

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ УКРАИНЫ
MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE

Національний технічний університет
«Харківський політехнічний інститут»
Национальный технический университет
«Харьковский политехнический институт»
National Technical University "Kharkiv Polytechnic Institute"

Вісник Національного технічного університету «ХПІ».
Серія: Системний аналіз, управління та інформаційні
технології

Збірник наукових праць
№ 1'2019
Видання засноване у 1961 р.

Вестник Национального технического университета «ХПИ».
Серия: Системный анализ, управление и информационные
технологии

Сборник научных работ
№ 1'2019
Издание основано в 1961 г.

Bulletin of the National Technical University "KhPI".
Series: System analysis, control and information technology
Collection of Scientific papers
No. 1'2019
The edition was founded in 1961

Харків
НТУ «ХПІ», 2019

Вісник Національного технічного університету «ХПІ». Серія: Системний аналіз, управління та інформаційні технології = Вестник Национального технического университета «ХПИ». Серия: Системный анализ, управление и информационные технологии = Bulletin of the National Technical University "KhPI". Series: System analysis, control and information technology : зб. наук. пр. / Нац. техн. ун-т «Харків. політехн. ін-т». — Харків : НТУ «ХПІ», 2019. — № 1'2019. — 73 с. — ISSN 2079-0023.

Видання публікує нові наукові результати в області системного аналізу та управління складними системами, отримані на основі сучасних прикладних математичних методів і прогресивних інформаційних технологій. Публікуються роботи, пов'язані зі штучним інтелектом, аналізом великих даних, сучасними методами високопродуктивних обчислень у системах підтримки прийняття рішень.

Для науковців, викладачів вищої школи, аспірантів, студентів і фахівців у галузі системного аналізу, управління і комп'ютерних технологій.

Издание публикует новые научные результаты в области системного анализа и управления сложными системами, полученные на основе современных прикладных математических методов и прогрессивных информационных технологий. Публикуются работы, связанные с искусственным интеллектом, анализом больших данных, современными методами высокопродуктивных вычислений в системах поддержки принятия решений.

Для научных работников, преподавателей высшей школы, аспирантов, студентов и специалистов в области системного анализа, управления и компьютерных технологий.

Edition publishes new scientific results in the field of system analysis and control of complex systems, based on the application of modern mathematical methods and advanced information technology. Works related to artificial intelligence, big data analysis and modern methods of high-performance computing in decision support systems are publishing.

For scientists, teachers of higher education, post-graduate students, students and specialists in the field of systems analysis, management and computer technology.

Свідоцтво Міністерства юстиції України
КВ № 23778-13618Р від 14 лютого 2019 р.

Мова статей – українська, російська, англійська.

Вісник Національного технічного університету «ХПІ». Серія: Системний аналіз, управління та інформаційні технології включений до зовнішніх інформаційних систем, у тому числі в базу журналів відкритого доступу DOAJ (Directory of Open Access Journals), бібліографічну базу даних OCLC WorldCat (США), індексується пошуковими системами Google Scholar і Crossref; зареєстрований у світовому каталозі періодичних видань бази даних Ulrich's Periodicals Directory (New Jersey, USA).

Офіційний сайт видання: <http://samit.khpi.edu.ua/>

Засновник
Національний технічний університет
«Харківський політехнічний інститут»

Founder
National Technical University
"Kharkiv Polytechnic Institute"

Редакційна колегія

Головний редактор:

Годлевський М. Д., проф., НТУ «ХПІ», Україна

Заступник головного редактора

Куценко О. С., проф., НТУ «ХПІ», Україна

Члени редколегії:

Александрова Т. Е., проф., НТУ «ХПІ», Україна

Безменов М. І., проф., НТУ «ХПІ», Україна

Бентаєб Ф., доц., Ліонський університет-2, Франція

Богомолів С., доц., Австралійський національний університет, Австралія

Гамаюн І. П., проф., НТУ «ХПІ», Україна

Дідманідзе І., проф., Батумський держ. ун-т ім. Шота Руставелі, Грузія

Дорофєєв Ю. І., проф., НТУ «ХПІ», Україна

Керстен В., проф., Гамбурзький технологічний університет, Німеччина

Любчик Л. М., проф., НТУ «ХПІ», Україна

Павлов О. А., проф., НТУУ «КПІ», Україна

Раскін Л. Г., проф., НТУ «ХПІ», Україна

Северин В. П., проф., НТУ «ХПІ», Україна

Ткачук М. В., проф., НТУ «ХПІ», Україна

Хайрова Н. Ф., проф., НТУ «ХПІ», Україна

Шаронова Н. В., проф., НТУ «ХПІ», Україна

Відповідальний секретар

Безменов М. І., проф., НТУ «ХПІ», Україна

Рекомендовано до друку Вченою радою НТУ «ХПІ».

Протокол № 7 від 5 липня 2019 р.

Editorial

Editor-in-chief

Godlevskiy M. D., prof., NTU "KhPI", Ukraine

Deputy editor-in-chief

Kutsenko O. S., prof., NTU "KhPI", Ukraine

Editorial staff members:

Alexandrova T. E., prof., NTU "KhPI", Ukraine

Bezmenov M. I., prof., NTU "KhPI", Ukraine

Bentayeb F., Associate Professor, University of Lyon-2, France

Bogomolov S., Assistant Professor, Australian National University, Australia

Gamayun I. P., prof., NTU "KhPI", Ukraine

Didmanidze I., prof., Batumi Shota Rustaveli State University, Georgia

Dorofiev Yu. I., prof., NTU "KhPI", Ukraine

Kersten Wolfgang, Prof., Hamburg University of Technology, Germany

Lyubchik L. M., prof., NTU "KhPI", Ukraine

Pavlov O. A., prof., NTUU "KPI", Ukraine

Raskin L. G., prof., NTU "KhPI", Ukraine

Severin V. P., prof., NTU "KhPI", Ukraine

Tkachuk M. V., prof., NTU "KhPI", Ukraine

Khairova N. F., prof., NTU "KhPI", Ukraine

Sharonova N. V., prof., NTU "KhPI", Ukraine

Executive secretary:

Bezmenov M. I., prof., NTU "KhPI", Ukraine

СИСТЕМНИЙ АНАЛІЗ І ТЕОРІЯ ПРИЙНЯТТЯ РІШЕНЬ

СИСТЕМНЫЙ АНАЛИЗ И ТЕОРИЯ ПРИНЯТИЯ РЕШЕНИЙ

SYSTEM ANALYSIS AND DECISION-MAKING THEORY

UDC 519.[816+854.2]

DOI: 10.20998/2079-0023.2019.01.01

A. A. PAVLOV

COMBINATORIAL OPTIMIZATION UNDER UNCERTAINTY AND FORMAL MODELS OF EXPERT ESTIMATION

Previously, the author formalized the concepts of uncertainty, compromise solution, compromise criteria and conditions for a quite general class of combinatorial optimization problems. The functional of the class' problems contains linear convolution of weights and arbitrary numerical characteristics of a feasible solution. It was shown that the efficiency of the presented algorithms for the uncertainty resolution is largely determined by the efficiency of solving the combinatorial optimization problem in a deterministic formulation. A part of the formulated compromise criteria and conditions uses expert weights. Previously, the author and his disciples also formulated combinatorial optimization models, optimality criteria, criteria for decisions' consistency. The models allow to evaluate and justify the degree of stability and reliability of the estimated values of empirical coefficients using a formally ill-conditioned empirical pairwise comparison matrix of arbitrary dimension. The matrix may contain zero elements. The theoretical research and statistical experiments allowed to choose the most efficient of these optimization models. In this article, on the base of earlier results by the author and his disciples, we formalize and substantiate the efficiency of the proposed sequential procedure for expert estimation of weights that determine compromise criteria and conditions. The procedure is an integral part of the algorithm introduced by the author to solve combinatorial optimization problems under uncertainty of the mentioned class. We give unified algorithm for efficient uncertainty resolution that includes original and efficient formal procedure for expert coefficients' estimation using empirical matrices of pairwise comparisons.

