of wagons and containers by their numbers;

– operational control of compliance with routes, established deadlines for the use of rolling stock and cargo delivery;

– numbered accounting of delivery and reception of transport objects at border stations, sea and river ports in international traffic, inter-railway and inter-state nodes, approach tracks;

– formation of data for mutual settlements for the use of wagons;

– calculation of the actual mileage of wagons, etc.

Project mission:

– increasing the efficiency, reliability and reliability of information about transport objects in the AFT CS UZ;

– ensuring integration into the pan-European identification system of transport objects.

System *efficiency* is achieved due to:

1) reduction of labor costs by:

– reading identification information from the transport facility;

– manual entry of information into the ACS;

– organization of control of entered information;

– correction of manual input errors;

2) increasing the reliability and efficiency of information on cargo transportation, which will provide opportunities:

– improve the technology of processing trains, wagons, containers and cargo at stations;

– improve the organization (procedures) of delivery-reception of transport objects at border stations, sea and river ports in international traffic, inter-railway and inter-state nodes, approach tracks;

– speed up the circulation of wagons and reduce the empty mileage of wagons;

– reduce costs for the repair of freight wagons as a result of the implementation of the program of transition to a system of repairs and maintenance according to the standards of the mileage actually performed.

Phase 2

The list of documents on the basis of which the system was created:

– comprehensive program of SAIRS implementation on the railways of the Commonwealth;

– concept of the construction of the "System of automatic identification of rolling stock and large-tonnage containers" was approved on May 8, 2002;

– automated control system of cargo transportation on the railway transport of Ukraine, Terms of Reference, 16289267.184154.201.T3;

– DSTU 3918–99 (ISO/IEC 12207:1995) Software life cycle processes;

– DSTU 3396.0–96 Protection of information. Technical protection of information. Basic Provisions;

– ISO 10374:1991/AMD.1:1995 Freight containers – Automatic identification.

The generalized structural diagram of SAIRS-UZ is shown in **Fig. 3.5**. The system consists of the following functional components (subsystems):

1. Coded on-board sensors (COS).
2. COS programming tools (COSRT).
3. Stationary equipment for exposure and reading information from on-board sensors (ERE).
4. Software-hardware complex "ARM SAIRS-UZ".
5. Complex of telecommunication equipment SAIRS-UZ.

A technical task (TT) was developed for SAIRS-UZ, which also formulated requirements for: reliability, safety, ergonomics and technical aesthetics, operation and maintenance, information security, patent purity, standardization and unification, etc.

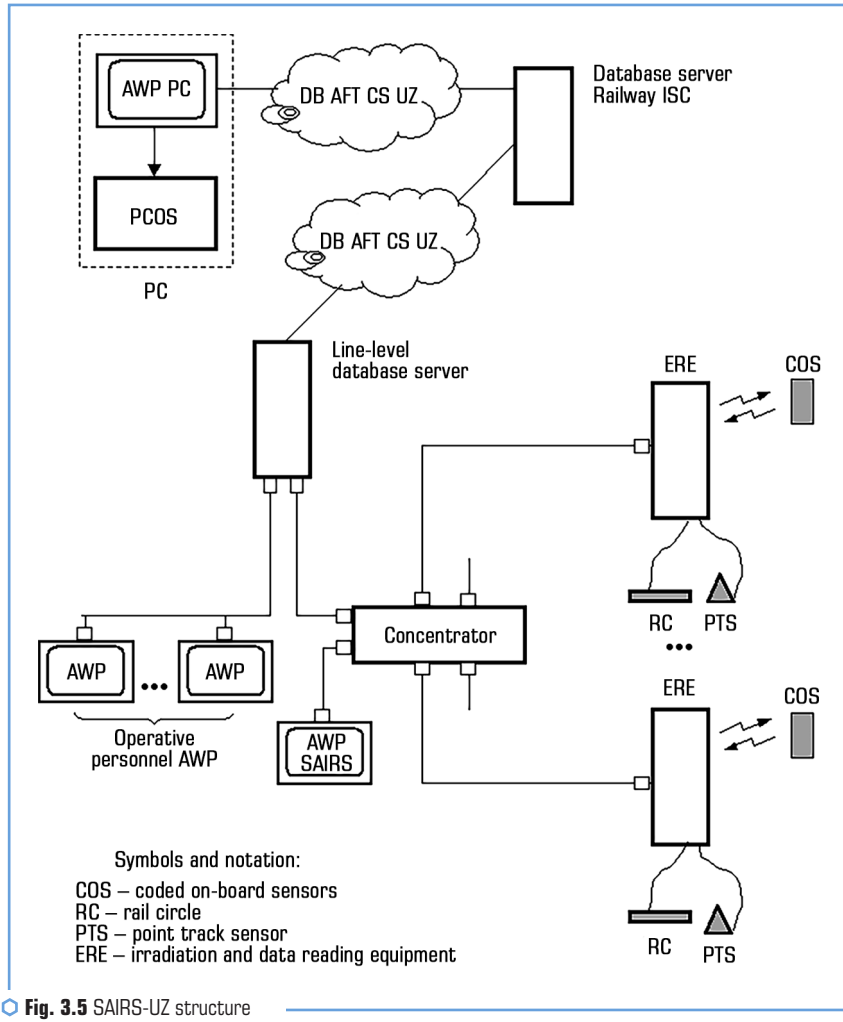


Fig. 3.5 SAIRS-UZ structure

TT has sections: mathematical support, information support, linguistic support, software, technical support. The system has an open architecture for the expansion of software and technical tools without changes to the software and information support, as well as additions and updates of functions. The basis for building a complex of technical means is a complex of mutually agreed and certified equipment in the UkrSEPRO state system, which provides the necessary technical characteristics of the system.

The functions of training and attestation of specialists should be performed by a specialized training center for AFT CS UZ users, created on the basis of the National Center of the Ukrainian Academy of Sciences. The qualification, number and mode of operation of SAIRS-UZ users are determined depending on real needs (specialization, features of the infrastructure and scope of work of the line enterprise).

The composition and content of the work on the creation of the system is given in **Table 3.1**.

Table 3.1 List of stages in the development and implementation of SAIRS-UZ

Stage No.	Name
1	Development of the technology of the sensor coding point in wagon repair and locomotive depots. Software implementation of the set of tasks "Programming code onboard sensors"
2	Development of technology for the point of reading information from sensors. Software implementation of the set of tasks "Reception and processing of identification information from coded on-board sensors"
3	Development of information interaction technology of SAIRS-UZ with the database of the railway ISC server. Software implementation of the set of tasks "Formation and transfer of information to the main models of the database of the railway level in the AFT CS UZ hierarchy
4	Development of technology for monitoring the technical condition of the SAIRS-UZ subsystems. Software implementation of the set of "System Testing" tasks
5	Development of technology for monitoring the functioning of SAIRS-UZ. Software implementation of sets of tasks: "Maintenance of an operational log of event registration in the system"; "Maintaining an archive log of event registration in the system"; "Audit of registered events in the system"
6	Development of technology and instructions for the operation of the complex of technical means SAIRS-UZ
7	Development of technology and instructions for forming and maintaining the SAIRS-UZ database
8	Development, coordination and approval of a set of operational documentation for SAIRS-UZ (ED, C1, B6, B7, I3, PD)
9	Equipment of the test site. Experimental operation of SAIRS-UZ at a designated training ground
10	Development and coordination of information exchange technology between neighboring states on the crossing of trains (locomotives, wagons, containers) at interstate node points within the CCRT
11	Development and implementation of the Partial technical task for the creation of AFT CS UZ unified secure workstation
12	PTT development for the creation of a COS and stationary equipment for exposure and reading information from on-board sensors (ERE). Organization of a tender for the PTT implementation
13	Program implementation of the set of tasks "Providing informative and reference information about numbered accounting, location and route of movement of locomotives"

Phase 3

The SAIRS-UZ implementation program on Ukrainian railways should go through two stages, as approved by the Decision of the thirtieth meeting of the Council on Railway Transport of the Commonwealth of Nations.

The 1st stage (2002-2003) provided for the creation of an internal experimental training ground and an international limited training ground for field tests of domestic and foreign components

of SAIRS-UZ, as well as working out the tasks of transferring and using information received from the SAIRS-UZ system.

In 2002 the creation of an experimental training ground began (project work on the installation of stationary irradiation and reading equipment (ERE) at the Koziatyn locomotive depot and the Khutir-Mykhaylivsky station was completed). At this training ground, the components of the SAIRS were tested for efficiency and interoperability, as well as to work out the accounting of the control of the location and use of freight locomotives and locomotive crews. Separate car depots of the Prydniprovska and Donetsk railways have been equipped with sensor coding points, and freight cars have been equipped with sensors.

In 2003 it was supposed to complete the creation of an internal experimental training ground and create a limited training ground for international communication. Conduct research on the use of domestic components in the SAIRS-UZ system; experimental operation of the first phase of hardware and software complexes; development of software of the second phase, and also, at the experimental range based on the Koziatyn locomotive depot and the points of turnover of locomotives of the Koziatyn depot at the stations of the South-Western Railway, to work out the tasks related to the rational use of locomotive crews and the locomotive park, using the SAIRS-UZ information.

Start conducting scientific research on improving the organization of the transportation process and its information support, in connection with the SAIRS-UZ implementation.

Conduct a structural analysis and develop recommendations for the purchase of channel-forming equipment, concentrators, modems, which are additionally necessary for the transmission of information from SAIRS-UZ.

Equip the locomotives of the Koziatyn depot, an experimental training ground, and the locomotives of the Kupiansk-Sortuvalny depot, which serve the directions of the limited testing ground "SAIRS" of international traffic.

In order to equip freight cars with coded on-board sensors, it is planned to equip the car depot of the international training ground with a sensor coding point. Kupiansk-Sortuvalny station (Southern Railway) and to continue equipping freight cars with sensors on the Donetsk, Prydniprovska, and Southern Railways.

The 2nd stage (2004–2006) provided for the provision of automatic reading of information at interstate crossing points, inter-railway crossing points, main sorting stations, main freight and port stations, due to the equipment of their reading points, as well as rolling stock that participates in transportation, – information carriers (COS).

In 2004, it was foreseen:

– equip interstate node points and transmission stations with reading points;

– finish the work on the COS equipment of all railway locomotives;

– start work on the COS equipment of electric and diesel trains;

– continue work on the equipment of freight wagons with the COS means and wagon depots and railway stations with sensor coding points;

– oblige the enterprises adjacent to the stations (CMC, GOK, etc.) to equip approach tracks by ERE and to equip COS with its own rolling stock with numbering starting with the number "5".

In 2005:

– equip inter-railway node points and main sorting stations with reading points and to continue work on equipping freight cars with COS means.

In 2006:

– equip the main cargo and port stations with reading points;

– complete the equipment with COS devices of all rolling stock, including passenger cars;

– complete the modernization of all information systems taking into account the use of information from SAIRS-UZ;

– put SAIRS-UZ into industrial operation for the customer.

The implementation of the SAIRS-UZ project was carried out on the basis of annual planning and approval of plans at the level of the deputy general director of the UZ. **Table 3.2** presents a fragment of the plan for 2003 in an abbreviated version.

Table 3.2 Implementation plan of SAIRS-UZ in 2003

No.	Name of work, stages of work	Deadline	Volumes of financing for the current year, UAH	Expected result
1	Scientific and technical support		976 000	
2	Production tasks		12 482 625	
2.1	Creation of a limited training ground "SAIRS" of international traffic	01.07.2003	3 467 772	International restricted range
2.2	Creation of an internal experimental training ground	30.09.2003	4 966 159	Experimental training ground
2.3	In accordance with the planned types of repair work, equipment with coded on-board sensors of freight cars	31.12.2003	3 925 800	The main types of freight cars are equipped with SAIRS-UZ coded on-board sensors
2.4	Equipment for information integrators of ISC	3rd quarter 2003	22 894	Equipped with ISC information integrators of railways
2.5	Obtaining permission to use the radio frequency range for the operation of ERE equipment	2nd quarter 2003	100 000	Permission
3	Regulatory documentation		30 000	
Total for 2003:				13 488 625
Responsible for financing and sources of financing				Financing, UAH
CTech, CF, at the expense of centralized financing according to indicative R&D plans				976 000
Head of railway, CF, CID at the expense of the railways' own funds				12 482 625
CID, CF, at the expense of centralized financing according to indicative plans for the development of regulatory documentation				30 000

Phase 4

Unfortunately, the project was discontinued in 2006 for reasons beyond the control of its developers. The reasons are lack of funding. Although more than 40 % of all freight cars and more than 50 locomotives were equipped with COS at that time. 2 border training grounds were working, information was being sent to the CD UZ in an operational mode. SAIRS is still working.

Attempts to restore this project using GPS sensors instead of RFID technology ended in nothing.

Despite the fact that the SAIRS-UZ project did not end with complete success, the results achieved during its implementation are definitely positive. In the process of its development and implementation, a lot of experience in managing projects of this scale has been accumulated. As an example of the complex organization of work in different areas of the project, **Table 3.2** is given (the structure or fragment of the implementation plan of SAIRS-UZ for 2003).

CONCLUSIONS

Chapter 3 "Project management of Ukraine's integration into the Trans-European transport network" (TEN-T) focuses on the analysis and systematization of high-speed railway project management experience, both in Ukraine and abroad, on projects related to the phased transition of railways of Ukraine on the standards of TEN-T railway tracks, on the practical aspects of the large-scale project SAIRS-UZ (System of Automated Identification of Rolling Stock of Ukrzaliznytsia), as well as on the determination of prospects and challenges in the field of project management facing Ukraine in this context:

1. Successful experience of international projects: the Beijing-Shanghai high-speed railway project discussed in the section serves as an important example of effective management of large-scale infrastructure initiatives. It demonstrated the advantages of integrated management, innovative technologies, economic benefits and environmental benefits, which can be useful for the implementation of similar projects in Ukraine. The monitoring systems and modern technologies that were used became a model for further projects in the transport industry.

2. Planning and feasibility study: for the successful implementation of high-speed railway projects in Ukraine, a thorough feasibility study is required, which includes an assessment of the balancing of passenger and cargo transportation, the use of economic and mathematical models, and the introduction of the latest technologies. Since passenger transportation may not be economically viable due to high costs, it is important to ensure efficient use of freight infrastructure and find optimal financial solutions for subsidies.

3. Adaptation to European railway track standards: the transition to the European track width standard is critical for the integration of Ukraine into the Trans-European transport network. This will require significant investment in the reconstruction of infrastructure and rolling stock, as well as solving problems related to transshipment technologies. The Lviv Railway Node plays a key role in this process, providing connections between different track gauge standards.

4. Staged approach and international cooperation: the integration of Ukrainian railways into the Trans-European transport network should be carried out in stages, using combined solutions and technologies. Effective project management and international cooperation are essential to ensure financial and technical stability. International donors, such as USAID, can play an important role in project financing and coordination.

5. Lessons from project implementation: the experience of project implementation, such as SAIRS-UZ, shows the importance of stable financing, an integrated approach to resource management, the interest of management and the integration of projects with the overall development strategy. The accumulated experience emphasizes the need to adapt to changing conditions and improve technologies and management practices.

Thus, this chapter highlights key aspects of high-speed rail project management and integration into the Trans-European Transport Network, infrastructure and technology upgrades, showing that success in the implementation of such large-scale initiatives requires an integrated approach, effective management and close international cooperation.

The successful integration of Ukraine into the European transport system opens up new opportunities for the development of the economy and the improvement of transport infrastructure.

REFERENCES

1. A Guide to the Project Management Body of Knowledge (PMBOK Guide) (2017). Project Management Institute.
2. Kerzner, H. (2017). Project Management: A Systems Approach to Planning, Scheduling, and Controlling. Wiley.
3. Larson, E. W., Gray, C. F. (2017). Project Management: The Managerial Process. McGraw-Hill Education.
4. Burke, R. (2013). Project Management: Planning and Control Techniques. Wiley.
5. Schwaber, K., Sutherland, J. (2020). The Scrum Guide. Scrum.org.
6. Meredith, J. R., Mantel, S. J. (2018). Project Management: A Managerial Approach. Wiley.
7. Chen, L. (2014). Economic Impacts of High-Speed Rail on Regional Development. Routledge.
8. Liu, H. (2012). Project Management in Infrastructure: Lessons from the Beijing-Shanghai High-Speed Rail. Academic Press.
9. Wang, Y. (2015). Technological Innovations in High-Speed Rail Systems. Springer.
10. Zhang, X. (2013). High-Speed Rail Development in China: A Case Study of the Beijing-Shanghai Line. Beijing University Press.
11. Natsionalna transportna stratehiia Ukrainy na period do 2030 roku (2018). Rozporiadzhennia Kabinetu Ministriv Ukrainy No. 430-r. 30.05.2018. Available at: https://zakon.rada.gov.ua/laws/show/430-2018-p

12. Pozdniakov, A. A., Myronenko, V. K., Pozdniakova, O. O. (2019). Information model of development of railway transport infrastructure in the system of multimodal passenger transportation. Science-Based Technologies, 42 (2), 280–287. https://doi.org/10.18372/2310-5461.42.13521
13. Matsiuk, V., Myronenko, V., Samsonkin, V., Pozdniakov, A. (2020). Conceptual model of multimodal transportation on the caspian-black sea route of the new silk road (via Ukraine and Poland). XII International Conference Transport Problems 2020, 476–481.
14. Pozdniakov, A. A. (2024). Modelling of future Ukraines high-speed railways with shared use for passengers and cargo service. Science-Based Technologies, 61 (1), 85–93.
15. Pozdniakov, A., Myronenko, V., Pozdniakova, O. (2019). Factors affecting the capacity of high-speed mainlines at different speeds for passenger and freight trains. Materialy mizhnarodnoi naukovo-praktychnoi konferentsii. Lvivska filiia DNUZT, 18–20.
16. Pozdniakov, A., Myronenko, V., Pozdniakova, O. (2023). Determination modelling of future Ukraine's high-speed railways with shared use for passengers and cargo service. TRANSBALTICA XIV: Transportation Science and Technology. Proceedings of the 14th International Conference TRANSBALTICA. Vilnius, 574–585. https://doi.org/10.1007/978-3-031-52652-7_57
17. Shyryna zaliznychnoi kolii: de, yak, chomu? (2020). Available at: https://www.railway.supply/uk/shirina-zaliznichno%D1%97-koli%D1%97-de-yak-chomu/
18. Shyryna zaliznychnoi kolii. Available at: https://mskukraine.com/uk/katalog-produkcziyi/kolorovyj-metaloprokat/nihrom-prokat/shyryna-zaliznychnoyi-koliyi/
19. Transparency International Ukraine. Available at: https://www.ti-ukraine.org/
20. Rybchuk, A. V. (2004). Transportni systemy svitu – vazhlyvyi element hlobalnoi vyrobnychoi infrastruktury. Aktualni problemy ekonomiky, 7 (37), 99–104.
21. Shchuklin, Y. (2023). Yaka misiia zaliznychnoho transportu v konteksti yevropeiskoi intehratsii Ukrainy? Available at: https://interfax.com.ua/news/blog/888842.html
22. Hanailiuk, B. (2022). Lvivskyi zaliznychnyi vuzol. Lvivski linii Lemberger Linien. Available at: https://railexpoua.com/wp-content/uploads/2022/11/Презентація-Борис-Ганайлюк.pdf
23. USAID planuie vydilyty 225 mln dolariv na budivnytstvo yevrokoli vid stantsii Mostyska do Lvova (2023). Tsentr transportnykh stratehii. Available at: https://cfts.org.ua/news/2023/12/06/usaid_planue_vidiliti_225_mln_dolariv_na_budivnitstvo_evrokoli_vid_stantsi_mostiska_do_lvova_77414

DOI: 10.15587/978-617-8360-03-0.CH4

Oksana Malanchuk, Anatoliy Tryhuba, Ivan Rogovskii,
Liudmyla Titova, Liudmyla Berezova, Mykola Korobko

CHAPTER 4

DIFFERENTIAL-SYMBOLIC APPROACH AND TOOLS FOR MANAGEMENT OF MEDICAL SUPPORT PROJECTS FOR THE POPULATION OF COMMUNITIES

ABSTRACT

The aim of the study is to propose a differential-symbolic approach to managing community health support projects, to develop algorithms and computer models on its basis, and to use them to conduct a study of the impact of project environment components on the choice of the optimal project implementation scenario and risk assessment.

The work uses project management methodology, system and differential-symbolic approaches, which underlie the developed algorithms and computer models for planning community health improvement projects and assessing their risks. To implement the proposed models, code was written in the Python programming language using libraries for solving differential equations, optimizing and visualizing results. NumPy libraries were used to work with numerical data and vectors, SciPy for numerically solving differential equations and optimizing the objective function, and Matplotlib for visualizing the results.