Keywords: combinatorial optimization, uncertainty, compromise criteria, compromise conditions, empirical matrix of pairwise comparisons, consistent decision.

О. А. ПАВЛОВ

КОМБІНАТОРНА ОПТИМІЗАЦІЯ В УМОВАХ НЕВИЗНАЧЕНОСТІ ТА ФОРМАЛЬНІ МОДЕЛІ ЕКСПЕРТНОГО ОЦІНЮВАННЯ

Раніше автором для досить загального класу задач комбінаторної оптимізації, функціонал яких містить лінійну згортку ваг і довільних числових характеристик допустимого розв'язку, були формалізовані поняття невизначеності, компромісного розв'язку, компромісних критеріїв та умов. Було показано, що ефективність наведених алгоритмів розв'язання невизначеності в значній мірі визначається ефективністю розв'язання задачі комбінаторної оптимізації в детермінованій постановці. Частина сформульованих компромісних критеріїв і умов використовує експертні вагові коефіцієнти. Раніше також автором і його учнями були сформульовані моделі комбінаторної оптимізації, критерії оптимальності, критерії узгодженості рішень, що дозволяють за формально погано узгодженою емпіричною матрицею парних порівнянь довільної розмірності, яка, можливо, містить нульові елементи, знаходити та обґрунтовувати ступінь стійкості та достовірності знайдених значень емпіричних коефіцієнтів. Проведені теоретичні дослідження та статистичні експерименти дозволили виділити з цих моделей оптимізації найбільш ефективні. У даній статті на основі отриманих раніше автором і його учнями результатів формалізовано і обґрунтовано ефективність запропонованої послідовної процедури знаходження експертних вагових коефіцієнтів, що визначають компромісні критерії та умови, як складової частини алгоритму розв'язання для введеного автором класу задач комбінаторної оптимізації в умовах невизначеності. Наводиться єдиний алгоритм ефективного розв'язання невизначеності, який включає в себе оригінальну ефективну формальну процедуру знаходження експертних коефіцієнтів за емпіричними матрицями парних порівнянь.

Ключові слова: комбінаторна оптимізація, невизначеність, компромісні критерії, компромісні умови, емпірична матриця парних порівнянь, узгоджене рішення.

A. A. PAVLOV

КОМБІНАТОРНАЯ ОПТИМИЗАЦИЯ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ И ФОРМАЛЬНЫЕ МОДЕЛИ ЭКСПЕРТНОГО ОЦЕНИВАНИЯ

Ранее автором для достаточно общего класса задач комбинаторной оптимизации, функционал которых содержит линейную свертку весов и произвольных числовых характеристик допустимого решения, были формализованы понятия неопределенности, компромиссного решения, компромиссных критериев и условий. Было показано, что эффективность приведенных алгоритмов разрешения неопределенности в значительной степени определяется эффективностью решения задачи комбинаторной оптимизации в детерминированной постановке. Часть сформулированных компромиссных критериев и условий использует экспертные весовые коэффициенты. Ранее также автором и его учениками были сформулированы модели комбинаторной оптимизации, критерии оптимальности, критерии согласованности решений, позволяющие по формально плохо согласованной эмпирической матрице парных сравнений произвольной размерности, возможно, содержащей нулевые элементы, находить и обосновывать степень устойчивости, достоверности найденных значений эмпирических коэффициентов. Проведенные теоретические исследования и статистические эксперименты позволили выделить из этих моделей оптимизации наиболее эффективные. В данной статье на основе полученных ранее автором и его учениками результатов формализована и обоснована эффективность предложенной последовательной процедуры нахождения экспертных весовых коэффициентов, определяющих

© A. A. Pavlov, 2019

компромиссные критерии и условия, как составной части алгоритма решения для введенного автором класса задач комбинаторной оптимизации в условиях неопределенности. Приводится единый алгоритм эффективного разрешения неопределенности, который включает в себя оригинальную эффективную формальную процедуру нахождения экспертных коэффициентов по эмпирическим матрицам парных сравнений.

Ключевые слова: комбинаторная оптимизация, неопределенность, компромиссные критерии, компромиссные условия, эмпирическая матрица парных сравнений, согласованное решение.

Introduction. We studied in [1] a class of combinatorial optimization problems of the following form:

$$\min_{\sigma \in \Omega} (\max) \sum_{i=1}^s \omega_i k_i(\sigma). \quad (1)$$

Here, ω_i are numbers, Ω the set of feasible solutions; σ an arbitrary feasible solution, $k_i(\sigma)$ is i -th arbitrary numerical characteristic of σ . We recommended in [1] to use the proposed approach to solve problems of the form (1) under uncertainty only for the following case: for the problems of the form (1) there exists an exact (or approximation) algorithm which is qualitatively more efficient in terms of speed and/or accuracy than an arbitrary solving method for the problem (1) in which the structure of the feasible solutions domain differs from Ω of (1), e.g., in additional constraints on inequalities. For a problem of the form (1), solution of which uses expert coefficients, we formulated in [1] the combinatorial optimization problem statement under uncertainty as follows.

There are L sets of weights $\{\omega_i^l, i = \overline{1, s}\}$, $l = \overline{1, L}$. Here, $\overline{a, b}$ denotes the interval of integer numbers from a to b . Each of the sets may be a set of coefficients $\omega_1, \dots, \omega_s$ of the problem (1) at the stage of implementation of its solution. Probabilities $P_l > 0$, $l = \overline{1, L}$, $\sum_l P_l = 1$, may be specified for each possible set of weights (such probabilities do not exist if the uncertainty is not described by a probabilistic model). We need to find a feasible solution satisfying one of the conditions given below.

Here we investigate the problem (1) for the case of $\min_{\sigma \in \Omega} \sum_{i=1}^s \omega_i k_i(\sigma)$. We consider the case of $\max_{\sigma \in \Omega} \sum_{i=1}^s \omega_i k_i(\sigma)$ in [1]. Let us introduce the notation:

$$f_{opt}^l = \min_{\sigma \in \Omega} \sum_{i=1}^s \omega_i^l k_i(\sigma),$$

$$\{\sigma_l\} = \arg \min_{\sigma \in \Omega} \sum_{i=1}^s \omega_i^l k_i(\sigma),$$

$$L_l = \sum_{m=1}^L (\sum_{i=1}^s \omega_i^m k_i(\sigma_l) - f_{opt}^m), l = \overline{1, L}.$$

Remark 1. If $\{\sigma_l\}$ consists of more than one solution, then we leave the one on which we achieve $\min_{\{\sigma_l\}} L_l$ and denote this solution as σ_l . We consider in [1] obtaining σ_l for the case when Ω is finite.

Suppose that

$$\Delta_l = \sum_{i=1}^s \omega_i^l k_i(\sigma) - f_{opt}^l \geq 0,$$

$$f_{opt}^l = \min_{\sigma \in \Omega} \sum_{i=1}^s \omega_i^l k_i(\sigma). \quad (2)$$

We need to find:

1) a feasible solution $\sigma(\Delta_1, \dots, \Delta_L)$ that has

$$\Delta_i \leq l_i, i = \overline{1, L} \quad (3)$$

where l_i are given numbers;

2) a feasible solution that satisfies the condition:

$$\min_{\sigma \in \Omega} \sum_{l=1}^L a_l (\sum_{i=1}^s \omega_i^l k_i(\sigma) - f_{opt}^l) \quad (4)$$

where $\forall a_l > 0$ are the coefficients specified by an expert.

Remark 2. For the general case, we formulated in [1] five compromise criteria and conditions for $L = 2$ and four compromise criteria and conditions for $L > 2$.

Remark 3. We formulated similar criteria, conditions, and algorithms for the case of the problem (1) when we reach the optimal functional value at $\max_{\sigma \in \Omega}$.

The compromise solutions proposed in [1] are based on the following statement and its corollaries.

Statement 1 [1]. The following is true for arbitrary $a_l > 0$, $l = \overline{1, L}$:

$$\begin{aligned} & \arg \min_{\sigma \in \Omega} \sum_{l=1}^L a_l [\sum_{i=1}^s \omega_i^l k_i(\sigma) - f_{opt}^l] \\ & = \arg \min_{\sigma \in \Omega} \sum_{i=1}^s (\sum_{l=1}^L a_l \omega_i^l) k_i(\sigma). \end{aligned} \quad (5)$$

Corollary 1. Solving the problem $\min_{\sigma \in \Omega} \sum_{l=1}^L a_l \times (\sum_{i=1}^s \omega_i^l k_i(\sigma) - f_{opt}^l)$ reduces to solving one problem of the form (1):

$$\min_{\sigma \in \Omega} \sum_{i=1}^s (\sum_{l=1}^L a_l \omega_i^l) k_i(\sigma). \quad (6)$$

Corollary 2. Suppose that $\sigma(a_1, \dots, a_L, \Delta_1, \dots, \Delta_L) \neq \sigma_1 \vee \dots \vee \sigma_L$; $\exists \sigma(a_1, a_2, \dots, a_{i-1}, a'_i, a_{i+1}, \dots, a_L, \Delta'_1, \dots, \Delta'_L) \neq \sigma_1 \vee \dots \vee \sigma_L, a'_i \neq a_i$. Then the following holds:

$$\begin{aligned} & (a'_i - a_i)(\Delta'_i - \Delta_i) < 0, \\ & (\Delta'_i - \Delta_i)[(a_1 \Delta'_1 + \dots + a_{i-1} \Delta'_{i-1} + \\ & + a_{i+1} \Delta'_{i+1} + \dots + a_L \Delta'_L) - (a_1 \Delta_1 + \dots + \\ & + a_{i-1} \Delta_{i-1} + a_{i+1} \Delta_{i+1} + \dots + a_L \Delta_L)] < 0. \end{aligned} \quad (7)$$

where $\sigma(a_1, \dots, a_L, \Delta_1, \dots, \Delta_L)$ is an optimal solution of the problem (6) with given expert coefficients $a_i > 0, i = \overline{1, L}$. $\Delta_i(\Delta'_i), i = \overline{1, L}$ are specified in (2) for sets $a_1, \dots, a_L; a_1, \dots, a_{i-1}, a'_i, a_{i+1}, \dots, a_L$, respectively.

It is shown in [1] that if $\sigma(1, \dots, 1, \Delta_1, \dots, \Delta_L) \neq \min_{j=\overline{1, L}} \{\sigma_j\}$, then, by virtue of Corollary 2 (by logic of the inequalities (7)), we can organize a sequential procedure for solving a problem of the form (6), increasing $\forall a_i$ at each iteration if $\Delta_i > l_i$ or decreasing $\forall a_j$ if $\Delta_j < l_j$ (here, $\sigma(\Delta_1, \dots, \Delta_L)$ is the current solution). As a result, we either find a solution $\sigma(\Delta_1, \dots, \Delta_L)$ that satisfies the condition $\Delta_i \leq l_i, i = \overline{1, L}$ or obtain a set of solutions $\{\sigma\}^1$, each of which violates the compromise condition (1). In [1], we proposed

to choose the solution from the set $\{\sigma_1, \dots, \sigma_L\} \cup \{\sigma\}^1$ as the compromise solution if it reaches

$$\min \sum_t C_{jt} (\Delta_{jt} - l_{jt}), \forall t \Delta_{jt} > l_{jt}, \quad (8)$$

where $C_j > 0, j = \overline{1, L}$ are expert coefficients.

Thus, the efficiency of a compromise solution that meets the compromise condition 1) or the compromise criterion 2) essentially depends on the relevance of expert coefficients $a_j > 0, C_j > 0, j = \overline{1, L}$, used in optimization models, to the essence of the practical problem, the formal model of which is the combinatorial optimization problem under uncertainty.

Formal models for expert coefficients' estimation by an empirical matrix of pairwise comparisons. The author and his disciples formulated in [2–4] combinatorial optimization models, optimality criteria, criteria for decisions' consistency. The models allow to evaluate and justify the degree of stability and reliability of the estimated values of empirical coefficients using a formally ill-conditioned empirical pairwise comparison matrix of arbitrary dimension. The matrix may contain zero elements. Theoretical research and statistical experiments [2–4] allow to use the following formal procedures to estimate the empirical coefficients $a_j > 0, C_j > 0, j = \overline{1, L}$.

1. The initial data is an empirical matrix Γ (for estimation of weights $a_j > 0, j = \overline{1, L}$, or $C_j > 0, j = \overline{1, L}$). In general case, we denote weights by $\omega_j > 0, j = \overline{1, L}$. γ_{ij} is (ij) -th coefficient of the matrix Γ . It determines, according to the expert, how many times the value of ω_i is greater than the value of ω_j .

To construct and substantiate optimization criteria and models for estimation of weights $\omega_i, i = \overline{1, L}$, by an empirical matrix of pairwise comparisons Γ , we use the set A of coefficients $\gamma_{ij}, i, j = \overline{1, L}$, of the matrix of pairwise comparisons Γ which satisfy the conditions:

- 1) $\{\gamma_{ij}, i = \overline{1, L}\} \subset A$;
- 2) $\forall i \neq j \gamma_{ij} \in A$ if $\gamma_{ij} \geq 1, \gamma_{ij} \notin A$.

We propose to use the measure

$$\frac{1}{\gamma_{ij}} \left| \frac{\omega_i}{\omega_j} - \gamma_{ij} \right| \quad (9)$$

as the consistency measure for estimated $\omega_i, i = \overline{1, L}$.

2. We propose to use the following models as optimization models for weights' estimation.

Model 1 requires sequential solving of the following linear programming (LP) problems:

$$\min_{\substack{\Delta_{ij}^1, \Delta_{ij}^2 \\ \forall (ij) \in |A|}} \left(\sum_{(ij) \in |A|} \sum \Delta_{ij}^1 - C_{lm} \sum_{(ij) \in |A|} \sum \Delta_{ij}^2 \right), \quad (10)$$

$$\ln \gamma_{ij} + \Delta_{ij}^2 \leq W_i - W_j \leq \ln \gamma_{ij} + \Delta_{ij}^1,$$

$$\forall (ij) \in |A|,$$

$$0 \leq \Delta_{ij}^1 \leq \ln(1 + l \cdot \Delta_1(l)), \forall (ij) \in |A|, \quad (11)$$

$$0 \geq \Delta_{ij}^2 \geq \ln(1 - m \cdot \Delta_2(m)), \forall (ij) \in |A|,$$

$$W_i \geq 0, i = \overline{1, L}, \quad (12)$$

where $W_i, \Delta_{ij}^1, \Delta_{ij}^2$ are the LP problem coefficients;

l, m natural numbers successively taking values (1; 1), (1; 2), (2; 1), (2; 2), etc.;

$\Delta_1(x), \Delta_2(x)$ given numeric scalar functions taking non-negative values, of a natural argument;

C_{lm} coefficient determined from the formula:

$$\ln(1 + l \cdot \Delta_1(l)) = C_{lm} \ln \frac{1}{1 - m \cdot \Delta_2(m)}; \quad (13)$$

$|A|$ the set of pairs $(ij), i \neq j$, for each of which $\gamma_{ij} \geq 1 (\gamma_{ij} \in |A|)$.