The main stages of the proposed differential-symbolic approach to managing community health support projects are presented. Mathematical models of differential-symbolic planning of projects for planning projects for improving the health of the community population and risk assessment of projects for medical support of the community population have been developed. They involve the use of differential equations to describe the dynamics of projects as a separate system and the use of symbolic expressions to represent individual parameters and their description. Algorithms of differential-symbolic management of projects for improving the health of the community population and risk assessment of projects for medical support of the community population have been developed, the block diagram of which involves the implementation of 16 and 9 interconnected steps, respectively. Based on the proposed algorithms, computer models of differential-symbolic planning of projects for improving the health of the community population and risk assessment of projects for medical support of the community population have been developed. Based on the use of computer models for given conditions of the project environment, the results of optimizing the configuration of projects for improving the health of the community population and risk assessment of projects for medical support of the community population have been obtained.

The prospect of further research is to expand the functionality of computer models, adding modules for the analysis of other component projects.

For the first time, a differential-symbolic approach to managing projects for medical support of the population of communities has been proposed, which is based on methods of mathematical modeling, numerical analysis and optimization, which ensure the determination of a rational configuration of these projects and the assessment of risks for the given characteristics of the project environment. Based on the substantiated stages of the differential-symbolic approach, mathematical models, algorithms and computer models have been developed. The use of the proposed computer models makes it possible to obtain the dependence of the growth rate of the percentage of the healthy population participating in educational activities on the configuration of projects for improving the health of the population of communities, as well as to determine the optimal scenarios for the implementation of these projects in the community and the risks of projects for medical support of the population of communities.

The proposed computer models are a tool for project managers, which allows to perform labor-intensive calculations to form possible scenarios for the implementation of projects to improve the health of the population in the community, determine the optimal one among them, as well as assess the risks of projects for medical support of the population of communities.

KEYWORDS

Differential-symbolic approach, computer models, modeling, planning, projects, medicine, population of communities.

CHAPTER 4

Currently, the development of various industries is taking place thanks to project-oriented management, which from year to year is becoming an increasingly effective tool for management activities [1, 2]. At the same time, changes are observed in the project environment of organizations in various industries, which significantly affect the implementation of their development projects and, accordingly, their efficiency [3]. Without the use of tools that take into account dynamic changes in the components of the project environment, it is impossible to successfully implement the relevant projects [4]. One of the industries in Ukraine that is undergoing reform is the medical industry. Today, medicine faces challenges that require effective project management to ensure optimal development and the provision of high-quality medical services. However, medical project management is a complex task due to the need to take into account not only technical aspects, but also the organizational structure, the supply chain of resources and services, regulatory and stringent requirements, as well as flexibility in making management decisions to resolve unforeseen situations during project implementation [5, 6]. Projects of medical support of the population of communities deserve attention, which ensure the promotion of a healthy lifestyle and activity among the population [7]. During their implementation, a number of scientific and applied tasks arise, among which effective planning and risk assessment are significant.

In this context, computer models of differential-symbolic planning for improving the psychological state and health of the population of communities and risk assessment of projects of medical support of the population of communities are an important tool that allows to systematize and optimize the process of managing these projects, contributing to their successful implementation and adaptation to a changing project environment. Our work reveals the features of the development and use of computer models of differential-symbolic planning and risk assessment of projects of medical support of the population of communities, which underlies the increase in the efficiency of management of these projects and, accordingly, the reduction of morbidity and increase in the activity of the population of communities.

Substantiation of effective scenarios and risk assessment are important processes that determine the effectiveness of project management [8]. Changes in the components of the project environment can occur at any time during the life cycle of projects, and they can have a significant impact on the outcome of the project [9, 10]. Effective project planning can help minimize the negative impact of changes on the project and ensure its successful completion [11].

In well-known scientific works, their authors identify separate scientific and applied tasks regarding project risk management [12] and management of medical development projects [13, 14]. These include the opacity of relationships between project participants and ineffective communication between them. There is also a lack of adaptation of each of the participants to changes in the project environment, as they do not have a reasoned justification for such changes. In addition, some of the participants do not perceive the appropriateness of changes, as they do not have a sufficient quantitative assessment of the impact of changes on the effectiveness of project implementation.

Some authors of scientific papers [15, 16] justify the feasibility of creating computer models, which are the basis for accelerating and making accurate management decisions. Computer models allow to quantitatively assess the impact of various decision alternatives on target indicators [17]. This can be especially useful in conditions of variability of the components of the project environment, when expert assessments are inaccurate. Computer models provide the generation of alternative project implementation scenarios and the assessment of their risks, which cannot be described intuitively. This allows project managers to consider a wide range of factors in the project environment and find a rational solution. Nevertheless, computer models provide automation of the process of making management decisions and its acceleration.

The development and use of computer models for managerial decision-making encounters certain difficulties [18, 19]. In particular, it is necessary to select a simple and sufficiently accurate mathematical apparatus that will underlie the developed computer model [20]. In particular, the creation of computer models that are too complex leads to the fact that users are unable to understand and use them in practice [21]. In addition, computer models with incorrect knowledge distort the result, lead to errors in the justification of management decisions or software operation [22]. Despite these shortcomings, computer models have significant potential to improve the efficiency and quality of management decision-making in projects [23]. They can be particularly useful in complex or uncertain situations where manual decision-making may be impossible or inefficient [24, 25].

Existing studies [26] on community projects confirm the possibility of benefiting clinics, patients and communities. The differential-symbolic approach to the management of community health support projects is noteworthy [27, 28]. Differential-symbolic equations allow modeling complex dependencies between variables, which can be useful for assessing the impact of changes on performance indicators, such as cost, timing and quality of project implementation [29]. In addition, differential-symbolic equations can be used to develop algorithms that automatically assess the impact of changes on the project and generate alternative solutions [30]. Therefore, there is a need to develop a differential-symbolic approach to the management of community health support projects, and to develop algorithms and computer models based on it. Such models will make it possible to substantiate the optimal scenario for the implementation of community health improvement projects and assess their risks, as well as visualize the impact of changes in the project environment on the performance indicators of these projects. The availability of such tools for project managers will increase the accuracy of determining an effective project implementation scenario, qualitatively assess risks, and establish communication between stakeholders [31]. The results of computer modeling increase the effectiveness of developing plans for responding to changes in the project environment.

The aim of research is to propose a differential-symbolic approach to managing community health support projects, develop algorithms and computer models on its basis, and also use them to conduct a study of the influence of project environment components on the choice of the optimal project implementation scenario and risk assessment.

In accordance with the aim, the following objectives should be solved in the work:

1. Substantiate the main components and stages of the differential-symbolic approach to managing community health support projects.

2. Propose and describe mathematical models and, on their basis, develop algorithms for differential-symbolic planning of community health support projects and assessment of their risks.

3. Develop computer models of differential-symbolic planning of community health support projects and assessment of their risks.

4. Based on the developed computer models, determine the optimal scenario for the implementation of the project to improve the health of the community for the given conditions of the project environment and assess the risks.

4.1 DIFFERENTIAL-SYMBOLIC APPROACH TO MANAGING COMMUNITY HEALTH SUPPORT PROJECTS

Community health support projects are understood as temporary actions aimed at implementing preventive measures related to the promotion of a safe and healthy lifestyle, as well as activities to prevent diseases among the community population and increase its activity. To increase the accuracy and quality of management of community health support projects, it is proposed to use

a differential-symbolic approach, which ensures the consideration of a number of necessary components, as presented in **Fig. 4.1**.

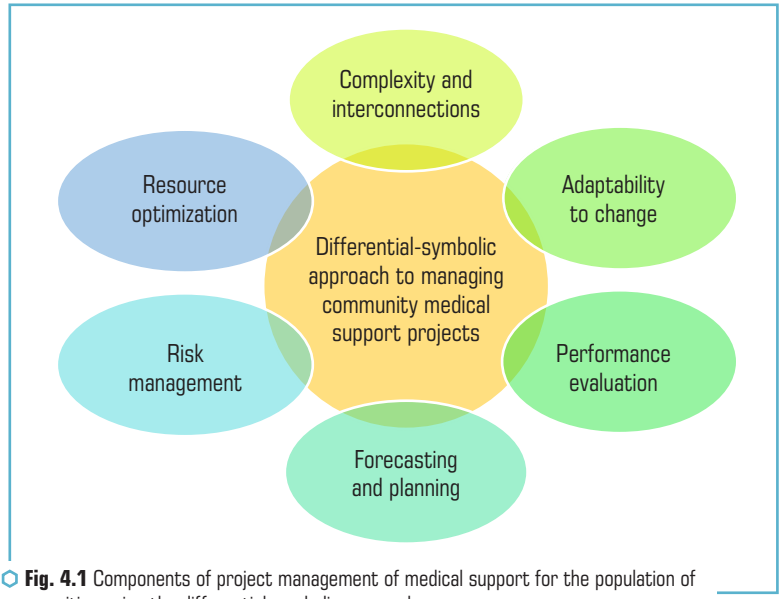


Fig. 4.1 Components of project management of medical support for the population of communities using the differential-symbolic approach

The proposed differential-symbolic approach makes it possible to take into account the complexity and interrelationships between both the components of the project environment and the components of community health support projects. In particular, health support projects often have a complex structure and include numerous relationships between different components. These components include medical services, resources, patients and their health status, medical personnel, and finances. The differential-symbolic approach allows to model these relationships, track the impact of changes in the project environment on various aspects of the project, and conduct a comprehensive analysis of their implementation scenarios.

Community health support projects can be subject to numerous changes caused by a changing project environment (changes in the health of the population, the emergence of new medical technologies, political and economic factors, etc.). The differential-symbolic approach allows to effectively manage these changes by modeling their impact on the implementation of the community health support project and quickly adapt already developed strategies to these changes.

By using the differential-symbolic approach, it is possible to analyze in detail the effectiveness of various aspects of a community health support project, such as resource allocation, costs

and results. This allows for a more accurate assessment of costs and benefits (value for stakeholders), as well as to determine the most effective management decisions.

Community health support projects are often long-term in nature, which requires accurate forecasting and planning. The differential-symbolic approach helps to create models that take into account various factors and their interaction, which provides more accurate forecasts and plans for the future development of community health security.

The implementation of community health support projects is accompanied by the emergence of risks that belong to their financial, technical and social components. The differential-symbolic approach allows to identify, analyze and assess these risks, as well as develop strategies to minimize them.

Resource management is a very important component of managing community health support projects, since resources are in most cases limited. The use of the differential-symbolic approach allows for effective planning and allocation of resources, ensuring their rational use and optimization.

In general, the use of the differential-symbolic approach to the management of community health support projects provides a comprehensive analysis and optimization of these projects, which allows for increasing their efficiency and achieving better results in providing medical care.

To present the features of the proposed differential-symbolic approach to the management of community health support projects, its scheme was constructed (**Fig. 4.2**). It illustrates the main stages and components of the proposed approach, which involves the use of differential equations to describe the dynamics of community health support projects as a separate system and the use of symbolic expressions to represent parameters and their description.

The scheme of the proposed differential-symbolic approach to the management of community health support projects displays eight main blocks. First of all, there is a block (Data input) that provides for the collection and input of the necessary initial data about the community, the medical needs of the population, resources and factors affecting the health of the population. The next block (Symbolic expressions for project parameters) provides for the formation of symbolic expressions that describe the components of community health support projects and their impact on the population. After this, there is a block (Differential equations for dynamics modeling), which uses differential equations for mathematical modeling of the dynamics of changes in the health of the community and their well-being.

The block (Numerical solution of differential equations (SciPy)) provides for the numerical solution of differential equations using the open SciPy library with high-quality scientific tools in the Python programming language. The block (Optimization (SciPy)) is also based on the SciPy library, which provides for the use of optimization methods to determine the optimal configuration of community health support projects.

The next block (Results analysis and visualization (Matplotlib)) involves analyzing the obtained results and visualizing them using the Matplotlib library in the Python programming language. The block (Selection of optimal scenario) ensures the establishment of the optimal scenario for project implementation based on the analysis of the obtained results, forecasting their indicators and taking into account the conditions of the project environment and constraints.

The last block (Report generation and recommendations) provides for the preparation of reports and analysis of the results of community health support projects, as well as the development of recommendations for their implementation and further implementation in the community.

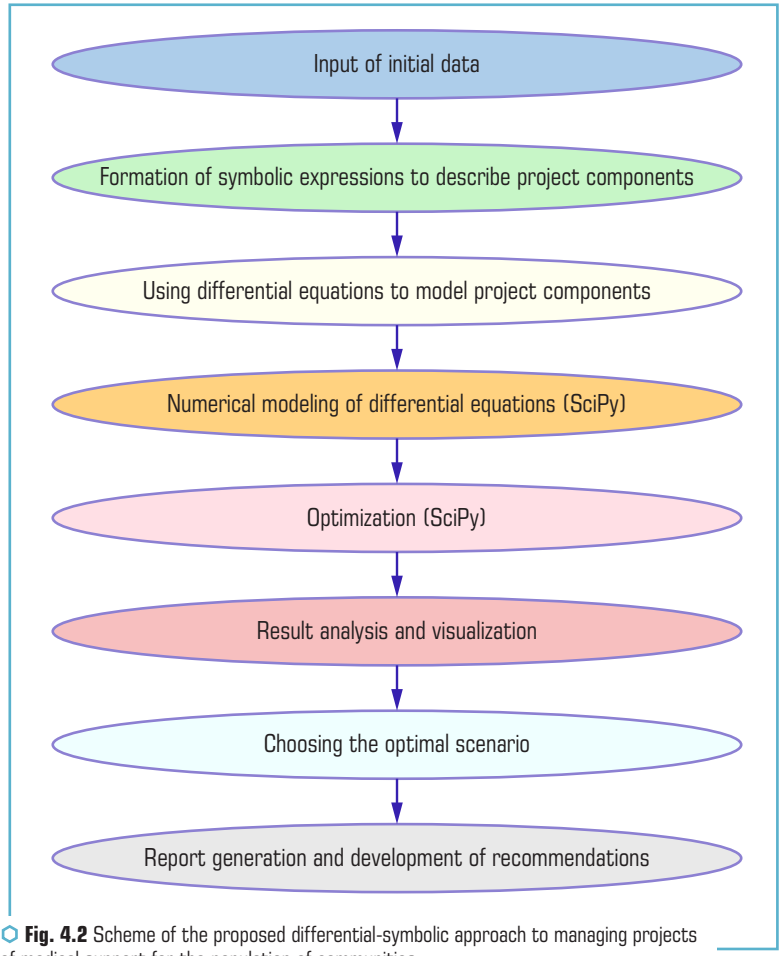


Fig. 4.2 Scheme of the proposed differential-symbolic approach to managing projects of medical support for the population of communities

The presented scheme (**Fig. 4.2**) makes it possible to outline the stages of the process of the differential-symbolic approach to the management of community health support projects, and also demonstrates the need to use mathematical modeling, numerical analysis and optimization

to assess risks and determine the rational configuration of these projects. Based on the proposed scheme for using the differential-symbolic approach to the management of community health support projects, appropriate mathematical models have been developed to solve scientific and applied problems of planning and assessing the risks of these projects.

To implement the proposed approach, our work uses the project management methodology, system and differential-symbolic approaches, which are the basis of the developed algorithms and computer models for planning projects to improve the health of the community population and assessing their risk. To implement the proposed models, code was written in the Python programming language using libraries for solving differential equations, optimizing, and visualizing results. NumPy libraries were used to work with numerical data and vectors, SciPy for numerically solving differential equations and optimizing the objective function, Matplotlib for data visualization, in particular for creating a 3D graph displaying the optimal scenario for implementing community health improvement projects and their quantitative risk assessment.

4.2 MATHEMATICAL MODEL OF DIFFERENTIAL-SYMBOLIC PLANNING OF PROJECTS TO IMPROVE THE HEALTH OF THE POPULATION

The mathematical model of differential-symbolic planning of projects to improve the health of the population involves the use of differential equations to describe the dynamics of projects as a separate system and the use of symbolic expressions to represent individual parameters and their description. Let's consider a general description of this mathematical model. Let's define differential equations that describe the implementation of medical projects to support the population of communities as a separate system.

Let $Y(t)$ – the vector of the state of project implementation in a given community; $P(t)$ – the vector of project configuration (system parameters); $U(t)$ – the vector of project management; $F(Y, P, U, t)$ – the function that determines the dynamics of the implementation of medical projects to support the population of communities. Then the mathematical expression of the differential equation has the form:

$$\frac{dY}{dt} = F\left(Y,P,U,t\right), \tag{4.1}$$

Let's introduce symbolic expressions for the existing state of the population in a given community and its changes. Let P_0 – the initial value of the share of morbidity in the population of the community, and ΔP changes in morbidity in the population of the community. Then the symbolic representation of changes in the population status in a given community looks like this:

$$P\left(t\right) = P_0 + \Delta P\left(t\right). \tag{4.2}$$

Here $P(t)$ represents a vector that characterizes the formation of the product of medical projects supporting the population of communities, which changes over time. It is determined by the budget spent $B(t)$, resources spent $E(t)$, project implementation duration $T(t)$, etc.:

$$P(t)=\begin{bmatrix} B(t) \\ E(t) \\ T(t) \end{bmatrix}. \tag{4.3}$$

Project management for improving the condition of the population of communities is determined by the influence of the project environment on its configuration (system parameters). In particular, an insufficient project budget over time $B(t)$ leads to a transition to another project implementation scenario with a change in the configuration of the desired product. The lack of available resources $E(t)$ at a particular point in time t leads to a change in the implementation scenario for medical projects to support the population of communities. Deviation from the work plan $T(t)$at a particular point in time t leads to changes in the project implementation scenario.

Let $Y(t)$ determine the state of the project at time t. Project implementation can be described by the differential equation:

$$\frac{dY}{dt}=F(Y,P(t),U(t),t), \tag{4.4}$$

where F – a function that determines the dynamics of the project depending on its current state Y, configuration $P(t)$, change management $U(t)$ and time t.

The project management vector $U(t)$ includes individual possible scenarios for its implementation. It concerns changes in the resources used, the work schedule or other important components of the project:

$$U(t)=\begin{bmatrix} U_1(t) \\ U_2(t) \\ \dots \\ U_p(t) \end{bmatrix}. \tag{4.5}$$

So, the vector $U(t)$ defines the processes of managing medical and social projects of community support in the form of justified scenarios of actions. Thus, the product of the project (system parameters) changes based on the ratio:

$$\Delta P(t)=U(t). \tag{4.6}$$

The objective function J can be defined to assess the effectiveness of management decisions and justify scenarios of actions in medical and social projects of community support:

$$J\left(Y,P\left(t\right),U\left(t\right),t\right). \tag{4.7}$$

Optimal management solutions can be found by maximizing or minimizing the objective function taking into account the constraints and conditions of the problem. To introduce constraints in the community support project regarding the budget $B(t)$, available resources $E(t)$, duration of work $T(t)$, let's use conditions that take into account the maximum values of these parameters. Let B_{max}, E_{max}, and T_{max} represent the maximum allowable values for the budget, resources involved and duration of work in the project, respectively. Then the constraints can be expressed as follows:

$$\begin{aligned} &B\left(t\right)\le B_{max};\\ &\left(t\right)\le \ _{\max};\\ &T\left(t\right)\le T_{max}. \end{aligned} \tag{4.8}$$

The objective function J is the minimum cost of funds for the implementation of the project with the maximum increase in the percentage of healthy population. Minimizing the objective function J, let's search for the optimal scenario for the implementation of the public health improvement project, which will ensure high-quality management:

$$min_{U(t)}J\left(Y,P_0+U,t\right). \tag{4.9}$$

Solving the optimization problem, the objective function J, taking into account the specified restrictions in expression (4.8), is written as follows:

$$\begin{aligned} &min_{U(t)}J\left(Y,P_0+U,t\right),\\ &\text{provided } B\left(t\right)\le B_{max},E\left(t\right)\le E_{\max},T\left(t\right)\le T_{max}, \end{aligned} \tag{4.10}$$

where J – the objective function for optimizing the scenario for the implementation of the medical project for supporting the population of communities; $U(t)$ – the management vector (implementation scenario) of the medical project for supporting the population of communities; Y – the vector of the state of the medical project implementation for supporting the population of communities in a given community; P_0 – the initial value of the share of morbidity in the community population; t – time.

If there is a need to provide for other conditions or restrictions in the medical project for supporting the population of communities, they can be added to the described system, taking into account all aspects of the task of managing the specified projects.

4.3 MATHEMATICAL MODEL OF DIFFERENTIAL-SYMBOLIC RISK ASSESSMENT OF PROJECTS TO IMPROVE THE HEALTH OF THE POPULATION

Public health improvement projects are an important component of community development and quality of life. The products of such projects can be various activities, such as educational programs, vaccinations and educational initiatives, which are aimed at improving the overall health of the community population. However, the implementation of such projects is associated with numerous challenges and risks that can affect their success. Assessing and managing these risks is key to ensuring the effectiveness and sustainability of projects.

One approach to risk assessment is the use of mathematical models that allow predicting and analyzing possible scenarios. The differential-symbolic approach to risk modeling allows taking into account the dynamic implementation of public health improvement projects, which include changes in their components over time and the impact of various factors of the project environment.

Let's propose a mathematical model for differential-symbolic risk assessment of public health improvement projects. This model is based on a system of differential equations that describe the dynamics of basic project indicators, such as the percentage of the population that participated in educational activities, vaccinations, and educational programs. The model also includes an assessment of changes in the project budget and takes into account the impact of these changes on the overall risk of the project.

The main goal of creating this model is to provide project managers with tools for quantifying risks and supporting decision-making in the process of managing projects to improve the health of the population. This will optimize the use of resources, minimize risks, and increase the efficiency of project implementation, which will ultimately contribute to improving the health of the population in communities.

To assess changes in the health of the population in communities, let's use differential equations that describe the change in key project indicators over time. In particular, let's consider indicators such as the percentage of the healthy population that participated in educational activities, vaccinations, and received health education.

It is possible to assume that the percentage of the healthy population $Y_1(t)$, $Y_2(t)$ and $Y_3(t)$, that participated in the relevant activities at time t, is known in advance. It is possible to write the dynamics equation for the implementation of educational activities:

$$\frac{dY_1(t)}{dt} = \alpha_1(1 - Y_1(t)) - \beta_1 Y_1(t), \tag{4.11}$$

where α_1 – the rate of involvement of the healthy population in educational activities; β_1 – the rate of decrease in the percentage of the involved population due to various factors (for example, loss of interest, etc.).

Regarding the dynamics equation for vaccination, it can be written as follows:

$$\frac{dY_2(t)}{dt} = \alpha_2(1 - Y_2(t)) - \beta_2 Y_2(t), \tag{4.12}$$

where α_2 – the rate of involvement of the healthy population in vaccination; β_2 – the rate of decrease in the percentage of the involved population in vaccination.

The dynamics equation for educational programs can be written as follows:

$$\frac{dY_3(t)}{dt} = \alpha_3(1 - Y_3(t)) - \beta_3 Y_3(t), \tag{4.13}$$

where α_3 – the rate of involvement of the healthy population in health educational programs; β_3 – the rate of decrease in the percentage of the involved population in educational programs.

At the beginning of the simulation, the condition is assumed that $t=0$:

$$\begin{aligned} Y_1(0) &= initial(Y_1); \\ Y_2(0) &= initial(Y_2); \\ Y_3(0) &= initial(Y_3), \end{aligned} \tag{4.14}$$

where $initial(Y_1)$ – the initial percentage of the healthy population that participated in educational activities; $initial(Y_2)$ – the initial percentage of the healthy population that participated in vaccination; $initial(Y_3)$ – the initial percentage of the healthy population that received health education.

The total budget of the population health improvement project $B(t)$ includes a baseline value and can change over time depending on the costs of the activities:

$$B(t) = initial_B - c_1\int_0^t Y_1(\tau)d\tau - c_2\int_0^t Y_2(\tau)d\tau - c_3\int_0^t Y_3(\tau)d\tau, \tag{4.15}$$

where c_1, c_2, c_3 – the costs per participant for educational activities, vaccination and education, respectively.

The risk of the population health improvement project $R(t)$ can be assessed based on the deviation from the planned indicators:

$$R(t) = \gamma_1 \left| initial(Y_1) - Y_1(t) \right| + \gamma_2 \left| iinitial(Y_2) - Y_2(t) \right| + \gamma_3 \left| initial(Y_3) - Y_3(t) \right|, \tag{4.16}$$

where γ_i – the risk weighting coefficients for each of the indicators.

For the numerical solution of differential equations, the Euler method can be used. Let the discrete time points t_i, $i=0, 1, \ldots, n$, be known, then:

$$\begin{aligned} Y_1(t_{i+1}) &= Y_1(t_i) + h\left(\alpha_1(1 - Y_1(t_i)) - \beta_1 Y_1(t_i)\right); \\ Y_2(t_{i+1}) &= Y_2(t_i) + h\left(\alpha_2(1 - Y_2(t_i)) - \beta_2 Y_2(t_i)\right); \\ Y_3(t_{i+1}) &= Y_3(t_i) + h\left(\alpha_3(1 - Y_3(t_i)) - \beta_3 Y_3(t_i)\right), \end{aligned} \quad (4.17)$$

where h – the solution step.

The solution step is determined by the formula:

$$h = \frac{e_t - s_t}{n}, \quad (4.18)$$

where e_t – the completion time of the project status assessment; s_t – time of project assessment; n – number of stages.

The presented equations (4.17) and (4.18) allow to assess the dynamics of changes in the health status of the population during the implementation of projects to improve the health status of the population, taking into account the impact of educational activities, vaccination and educational programs.

The proposed mathematical model allows to assess the dynamics of changes in the health status of the population, the project budget and risks throughout the entire period of implementation of the project to improve the health status of the population.

4.4 ALGORITHM AND COMPUTER MODEL OF DIFFERENTIAL-SYMBOLIC PLANNING OF PROJECTS TO IMPROVE THE HEALTH OF THE POPULATION

We have developed an algorithm for creating a computer model of differential-symbolic planning of projects to improve the health of the population, the block diagram of which is presented in **Fig. 4.3**.

Differential-symbolic planning of projects to improve the health of the population involves the implementation of 16 steps, which involve the use of formulas (4.1)–(4.10):

1. Data collection for planning medical projects to support the population of communities.

2. Determination of the main parameters, which include the current health and activity of the community population, the project budget, available resources, the duration of the project implementation and other important characteristics.

3. Formation of differential equations involves writing the differential equation (4.1) based on the function $F(Y, P, U, t)$, which determines the dynamics of project implementation.

4. Symbolic definition of the project configuration, which provides a representation of the project configuration $P(t)$ using equation (4.3).

5. Formation of the project management vector $U(t)$ using equation (4.5) and conditions (4.6).

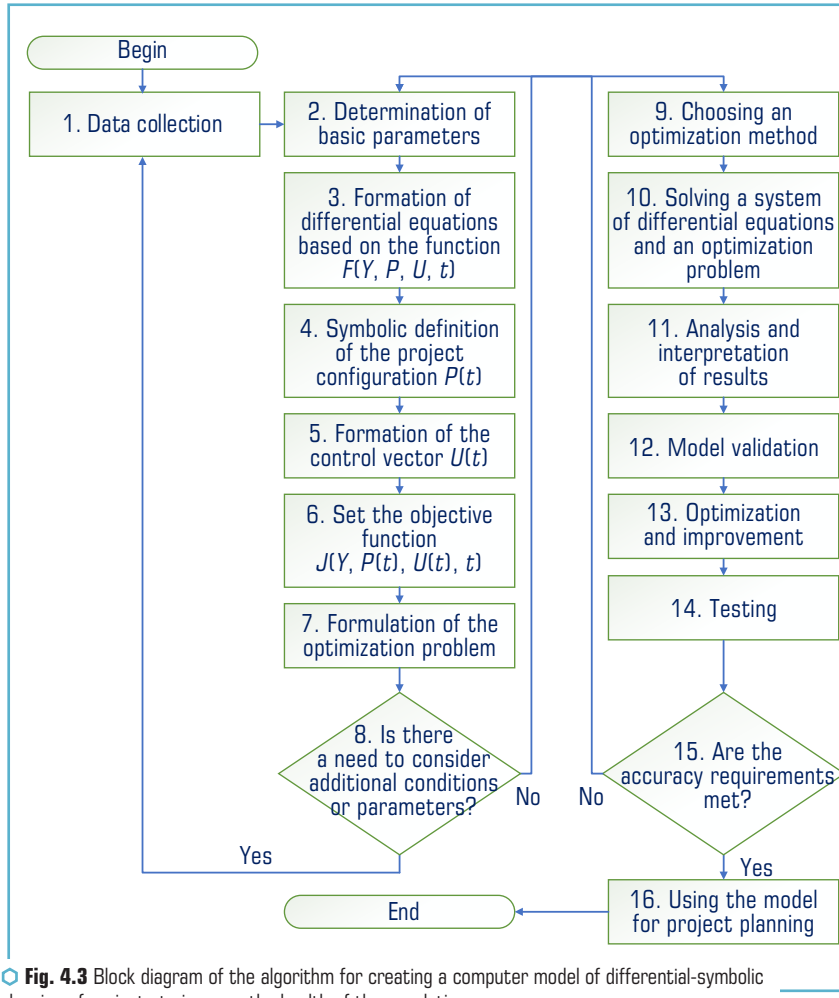


Fig. 4.3 Block diagram of the algorithm for creating a computer model of differential-symbolic planning of projects to improve the health of the population

6. Set the objective function $J(Y, P(t), U(t), t)$ in accordance with the defined goals and constraints.

7. Formulation of the optimization problem using equation (4.10) and condition (4.8).

8. Checking the condition whether there is a need to take into account additional conditions or parameters. If yes, then go to Step 1. In this case, additional conditions or parameters should

be added that may affect the dynamics of the project (for example, restrictions on the maximum values of the budget, resources and duration). If not, then go to Step 9.

9. Select an optimization method to solve the optimization problem. This can be a numerical method, metaheuristics or analytical method, depending on the complexity of the project and the model.

10. Solving the system of differential equations and the optimization problem. In this case, the selected method is used to solve the system of differential equations and the optimization problem.

11. Analysis and interpretation of the results, which provides an assessment of the results obtained, an analysis of the impact of various parameters and the quality of management decisions on the implementation of the project.

12. Model validation ensures verification of the correctness and adequacy of the model, its results are compared with empirical data or literature sources.

13. Optimization and improvement ensures the implementation of necessary adjustments and, if necessary, optimization of the model. The model is documented, including the introduced assumptions, mathematical equations, optimization methods and other important aspects.

14. Testing the model with various input parameters and checking its stability and reliability.

15. Checking the condition whether the accuracy requirements are met. If yes, then go to Step 16. If not, go to Step 2.

15. Using the model for planning medical projects to support the population of communities.

Based on the proposed algorithm, we created a computer model of differential-symbolic planning of medical projects to support the population of communities. The code is written in the Python programming language using libraries for solving differential equations, optimizing and visualizing results. In particular, NumPy was used to work with numerical data and vectors. SciPy for numerical solution of differential equations and optimization of the objective function, Matplotlib for data visualization, in particular for creating 3D graphs. The SLSQP optimization method was used to find the optimal values of control parameters that minimize the objective function under given constraints.

4.5 ALGORITHM AND COMPUTER MODEL FOR DIFFERENTIAL-SYMBOLIC RISK ASSESSMENT OF PROJECTS TO IMPROVE THE HEALTH OF THE POPULATION

We have developed an algorithm for creating a computer model for differential-symbolic risk assessment of projects to improve the health of the population, the block diagram of which is presented in **Fig. 4.4**.

Differential-symbolic risk assessment of projects to improve the health of the population involves the implementation of 9 steps, which involve the use of formulas (4.11)–(4.18):

1. Initialization of variables to perform risk assessment of projects to improve the health of the population. To do this, let's assume that the percentage of the healthy population is previously

known, and then, based on formulas (4.11)–(4.13), write the dynamics equation for the implementation of measures to improve the health of the population. Let's fix the initial percentage of the healthy population that participated in various measures using formula (4.14). Let's set the initial value of the budget and duration of the project implementation.

2. Determination of the solution step using formula (4.18).

3. Checking the compliance condition of all initial data. If yes, then go to Step 4. If no, then go to Step 1 and make changes to the initial data to perform risk assessment of projects to improve the health of the population.

4. For numerical solution of differential equations use Euler's method using formulas (4.17).

5. Calculate the budget of the project to improve the health of the population using formula (4.15) and conditions: $\int_0^t Y_1(\tau)d\tau \approx h\sum_{j=0}^{i} Y_1(t_j)$, $\int_0^t Y_2(\tau)d\tau \approx h\sum_{j=0}^{i} Y_2(t_j)$ and $\int_0^t Y_3(\tau)d\tau \approx h\sum_{j=0}^{i} Y_3(t_j)$.

6. Assess the risks of the project to improve the health of the population $R(t)$ based on the deviations found from the planned indicators using equation (4.16).

7. Update the counter of the time of implementation of the project to improve the health of the population – $t=t+h$.

8. Check the condition whether the specified stage of modeling does not exceed the completion time of the project to improve the health of the population. If yes, then go to Step 4. In this case, it is possible to update the values of the indicators for the numerical solution of differential equations. If not, then go to Step 9.

9. Using the model to assess the risks of projects to improve the health of the population.

Based on the proposed algorithm, we created a computer model for differential-symbolic risk assessment of population health improvement projects. The code is written in the Python programming language using libraries for solving differential equations, optimizing and visualizing results. The NumPy library is used to work with numerical arrays and perform mathematical operations. It is used to implement numerical solutions of differential equations and calculations in the model. The Matplotlib library was used to visualize the modeling results. This allowed to create graphs of changes in the percentage of healthy population, budget and risk during the project. It is also used to add a vertical line to the graphs, indicating the optimal risk value. The Pandas library is used to create and process data tables. It is used to conveniently display the initial data, modeling results and indicators for a given risk level in the form of tables.

The structure of the computer model code consists of the following main blocks:

1) initialization of the initial data and model parameters;

2) checking the initial data for correctness;

3) simulation cycle that performs numerical solution of differential equations and records the results;

4) output and storage of results in the form of graphs and tables;

5) addition of analysis and visualization of the optimal risk value.

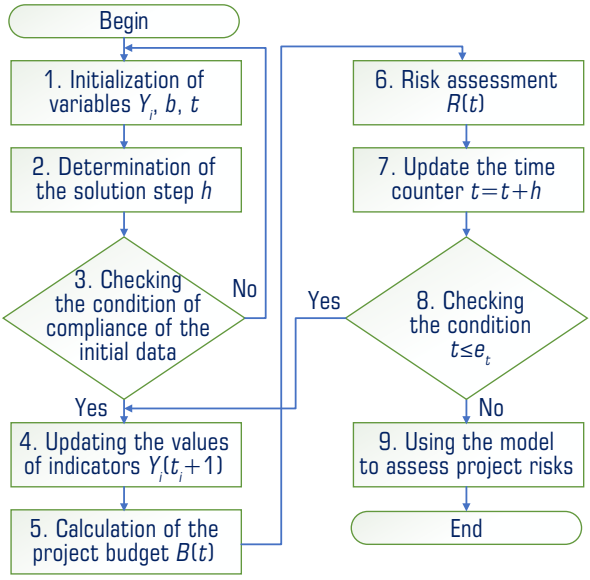


Fig. 4.4 Block diagram of the algorithm for creating a computer model for differential-symbolic risk assessment of projects to improve the health of the population

4.6 RESULTS OF DIFFERENTIAL-SYMBOLIC PLANNING OF PROJECTS TO IMPROVE THE HEALTH OF THE POPULATION

The developed computer model of differential-symbolic planning was tested on the example of projects to improve the health of the population of communities. The initial data for optimizing the implementation scenario of the project to improve the health of the population of communities are presented in **Table 4.1**.

The constraints are set taking into account expression (4.8) in the form of a vector as the maximum values of the project budget B_{max} = 100000 USD, available resources E_{max} =30000 USD (part of the budget allocated for information campaigns, training and educational activities, purchase of necessary equipment and resources to ensure the successful implementation of population activities in the project) and the duration of the project T_{max} = 24 months.

Based on the conducted research, we have constructed dependences of the growth rate of the percentage of the healthy population who participated in educational activities during the implementation of the project in different configurations (**Fig. 4.5**).

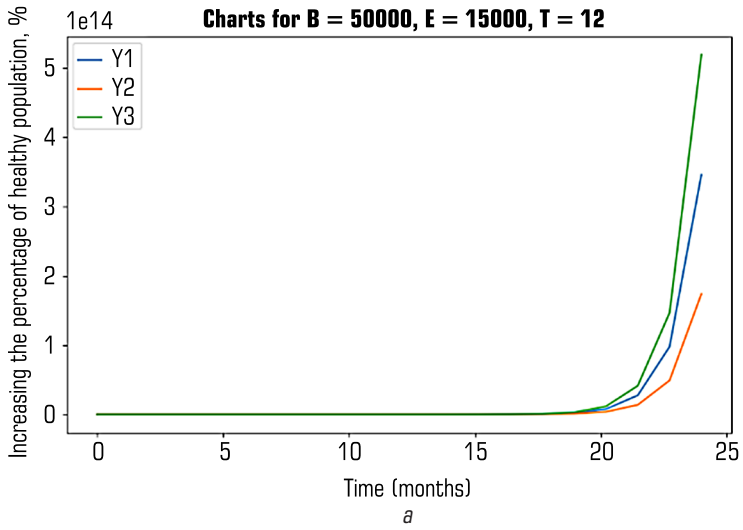
Table 4.1 Initial data for optimizing the implementation scenario of the project to improve the health of the population of communities

Indicator	Designation	Unit of measurement	Value
Percentage of healthy population that participated in educational activities	initial_Y1	%	41
Percentage of healthy population that participated in vaccination	initial_Y2	%	55
Percentage of healthy population that received health education	initial_Y3	%	65
Baseline value of project budget	initial_B	USD	50000
Baseline value of available human resources allocated to implement activities in the community	initial_E	USD	15000
Initial duration of work in the project	initial_T	month	12
Initial simulation time	start_time	month	0
Final simulation time	end_time	month	24
Number of points for numerical solution of differential equations	num_points	pcs	20

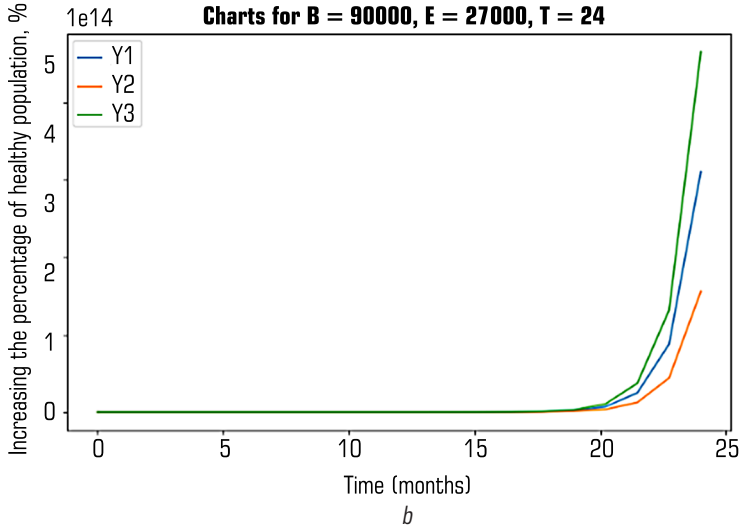
It was found that regardless of the basic value of the project budget and available human resources allocated for the implementation of activities in the community, the percentage of the healthy population who participated in educational activities begins to increase only after 18 months of implementation of the specified projects.

Subsequently, using the developed computer model of differential-symbolic project planning, which provides mathematical modeling of the dynamics of changes in the health status of the community population, a numerical solution of differential equations was performed using the open library SciPy. The results of the numerical solution of differential equations are presented in **Table 4.2**.