When using problems (10)–(12) for weights' estimation, we set the functions $\Delta_1(x), \Delta_2(x)$ in such a way that the values of $l \cdot \Delta_1(l)$ and $m \cdot \Delta_2(m)$ at each iteration (for each attempt of solving the LP problem (10)–(12)) increase in a small ratio to their previous value. At the first iteration they take the minimum possible values. Due to this method of functions $\Delta_1(x), \Delta_2(x)$ specification, we can achieve the most consistent weights, with respect to measure (9), for each specific problem. This is possible due to the peculiarities of the current matrix of pairwise comparisons.

Iterations stop at the first successful solution of the LP problem (10)–(12) (the first LP problem with a non-empty set of constraints). After that, we find the weights of the objects $\omega_i^*, i = \overline{1, L}$, from the relationship:

$$\omega_i^* = e^{W_i^*}, i = \overline{1, L}. \quad (14)$$

Model 2 is the LP problem of the form:

$$\min \sum_{(ij) \in |A|} \sum r_{ij} \cdot \gamma_{ij},$$

$$-\gamma_{ij} \leq \omega_i - \gamma_{ij} \omega_j \leq \gamma_{ij}, \gamma_{ij} \geq 0, \quad (i, j) \in |A|$$

$$\omega_i \geq a \geq 1, i = \overline{1, L} \quad (15)$$

where a is a specified number;

$\omega_i, i = \overline{1, L}, \gamma_{ij}, (ij) \in |A|$ the LP problem variables;

r_{ij} weighted coefficient.

We propose to solve the LP problem (15) for coefficients r_{ij} that take the following values:

$$1) \quad r_{ij} = \frac{1}{\gamma_{ij} - 1}, \forall (ij) \in |A|; \quad (16)$$

$$2) \quad r_{ij} = \frac{1}{\sqrt{\gamma_{ij} - 1}}, \forall (ij) \in |A|; \quad (17)$$

$$3) \quad -\gamma_{ij} \leq \omega_i - \gamma_{ij} \omega_j \leq \gamma_{ij}, \gamma_{ij} \geq 0. \quad (18)$$

Remark 4. For cases $\gamma_{ij} - 1 \in [0, \varepsilon]$ where $\varepsilon > 0$ is a given number, we replace $\gamma_{ij} - 1$ by ε in expressions (16)–(18).

Remark 5. We can also use other analytical expressions for coefficients r_{ij} that satisfy the condition: $r_{ij}^1 < r_{ij}^2$

follows from $\forall \gamma_{ij}^1 > \gamma_{ij}^2 > 1 + \varepsilon$ (where $\varepsilon > 0$ is a given number).

Remark 6. An analysis of statistical experiments in [4] concluded that the following analytical expression for r_{ij} is statistically efficient:

$$r_{ij} = \frac{1}{\sqrt[3]{(\gamma_{ij} - 1)^2}}, \gamma_{ij} > 1 \ (\varepsilon = 0.1).$$

Nevertheless, in each individual case, we recommend to solve a set of LP problems (15) in which r_{ij} are determined from expressions (16)–(18).

3.1. *Choosing the best solution.* As a result of implementing item 2, we obtain a set of weights $\{\omega_i^p, i = \overline{1, L}\}$, $p = \overline{1, k}$, where k is the number of different found sets of weights $\omega_i, i = \overline{1, L}$. The solution is a set of weights $\omega_i^m, i = \overline{1, L}$, on which we reach the minimum:

$$\min_{p=\overline{1,k}} \sum_{(ij) \in |A|} \sum \frac{1}{\gamma_{ij}} \left| \frac{\omega_i^p}{\omega_j^p} - \gamma_{ij} \right|. \quad (19)$$

3.2. We need to perform the following operations [2, 4] to determine whether the set of weights $\omega_i^m, i = \overline{1, L}$, is a stable solution obtained by the empirical matrix of pairwise comparisons Γ .

We build M combinatorial optimization problems, in each of which we replace $\gamma_{ij}, \forall (ij) \in |A|$, with arbitrary $\gamma_{ij}^l, \forall (ij) \in |A|, l = \overline{1, M}$, that satisfy the conditions

$$\gamma_{ij}^l - \gamma_{ij} = \pm \left| \frac{\hat{\omega}_i}{\hat{\omega}_j} - \gamma_{ij} \right|, \forall (ij) \in |A|,$$

where M is a sufficiently large natural number.

Remark 7. Combinatorial optimization problem to obtain $\omega_i^l, l = \overline{1, M}$, coincides with the one for which we obtained $\omega_i^m, i = \overline{1, L}$, up to a replacement of γ_{ij} by γ_{ij}^l .

Let $\hat{\omega}_i^l, i = \overline{1, L}$, be the estimates of the alternatives' weights obtained by l -th optimization model (l -th set of optimization models), $l = \overline{1, M}$. The weights $\omega_i^m, \omega_i^l, l = \overline{1, M}, i = \overline{1, L}$, are equally normalized.

Weights' estimates $\hat{\omega}_i^m, i = \overline{1, L}$, are statistically justified, stable, and well consistent, if numbers $\rho(\hat{\omega}^m, \hat{\omega}^l), l = \overline{1, M}$, are acceptable. Here, $\hat{\omega}^m = (\omega_1^m, \dots, \omega_L^m)^T, \hat{\omega}^l = (\omega_1^l, \dots, \omega_L^l)^T, l = \overline{1, M}$, and $\rho(\hat{\omega}^m, \hat{\omega}^l)$ can be interpreted as the degree of difference between vectors $\hat{\omega}^m$ and $\hat{\omega}^l$ [3].

We proposed in [3] to use this measure:

$$\begin{aligned} \rho(\hat{\omega}^m, \hat{\omega}^l) &= \\ &= \sqrt{\sum_{i=1}^L \left(\frac{\omega_i^m}{\sqrt{\sum_{j=1}^L (\omega_j^m)^2}} - \frac{\omega_i^l}{\sqrt{\sum_{j=1}^L (\omega_j^l)^2}} \right)^2}, l = \overline{1, M}. \quad (20) \end{aligned}$$

We studied a well-conditioned empirical matrix of pairwise comparisons in [4], according to the computational procedure proposed by Saaty [5–12]. We used Model

1 to estimate the weights. It turned out, as a result of simulation, that $\rho(\hat{\omega}, \hat{\omega}^l) \in (0, 0.12)$ in 100 % of cases, where $l = \overline{1, M}$ and $\hat{\omega} = (\omega_1, \dots, \omega_L)^T$ are the weights estimated by the simulated well-conditioned matrices of pairwise comparisons.

Thus, if $\rho(\hat{\omega}^m, \hat{\omega}^l) \in (0, 0.12) \ \forall l = \overline{1, M}$, then the estimated values ω_i^m of weights $\omega_i, i = \overline{1, L}$, are stable, and the empirical matrix of pairwise comparisons Γ is equivalent to a well-conditioned empirical matrix of pairwise comparisons by the efficiency of the weights' estimation in the sense of [5–12].

Remark 8. The implementation of item 3.2 is laborious and can be omitted if the guaranteed obtaining of stable and well consistent values of $\omega_i, i = \overline{1, L}$, is not mandatory.

Conclusions. In this article, we have proposed an efficient algorithm for uncertainty resolution for a quite general class of combinatorial optimization problems under uncertainty which use expert weights. The algorithm includes efficient algorithmic procedure for expert weights' estimation based on previous results of the author and his disciples.