We conducted a study of the impact of changing parameters on the model results – on the percentage of healthy population in the community (initial_Y1). To do this, we constructed sensitivity graphs for each of the considered parameters (initial_Y1, initial_Y2, initial_Y3, initial_B, initial_E, initial_T) (**Fig. 4.6**). This graph presents a sensitivity analysis of each of the considered parameters to the parameter initial_Y1, which corresponds to the percentage of the healthy population participating in educational activities. The curves in the graph show changes in the values of the models when the parameter initial_Y1 is increased and decreased by 10 %. The blue line reflects the baseline value of the parameter. The green line represents the parameter value increased by 10 % from the baseline value. The red line shows the parameter value decreased by 10 % from the baseline value. It was found that the greatest impact on the percentage of the healthy population participating in educational activities initial_Y1 is the initial value of the project budget initial_B and the initial amount of financial and human resources initial_E, which are intended for the implementation of activities in the community. They have the largest deviation between the baseline values and the changed values. This indicates a significant impact on the parameter initial_Y1 of the model result.



a



b

Fig. 4.5 Dependences of the increasing the percentage of healthy population who participated in educational activities during the project implementation for the following configurations: *a* – B = 50000 USD, E = 15000 USD, T = 12 months; *b* – B = 90000 USD, E = 27000 USD, T = 24 months

Table 4.2 Results of the numerical solution of differential equations

Point number for numerical solution	Time (months)	initial_ Y1	initial_ Y2	initial_ Y3	Point number for numerical solution	Time (months)	initial_ Y1	initial_ Y2	initial_ Y3
0	0.00 0000	4.1000 00e+01	5.5000 00e+01	6.5000 00e+01	10	12.63 1579	3.9922 36e+09	2.0066 97e+09	5.9894 14e+09
1	1.26 3158	3.3126 65e+04	1.6685 50e+04	4.9702 27e+04	11	13.89 4737	1.4118 86e+10	7.0968 46e+09	2.1182 04e+10
2	2.52 6316	1.5013 64e+05	7.5500 47e+04	2.2524 80e+05	12	15.15 7895	4.9932 41e+10	2.5098 52e+10	7.4911 87e+10
3	3.78 9474	5.6395 00e+05	2.8350 38e+05	8.4607 83e+05	13	16.42 1053	1.7658 96e+11	8.8762 75e+10	2.6493 13e+11
4	5.05 2632	2.0274 31e+06	1.0191 23e+06	3.0416 89e+06	14	17.68 4211	6.2452 18e+11	3.1391 59e+11	9.3694 87e+11
5	6.31 5789	7.2031 39e+06	3.6206 92e+06	1.0806 63e+07	15	18.94 7368	2.2086 66e+12	1.1101 86e+12	3.3135 86e+12
6	7.57 8947	2.5507 40e+07	1.2821 33e+07	3.8267 88e+07	16	20.21 0526	7.8111 08e+12	3.9262 53e+12	1.1718 74e+13
7	8.84 2105	9.0241 74e+07	4.5360 05e+07	1.3538 66e+08	17	21.47 3684	2.7624 55e+13	1.3885 48e+13	4.1444 16e+13
8	10.10 5263	3.1917 94e+08	1.6043 56e+08	4.7885 39e+08	18	22.73 6842	9.7696 20e+13	4.9106 99e+13	1.4657 03e+14
9	11.36 8421	1.1288 34e+09	5.6740 84e+08	1.6935 51e+09	19	24.00 0000	3.4550 97e+14	1.7367 04e+14	5.1835 63e+14

The created model allowed to identify possible scenarios (**Fig. 4.7**) for the implementation of projects to improve the health of the community population, as well as to choose the optimal scenario from among a set of alternative ones (**Table 4.3**).

The obtained optimization results show that the values of the objective function (J), which concerns the minimization of the cost of funds for the implementation of the project with the maximum increase in the percentage of healthy population, fall on the scenario that assumes the following optimal values of parameters:

1) project budget $B^{opt} = 45000$ USD;

2) part of the project budget, which is provided for the purchase of necessary equipment and resources to ensure the successful implementation of population measures in the project $E^{opt} = 14250$ USD;

3) duration of the project after its initiation $T^{opt} = 9.6$ months.

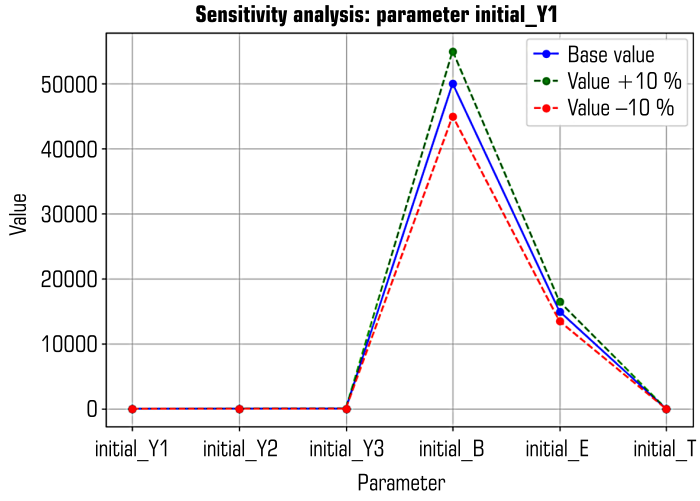


Fig. 4.6 Sensitivity graphs of each of the considered parameters relative to the percentage of healthy population in the community

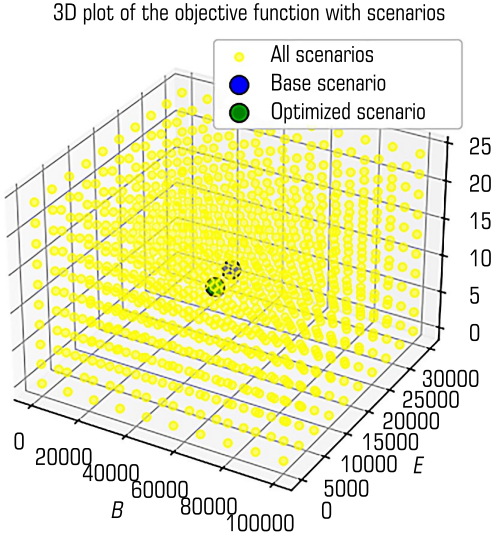


Fig. 4.7 Results of determining the optimal scenario for implementing a project to improve the health of the population in the community

Table 4.3 Results of optimizing the scenario for implementing a project to improve the health of the population in the community

Scenario	Y1	Y2	Y3	Budget used	Resources used	Duration (months)
Basic	41	55	65	50000.0	15000.0	12.0
Optimal	51	67	80	45000.0	14250.0	9.6

4.7 RESULTS OF DIFFERENTIAL-SYMBOLIC RISK ASSESSMENT OF PROJECTS TO IMPROVE THE HEALTH OF THE POPULATION

The developed computer model for differential-symbolic risk assessment of community health improvement projects was tested on the example of community health improvement projects. The initial data for risk assessment of community health improvement projects are presented in **Table 4.4**.

Table 4.4 Initial data for risk assessment of community health improvement projects

Initial percentage of healthy population, %	Initial budget, UAH	Project duration, months	Time step, months	Coefficient of impact of measures on health, α	Coefficient of budget expenditure on health improvement, β	Coefficient of risk reduction, γ
60	1000000	24	1	0.1	0.05	0.02

It is assumed that the initial percentage of healthy population is 60 %. This value indicates that 60 % of the community population is healthy at the beginning of the project. The available budget of the community health improvement project is 1 million UAH. This is the main resource that will be spent on the implementation of the health improvement project. The project is planned to be implemented within 24 months. This determines the total time during which measures to improve the health of the population will be implemented. The selected time step for modeling and assessing the project is 1 month. This means that the model will be updated monthly.

Further, using the developed computer model of differential-symbolic risk assessment of health improvement projects, which provides mathematical modeling of the dynamics of changes in the health of the community population, a numerical solution of differential equations was performed. The results of the simulation based on the numerical solution of differential equations are presented in **Table 4.5**.

We conducted a study of the impact of the project implementation period and the duration of its life cycle on risk. For this purpose, let's develop the dependencies (**Fig. 4.8–4.10**).

It was found that the percentage of healthy population gradually increases with the duration of the project to improve the health of the population (**Fig. 4.8**). This indicates the effectiveness of the measures implemented within the project [32]. The increase in the percentage of healthy

population of the community by 21.2 % during the 24 months of project implementation demonstrates the positive impact of the implemented measures on improving the health of the population.

Table 4.5 Results of risk simulation based on the numerical solution of differential equations

Time, months	Percentage of healthy population, %	Project budget, UAH	Risk
1	60.0	1000000	0.80
2	60.6	998600	0.79
3	61.1	997200	0.78
4	61.7	995800	0.77
5	62.2	994400	0.76
...	...	...	...
22	80.5	978600	0.39
23	80.9	977200	0.38
24	81.2	975800	0.37

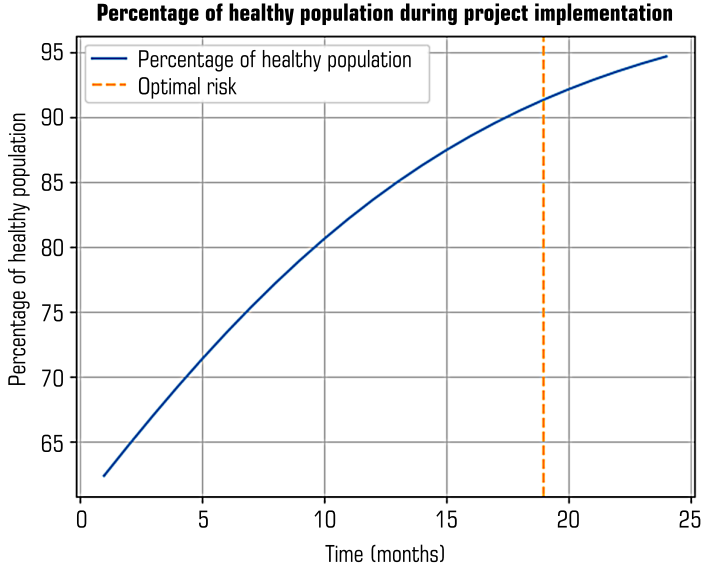


Fig. 4.8 Dependence of the percentage of healthy population of the community on the duration of the project implementation to improve the health of the population

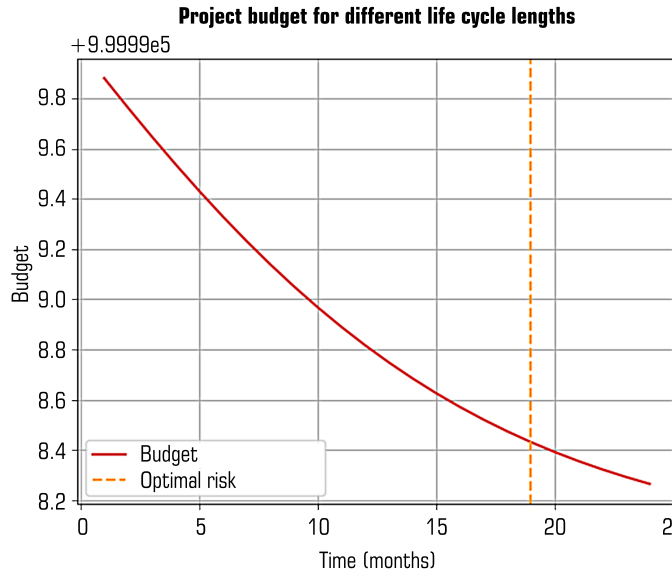


Fig. 4.9 Dependence of the change in the budget requirement of the project to improve the health of the population on the duration of its implementation

The required budget of the project to improve the health of the population decreases over time (**Fig. 4.9**), which may be the result of the costs of implementing the measures or less funding due to the achievement of the set goals [33]. It was found that the risk decreases during the project implementation (**Fig. 4.10**), which indicates the successful reduction of risks associated with the health of the population. A risk reduction of 0.43 (or 53.75 %) over 24 months is a significant achievement and confirms the positive impact of the project implementation on the health of the population of the community.

Overall, the implementation of the project to improve the health of the population demonstrates positive results, as the percentage of healthy population increases and the risk decreases [34]. This indicates that the implemented measures have a significant impact on improving the health of the population [35].

The created model allowed to identify possible risks (**Fig. 4.4**) for different durations and budgets of the implementation of projects to improve the health of the community population, as well as determine the optimal risk from a set of alternative project implementation options (**Table 4.6**).

It was established that the optimal risk of the project to improve the health of the community population is at the level of 0.20. At the same time, the project requires 16 months to implement, which confirms its high effectiveness in controlling and reducing risks [36]. This also shows the

effectiveness of management measures and the economic benefit from implementing the project according to the proposed scenario with the minimum possible risk and maximum benefits [37].

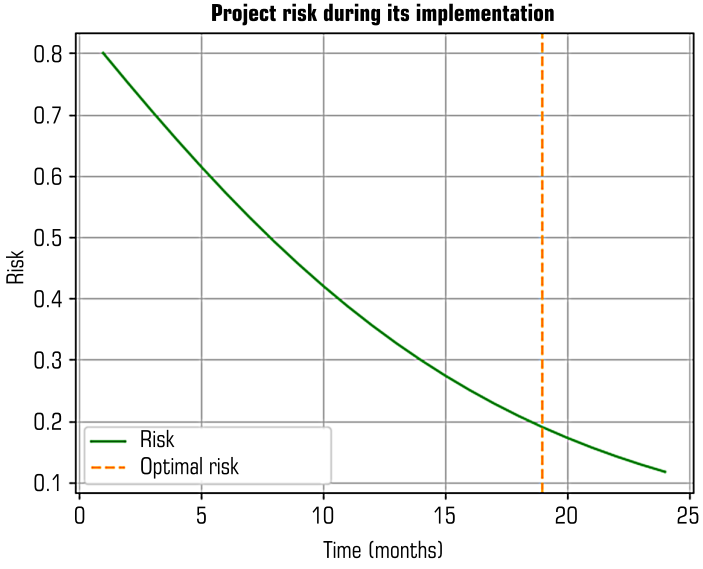


Fig. 4.10 Dependence of the risk of the project to improve the health of the population on the duration of its implementation

Table 4.6 Results of risk optimization during the implementation of the project to improve the health of the community population

Duration of project implementation, months	Percentage of healthy population, %	Project budget, UAH	Risk
16	75.0	987000	0.20

Thus, the proposed differential-symbolic approach, which is implemented in the presented algorithms and computer models, is a fairly simple and effective tool for project managers [38], which ensures accurate planning of community health improvement projects and risk assessment of these projects [39]. This is confirmed by the results obtained based on the conducted study of the influence of project environment components on the choice of the optimal scenario for the implementation of community health improvement projects and the quantitative value of the risks of these projects. For further research, all possible constraints, the function of the object and internal optimization

parameters should be reviewed. The optimization method can also be changed, which will ensure the formation of real constraints and take into account additional factors to improve the stability of optimization [40]. Based on computer models of differential-symbolic planning of community health improvement projects and risk assessment, management plans for these projects can be developed [41]. Quantitative assessment of the effectiveness of the models has shown that it can be an effective tool for project managers. Models allow to effectively manage the configuration and risks of projects and minimize the cost of their implementation while maximizing the percentage of healthy population.

CONCLUSIONS

Analysis of recent studies and publications shows that planning of medical projects for community support is an important aspect of managing these projects. Given the growth of medical data, rapid technological changes and expansion of the capabilities of analytical tools, computer models are becoming an integral part of planning the implementation of these projects.

The main stages of the differential-symbolic approach to managing community medical support projects involve the use of mathematical modeling, numerical analysis and optimization to determine the rational configuration of these projects. Based on the scheme of the proposed differential-symbolic approach to managing community medical support projects, appropriate mathematical models have been developed, which made it possible to develop algorithms for differential-symbolic planning of community medical support projects and risk assessment of these projects.

Based on the proposed algorithms, computer models for differential-symbolic planning of medical projects to support the population of communities and risk assessment of these projects have been developed. They are based on the mathematical models we have proposed, written in the Python programming language using libraries for solving differential equations, optimizing and visualizing results. The models allow to perform labor-intensive calculations to form possible scenarios for the implementation of medical projects to support the population of communities and assess the risk of these projects, as well as visualize possible configurations of projects and determine their optimal scenario.

Using the proposed computer models on the example of planning a project to improve the health of the population in a community made it possible to obtain the dependence of the growth rate of the percentage of the healthy population who participated in educational activities on the configuration of projects to improve the health of the population of communities. In addition, it made it possible to determine the optimal scenario for the implementation of a project to improve the health of the population in a community, which demonstrates its effectiveness and practical use.

The prospect of further research is to expand the functionality of computer models, adding modules for analyzing other components of medical projects to support the population of communities. Based on the use of computer models, it is possible to conduct research on the effectiveness of the implementation of various types of medical projects, taking into account the real conditions of the project environment.

REFERENCES

1. Bushuyev, S., Bushuiev, D., Bushuieva, V. (2020). Interaction Multilayer model of Emotional Infection with the Earn Value Method in the Project Management Process. 2020 IEEE 15th International Conference on Computer Sciences and Information Technologies (CSIT), 146–150. https://doi.org/10.1109/csit49958.2020.9321949
2. Gogunskii, V. D., Kolesnikova, K. V., Lukianov, D. V. (2022). Entropy analysis of organizations' knowledge systems on the example of project management standards. Applied Aspects of Information Technology, 5 (2), 91–104. https://doi.org/10.15276/aait.05.2022.7
3. Tryhuba, A., Ratushny, R., Tryhuba, I., Koval, N., Androshchuk, I. (2020). The Model of Projects Creation of the Fire Extinguishing Systems in Community Territories. Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis, 68 (2), 419–431. https://doi.org/10.11118/actaun202068020419
4. Koval, N., Kondysiuk, I., Tryhuba, I., Boiarchuk, O., Rudynets, M., Grabovets, V., Onyshchuk, V. (2021). Forecasting the fund of time for performance of works in hybrid projects using machine training technologies. Proceedings of the 3nd International Workshop on Modern Machine Learning Technologies and Data Science Workshop. Proc. 3rd International Workshop (MoMLeT&DS 2021), I, 96–206. Available at: https://ceur-ws.org/Vol-2917/paper18.pdf
5. Tryhuba, A., Boyarchuk, V., Koval, N., Tryhuba, I., Boiarchuk, O., Pavlikha, N. (2021). Risk-adapted model of the lifecycle of the technologically integrated programs of dairy cattle breeding. 2021 IEEE 16th International Conference on Computer Sciences and Information Technologies (CSIT), 2, 307–310. https://doi.org/10.1109/csit52700.2021.9648672
6. Malanchuk, O., Tryhuba, A., Tryhuba, I., Bandura, I. (2023). A conceptual model of adaptive value management of project portfolios of creation of hospital districts in Ukraine. CEUR Workshop Proceedings. 3453, 82–95. Available at: https://ceur-ws.org/Vol-3453/paper8.pdf
7. Tryhuba, A., Padyuka, R., Tymochko, V., Lub, P. (2022). Mathematical model for forecasting product losses in crop production projects. CEUR Workshop Proceedings, 3109, 25–31. Available at: https://ceur-ws.org/Vol-3109/paper8.pdf
8. Kovalchuk, O., Zachko, O., Kobylkin, D. (2022). Criteria for intellectual forming a project teams in safety oriented system. 2022 IEEE 17th International Conference on Computer Sciences and Information Technologies (CSIT), 430–433. https://doi.org/10.1109/csit56902.2022.10000494
9. Mainga, W. (2017). Examining project learning, project management competencies, and project efficiency in project-based firms (PBFs). International Journal of Managing Projects in Business, 10 (3), 454–504. https://doi.org/10.1108/ijmpb-04-2016-0035
10. Rodionov, D. G., Pashinina, P. A., Konnikov, E. A., Konnikova, O. A. (2022). Information Environment Quantifiers as Investment Analysis Basis. Economies, 10(10), 232. https://doi.org/10.3390/economies10100232

11. Daemi, A., Chugh, R., Kanagarajoo, M. V. (2021). Social media in project management: A systematic narrative literature review. International Journal of Information Systems and Project Management, 8 (4), 5–21. https://doi.org/10.12821/ijispm080401
12. Boyarchuk, V., Tryhuba, I., Pavlikha, N., Kovalchuk, N. (2021). Study of the impact of the volume of investments in agrarian projects on the risk of their value. CEUR Workshop Proceedings, 2851, 303–313. Available at: https://ceur-ws.org/Vol-2851/paper28.pdf
13. Malanchuk, O. M., Tryhuba, A. M., Pankiv, O. V., Sholudko, R. Ya. (2023). Architecture of an intelligent information system for forecasting components of medical projects. Applied Aspects of Information Technology, 6 (4), 376–390. https://doi.org/10.15276aait.06.2023.25
14. Tryhuba, A., Malanchuk, O., Tryhuba, I. (2023). Prediction of the duration of inpatient treatment of diabetes in children based on neural networks. Proceedings of the Modern Machine Learning Technologies and Data Science Workshop (MoMLeT&DS 2023), 3426, 122–135. Available at: https://ceur-ws.org/Vol-3426/paper10.pdf
15. Antoshchuk, S. G., Breskina, A. A. (2023). Human action analysis models in artificial intelligence based proctoring systems and dataset for them. Applied Aspects of Information Technology, 6 (2), 190–200. https://doi.org/10.15276/aait.06.2023.14
16. Ratushny, R., Tryhuba, A., Bashynsky, O., Ptashnyk, V. (2019). Development and Usage of a Computer Model of Evaluating the Scenarios of Projects for the Creation of Fire Fighting Systems of Rural Communities. 2019 XIth International Scientific and Practical Conference on Electronics and Information Technologies (ELIT), 34–39. https://doi.org/10.1109/elit.2019.8892320
17. Tryhuba, A., Boyarchuk, V., Tryhuba, I., Ftoma, O., Tymochko, V., Bondarchuk, S. (2020). Model of Assessment of the Risk of Investing in the Projects of Production of Biofuel Raw Materials. 2020 IEEE 15th International Conference on Computer Sciences and Information Technologies (CSIT), 151–154. https://doi.org/10.1109/csit49958.2020.9322024
18. Rogovskii, I., Sivak, I., Shatrov, R., Nadtochiy, O. (2024). Agroengineering studies of tillage and harvesting parameters in soybean cultivation. 23rd International Scientific Conference Engineering for Rural Development Proceedings, 23, 965–970. https://doi.org/10.22616/erdev.2024.23.tf195
19. Rogovskii, I. L., Reznik, N. P., Osadchuk, N. V., Ivanova, T. M., Zinchenko, M. M., Melnyk, L. Yu., Ryzhakova, H. (2024). Institutional Aspects of Development of Budget System: Theory and Practice of Ukraine. Intelligent Systems, Business, and Innovation Research, 925–937. https://doi.org/10.1007/978-3-031-36895-0_78
20. Tryhuba, A., Komarnitskyi, S., Tryhuba, I., Hutsol, T., Yermakov, S., Muzychenko, A. et al. (2022). Planning and Risk Analysis in Projects of Procurement of Agricultural Raw Materials for the Production of Environmentally Friendly Fuel. International Journal of Renewable Energy Development, 11 (2), 569–580. https://doi.org/10.14710/ijred.2022.43011
21. Pons, N. L., Pérez, Y. P., Stiven, E. R., Quintero, L. P. (2014). Design of a knowledge management model for improving the development of computer projects' teams. Revista Espanola de Documentacion Cientifica, 37 (2), e044. https://doi.org/10.3989/redc.2014.2.1036

22. Tryhuba, A., Kondysiuk, I., Tryhuba, I., Koval, N., Boiarchuk, O., Tatomyr, A. (2020). Intellectual information system for formation of portfolio projects of motor transport enterprises. CEUR Workshop Proceedings, 3109, 44–52. Available at: http://ceur-ws.org/Vol-3109/paper7.pdf
23. Sheichenko, V., Petrachenko, D., Rogovskii, I., Dudnikov, I., Shevchuk, V., Sheichenko, D. et al. (2024). Determining patterns in the separation of hemp seed hulls. Engineering Technological Systems, 4 (1 (130)), 54–68. https://doi.org/10.15587/1729-4061.2024.309869
24. Aulin, V., Rogovskii, I., Lyashuk, O., Tykhyi, A., Kuzyk, A., Dvornyk, A. et al. (2024). Revealing patterns of change in the tribological efficiency of composite materials for machine parts based on phenylone and polyamide reinforced with arimide-t and fullerene. Eastern-European Journal of Enterprise Technologies, 3 (12 (129)), 6–18. https://doi.org/10.15587/1729-4061.2024.304719
25. Rogovskii, I. L., Reznik, N. P., Druzhynin, M. A., Titova, L. L., Nychay, I. M., Nikulina, O. V. (2024). Non-uniform Field of Concrete Deformations of Circular Cross-Section Columns Under Cross Bending Applying Digital Image Correlation Method. Intelligent Systems, Business, and Innovation Research, 939–951. https://doi.org/10.1007/978-3-031-36895-0_79
26. Gomez, I., Fowkes, V., Price, I. (2024). Teaching community engagement for health professions students in underserved areas. Discover Education, 3 (1). https://doi.org/10.1007/s44217-024-00323-3
27. Romaniuk, W., Rogovskii, I., Polishchuk, V., Titova, L., Borek, K., Shvorov, S. et al. (2022). Study of Technological Process of Fermentation of Molasses Vinasse in Biogas Plants. Processes, 10 (10), 2011. https://doi.org/10.3390/pr10102011
28. Rogovskii, I., Titova, L., Sivak, I., Berezova, L., Vyhovskyi, A. (2022). Technological effectiveness of tillage unit with working bodies of parquet type in technologies of cultivation of grain crops. 21st International Scientific Conference Engineering for Rural Development Proceedings, 21, 884–890. https://doi.org/10.22616/erdev.2022.21.tf279
29. Rogovskii, I., Titova, L., Shatrov, R., Bannyi, O., Nadtochiy, O. (2022). Technological effectiveness of machine for digging seedlings in nursery grown on vegetative rootstocks. 21st International Scientific Conference Engineering for Rural Development Proceedings, 21, 924–929. https://doi.org/10.22616/erdev.2022.21.tf290
30. Romaniuk, W., Rogovskii, I., Polishchuk, V., Titova, L., Borek, K., Wardal, W. J. et al. (2022). Study of Methane Fermentation of Cattle Manure in the Mesophilic Regime with the Addition of Crude Glycerine. Energies, 15 (9), 3439. https://doi.org/10.3390/en15093439
31. Nazarenko, I., Bernyk, I., Dedov, O., Rogovskii, I., Ruchynskyi, M., Pereginets, I., Titova, L. (2021). Research of technical systems of processes of mixing materials. Dynamic processes in technological technical systems, 57–76. https://doi.org/10.15587/978-617-7319-49-7.ch4
32. Nazarenko, I., Mishchuk, Y., Mishchuk, D., Ruchynskyi, M., Rogovskii, I., Mikhailova, L. et al. (2021). Determiantion of energy characteristics of material destruction in the crushing chamber of the vibration crusher. Eastern-European Journal of Enterprise Technologies, 4 (7 (112)), 41–49. https://doi.org/10.15587/1729-4061.2021.239292

33. Rogovskii, I. L., Titova, L. L., Gumenyuk, Y. O., Nadtochiy, O. V. (2021). Technological effectiveness of formation of planting furrow by working body of passive type of orchard planting machine. IOP Conference Series: Earth and Environmental Science, 839 (5), 052055. https://doi.org/10.1088/1755-1315/839/5/052055
34. Rogovskii, I. L., Titova, L. L., Trokhaniak, V. I., Borak, K. V., Lavrinenko, O. T., Bannyi, O. O. (2021). Research on a grain cultiseeder for subsoil-broadcast sowing. INMATEH. Agricultural Engineering, 63 (1), 385–396. https://doi.org/10.35633/INMATEH-63-39
35. Wang, L., Xiong, H. (2018). Research on the risk evaluation of the medical care combined PPP project based on the fuzzy comprehensive evaluation method. Conference Proceedings of the 6th International Symposium on Project Management, ISPM 2018, 1287–1293.
36. Quioc, M. A. F., Ambat, S. C., Lagman, A. C., Ramos, R. F., Maaliw, R. R. (2022). Analysis of exponential smoothing forecasting model of medical cases for resource allocation recommender system. 10th International Conference on Information and Education Technology (ICIET 2022), 390–397. https://doi.org/10.1109/ICIET55102.2022.9778987
37. Chernov, S., Titov, S., Chernova, L., Kunanets, N., Piterska, V., Shcherbynac, Y., Petryshyn, L. (2021). Efficient algorithms of linear optimization problem solution. CEUR Workshop Proceedings, 2851. Available at: https://ceur-ws.org/Vol-2851/paper11.pdf
38. Politansky, R. L., Nytrebych, Z. M., Petryshyn, R. I., Kogut, I. T., Malanchuk, O. M., Vistak, M. V. (2021). Simulation of the Propagation of Electromagnetic Oscillations by the Method of the Modified Equation of the Telegraph Line. Physics and Chemistry of Solid State, 22 (1), 168–174. https://doi.org/10.15330/pcss.22.1.168-174
39. Aulin, V., Rogovskii, I., Lyashuk, O., Titova, L., Hrynkiv, A., Mironov, D. et al. (2024). Comprehensive assessment of technical condition of vehicles during operation based on Harrington's desirability function. Eastern-European Journal of Enterprise Technologies, 1 (3 (127)), 37–46. https://doi.org/10.15587/1729-4061.2024.298567
40. Tryhuba, A., Vovk, M., Batyuk, B., Bogdanova, N., Proskurovych, O., Golomsha, N. et al. (2022). Improving the quality of management in the system of forecasting milk procurement in communities usage the technology of neutron networks. Journal of Hygienic Engineering and Design, 40, 201–209. Avaiable at: https://keypublishing.org/jhed/wp-content/uploads/2022/12/13.-Abstract-Anatoliy-Tryhuba.pdf
41. Tryhuba, I., Tryhuba, A., Grabovets, V., Bodak, V., Horodetska, N. (2023). Forecasting the duration of work in plant protection projects. CEUR Workshop Proceedings, 3453, 96–105. Avaiable at: https://ceur-ws.org/Vol-3453/paper9.pdf

DOI: 10.15587/978-617-8360-03-0.CH5

Georgii Prokudin, Viktoriia Lebid, Oleksii Chupaylenko,
Tetiana Khobotnia, Maksym Roi, Mykhailo Holovatiuk

CHAPTER 5

APPLICATION OF PROJECT ANALYSIS TO IMPROVE THE QUALITY OF TRANSPORT SERVICES IN INTERNATIONAL ROAD CARGO TRANSPORTATION

ABSTRACT

The paper considers theoretical approaches, models and methods for assessing the quality of transportation projects as a means of increasing the efficiency of providing transport services in cargo transportation projects of a project-oriented enterprise. N-model for making an optimal decision on the importance of a set of criteria that determine the quality of transport services, taking into account expert information, has been developed. It allows determining the advantages of one criterion over another based on the theory of the importance of criteria, which can be applied in project process management. A model for ensuring the relationship between quality indicators of cargo transportation projects and determining the attractiveness of international routes has been developed. The effectiveness of the application of the developed methods and models at project-oriented enterprises of the transport industry has been proven by testing them at motor transport enterprises.

KEYWORDS

Project, project management, knowledge base, fuzzy production rules, expert information, linguistic variable, attractiveness of the traffic route, transport services, international transport corridor.

The development of cargo transportation and the intensification of the provision of services in international traffic (hereinafter referred to as IT), including along international transport corridors (hereinafter referred to as ITC), increases competition among Ukrainian carriers. It is precisely the growing competition in the international road transportation market (hereinafter referred to as IRT) that forces project-oriented enterprises in the transport industry to look for new opportunities to reduce transport costs.

The analysis of the transport support of cargo transportation by a project-oriented enterprise shows that each transportation of a project-oriented enterprise can be considered as a separate

project, since it demonstrates the phased implementation of individual actions similar to the sequence of stages of project implementation.

For the transport support of cargo transportation projects (hereinafter referred to as CTP), which are implemented in a competitive environment of motor transport services, the tools of the project management theory should be used. When designing various types of transport support projects (hereinafter referred to as TSP), using project management elements, such stages as formulating a project idea, setting goals and objectives, their phased implementation and creating a project product (transport service – TS) should be followed. Providing high-quality TS for cargo transportation as a project product implementation is a dynamic process that occurs under conditions of uncertainty (the influence of internal and external factors), has time limitations, and is characterized by available resources and features of the operation of the project product.

In the conditions of dynamic development of transport and logistics services, the issue of cargo transportation in the international aspect is considered in the works of A. Vorkut, T. Vorkut, V. Racha, G. Prokudin, Y. Tsvetov, V. Dykan, K. Koshkin, A. Novikova. Scientists also comprehensively consider the issue of project management, which is covered in the works of S. Bushuyev, N. Bushueva, T. Vorkut, S. Tsyutsyura, V. Racha, Y. Teslia, I. Chumachenko, M. Dergausov, V. Kreymer, D. Montgomery, P. Crosby, W. Deming, I. Durand [1–5].

As a result of the analysis, it was found that the assessment of the satisfaction of the requirements of transportation market participants depends on the TS provided in the CTP. Transport services must comply not only with the mandatory accepted standards, but also with the principles of quality management as a product of the project, which reflect the planning of quality management, quality assurance and quality control of the transportation process. However, approaches to solving the issue of project management in the ICT field do not provide an answer to the question of how to assess the TS quality as a product of a transportation project, taking into account qualitative, quantitative and relay information received from participants in the transport process (hereinafter referred to as PTP) [6, 7].

It has been established that determining the relationship between the project product and the satisfaction of customer needs is one of the most difficult tasks: the needs of customers in the transport industry are diverse, dynamic, alternative, uneven in time and space, and are accepted in conditions of uncertainty of the ICT market. Therefore, it is practically impossible to provide them only by controlling the TS quality in the CTP. There should be a comprehensive approach, the implementation of which is possible only within the framework of a quality management system.

It is worth noting that in the term "quality management system" the main emphasis is not on the word "quality", but on the word "management". Therefore, the system is aimed not so much at quality control, but, first of all, at managing the quality of the project product, that is, TS quality management [8, 9].

This approach is a complex task given the set of assessment criteria that PTP must satisfy. The complexity is also determined by the TS specificity, which is that it cannot be withdrawn, corrected or reworked at the implementation phase of the transportation project life cycle.

It is worth noting that the TSP problems of freight transport on ITC routes have not yet received a comprehensive scientific analysis. Paying tribute to the achievements of specialists in the field of freight transport development, it should be noted that the issues of managing freight transport projects on ITC road routes require further scientific research. Despite the wide range of research conducted, the problem of assessing individual types of CTP has a number of unresolved issues, and there is no effective methodology for assessing such projects [10, 11].

In particular, insufficient attention has been paid to the assessment of projects for the quality of providing TS in the transport chain of cargo delivery along ITC road routes, taking into account the heterogeneous information received from participants in the transport process (hereinafter referred to as PTP).

In conditions of fierce competition among motor transport enterprises, it is necessary to apply new approaches to project management and the quality of services provided in them. This may be a project approach focused on qualitatively new levels of project management, that is, at project-oriented enterprises, the main attention should be paid to projects that should include other, shorter routes of vehicles, an appropriate level of transportation service and an attractive competitive tariff, namely a rational "commercial triangle" delivery time-delivery service-tariff.

Therefore, there is a need for scientific research into the assessment of the TS quality as a project product in the CTP, taking into account a set of factors such as cargo delivery time, the speed of movement of vehicles across the customs border of Ukraine and the tariff, which are determined by customer requirements. Such a task requires a detailed disclosure of their essence and relationships based on deep theoretical research using mathematical modeling and other scientific methods.

5.1 THE RELEVANCE OF IMPROVING THE QUALITY OF TRANSPORT SERVICES IN INTERNATIONAL ROAD FREIGHT TRANSPORTATION

Modern challenges facing the transport industry require new approaches to organizing transportation, especially in the context of meeting growing customer demands. Customer orientation is becoming a key factor in competitiveness, which necessitates the revision of traditional methods of route selection in transportation projects.

The paper focuses on improving approaches to managing cargo transportation routes, taking into account customer needs. Applying a systematic approach to route selection allows not only to optimize logistics processes, but also to ensure the quality of transport services at all stages of the transportation project. Underestimating these aspects can lead to breach of contractual obligations, delays in delivery, or a decrease in customer satisfaction, which, in turn, negatively affects the reputation of transport operators.

Applying a systems approach and systems analysis in managing cargo transportation projects allows to identify key problems that arise for the project team and assess the consequences of decisions made at all stages of the project life cycle.

Underestimating the importance of a comprehensive approach to quality management can lead to a decrease in the efficiency of transport services and a violation of contractual obligations to customers. This situation increases the risks of delays in cargo delivery or a decrease in the quality of service.

The work uses a systemic approach to project management, a process approach in the development of project models, qualimetry methods and elements of the theory of the importance of criteria. The theoretical basis of the work is the basic provisions of project management and quality management, the theory of fuzzy sets and fuzzy logic. A number of mathematical methods are also used, namely: the method of analyzing expert assessments for selecting criteria for assessing projects and programs, determining their hierarchy; the mathematical apparatus of decision-making theory for determining the influence of individual parameters on the quality of transport services in the CTP; simulation modeling methods for modeling the integral quality indicator of making an optimal decision on choosing a route for a specific ITC. Using the theory of fuzzy sets, the possibility of multivariate choice is provided and a linguistic model of transportation project management is proposed. The information base of the study is statistical data on the implementation of transportation projects and the results of our own scientific research.

It is worth noting that the quality of transport services can be determined by the functional-instrumental model proposed by K. Grönss [6]. Instrumental quality reflects the final result of the project for the client, that is, what he/she receives, while functional quality characterizes the service process itself at all stages of the transportation project.

The input parameters in the quality system are key indicators that affect the level of service. This includes the quality of service at the consignor, the efficiency of the transportation process, the quality of service on international routes, the level of service at the consignee.

Thus, a customer-oriented approach to quality management is aimed at ensuring a long-term partnership between transport operators and their customers by adapting services to customer needs and increasing competitiveness in the market.

5.2 THE SYSTEM OF ENSURING THE QUALITY OF TRANSPORT SERVICES IN TRANSPORTATION PROJECTS

Let's present the quality system of transport services in transportation projects (**Fig. 5.1**). The input parameters are indicators that affect the level of quality of transport services in transportation projects. Such indicators include the quality of service provided by shippers, the quality of the cargo transportation process, the quality of service provided by carriers at car service points on international routes, and the quality of service provided by consignees, the quality of preparation of transport documents, and the quality of provision of additional services (as an example of informing customers).

Effective management of such elements as planning the transportation route, monitoring and improving the quality of transport services on international routes is possible only if their properties are comprehensively assessed. This assessment requires the availability of complete, reliable and

quantitative information on the quality of transport services within the framework of transportation projects. It is quantitative indicators that allow using data for making management decisions.

In modern management of cargo transportation projects, achieving a high level of transportation quality is based on the application of the science of "qualimetry". Qualimetry, which is developed on the basis of applied and theoretical approaches, allows for the calculation of complex quantitative quality indicators. This contributes to more accurate analysis, forecasting and optimization of processes.

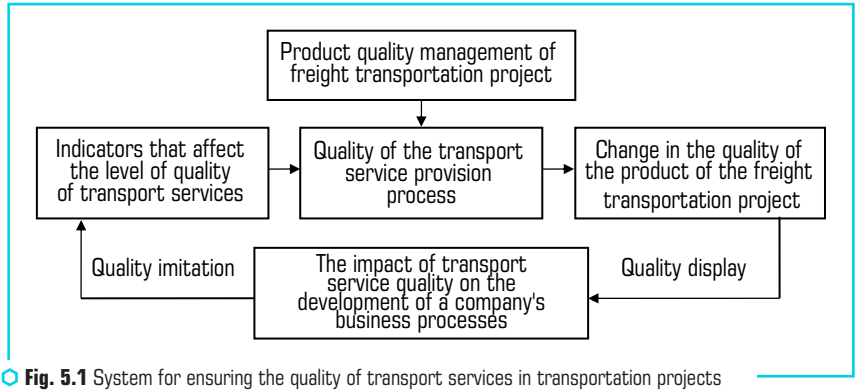


Fig. 5.1 System for ensuring the quality of transport services in transportation projects

Theoretical qualimetry studies general patterns and mathematical models of services and processes that are related to the assessment of project quality. The object of theoretical qualimetry is the philosophical and methodological problems of service development. The task of applied qualimetry is to develop algorithms and mathematical models for assessing the quality of various types of objects [7].

To assess the quality of transport services of carriers in transportation projects, it is necessary to take into account the main functions of management objects:

1. Target – according to this function, quality assessment is aimed at controlling the achievement of improving standards in transportation projects. The control goal is formed by determining key indicators that determine the quality of transport services. The target function provides the principle for building an assessment methodology.

2. Classification – this function assumes that the results of quality assessment become the basis for creating categories and classes of transport service quality, which allows for a more clear segmentation of services within projects.

3. Stimulating – quality assessment mechanisms stimulate project executors both through moral recognition and through material incentives aimed at increasing the efficiency of their work in fulfilling various types of transportation orders.

4. Information function – the quality assessment system is a key source of information for quality management in transportation projects. The information function is aimed at ensuring transparency and efficiency of management.

5. Aggregating – generalization of quality assessments is a necessary component in multi-level management projects, which determines the reliability of assessment information for making final management decisions.

6. Analytical – quality assessment is based on a deep analysis of transport service processes, which allows making informed decisions during the implementation of projects.

7. Predictive. Used both when determining assessment criteria and in the process of using assessments directly in the forecasting system.

To assess the quality of transport services in transportation projects, clear management of transportation organization functions is necessary so that all departments of a project-oriented enterprise have management and control skills, as well as methods for assessing quality and ensure the necessary level of responsibility for it.