References

1. Pavlov A.A. Optimization for one class of combinatorial problems under uncertainty. *Адаптивні системи автоматичного управління: міжвідомчий наук.-техн. збірник*. Київ: НТУУ «КПІ», 2019. Т. 1. № 34 (in press)
2. Zgurovsky M. Z., Pavlov A. A. *Combinatorial Optimization Problems in Planning and Decision Making: Theory and Applications*. Cham (Switzerland): Springer, 2019. 526 p. Chapter 8, The four-level model of planning and decision making, pp. 347–406. doi: 10.1007/978-3-319-98977-8_8
3. Згуровский М. З., Павлов А. А. *Принятие решений в сетевых системах с ограниченными ресурсами*. Київ: Наук. думка, 2010. 573 с.
4. Бурков В. Н., Бушуев С. Д., Ввозный А. М., Гайда А. Ю., Григорян Т. Г., Иванова А. А., Кнырик Н. Р., Колесник М. Э., Кононенко И. В., Кошкин К. В., Павлов А. А., Рыжков С. С., Рыжков А. С., Слободян С. О., Танака Х., Чернов С. К. *Управление ресурсами распределенных проектов и программ / Management of the distribute projects and programmes resources*. Николаев: издатель Торубара В. В., 2015. 386 с.
5. Saaty T. L. *The Analytic Hierarchy Process*. New York: McGraw Hill, 1980
6. Saaty T. L., Kearns K. *Analytical Planning: The Organization of Systems*. Oxford: Pergamon Press, 1985. 216 p. doi: 10.1016/C2013-0-03782-6
7. Saaty T. L. How to make a decision: The analytic hierarchy process. *European Journal of Operational Research*, 1990. Vol. 48, Iss. 1, pp. 9–26. doi: 10.1016/0377-2217(90)90057-1
8. Saaty T. L. *Decision Making with Dependence and Feedback: The Analytic Network Process*. Pittsburgh: RWS Publications, 1996.
9. Saaty T. L. *Multicriteria Decision Making: The Analytic Hierarchy Process*. Pittsburgh: RWS Publications, 1996. 479 p.
10. Saaty T. L. Analytic Hierarchy Process. *Encyclopedia of Biostatistics / eds. P. Armitage, T. Colton*. New Jersey: John Wiley & Sons, 2005. doi: 10.1002/0470011815.b2a4a002
11. Saaty T. L. Decision making with the analytic hierarchy process. *International Journal of Services Sciences*. Vol. 1. No. 1, 2008. P. 83–98. doi: 10.1504/IJSSci.2008.01759
12. Андрейчиков А. В., Андрейчикова О. Н. *Анализ, синтез, планирование решений в экономике*. Москва: Финансы и статистика, 2000. 368 с.

References (transliterated)

1. Pavlov A.A. Optimization for one class of combinatorial problems under uncertainty. *Adaptyvni systemy avtomatychnoho upravlinnya: mizhvidomchyy nauk.-tekhn. zbirnyk* [Adaptive systems of automatic

- control: interdepartmental scientific-technical collection of papers]. Kiev, NTUU "KPI" Publ. Vol. 1, no. 34 (in press)
- Zgurovsky M. Z., Pavlov A. A. *Combinatorial Optimization Problems in Planning and Decision Making: Theory and Applications*. Cham, Springer, 2019. 526 p. Chapter 8. The four-level model of planning and decision making, pp. 347–406. doi: 10.1007/978-3-319-98977-8_8
 - Zgurovsky M. Z., Pavlov A. A. *Prinyatie resheniy v setevykh sistemakh s ogranichenymi resursami* [Decision making in network systems with limited resources]. Kiev, Nauk. dumka Publ., 573 p.
 - Burkov V. N., Bushuev S. D., Voznyy A. M., Gayda A. Yu., Gri-goryan T. G., Ivanova A. A., Knyrik N. R., Kolesnik M. E., Kono-nenko I. V., Koshkin K. V., Pavlov A. A., Ryzhkov S. S., Ryzhkov A. S., Slobodyan S. O., Tanaka Kh., Chernov S. K. *Upravlenie resursami raspredelennykh projektov i programm* [Management of the distribute projects and programmes resources]. Nikolaev, Torubara V. V. Publ., 2015. 386 p.
 - Saaty T. L. *The Analytic Hierarchy Process*. New York, McGraw Hil Publ., 1980.
 - Saaty T. L. Kearns K. *Analytical Planning: The Organization of Systems*. Oxford, Pergamon Press, 1985. 216 p. doi: 10.1016/C2013-0-03782-6
 - Saaty T. L. How to make a decision: The analytic hierarchy process. *European Journal of Operational Research*. 1990, vol. 48, iss. 1, pp. 9–26. doi: 10.1016/0377-2217(90)90057-1
 - Saaty T. L. *Decision Making with Dependence and Feedback: The Analytic Network Process*. Pittsburgh: RWS Publications, 1996.
 - Saaty T. L. *Multicriteria Decision Making: The Analytic Hierarchy Process*. Pittsburgh, RWS Publ., 1996. 479 p.
 - Saaty T. L. Analytic Hierarchy Process. *Encyclopedia of Biostatistics* / eds. P. Armitage and T. Colton. New Jersey, John Wiley & Sons Publ., 2005. doi: 10.1002/0470011815.b2a4a002
 - Saaty T. L. Decision making with the analytic hierarchy process. *International Journal of Services Sciences*. vol. 1, no. 1, 2008, pp. 83–98. doi: 10.1504/IJSSci.2008.01759
 - Andreychikov A. V., Andreychikova O. N. *Analiz, sintez, planirovanie resheniy v ekonomike* [Analysis, Synthesis, Decisions Planning in the Economy]. Moscow, Finansy i statistika Publ., 2000. 368 p.

Received 07.04.2019

Відомості про авторів / Сведения об авторах / About the Authors

Павлов Олександр Анатолійович (Павлов Александр Анатольевич, Pavlov Alexander Anatolievich) – доктор технічних наук, професор, завідувач кафедри автоматизованих систем обробки інформації та управління, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»; м. Київ, Україна; ORCID: <http://orcid.org/0000-0002-6524-6410>; e-mail: pavlov.fiot@gmail.com

Т. В. ХОМ'ЯК, А. В. МАЛІЄНКО, Г. В. СИМОНЕЦЬ

ЗАСТОСУВАННЯ МЕТОДІВ ЗГЛАДЖУВАННЯ ДЛЯ ПРОГНОЗУВАННЯ ОБСЯГУ ВИРОБНИЦТВА

Об'єктом дослідження є підприємство ТОВ ТПК «Терра» – провідний вітчизняний виробник високоякісних захисних покриттів промислового призначення для антикорозійного захисту металоконструкцій і залізобетонних конструкцій. Наразі на підприємстві відкриті питання планування майбутніх замовлень, доцільного використання потужностей і розширення виробничих можливостей. В роботі проведено системний аналіз цього підприємства шляхом розробки функціональної моделі процесів та її декомпозиції з розкладанням до другого рівня деталізації. В результаті отримано задачу з прогнозування обсягу випуску антикорозійних покриттів, які запобігають руйнації металевих конструкцій та механізмів. Ця задача є актуальною як для самого підприємства, так і для задоволення попиту на ринку збуту. Найбільш ефективними математичними моделями, за допомогою яких можна прогнозувати розвиток процесів виробництва – це моделі на основі часових рядів. Одними з найпоширеніших методів прогнозування показників таких рядів є методи згладжування, які використовуються для зменшення впливу випадкових коливань. Поставлену задачу вирішено методами експоненціального та адаптивного згладжування. Для перевірки адекватності отриманих моделей застосовано критерій серій (визначення випадковості відхилень від тренду), критерій піків (перевірка рівності нулю математичного сподівання), R/S-критерій (визначення відповідності розподілу залишкової компоненти нормальному закону), критерій Дарбіна-Уотсона (визначення незалежності значень залишкової компоненти). Проведено аналіз кожної отриманої моделі прогнозування, а також оцінено якість прогнозів. Зроблено висновки щодо подальшого виготовлення продукції на основі прогнозованих значень. Запропонована модель має практичну спрямованість і може використовуватися у задачах, пов'язаних з прогнозуванням в умовах виробничих підприємств.