The implementation of quality assessment functions is closely related to the tasks of managing the quality of services in freight transport projects. Quality assessment at the enterprise is an important tool for ensuring the efficiency and safety of transportation, allowing to identify weaknesses in logistics processes and adjust service delivery strategies. It includes monitoring such aspects as timeliness of delivery, compliance with safety requirements, accuracy of document flow and the level of customer satisfaction. The use of a quality management system allows to respond in a timely manner to changes in external and internal conditions, ensuring stability and improvement of transport processes. The management sphere in transportation projects takes into account the cost indicators of business operations and their impact on the project. Therefore, with sustainable management in project-oriented enterprises, the project manager, taking into account the main functions of the management object, makes rational decisions. It is worth noting here that quality is a multi-criteria and multi-factor concept, and therefore it is necessary to take into account the most influential indicators for assessment. Quality management in projects is a philosophy that can and should be the basis of any activity for the continuous improvement of all processes of the organization of management on various issues that may arise in the implementation of projects [7, 8].

Based on this, the concepts of quality of the service provision process and quality of results in transportation projects are essentially different and are defined separately from each other, that is, practically not interconnected. As a result, there are clashes of interests between workers in the automotive industry and consumers of transport services. Therefore, it is necessary to distinguish two classes of concepts in "transport qualiology" (the science of quality):

– the class of concepts of "quality of result", which includes the quality of transport products, quality of services;

– the class of concepts of "quality of process" – quality of transport service; quality of planning of the transport process in transportation projects; quality of rolling stock used in transportation; quality of preparation of cargo for transportation, quality of provision of related transport services.

One of the modern trends in the quality of transport service is the formalization of the concept of quality of service to consumers of services. In the context of international cargo transportation, this trend is particularly relevant, since ensuring high quality service at different stages of the logistics chain directly affects the efficiency of transportation and customer satisfaction.

An important aspect is the integration of a customer-oriented approach, which involves focusing on the needs and requirements of customers, as well as the ability to adapt to the specifics of each individual international route.

Formalization of this approach includes the creation of clear standards and procedures that allow assessing the effectiveness of service provision in different countries and regions, as well as taking into account individual customer wishes regarding delivery times, storage conditions and cargo safety. Thus, the process of improving the quality of any services (not only transport) must begin with "quality control", gradually defining "guaranteed quality" (quality assurance), then moving on to "standardized quality control", and the final result should be "ensuring the requirements of service consumers" (customer value) [9, 10].

In the context of military operations on the customs territory of Ukraine, it is important to take into account additional factors that may affect the quality of service, in particular, changes in the safety of transport routes, risks of damage to cargo or delays at customs. In such conditions, the strategic task becomes not only to ensure the safety of cargo, but also to promptly respond to changes in the situation, which allows minimizing possible losses and ensuring continuity of supply. This requires flexibility in approaches to organizing transportation and the use of the latest technologies for monitoring and managing quality at all stages of transportation.

It is worth noting that the assessment of the quality of transport services in transportation projects requires detailing, that is, it is necessary to establish which quality indicators should be selected for consideration, by what methods and with what accuracy to determine the results, how to process them experimentally and in what form to present the assessment result.

The criteria for the efficiency of cargo delivery are closely interrelated with the assessment of the quality of transport services, since quality allows the enterprise, on the one hand, to reduce the costs of cargo delivery in international traffic (thereby reducing the cost of services), and on the other hand, to increase its own income and customer base by increasing the attractiveness of transportation projects for customers [11, 12].

The structural scheme of quality management of cargo transportation in international traffic in accordance with quality management projects is presented in **Fig. 5.2**.

The structure of indicators of quality of transport products, from the point of view of providing services in transportation projects along international transport corridors, is characterized by the following indicators (**Fig. 5.3**):

– environment – the presence of necessary service points within the TC routes, amenities, equipment and the presence of qualified personnel at these points. The services of service facilities should include parking of vehicles; their security; cargo operations necessary for customs processing of cargo and vehicles; servicing of vehicle crews (in accordance with the

European Agreement concerning the work of crews of vehicles engaged in international road transport, AETR) and others;

– reliability – the totality of trust in the results of the performed transportation;

– psychological properties, otherwise sociability, which determines the possibility of finding contact between participants in the transportation process.

Analysis of works devoted to the assessment of the quality of transport services indicates the importance of taking into account the competitiveness of transport services when servicing cargo owners. The competitiveness of a transport service affects the competitiveness of the development of the object under study.

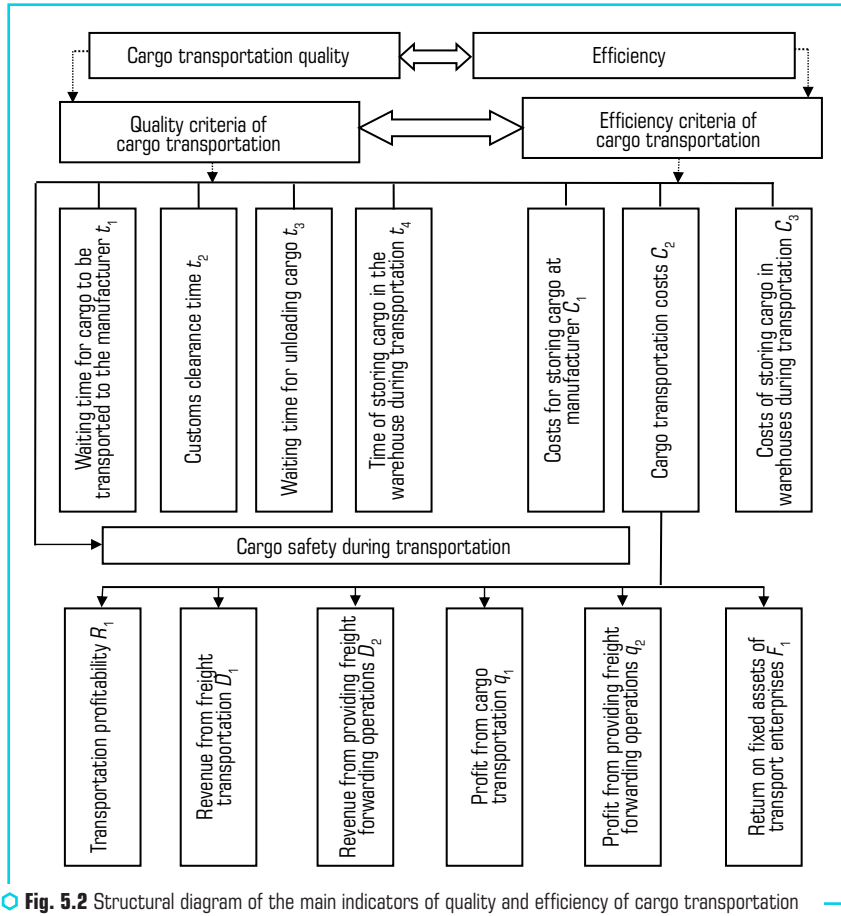


Fig. 5.2 Structural diagram of the main indicators of quality and efficiency of cargo transportation

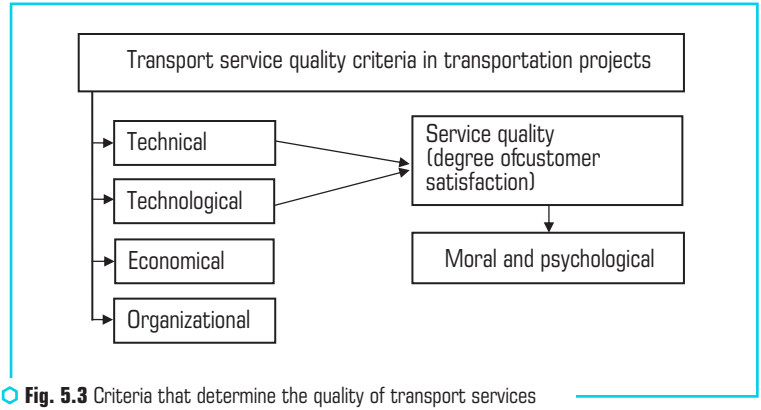


Fig. 5.3 Criteria that determine the quality of transport services

5.3 DETERMINING THE COMPETITIVENESS ASSESSMENT OF TRANSPORT SERVICES

Taking into account the improvement of the quality characteristics of a transport service, its competitiveness can vary in a wide range, taking into account the dynamics of the transport market, changes in tariffs, changes in the influence of other external factors. Since competitiveness studies should be conducted constantly to determine the level of satisfaction of participants in transportation projects, let's propose a structural scheme for determining the competitiveness of a transport service, in which K is an indicator of the competitiveness of transport services (**Fig. 5.4**).

If when calculating the competitiveness of transport services $K \leq 1$, this means that the transport service does not meet the requirements of the participants in the transportation process in transportation projects, and in the case $K \geq 1$ – all the requirements for the quality of transport services satisfy the carriers and this service meets the competitive quality.

In many works, scientists argue that improving the quality of transport services in transportation projects can occur by improving the norms of the technological process of delivering goods in international traffic. This means that scientists recognize the importance of improving the norms and standards that regulate the technological processes of transporting goods. "Technological process" in this context is considered as a set of all stages and procedures through which the cargo passes from the sender to the recipient.

These stages include route planning (i.e. choosing the optimal routes for transporting cargo, taking into account all external factors, such as weather conditions, road conditions, customs procedures, safety of the transport process, etc.); customs procedures (including timely and correct documentation for the unhindered passage of cargo across borders, which is important for compliance with international standards in modern conditions and reducing delays at the border);

transportation technologies (choosing the most effective means of transport for transporting cargo, as well as the use of modern logistics technologies, such as real-time cargo monitoring); risk management (in international transportation, it is important to anticipate possible risks (for example, military actions in the customs territory of Ukraine as a result of the Russian invasion, natural disasters) and adapt the technological process to new conditions.

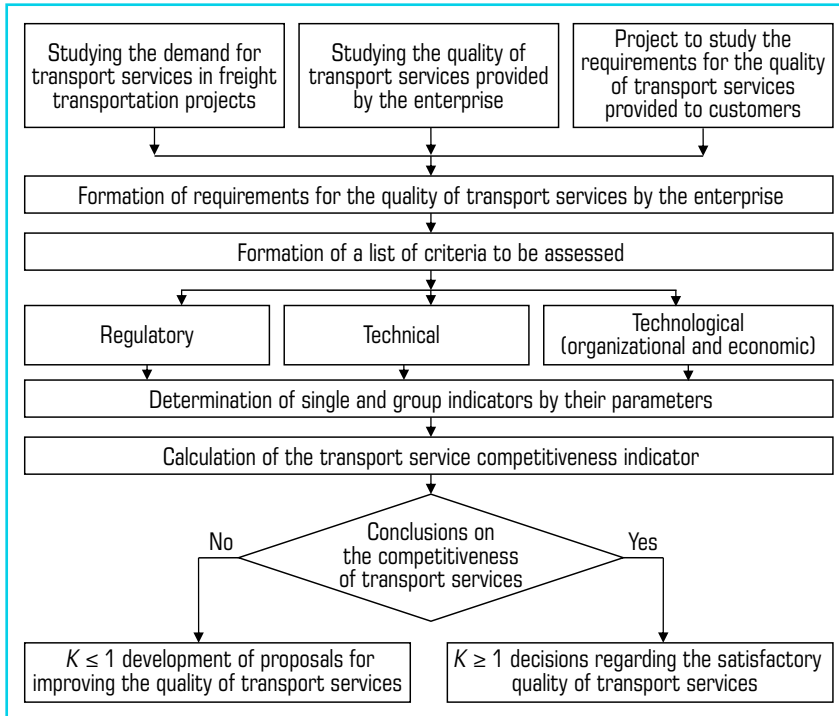


Fig. 5.4 Structural diagram for determining the assessment of the competitiveness of transport services

Thanks to the improvement of the norms of the technological process (this may be the improvement of existing procedures or the introduction of new standards), the quality of transport services can be significantly improved. This leads to a reduction in delays at the customs border (i.e., clearly defined and improved customs procedures allow to reduce the time spent on transporting goods and avoid unnecessary delays at customs or at border crossing points); ensuring cargo safety (the introduction of norms that regulate the processes of packaging, transportation and storage of cargo helps to minimize the risks of damage to cargo or its loss during transportation); improving communication between participants in the process (it is precisely the improved

procedures that ensure more effective interaction between carriers, customs authorities and other participants in the transport chain).

Thus, improving technological processes is a key factor in improving the overall quality of transportation and the competitiveness of transport services, which includes not only reducing costs and time, but also increasing the level of safety and customer satisfaction.

The competitiveness indicator of a transport service is calculated as follows:

$$K_{trans.serv.} = \frac{J_r \cdot J_t}{J_{tl}}, \quad (5.1)$$

where J_r, J_t, J_{tl} – group criteria according to regulatory, technical and technological parameters.

However, existing methods for assessing the competitiveness of transport services have some limitations. As a rule, they are focused on calculating the assessment of the actual level of quality, that is, existing at the time of the study, and are not designed for the future.

The indicator $K_{trans.serv.}$ should be critical in relation to the parameters under study, relatively simple to quantify, universal and allow for comparative analysis.

Ensuring the improvement of the quality of transport services is one of the indicators of increasing the volume of cargo transportation. Thus, studies [13, 14] confirm the relationship of the category "quality" with other market categories, and therefore the following patterns are obvious: an increase in the volume of transportation, income from transportation, a decrease in cost and operating costs, which are the final results in transportation projects, are directly related to the quality of transport services.

A developed transport network contributes to the safety of cargo during delivery and the reduction of the transport component in the cost of goods, reducing downtime at the border, ensuring timely delivery of cargo "just in time". Therefore, it is proposed to include the transport accessibility indicator, which depends on the length of the communication routes, the corresponding geographical configuration of highways, throughput and intensity of vehicle traffic, in the indicators of the quality of transport service for cargo owners and carriers in the ITC development programs.

Transport accessibility is determined by the transport network density indicator d_T per 1000 km^2 of the area of the territory along which the highway passes, per 10,000 population and per 1000 tons of manufactured products in terms of weight:

$$d_T^S = \frac{L_E}{S};\ d_T^N = \frac{L_E}{N};\ d_T^Q = \frac{L_E}{Q}, \quad (5.2)$$

where L_E – operational length of highways, km; S – area of the studied territory, through which the ITCs pass, km^2/1000; N – population, people/10000; Q – volumes of manufactured products, t/1000 respectively.

The transport network density indicator d_T is defined as:

$$d_T = \frac{L}{\sqrt[3]{S \cdot N \cdot Q}}, \tag{5.3}$$

where L – the length of the connecting routes, km.

The length of the connecting routes L is calculated using the coefficients of the reduced transport routes, determined taking into account the capacity and intensity of vehicle traffic on the studied sections of the transport corridors. Thus, the most difficult moment when taking into account this indicator is determining its optimal values. Usually, its comparison with the reference value in developed European countries is used. In this case, not only the length of the connecting routes is taken into account, but also the intensity of the use of road transport on the ITC routes and the carrying capacity of the routes.

The transport accessibility indicator $d_{A(ITC)}$ on the routes of international transport corridors can be defined as the weighted average value of the time required by carriers to deliver cargo [14]:

$$d_{A(ITC)} = \frac{\phi\left[1-(t_1+t_2)\right]+Z}{V_{av}}, \tag{5.4}$$

where ϕ – directions of secondary roads, characterizing the transport accessibility of carriers to highways, km; t_1 – coefficient characterizing the non-isolation of the departure point from the entire transport network; t_2 – coefficient characterizing the reserve of the transport network configuration; Z – transport focus of the territory, which is the shortest distance that must be overcome by the best routes when delivering cargo in international traffic, km; V_{av} – average speed of the vehicle along the route, km/h.

The transport accessibility indicator allows to determine the time of delivery of goods "from door to door" and allows to take into account the reliability of cargo delivery in transportation projects.

The quality of transport service in transportation projects will be assessed by a complex indicator of the quality assessment of the $Q_{A(trans.serv.)}$, which includes aggregated criteria, block criteria and single criteria. Aggregated indicators are represented by the set [15]:

$$Q_{A(trans.serv.)} = \left\{R_{qual}, C_{qual}, F_{qual}, S_{qual}\right\}. \tag{5.5}$$

Reliability (R_{qual}) is represented by the set:

$$R_{qual} = \left\{r_1, r_2, r_3\right\}, \tag{5.6}$$

where r_1 – savings in the transportation process (delivery of cargo without losses, damage, loss and contamination); r_2 – timeliness of transportation projects (speed of delivery, accuracy, execution of the delivery schedule in accordance with the AETR requirements); r_3 – fulfillment of contractual obligations (fulfillment of accepted applications, completeness of fulfillment of guarantees).

Complexity (C_{qual}) is represented by a set:

$$C_{qual} = \{c_1, c_2, c_3\}, \tag{5.7}$$

where c_1 – complexity of transport services (set of services by type of cargo, brand of rolling stock used); c_2 – range of transport services (availability of additional services for customs clearance, cargo insurance, loading and unloading, consulting services); c_3 – informativeness (completeness of information and its reliability, regularity of information receipt).

Flexibility (F_{qual}) is represented by the set:

$$F_{qual} = \{f_1, f_2, f_3, f_4\}, \tag{5.8}$$

where f_1 – convenience of the service provided (convenience and speed of application processing, goods and accompanying documents, convenience of cargo acceptance and delivery, availability of different levels of service, individual approach to each participant in the transportation process); f_2 – service culture of carriers (communicability, friendliness, ethics); f_3 – efficiency of service (competence and professionalism of personnel, speed of processing applications for the development of transportation projects, speed and quality of response to complaints, claims, efficiency of order fulfillment, possibility of delivering goods on demand (just-in-time delivery)); f_4 – aesthetics (politeness, responsiveness, accessibility and trust in personnel, level of skill, comfort and trust in project participants, effectiveness of communication between the contractor and the client).

Safety in transportation projects (S_{qual}) is represented by the set:

$$S_{qual} = \{s_1, s_2, s_3\}, \tag{5.9}$$

where s_1 – safety of the main service (ensuring safety during transportation in accordance with the AETR requirements); s_2 – safety of further service (ensuring the safety of the cargo and vehicle in the country of destination); s_3 – safety of the rolling stock selected for the transportation project (technical inspection of the vehicle, permit for the ICT implementation).

Thus, the proposed criteria can be used to assess the quality of transport services at all stages of the life cycle of the transportation project. The assessment of the quality of transport services in transportation projects according to the given criteria can be called complete and objective. After all, project-oriented enterprises in the field of motor transport services work to implement the final product – customer satisfaction, since it is the client who assesses the quality of the transport service by comparing the desired and obtained result.

Quality management requires a quantitative assessment of the quality of transport services as a project product in cargo transportation projects. For this, it is necessary to present a basic quality indicator, i.e. a reference value for comparative quality assessments. Depending on the

purpose of the assessment, the basic indicators can take on different values. The basic indicators can be the quality indicators of transportation projects passing through the territories of foreign countries, the quality indicators of the transportation project product for the past period, or indicators calculated using a simulation model.

According to the number of parameters studied in the process of determining the assessment of the quality of transport services in transportation projects, single, complex and integral quality indicators are distinguished.

Single indicators, as a rule, characterize only one property. This can be only reliability, or only the safety of cargo delivery in transportation projects or on other class E roads.

A complex assessment determines a service quality indicator that characterizes several properties of the service at the same time.

An integral quality assessment is a type of complex service quality indicator (not only transport), which reflects the ratio of the total useful effect of a given service to the costs calculated for the purchase or provision of the service to users:

$$J = \frac{E}{Z} = \frac{E}{Z_c + Z_{cr}}, \tag{5.10}$$

where E – the total useful effect of providing transport services; Z_c – the capital investment in creating services; Z_{cr} – the sum of current costs for the purchase or provision of transport services (money units).

Integral assessment is used in most cases and is carried out in two stages. First, simple properties are assessed, and then complex properties are assessed, in particular, quality as a whole (**Fig. 5.5**).

In [16], the shortcomings of the comprehensive quality assessment method are noted. The author notes that comprehensive quality assessment does not take into account ergonomic ("person-machine-environment"), aesthetic and other properties of the service. In addition, comprehensive quality assessment is defined only for those services for which the total useful effect is calculated only in natural or cost units. Therefore, in most practical problems, the final result of qualimetric calculations is not an absolute indicator P_{ij}, but a relative K_{ij} [16].

Then the integral indicator is a function of two absolute indicators – the measured P_{ij} and the one taken as the base indicator P_{ij}^{bas}:

$$K_{ij} = f(P_{ij}; P_{ij}^{bas}). \tag{5.11}$$

In general, the quality management indicator (integral indicator) [17] is the ratio of the consumer value (CV) to the cost (C) of the services provided.