Ключові слова: системний аналіз, задача прогнозування, експоненціальне згладжування, адаптивне згладжування, точність прогнозування, адекватність моделі.

Т. В. ХОМ'ЯК, А. В. МАЛІЄНКО, Г. В. СИМОНЕЦЬ

ПРИМЕНЕНИЕ МЕТОДОВ СГЛАЖИВАНИЯ ДЛЯ ПРОГНОЗИРОВАНИЯ ОБЪЕМА ПРОИЗВОДСТВА

Объектом исследования является предприятие ООО ТПК «Терра» – ведущий отечественный производитель высококачественных защитных покрытий промышленного назначения для антикоррозионной защиты металлоконструкций и железобетонных конструкций. Сейчас на предприятии открыты вопросы планирования будущих заказов, целесообразного использования мощностей и расширения производственных возможностей. В работе проведен системный анализ этого предприятия путем разработки функциональной модели процессов и ее декомпозиции с разложением до второго уровня детализации. В результате получено задачу прогнозирования объема выпуска антикоррозионных покрытий, предотвращающих разрушение металлических конструкции и механизмов. Эта задача является актуальной как для самого предприятия, так и для удовлетворения спроса на рынке сбыта. Наиболее эффективными математическими моделями, с помощью которых можно прогнозировать развитие процессов производства – это модели на основе временных рядов. Одними из самых распространенных методов прогнозирования показателей таких рядов являются методы сглаживания, которые используются для уменьшения влияния случайных колебаний. Поставленную задачу решено методами экспоненциального и адаптивного сглаживания. Для проверки адекватности полученных моделей применен критерий серий (определение случайности отклонений от тренда), критерий пиков (проверка равенства нулю математического ожидания), R/S-критерий (определение соответствия распределения остаточной компоненты нормальному закону), критерий Дарбина-Уотсона (определение независимости значений остаточной компоненты). Проведен анализ каждой полученной модели прогнозирования, а также оценено качество прогнозов. Сделаны выводы относительно дальнейшего изготовления продукции на основе прогнозируемых значений. Предложенная модель имеет практическую направленность и может использоваться в задачах, связанных с прогнозированием в условиях производственных предприятий.

Ключевые слова: системный анализ, задача прогнозирования, экспоненциальное сглаживание, адаптивное сглаживание, точность прогнозирования, адекватность модели.

Т. В. ХОМ'ЯК, А. В. МАЛІЄНКО, Г. В. СИМОНЕЦЬ

APPLICATION OF SMOOTHING METHODS FOR FORECASTING PRODUCTION VOLUME

The object of the research is the enterprise LLC TPK «Terra» – the leading domestic manufacturer of high-quality protective coatings for industrial use for corrosion protection of metal structures and reinforced concrete structures. Now the company is open to planning future orders, the appropriate use of capacity and the expansion of production capacity. The paper conducted a system analysis of this enterprise by developing a functional model of processes and its decomposition with decomposition to the second level of detail. As a result, the task of forecasting the volume of release of anti-corrosion coatings, preventing the destruction of metal structures and mechanisms. This task is relevant both for the enterprise itself and for meeting the demand in the sales market. The most effective mathematical models that can be used to predict the development of production processes are models based on time series. One of the most common methods for predicting the performance of such series are smoothing methods that are used to reduce the effect of random fluctuations. The problem is solved by the methods of exponential and adaptive smoothing. To check the adequacy of the models obtained, the test of the series (determination of randomness of deviations from the trend), the criterion of the peaks (checking the equality of the mathematical expectation), the R/S criterion (determining the correspondence of the distribution of the residual component to the normal law), the Durbin–Watson criterion (determining the independence of the residual components). The analysis of each prediction model obtained was carried out, and the quality of forecasts was also assessed. Conclusions are made regarding the further manufacture of products based on predicted values. The proposed model has a practical orientation and can be used in tasks related to forecasting in the conditions of industrial enterprises.

Keywords: system analysis, forecasting problem, exponential smoothing, adaptive smoothing, forecasting accuracy, model adequacy.

Вступ. У наш час рухливість економічних явищ та процесів, особливо на рівні галузей та підприємств, зростає, і накопичуються дані статистичних спостережень. На базі цих даних задача прогнозування

розвитку процесу у системі в майбутньому є провідною.

Актуальність теми дослідження впливає з потреби підприємств до розвитку. Для того, щоб бути

конкурентоспроможним підприємством потрібно удосконалювати процеси у відповідності до часу. Наразі існує велика кількість інструментів для оптимізації процесів – одними з них є системний аналіз. Керівництвом підприємства поставлено задачу аналізу роботи підприємства, в результаті якого було отримано задачу з прогнозування виготовлення антикорозійних покриттів – основної продукції підприємства, задля подальшого планування і оптимізації процесів.

На сьогоднішній день найбільш ефективними математичними моделями, за допомогою яких можна прогнозувати розвиток процесів, є ті, під час побудови яких використовуються часові ряди. Широке коло соціально-економічних, технічних і природних процесів часто зображають набором послідовних значень, що зафіксовані в рівновіддалені моменти часу. Такий набір значень називають часовим рядом.

Одними з найпоширенішими методами прогнозування показників часових рядів є методи згладжування. Суть різних прийомів згладжування часових рядів зводиться до заміни фактичних рівнів часового ряду розрахунковими, які в меншій мірі схильні до коливань. Методи згладжування використовуються для зменшення впливу випадкового компонента (випадкових коливань) у часових рядах. Вони дають можливість отримувати більш «чисті» значення, які складаються лише з детермінованих компонентів.

Метою роботи є прогнозування обсягу виготовлення продукції на подальший період.

Об'єктом дослідження в роботі є підприємство ТОВ ТПК «Терра», що займається виготовленням антикорозійних покриттів.

Предметом дослідження є методи системного аналізу, прогнозування рівнів часових рядів і системи з критеріїв для оцінки результатів в процесі прогнозування обсягів виробництва.

Методи дослідження. Для вирішення задачі спочатку проведено системний аналіз та складено функціональну модель з розкладанням до другого рівня деталізації. При проведенні системного аналізу отримано задачу з прогнозування обсягу випуску продукції, яку вирішено за допомогою методів згладжування: експоненціального та адаптивного. Проведено аналіз результатів з оцінкою методів і прийнято рішення про використання одного з них на практиці.

Наукова новизна отриманих результатів полягає у використанні системного підходу для аналізу задачі в цілому та розбиття її на підзадачі. Зазначені методи прогнозування перевірено на надійність.

Результатом роботи є розроблений комплекс системи для планування на підприємстві обсягів виробництва, рекомендовано до впровадження у задачах, пов'язаних з прогнозуванням в умовах виробничих підприємств.

Опис об'єкту. Компанія «Терра» заснована в 1998 році і наразі є провідним вітчизняним виробником високоякісних захисних покриттів промислового при-

значення для антикорозійного захисту металлоконструкцій і залізобетонних конструкцій NANO Industrial use™. Компанія налічує більш ніж 1000 найменувань продукції. Данні про замовлення, виготовлення, продаж кожного з них зберігаються протягом декілька років. Їх доцільно представляти у вигляді часових рядів, тобто наборів послідовних значень, зафіксованих в рівновіддалені моменти часу.

Наразі на підприємстві відкрито питання планування задля готовності до майбутніх замовлень, доцільного використання потужностей і розширення виробничих можливостей, як результат виконання цілей і місії компанії.

Функціональна модель системи. Для розуміння процесів на підприємстві побудовано модель в форматі IDEF0 за допомогою AllFusion Process Modeler [1, 2]. Спочатку функціональність підприємства описується в цілому, без подробиць. Такий опис називається контекстною діаграмою [3]. Для даного процесу контекстна діаграма представлена моделлю (див. рис. 1), де I_1 – сировина, I_2 – заявка на виготовлення, C_1 – закони України, C_2 – документація з техніки безпеки, C_3 – нормативна документація, M_1 – персонал, M_2 – технічне оснащення, O_1 – фарба, O_2 – покриття поверхні.

На контекстній діаграмі початкового рівня (див. рис. 1) показано, які фактори необхідні на вході (сировина та замовлення).

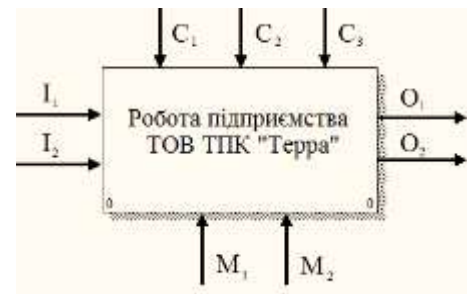


Рис. 1. Модель процесу роботи підприємства

Процес «робота підприємства» виконується під керівництвом нормативних документів, законів України та документації з техніки безпеки. Замовлення опрацьовуються, закуповується сировина для подальшого виготовлення та покриття поверхонь за участю будівельних бригад. Якщо всі умови виконуються, то на виході отримують результат у вигляді готової продукції або покритих на замовлення конструкцій.

Наступним рівнем функціональної моделі є декомпозиція контекстної діаграми (див. рис. 2) на такі підфункції:

- виробництво покриттів;
- будівельні роботи з покриття.

Функція «виробництво покриттів» (див. рис. 3) декомпозована на підфункції:

- формування замовлення;
- закупка сировини;
- виготовлення;
- відвантаження.

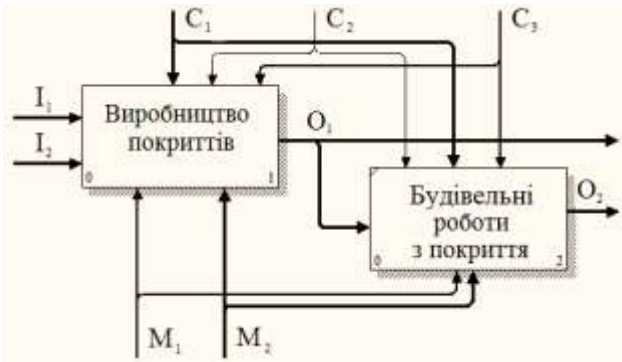


Рис. 2. Декомпозиція моделі

Після проведеної декомпозиції процесу «виробництво покриттів», бачимо, що даний процес є доволі трудомісткий особливо виходячи з того, що в процесі виготовлення задіяна велика кількість різного виду сировини та працівників.

Оскільки процес виробництва покриттів є доволі складним і є ядром підприємства – його основним джерелом прибутку, було прийнято рішення про доцільність проведення прогнозування обсягів виробництва для подальшого виготовлення. Прогнозування з подальшим плануванням на базі прогнозованих значень є важливою складовою в роботі підприємства адже допомагає зекономити ресурси і разом з тим збільшити прибуток.

Отже, для даної задачі проведено системний аналіз підприємства для конкретизації кожного етапу [4], на основі якого обрано методи прогнозування значень.

Експоненціальне згладжування. Це метод математичного перетворення, який застосовується при прогнозуванні часових рядів. Свою назву він отримав через те, що при кожній наступній ітерації враховуються всі попередні значення ряду, але ступінь врахування зменшується за експоненціальним законом [5, 6].

За цим методом нова послідовність будується за правилом:

$$S_t = \lambda y_t + (1 - \lambda)S_{t-1}, \quad S_1 = y_1, \quad t = \overline{2, T}, \quad (1)$$

де T – кількість наведених спостережень;
 λ – єдина вага, яка може обиратися кількома шляхами.

По-перше, значення λ обирається апріорним методом з проміжку (0;1). В такому випадку, якщо обирається значення близьке до 1, то будуть більш важливими при прогнозуванні останні дані часового ряду, при виборі λ близьким до 0, більш впливовими будуть минулі значення.

По-друге, можна покласти $\lambda = 2/(1 + T)$.

По-третє, λ обирається так, щоб мінімізувати один з критеріїв точності прогнозів.

Прогноз значень часового ряду дорівнює останньому члену послідовності:

$$\hat{y}_{t+p} = S_t, \quad p = 1, 2, \dots, \quad (2)$$

де p – рівень прогнозування.

Адаптивне згладжування. Цей метод базується на попередньому і дозволяє автоматично змінювати константу згладжування в процесі обчислень [7, 8, 9]. Нова послідовність будується за правилом:

$$S_{t+1} = \lambda_t y_{t+1} + (1 - \lambda_t)S_t, \quad t = \overline{1, T}, \quad (3)$$

де λ змінюється з часом в залежності від похибки прогнозування:

$$\lambda_t = \left| \frac{E_t}{M_t} \right|,$$

$$E_t = \beta(y_t - \hat{y}_t) + (1 - \beta)E_{t-1},$$

$$M_t = \beta|y_t - \hat{y}_t| + (1 - \beta)M_{t-1},$$

де $\beta \in (0; 1)$ – обирається апріорі.

Прогноз значень часового ряду дорівнює останньому члену послідовності (2).

Вирішення задачі прогнозування. В ході дослідження діяльності підприємства проаналізовано показники продажу одного з видів покриття за останні 2,5 роки.

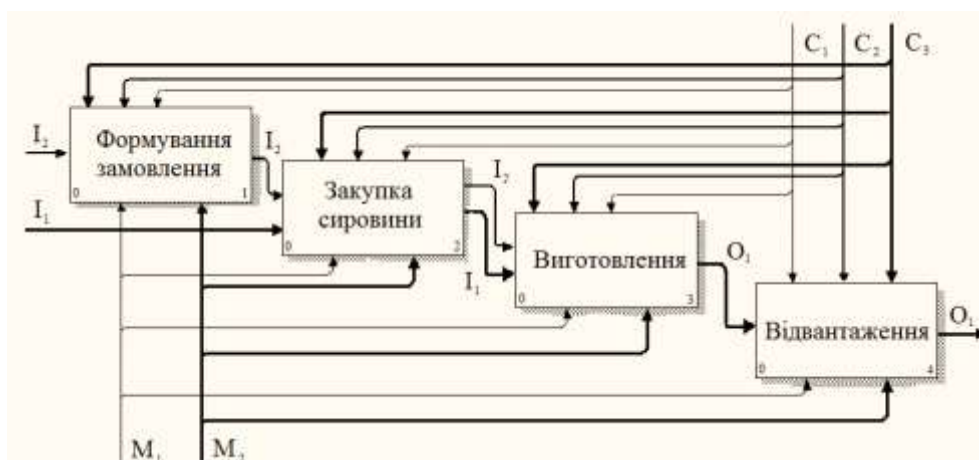


Рис. 3. Декомпозиція процесу «Виробництво покриттів»

В роботі для експоненціального згладжування обрано λ таке, при якому мінімізується середньоквадратична помилка прогнозу, що складає 10%, а прогнозне значення – 50 одиниць товару.