The formula for calculating the dynamics of integral quality and efficiency indicators for different levels and goals of freight transportation management is as follows [18]:

$$K_i^{tr} = \left[1 + \frac{\left(\pm \sum_{n=1}^{N} \Delta E_T \right)}{Z_{gr}} \right] \cdot 100, \qquad (5.12)$$

where ΔE_T – the total economic effect (+) or (–) of the change in simple natural quality indicators in the studied period, c.u.; Z_{gr} – the total costs of the carrier for the last studied year, c.u.; $n = 1, 2, 3$; N – the number of integrated simple natural quality indicators.

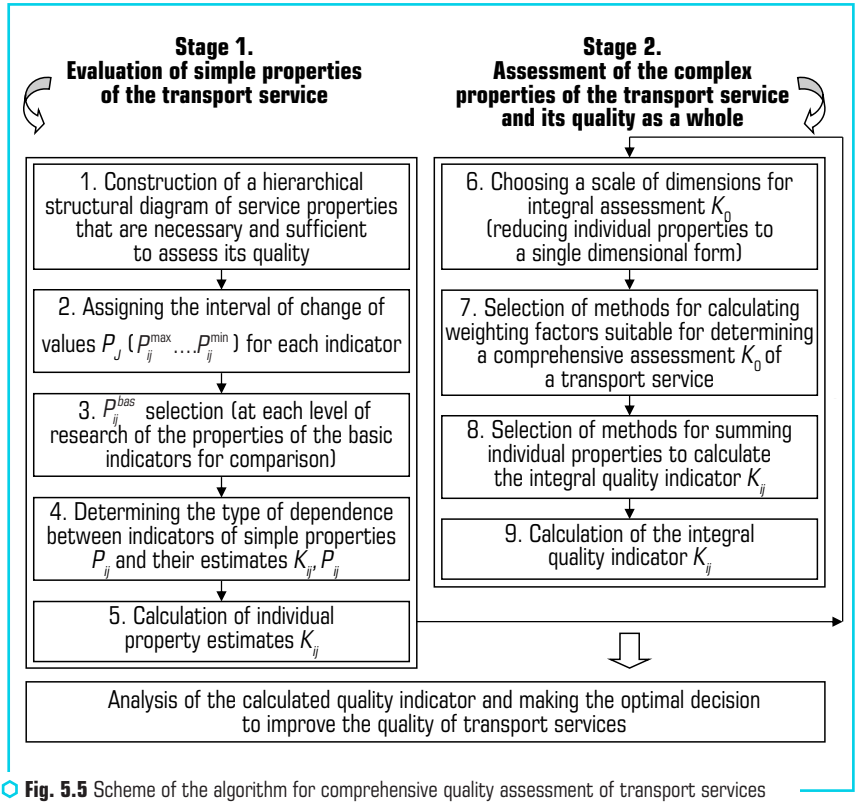


Fig. 5.5 Scheme of the algorithm for comprehensive quality assessment of transport services

An important approach in determining quality is the weighted average indicator, by which the complex indicator of the quality of transport service is determined:

$$K = \sum_{i=1}^{n} a_i \cdot g_i, \tag{5.13}$$

where a_i – the number of studied elements; g_i – the importance coefficient; $i = 1 \ldots n$.

However, it should be noted that the main mistake in developing a complex indicator is an attempt to replace the economic content of quality with formalistic methods of calculating indicators in the form of the sum of various economic, operational and environmental parameters, that is, this indicator does not reflect economic quality.

Using the method of scoring in qualimetry, it is possible to evaluate the work of all enterprises that develop transportation projects, according to the expression [19]:

$$M_S = \frac{\sum n}{S} \cdot (1 - K_{red}), \tag{5.14}$$

where $\sum n$ – the sum of the score ratings of the indicators; S – the number of quality indicators; K_{red} – the reduction coefficient of quality indicators depending on the degree of their implementation.

It is proposed to distribute the scores into three categories of transport service quality [20]:

– the highest category (with a score of 5–10);

– the first category (with a score of 3–5);

– the second category (with a score of 1–3). As an example, the distribution of scores by the indicator "fulfillment of the vehicle movement schedule along the route" in transportation projects is presented in **Table 5.1**.

Table 5.1 Determination of scores by the studied indicator

Completion percentage, %	**100**	**97–100**	**93–97**	**< 93**
Score	10	6	4	2

Therefore, the quantitative values of the integral quality assessment, which are suitable from the point of view of high-quality transport service, should be considered in the range of 5–0 points. This range is conditional and may vary depending on the requirements of carriers.

At the level of operating enterprises (for example, service enterprises, service stations, TIR parking lots), it is proposed to use a methodology widely used in construction, which involves the use of two indicators [21]:

– the average evaluation score, which is defined as a weighted average value;

– the dispersion, which is defined as the weighted average value of the deviation from its average value, and its derivatives – the mean square or standard deviation and the coefficient of variation.

The average evaluation score characterizes the level of quality and is defined as:

$$S_{av} = \frac{3m_1 + 4m_2 + 5m_3}{m_1 + m_2 + m_3}, \quad (5.15)$$

where S_{av} – average score; m_1 – number of elements for which the score is "satisfactory"; m_2 – number of elements for which the score is "good"; m_3 – number of elements for which the score is "excellent".

The variance characterizes the constancy of the quality level, that is, the distribution of scores among themselves:

$$S_v = \frac{(S-3)\cdot m_1 + (S-4)\cdot m_2 + (S-5)\cdot m_3}{m_1 + m_2 + m_3}. \quad (5.16)$$

However, the use of the evaluation score and dispersion is possible only when comparing individual properties over several years and, in particular, by types of services provided in transportation projects.

Therefore, a comprehensive assessment does not allow taking into account the criteria complicated by the fact that they are evaluated on different scales. After all, the criteria are divided into 3 groups by their nature: quantitative indicators; qualitative; relay (yes/no), which complicates the calculation of the integral indicator of the quality of the transport service. In addition, each quality criterion has a different weight in the totality of the studied elements.

Therefore, the indicator of the quality of transport service according to the integral assessment Q_{ij} can be understood as the fulfillment of the requirements of a set of individual parameters q_{ij} – the speed of cargo delivery, reliability, regularity, safety of cargo delivery and other factors that can affect the cost of transportation [22]:

$$Q_{ij} = (q_{ij}^1, q_{ij}^2, q^3, q^4 \ldots q_{ij}^n), \quad (5.17)$$

where n – the number of parameters that determine the quality of transport services of carriers.

The use of such an approach makes it possible, firstly, to make an integral assessment of the quality of transport services on the routes of international transport corridors, according to all properties, and secondly, to determine the most effective measures for improving significant evaluation parameters, always starting with the most significant.

The quality system "Q" can be considered as a set of 3 subsystems: quality assurance Q_{ij}, quality management U_q and development of the quality indicator R_q. The first subsystem of quality assurance Q_{ij} should be implemented as planning the life cycle of transportation projects. Consideration of the subsystem allows to highlight all technological components (at the stages of planning, management and control).

The second subsystem of quality management U_q determines the requirements for project management and control of cargo transportation along routes. The PDCA cycle is taken as a basis (P – preparation of transportation routes, D – execution of routes in accordance with the prepared ones, C – verification of the route traveled in accordance with the planned routes, A – taking necessary measures in case of certain deviations during cargo delivery). The assessment of transportation quality is based on the concept of the permissible spread of the parameters of the controlling value $qij(+)$ in relation to $q_{ij}^{H}(t)$; $(t) \in Q_{ij}^{H}(t)$, where $Q_{ij}^{H}(t)$ – the standard of quality of transport service in transportation projects. The development of the quality indicator Q_{ij} should be understood as increasing the stability of the implementation of the given quality standard $\Delta Q_{ij}(T_{pl}) \to \min$ in the planning period T_{pl} and reducing technological and other changes, i.e. reducing transport delays ($D_{ij}^{tot} \to \min$ at $Q_{ij}(T_{pl}) = const.$), by reducing the time for customs procedures when transporting goods in international traffic.

By performing all the above measures, it is possible to get the effect of implementing transportation projects, which is manifested in an increased number of orders for a project-oriented enterprise (**Fig. 5.6**).

The process of determining the level of transportation quality by quantitative assessment is complex and time-consuming. This is due to the lack of clear criteria for measuring and assessing quality, the inability in most cases to use quantitative methods for measuring the level of service quality, as well as the subjectivity of expectations and perceptions of the services provided. In addition, the assessment of service quality is complicated by the fact that each quality indicator has a different weight in the set of elements. Therefore, this task reflects the relevance of the work being conducted.

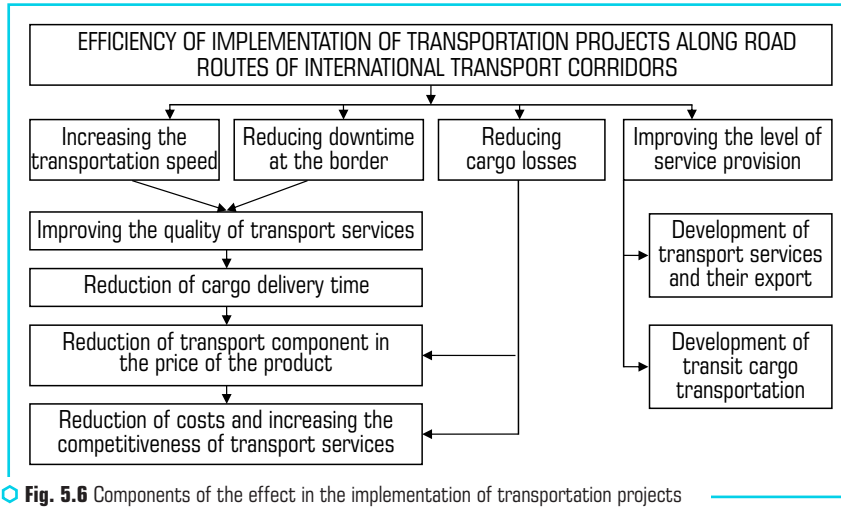


Fig. 5.6 Components of the effect in the implementation of transportation projects

The development of a method for assessing the quality of transport services will create a kind of legal field for competition in the market, granting the right to provide transportation to the most competitive carrier.

A significant number of components of transport service quality is due to the fact that breaking down indicators into small elements and assessing each separately, and then combining them can lead to a distortion of the average score of the group of indicators.

Thus, today, enough methods have been developed to assess the quality of various project management objects (transport service, operational work, cost indicators), at the same time, none of the considered methods is used in the practice of quality management for making management decisions when developing transportation projects [23].

Studies on the assessment of the quality of transport services in transportation projects prove that to increase the efficiency of a project-oriented enterprise, it is necessary to manage the quality of transport services throughout the entire life cycle of projects, and especially at the stage of project initiation.

This approach defines information about quality as the main component of managing international cargo transportation projects. For participants in the freight transportation market, the efficiency of cargo delivery in international traffic is determined by the quality of transport services in transportation projects, which is characterized by qualitative and quantitative indicators. The following quality indicators should be included [24, 25]:

– timeliness of cargo transportation;
– ensuring the safety of cargo transportation;
– reliability of delivery within a certain period;
– acceptable cost of transportation by international routes;
– regularity of cargo delivery;
– capacity of the transport corridor;
– speed of the route;
– quality of service provision to carriers.

Studies of the assessment of the quality of transport services prove that the practical formulation is reduced to a multi-criteria problem. Quite often, when studying the problem of multi-criteria, all criteria, except one, chosen as dominant, are taken as constraints. In this case, optimization is carried out according to the dominant criterion. This approach to solving practical problems greatly simplifies, but also reduces the accuracy of the decisions made [9, 10].

In the tasks of assessing efficiency by several criteria, it is necessary to determine the value of the objective function that corresponds, for example, to achieving the maximum throughput of the studied corridor at given transportation costs or achieving the given maximum effect at minimum costs to improve the quality of freight transportation in transportation projects.

The fundamental complexity of solving a multi-criteria problem is that there is usually no single solution that would be the best for all criteria at once. Therefore, it is proposed to apply elements of the theory of the importance of criteria, which is based on the mathematical justification of

the basic definition "one criterion is more important than another with a certain coefficient of relative importance".

When comparing criteria by importance, that is, finding out whether one criterion is more important than another, it is assumed that the criteria are homogeneous. This means that the criteria must have a single dimensional scale. In addition, the condition of homogeneity must be met, in which each gradation of the scale reflects the same level of preference for each of the criteria.

However, when assessing transport services, there is a set of heterogeneous criteria, the values of which are measured within certain scales and expressed in separate units of measurement.

Therefore, usually when developing a single complex quality indicator (Φ), all criteria must be reduced to a single (dimensionless) form, in other words, normalized [11, 26]:

$$\Phi = a_1 K_1 + a_2 K_2 + \ldots + a_n K_n, \tag{5.18}$$

where a_1, a_2,…, a_n – the importance coefficients, the values of which characterize the relative importance of the criteria.

But this approach to constructing a complex indicator is not always attractive, since it can lead to an unsuccessful choice of solution. When determining the importance of heterogeneous criteria, it is necessary to reduce them to a single ordinal (qualitative) scale. This is the main difference from the criterion normalization method, which assumes the quantitative superiority of each criterion when constructing a complex evaluation indicator in projects.

If the scale is single with an accuracy of an arbitrary monotonically increasing transformation, then it is ordinal. The numbers in this scale are compared with each other according to the usual numerical relations "greater than or equal to", "greater than" or "equals". At the same time, in an ordinal scale, the numbers do not answer the question: how much or how many times one criterion is more important than another. On this basis, the ordinal scale is often referred to the class of qualitative scales.

The mathematical model of making an optimal decision in the presence of many criteria includes three elements: a set of decisions (V), a vector criterion (K), a ratio of preference and indifference of the decision maker (hereinafter referred to as the DM). The criterion (K_i) is a function defined on the set (V), and takes values from the set of decisions (X_i), which is called a scale, or a set of estimates [6, 27].

The estimates can be numerical (for example, the throughput of a particular TC, thousand vehicles per day), verbal (for example, "high level of service provision within the transport corridor" or "low level of service provision") and symbolic (for example, the category of roads – I, II, etc.). In the following, let's consider only criteria with a numerical scale.

Thus, each option (v) is characterized by evaluations (m) according to the criteria $K_1(v)$, $K_2(v)$, …, $K_n(v)$, which make up the vector $K(v) = (K_1(v),\ldots, K_m(v))$, which is called the "vector evaluation of the option". Its designation can be either $K(v)$ or $x(v)$, i.e.:

$$K(v) = (K_1(v),\ldots,K_m(v)) = x(v) = (x_1(v),\ldots,x_m(v)). \quad (5.19)$$

To demonstrate the practical application of the theory of the importance of criteria, let's give a conditional example.

5.4 ASSESSING THE OPERATION QUALITY AND EFFICIENCY OF FOUR INTERNATIONAL TRANSPORT CORRIDORS

Let's assess the quality and efficiency of the functioning of four international transport corridors, which it is possible to propose for transportation projects passing through the territory of Ukraine, according to the following criteria:

1) transport corridor capacity, thousand vehicles per day (A) – this criterion assesses the ability of the transport corridor to handle a certain number of vehicles per day. A higher capacity indicator indicates higher efficiency and the ability of the corridor to cope with large volumes of transportation;

2) actual traffic intensity, thousand vehicles per day (B) – this criterion determines the actual load on the corridor, that is, the number of motor vehicles that actually carry out transportation. By comparing the actual traffic intensity with the capacity, it is possible to determine whether the corridor is overloaded and whether its resources are used efficiently;

3) ensuring the safety and reliability of cargo transportation along ITC routes (C) – this criterion assesses the level of safety on each route, taking into account the risks of road accidents, natural disasters, abuse, as well as the effectiveness of cargo control. Transportation reliability is an important aspect for ensuring the timely delivery of goods without loss or damage;

4) conditions for providing service services within the selected ITC (D) – this criterion assesses the level of service provided within the transport corridor. It includes the availability and quality of infrastructure for technical maintenance of transport, the availability of control points, customs offices, gas stations, as well as the level of information services and support for cargo owners.

This vector assessment method allows for a comprehensive comparison of transport corridors according to all criteria at the same time. Each corridor receives assessments for each of the criteria, and based on these assessments, a vector is formed that shows its overall efficiency. By comparing these vectors with each other, it is possible to draw conclusions about which corridor is the most effective for implementing transportation projects, in particular, in terms of throughput, traffic intensity, safety and service conditions. The initial data is given in **Table 5.2**, namely, the comparison of routes by the efficiency of implementing transportation projects is based on their vector assessments.

Conditionally, each criterion is evaluated with the usual scores 2, 3, 4, 5. As an option, conditional ITCs are used, that is, a set of solutions $\nu = \{v^1, v^2, v^3, v^4\}$. There are $m = 4$ criteria in total, which are represented by a single common scale and form a vector score for each ITC, for example $K(v^1) = (3, 5, 5, 4)$. Thus, there is a set of vector scores, which are called real or achievable.

Table 5.2 The value of the criteria for the effectiveness of the ITC operation functioning in Ukraine, proposed in transportation projects

ITC	Criteria			
	A	B	C	D
ITC I	3	5	5	4
ITC II	4	4	4	5
ITC III	3	5	5	4
ITC IV	3	5	3	5

Scores of the relative importance of criteria can be qualitative and quantitative. The qualitative importance of criteria is qualitative estimates, which are expressed by statements that one criterion is more important than another, or the criteria are equivalent. The statement "criterion K_i is more important than criterion K_j" is denoted as $i \succ j$, and the statement "criteria K_i and K_j are also equivalent" is denoted by $i \approx j$.

According to the data in **Table 5.2**, ITC I and ITC III have equivalent vector estimates, which is denoted as follows: $v^1 I_0 v^3$, and is identical to $v^3 I_0 v^1$. Here I_0 reflects the indifference relation, which means that when choosing an effectively functioning ITC, one can give preference to both ITC I and ITC III, which are equivalent to each other according to the selected criteria. Let's introduce the notation P^0, which means the preference relation between vector scores yP^0z, i.e. y prevails over z. This means that (4, 4, 4, 5) P^0 (4, 2, 4, 5). However, when comparing ITC I and ITC II, it is not possible to write:

Not (3, 5, 5, 4) P^0 (4, 4, 4, 5), not (4, 4, 4, 5) P^0 (3, 5, 5, 4), since such vector estimates are incomparable with respect to P^0.

If the statement $v'P^0v''$ is true for two options v', v'', then the option v'' cannot be considered the best and is called the dominant option. If for option v^* there is no such value of v that is the best with respect to P^0, that is, for which it would be correct to write vP^0v^*, then it is called non-dominant, or optimal according to Edgeworth-Pareto. In other words, the set of such options is the Edgeworth-Pareto set (V_0) [13, 28].

It becomes clear that only those options that belong to the set (V_0) can be optimal. Therefore, a preliminary analysis of all possible options allows to narrow the set of options (V) to the set (V_0).

In our example, as is usually the case in many applied multi-criteria problems, the set of options does not have a unique solution. Then the question arises, how exactly to choose the single best one from the set of heterogeneous options? To do this, it is proposed to introduce additional information from the DM (in our case, these are the participants in the transport process). The role of additional information is data on the relative importance of the criteria, as well as their scales. Additional information can reflect both the indifference ratio I_Ω for probable options and the preference ratio I^Ω for vector estimates [13, 29].

It is assumed that the TC capacity is more important than the actual intensity of vehicle traffic on the ITC routes according to information from the DM. In this case, the actual intensity of vehicle traffic and ensuring the safety and reliability of cargo transportation along the ITC routes are equivalent. Then this information can be written in the following form:

$$\Omega = \{1 > 2,2 \approx 3,3 \succ 4\}. \tag{5.20}$$

If the criterion K_3 is more important than the criterion K_4, then the vector estimate $x(v^1) = (3,5,5,4)$ prevails over y, then it is correct to write $x(v^1)P^{3\succ 4}y$. In particular, if K_1 and K_3 the criteria and are equally important in making the optimal decision, then their vector estimates $x(v^1) = (3,5,5,4)$ and $x(v^3) = (3,5,5,4)$ are equivalent, i.e. $(3,5,5,4)I^{1\approx 3}(3,5,5,4)$.

Analyzing this information, we would like to have effective methods with which we could build a chain of two arbitrary vector estimates x and y. Such methods exist and are called "combinatorial methods". They indicate that it is impossible to arbitrarily compare the vector estimates v^1, v^2, v^3 based on information Ω. Therefore, the variants v^1, v^2, v^3 are not dominant with respect to P_Ω. Thus, the information Ω makes it possible to narrow the set (V_0) to the set $V_\Omega = \{v^1, v^2, v^3, v^4\}$, in which two options for vector estimates are equivalent.

Therefore, the optimal solution from the set of possible ones is a solution that is non-dominant with respect to I^Ω and is determined by quantitative information about the importance of the criteria [30].

To check the relations xP^Ω and $xI^\Omega y$, there are also algebraic methods that are effective when comparing equivalent criteria. Information about the equal importance of two criteria is denoted by S. Let $x_\downarrow$ be a vector estimate formed from the values of x in decreasing order. For example, if $x = (3,4,2,3,5)$, then $x_\downarrow = (5,4,3,3,2)$. Therefore, the statements underlying this method can be described in the following form as $xP^\Omega y$ in the case if $x_\downarrow P^0 y_\downarrow$; $xI^S y$ in the case if $x_\downarrow = y_\downarrow$.

Unlike the qualitative importance of criteria, quantitative – can appear in two main forms:

1) in the degree of superiority of the importance of one criterion over another, that is, the criterion K_i is h times more important than the criterion K_j, if $h > 0$, however, if $h < 1$, then in fact the criterion K_j is $1/h > 1$ times more important than the criterion K_i, and under the condition $h = 1$ that the criteria under study are equivalent;

2) in the importance values of individual criteria, which are qualitatively measured on one general scale of importance, that is, the importance of the criterion K_i is expressed by the value β_i, where $\beta_j \geq 0$.

The degree of superiority (h) of the criterion over K_j is determined by the ratio of their importance values β_i and β_j:

$$h = \frac{\beta_i}{\beta_j}. \tag{5.21}$$

If a criterion K_i is more important than a criterion K_j by h times, then this statement is denoted by the expression $i \succ^h j$. Determining the superiority of the importance of one criterion over another by several times is based on the concept of the N-model, which takes into account only quantitative information about the importance of the criteria. This information is denoted by the Greek letter Θ (theta) and is formed on the basis of the DM experience about the advantages of some criteria over others.

If, when calculating the importance of the criteria β_i are summing to unity, they are called "importance coefficients". These coefficients determine the share of the "unit importance" of the set of all criteria that falls on each individual criterion K_i.

By the N-model let's mean a model with $n_1 + \ldots + n_m$ homogeneous criteria, and the first n_1 criteria can be obtained as a result of repeating the first criterion n_1 times, the next n_2 criteria – by repeating the first criterion n_2 times, etc. By analogy, the vector estimates of the initial model are formed into "extended" vector estimates of the N-model, which are called "N-fold estimates", or N-estimates. This means that each estimate for each criterion K_i is repeated n_i times [17, 31].

Information Θ corresponds not to one, but to a whole set of N-models, that is, it is enough to multiply all the numbers n_i by any natural number > 1 and it is possible to obtain a new N-model. Among the set of all N-models that correspond to information Θ, the simplest is the model in which all m numbers $n_1, \ldots, n_m$ are mutually simple in calculations [32].

Therefore, based on the so-called N-models, a new basic definition of the quantitative importance of criteria can be formed. A criterion K_i is h times more important than a criterion if the N-models meet or the following conditions are valid:

1) $\dfrac{n_1}{n_2} = h;$

2) each of the n_i criteria obtained from the criterion K_i is equivalent to any of the n_j criteria formed from the criterion K_j.

Let's consider the use of quantitative information about the importance of the criteria. The previously accumulated information about the quantitative importance of the criteria is presented in the following form [33, 34]:

$$\Theta = \left\{1 \succ^{3/2} 2,2 \approx 3,3 \succ^2 4\right\}. \tag{5.22}$$

In other words, let's consider the "extended" vector estimates of each ITC according to the information Θ. In particular, each estimate of the international transport corridor will be written out as many times as the equivalent criteria include this estimate:

$$\begin{aligned} x^{\Theta}(\upsilon^1) &= (3,3,3,5,5,4,4,4), \quad x^{\Theta}(\upsilon^2) = (4,4,4,4,4,5,5,5), \\ x^{\Theta}(\upsilon^3) &= (3,3,3,5,5,4,4,4), \quad x^{\Theta}(\upsilon^4) = (3,3,3,5,3,5,5,5). \end{aligned} \tag{5.23}$$

Let's note that all components of these vector estimates according to their formation can be considered as the values of eight equivalent criteria. Therefore, the developed *N*-model will have the form $N = (3,1,1,3)$. Next, let's present the estimates of the *N*-model, having previously arranged their components in descending order:

$$x^{\Theta}(V^1) = (5,5,4,4,4,3,3,3), \quad x^{\Theta}(V^2) = (5,5,5,4,4,4,4,4),$$
$$x^{\Theta}(V^3) = (5,5,4,4,4,3,3,3), \quad x^{\Theta}(V^4) = (5,5,5,5,3,3,3,3). \tag{5.24}$$

Comparing the *N*-estimates of the ITC II and ITC IV with respect to P^0, that is, comparing the components of the estimates by magnitude, let's obtain:

$$x \overset{\Theta}{\downarrow}(V^2) P^0 x \overset{\Theta}{\downarrow}(V^1) \text{ and } x \overset{\Theta}{\downarrow}(V^2) P^0 x \overset{\Theta}{\downarrow}(V^3). \tag{5.25}$$

This means that the ITC II, taking into account the information Θ, is a more effective option compared to the ITC I and ITC III.

When analyzing the *N*-estimates of the considered ITC II and ITC IV, ITC II is non-dominant compared to the ITC IV. Thus, the use of information Θ allows to narrow the set $V_{\Omega} = \{V^1, V^2, V^3, V^4\}$ to the set $V_{\Theta} = \{V^2, V^4\}$ in which the criterion V^2 prevails. However, the considered approach is idealized when it is possible to use for calculations the exact agreed values of the degree of superiority of one criterion over another, obtained from the DM.

But there are cases when, based on quantitative information about the importance of the criteria Θ, it is possible to obtain consistent interval estimates of the importance of the criteria.

For example, for the criteria K_i and K_j there is an unknown value h_{ij}, which is in the interval (l_{ij}, r_{ij}) and reflects the degree of superiority of the importance of the criterion K_i over the criterion K_j. If such information is available, it is impossible to apply the defined *N*-model. Therefore, the preference-indifference relationship must be determined using another approach, which assumes that the interval estimates of importance are consistent and that there exists an *N*-model that is consistent with the information Θ. This means that for each pair of criteria K_i and K_j there exists a set of values $\{h_{ij}\}$, the magnitude of which is equal to:

$$h = \frac{n_i}{n_j}, \tag{5.26}$$

and is in the interval (l_{ij}, r_{ij}), that is, the double inequality $l_{ij} < h_{ij} < r_{ij}$ is satisfied. Then the definition of the advantage P^{Θ}, formed by interval estimates of importance based on information Θ, is defined as follows: $xP^{\Theta}y$ is true if for each *N*-model consistent with Θ, the statement $xP^{\Theta}y$ is true.

Let's suppose that the degree of advantage of criterion V^1 over criterion V^2 is in the interval from 1.2 to 1.7, and the degree of advantage of criterion V^3 over criterion V^4 is in the range from 1.7 to 2.5. Then it turns out that with information Θ, criteria V^2 and V^1, as well as criteria V^2 and V^3,

are incomparable in terms of the degree of advantage, that is, it is impossible to choose such an ITC that is the best according to all criteria.

However, if the degree of superiority of criterion V^2 over criterion V^1 is within narrower limits, for example, in the range from 1.3 to 1.5, and the degree of superiority of criterion V^3 over criterion V is within the range from 1.7 to 2x, then the condition will be fulfilled:

$$V^2 P_{\ominus} V^1, V^2 P_{\ominus} V^4. \quad (5.27)$$

According to the studied criteria, the optimal one is the conditional transport corridor II. Determining the advantage of one criterion over another is based on the concept of the *N*-model, which allows to avoid performing cumbersome arithmetic operations when finding the optimal solution. The *N*-model involves using a set of weighting coefficients for each criterion, which allows to determine their importance relative to each other in conditions of uncertainty, in particular, when the information is interval. This allows to effectively process data that may have variations or shifts within given intervals, and thereby facilitates decision-making without the need to perform complex arithmetic operations. Interval information takes into account the variability of indicators, which can be important when assessing the real conditions of the functioning of transport corridors, when accurate data may be unavailable or incomplete. Using the *N*-model, it is possible to construct estimates for each criterion within given intervals and weigh these criteria to determine the overall efficiency of transport corridors. Given the comparison, the optimal one for the implementation of transportation projects is the transport corridor II, which demonstrates the best ratio between capacity, traffic intensity, safety and service conditions. The use of interval information and the *N*-model allows avoiding excessive computational complexity and provides more accurate and flexible decision-making, which is important in conditions of changing and uncertain circumstances, such as international transportation in conditions of war or unstable economic situation [35].

CONCLUSIONS

The main result of the study is to solve theoretical and practical problems in managing the quality of the project product (transport service) when performing international road transportation.

For the first time:

– a product management model for cargo transportation projects has been developed, which allows taking into account the significance of qualitative and quantitative characteristics of the transport service at each stage of the project life cycle;

– a comprehensive indicator for assessing the product of a cargo transportation project has been proposed, which takes into account the influence of indicators of a quantitative, qualitative and relay nature on the effectiveness of the project.

Improved:

– a model of the life cycle of a cargo transportation project taking into account the traditional phases of the project life cycle, which, unlike the existing one, takes into account the significance of the influence of quantitative and qualitative characteristics obtained experimentally from transport participants on the level of transport support for cargo transportation;

– a model for managing risks for transport support for cargo transportation projects in international traffic in the absence of complete and accurate information about the conditions of transportation.

Further development has been made:

– clarification of the principle of reflecting the quality of the project product, which determines the transfer of the quality of service provision to the quality of the final result;

– a terminological base for managing transport support projects by clarifying the concepts of "freight transport support project", "transportation project product", "result of the freight transportation project", "quality of transport service as a project product" by adapting them to the specifics of freight transportation projects.

The practical value of the results obtained lies in the development of a project analysis methodology for selecting freight transportation projects at project-oriented enterprises, the main activity of which is focused on international freight transportation, taking into account quality according to criteria that constitute a comprehensive indicator of the quality of the project product.

For the practical implementation of the proposed methodology for selecting a road route for transport corridors, a computer program has been developed that allows determining an integral assessment of the project product, which is based on information that is relevant in the transportation project. During the project implementation phase, the proposed program can be used to select the optimal route and enter new data or adjustments at the request of transport participants.

REFERENCES

1. Prokudin, H. S., Remekh, I. O. (2022). Theoretical fundamentals of the organization of freight transportation on transport networks. The National Transport University Bulletin, 3 (53), 312–321. https://doi.org/10.33744/2308-6645-2022-3-53-312-321
2. Chupailenko, O. A., Bilokur, M. V., Polishchuk, R. V., Kolesnyk, Yu. O. (2022). (2022). Logistics of functioning of multimodal transportation. The National Transport University Bulletin, 3 (53), 409–426. https://doi.org/10.33744/2308-6645-2022-3-53-409-426
3. Polishchuk, V. P., Guzhevska, L., Denys, O. (2021). Economic and mathematical model of cargo transportation in international piggyback connection. European Journal of Intelligent Transportation Systems, 1 (3). https://doi.org/10.31435/rsglobal_ejits/30032021/7371
4. Guzhevska, L., Denys, O. (2021). Modern problems of organization of multimodal transportation. Science Review, 1 (36). https://doi.org/10.31435/rsglobal_sr/30012021/7375

5. Prokudin, G., Chupaylenko, O., Prokudin, O., Khobotnia, T. (2021). Management decision support system freight transportation on transport networks. European Journal of Intelligent Transportation Systems, 1 (3). https://doi.org/10.31435/rsglobal_ejits/30032021/7352
6. Prokudin, G., Chupaylenko, O., Dudnik, O., Prokudin, O., Dudnik, A., & Svatko, V. (2018). Application of information technologies for the optimization of itinerary when delivering cargo by automobile transport. Eastern-European Journal of Enterprise Technologies, 2 (3 (92)), 51–59. https://doi.org/10.15587/1729-4061.2018.128907
7. Mazurenko, A., Kudriashov, A., Lebid, I., Luzhanska, N., Kravchenya, I., & Pitsyk, M. (2021). Development of a simulation model of a cargo customs complex operation as a link of a logistic supply chain. Eastern-European Journal of Enterprise Technologies, 5 (3 (113)), 19–29. https://doi.org/10.15587/1729-4061.2021.242915
8. Shoman, W., Yeh, S., Sprei, F., Köhler, J., Plötz, P., Todorov, Y., Rantala, S., Speth, D. (2023). A Review of Big Data in Road Freight Transport Modeling: Gaps and Potentials. Data Science for Transportation, 5 (1). https://doi.org/10.1007/s42421-023-00065-y
9. Wolff, M., Abreu, C., Caldas, M. A. F. (2019). Evaluation of road transport: a literature review. Brazilian Journal of Operations & Production Management, 16 (1), 96–103. https://doi.org/10.14488/bjopm.2019.v16.n1.a9
10. Chung, S.-H. (2021). Applications of smart technologies in logistics and transport: A review. Transportation Research Part E: Logistics and Transportation Review, 153, 102455. https://doi.org/10.1016/j.tre.2021.102455
11. Martin, B., Ortega, E., Cuevas-Wizner, R., Ledda, A., De Montis, A. (2021). Assessing road network resilience: An accessibility comparative analysis. Transportation Research Part D: Transport and Environment, 95, 102851. https://doi.org/10.1016/j.trd.2021.102851
12. Duna, N., Matviienko, A. (2022). Prospects for the Development of the Ukrainian Road Freight Transport Market: the European Integration Aspect. Herald UNU. International Economic Relations And World Economy, 44, 21–29. https://doi.org/10.32782/2413-9971/2022-44-4
13. Volynets, L. (2021). Liberalization of international road transportation as a new impulse for development of transport industry. Economics of the Transport Complex, 37, 161–176. https://doi.org/10.30977/etk.2225-2304.2021.37.161
14. Nitsenko, V., Sharapa, O., Burdeina, N., Hanzhurenko, I. (2018). Accounting and analytical information in the management system of a trading enterprise in Ukraine. Visnyk KhNAU im. V. V. Dokuchaieva. Seriia "Ekonomichni nauky", 2, 3–18.
15. Kunda N., Lebid V. (2019). Interconnection of Transport Services while Cargo Transportation, Journal of Advanced Research in Law and Economics. 10 (8 (46)), 2394–2406. https://doi.org/10.14505/jarle.v10.8(46).18
16. Kunda, N., Lebid, V. (2019). Indefinite-multiple model of risk management during freight transportation. Bulletin of Kharkov National Automobile and Highway University, 85, 117–124. https://doi.org/10.30977/bul.2219-5548.2019.85.0.117

17. Poliak, M., Poliakova, A., Svabova, L., Zhuravleva, N. A., Nica, E. (2021). Competitiveness of Price in International Road Freight Transport. Journal of Competitiveness, 13 (2), 83–98. https://doi.org/10.7441/joc.2021.02.05
18. Krasnyanskiy, M., Penshin, N. (2019). Quality criteria when assessing competitiveness in road transport services. Transport Problems, 11 (4), 15–20. https://doi.org/10.20858/tp.2016.11.4.2
19. Mitsakis, E., Iordanopoulos, P., Aifadopoulou, G., Tyrinopoulos, Y., Chatziathanasiou, M. (2019). Deployment of Intelligent Transportation Systems in South East Europe: Current Status and Future Prospects. Transportation Research Record: Journal of the Transportation Research Board, 2489 (1), 39–48. https://doi.org/10.3141/2489-05
20. Cancela, H., Mauttone, A., Urquhart, M. E. (2015). Mathematical programming formulations for transit network design. Transportation Research Part B: Methodological, 77, 17–37. https://doi.org/10.1016/j.trb.2015.03.006
21. Danchuk, V., Bakulich, O., Svatko, V. (2019). Identifying optimal location and necessary quantity of warehouses in logistic system using a radiation therapy method. Transport, 34 (3), 175–186. https://doi.org/10.3846/transport.2019.8546
22. Medvediev, I., Eliseyev, P., Lebid, I., Sakno, O. (2020). A modelling approach to the transport support for the harvesting and transportation complex under uncertain conditions. IOP Conference Series: Materials Science and Engineering. Transport, ecology – sustainable development. EKO. IOP Publishing, 977, 012003. https://doi.org/10.1088/1757-899x/977/1/012003
23. Gryshchuk, O., Petryk, A., Yerko, Y. (2022). Development of methods for formation of infrastructure of transport units for maintenance of transit and export freight flows. Technology Audit and Production Reserves, 1 (2 (63)), 26–30. https://doi.org/10.15587/2706-5448.2022.251505
24. Cassiano, D. R., Bertoncini, B. V., de Oliveira, L. K. (2021). A Conceptual Model Based on the Activity System and Transportation System for Sustainable Urban Freight Transport. Sustainability, 13 (10), 5642. https://doi.org/10.3390/su13105642
25. Cicconi, P., Landi, D., Germani, M. (2016). A Virtual Modelling of a Hybrid Road Tractor for Freight Delivery. Volume 12: Transportation Systems. https://doi.org/10.1115/imece2016-68013
26. Bekmagambetov, M., Kochetkov, A. (2022). Analysis of modern software of transport simulation of research, design, technology. Journal of Automotive Engineers, 6 (77), 25–34.
27. Prokudin, G., Oliskevych, M., Chupaylenko, O., Maidanik, K. (2020). Optimization model of freight transportation on the routes of international transport corridors. Journal of Sustainable Development of Transport and Logistics, 5 (1), 66–76. https://doi.org/10.14254/jsdtl.2020.5-1.7
28. Chukurna, O. P., Nitsenko, V. S., Hanzhurenko, I. V., Honcharuk, N. R. (2019). Directions of innovative development of transport logistics in Ukraine. Economic Innovations, 21 (1 (70)), 170–181. https://doi.org/10.31520/ei.2019.21.1(70).170-181

29. Roi, M. P. (2020). Method of optimization of the integrated transport process of cargo road transportation. Scientists of Notes of TNU named Vernadsky V. I. Series: Technical Sciences, 31 (5 (70)), 220–227. https://doi.org/10.32838/2663-5941/2020.5/36
30. Syniavs'ka, O., Repich, T. (2020). Optimization of freight transportation using the GPS monitoring system. Young Scientist, 2 (78), 356–359. https://doi.org/10.32839/2304-5809/2020-2-78-75
31. Moroz, M., Zahorianskyi, V., Haikova, T., Kuzev, I. (2022). The using of operations research methods to optimize freight road transportation in the agroindustrial complex. Bulletin of the National Technical University "KhPI" Series: New Solutions in Modern Technologies, 1 (11), 44–50. https://doi.org/10.20998/2413-4295.2022.01.07
32. Jourquin, B. (2019). Estimating Elasticities for Freight Transport Using a Network Model: An Applied Methodological Framework. Journal of Transportation Technologies, 9 (1), 1–13. https://doi.org/10.4236/jtts.2019.91001
33. Zahorianskyi, V., Zahorianska, O., Moroz, M., Moroz, O. (2022). Development of a Model for Minimizing the Energy Costs of the Transport and Technological Complex. 2022 IEEE 4th International Conference on Modern Electrical and Energy System (MEES), 1–5. https://doi.org/10.1109/mees58014.2022.10005635
34. Odoki, J. B., Di Graziano, A., Akena, R. (2015). A multi-criteria methodology for optimising road investments. Proceedings of the Institution of Civil Engineers – Transport, 168 (1), 34–47. https://doi.org/10.1680/tran.12.00053
35. Gnap, J., Konecny, V., Vajran, P. (2018). Research on Relationship between Freight Transport Performance and GDP in Slovakia and EU Countries. Naše more, 65 (1), 32–39. https://doi.org/10.17818/NM/2018/1.5

Edited by
Maksym Ievlanov

PROJECT MANAGEMENT: INDUSTRY SPECIFICS

Maksym Ievlanov, Nataliya Vasiltcova, Olga Neumyvakina, Iryna Panforova,
Maksym Naumenko, Iryna Hrashchenko, Tetiana Tsalko, Svitlana Nevmerzhytska,
Svitlana Krasniuk, Yurii Kulynych, Viktor Myronenko, Andrii Pozdniakov, Yaroslav Ziubryk,
Valerii Samsonkin, Ivan Riabushko, Oksana Malanchuk, Anatoliy Tryhuba, Ivan Rogovskii,
Liudmyla Titova, Liudmyla Berezova, Mykola Korobko, Georgii Prokudin, Viktoriia Lebid,
Oleksii Chupaylenko, Tetiana Khobotnia, Maksym Roi, Mykhailo Holovatiuk

Collective monograph

Technical editor I. Prudius
Desktop publishing T. Serhiienko

TECHNOLOGY CENTER PC®
Published in December 2024
Enlisting the subject of publishing No. 4452 – 10.12.2012
Address: Shatylova dacha str., 4, Kharkiv, Ukraine, 61165