Результати експоненціального згладжування зображено на рис.4.



Рис. 4. Діаграма залежності фактичних та розрахункових даних за експоненціальним методом

Результати адаптивного згладжування представлено на рис.5.



Рис. 5. Діаграма залежності фактичних та розрахункових даних за адаптивним методом

Середньоквадратична похибка прогнозу за даним методом складає 11%, а прогнозне значення – 52 одиниці товару.

Оцінка якості прогнозів. Для будь-якої моделі питання про можливість її застосування для аналізу та прогнозування може бути вирішено після встановлення адекватності, тобто відповідності досліджуваному процесу чи об'єкту [10, 11].

Для перевірки адекватності моделей застосовано наступні критерії:

- критерій серій – визначення випадковості відхилень від тренду;
- критерій піків – перевірка рівності нулю математичного сподівання;
- R/S-критерій – визначення відповідності розподілу залишкової компоненти нормальному закону;

- критерій Дарбіна – Уотсона – визначення незалежності значень залишкової компоненти [12].

У табл. 1 наведено результати перевірки моделей на адекватність для екстраполяції кількості виготовлення продукції.

Таблиця 1 – Результати перевірки на адекватність

| Критерій | Умова адекватності | Значення постійних показників | Експоненціальна модель | | Адаптивна модель | |
|--|---------------------------|----------------------------------|---------------------------|-------------|---------------------------|-------------|
| | | | адекватна | неадекватна | адекватна | неадекватна |
| Критерій серій | $k_{max} < k, v > v_1$ | $k = 8,19$ $v_1 = 10,22$ | $k_{max} = 6$ $v = 12$ | адекватна | $k_{max} = 5$ $v = 13$ | адекватна |
| Критерій піків | $M - S \leq p \leq M + S$ | $M = 18,67$ $S = 2,24$ | $p = 14$ | адекватна | $p = 13$ | адекватна |
| Рівність математичного сподівання залишкової компоненти нулю | $t_{роз} < t_{крит}$ | $t_{крит} = 2,042$ | $t_{роз} = 1,3$ | $\mu = 0$ | $t_{роз} = 1,239$ | $\mu = 0$ |
| Критерій Дарбіна-Уотсона | $d^u < d < 4 - d^u$ | $d^u = 1,26$ $4 - d^u = 2,74$ | $d = 1,4$ | адекватна | $d = 2,16$ | адекватна |
| Висновки | | | адекватна | | адекватна | |

Обидві моделі за отриманими результатами можна вважати адекватними і застосовувати для зазначених даних. Тому перейдемо до перерахунку надійності прогнозу задля вибору одного з методів прогнозування для застосування в ТОВ ТПК «Терра».

При визначенні точності прогнозу розглянуто основні характеристики моделей прогнозування для екстраполяції, які наведено у табл. 2.

Таблиця 2 – Результати визначення надійності прогнозів

| Модель | | Експоненціальна | Адаптивна |
|-----------------------------|----------------|-----------------|-----------|
| | | | |
| Точність прогнозу | r | 78% | 73% |
| Середньоквадратична похибка | S | 10% | 11% |
| Середня відносна похибка | $\delta_{сеп}$ | 22% | 27% |
| Коефіцієнт розбіжності | v | 0,106 | 0,114 |
| Коефіцієнт кореляції | R | 0,989 | 0,987 |
| Критерій Фішера | $F_{роз}$ | 1095,99 | 944,30 |
| | $F_{крит}$ | 7,64 | 7,64 |
| Критерій Стюдента | $t_{роз}$ | 36,06 | 32,97 |
| | $t_{крит}$ | 2,76 | 2,76 |

Отже, при аналізі показників рівня виготовленої продукції доцільно використовувати експоненціальну модель. Це підтверджено критеріями адекватності моделі фактичним даним та розрахованими критеріями надійності моделі.

Точність експоненціальної моделі приймає значення 78,3%. Цей показник є вищим, ніж у адаптивної моделі. Середньоквадратична похибка складає 10%. Значення коефіцієнта кореляції 0,989, що свідчить про високий лінійний зв'язок між ознаками. Коефіцієнт розбіжності (Тейла) наближений до 0, що свідчить про наближення прогнозу до ідеального. Стосовно критеріїв Фішера та Стюдента: розрахункові значення критеріїв перевищують критичні (табличні) значення, а це підтверджує статистичну значущість моделі [10].

Таким чином, для прогнозування розвитку підприємства необхідно використати метод експоненціального згладжування. Для більш наочного представлення реальних та розрахункових даних надано діаграми (див. рис. 4, 5).

Висновки. Для структурування та виділення конкретних етапів роботи підприємства проведено системний аналіз. Це дозволило розділити подане питання на дві основні підзадачі та виділити одну з них для подальшого опрацювання. Виявленою задачею після декомпозиції моделі стало прогнозування обсягів виробництва. Цю задачу вирішено за допомогою методів згладжування часових рядів, а саме методами експоненціального та адаптивного згладжування. Проведено аналіз результатів прогнозування і обрано експоненціальну модель для подальшого прогнозування на виробництві. У результаті отримано рекомендації щодо обсягу виробництва у наступному місяці в розмірі 50 одиниць «ЕП Мастик».

Отже, завдяки виконаній роботі відтепер можна якісно прогнозувати обсяги виробництва на підприємстві ТОВ ТПК «Терра», асортимент якого перевищує дві сотні найменувань товарів. Розроблений комплекс оцінки якості результатів дозволить в подальшому виходити з характеру часового ряду обсягу виробництва різних товарів і використовувати найбільш якісну модель прогнозування.

Список літератури

1. Маклаков С. В. *BPwin i ERwin. CASE-засоби розробки інформаційних систем*. Москва: Наука, 1999. 163 с.
2. Марка Д. А., МакГоуен К. *Методологія структурного аналізу і проектування SADT*. Москва: Наука, 1993. 231 с.
3. Хом'як Т. В., Суїма І. О. Розробка системи підтримки і прийняття рішень при відкритті закладу ресторанного господарства. *Вісник Нац. техн. ун-та «ХПІ»: зб. наук. пр. Темат. вип.: Системний аналіз, управління та інформаційні технології*. Харків: НТУ «ХПІ», 2017. № 51 (1272). С. 65–71.
4. Згуровский М. З., Панкратова Н. Д. *Системный анализ: проблемы, методология, приложения*. Київ: Наукова думка, 2005. 742 с.
5. Ставицький А. В. *Навчально-методичний комплекс з курсів «Прогнозування» та «Фінансове прогнозування»*. Київ, 2006. 107 с.
6. Enders W. *Applied econometric time series*. New York: John Wiley & Sons, Inc., 1995. 434 p.
7. Оліскевич М. О. *Основи економетрії часових рядів. Навчальний посібник*. Видавничий центр ЛНУ ім. І. Франка, 2008. 321 с.
8. Давнис В. В. *Адаптивные модели: анализ и прогноз в экономических системах*. Воронеж: Воронеж. гос. ун-т, 2006. 380 с.
9. Лукашин Ю. П. *Адаптивные методы краткосрочного прогнозирования временных рядов: учебное пособие*. Москва: Финансы и статистика, 2003. 416 с.

10. Гмурман В. Е. *Теория вероятности и математическая статистика: навчальний посібник*. Москва: Вища освіта, 2008. 479 с.
11. Черняк О. І., Ставицький А. В. *Динамічна економетрика: Навч. посібник*. Київ: Знання, 2000. 199 с.
12. Бокс Дж., Дженкинс Г. *Анализ временных рядов. Прогноз и управление*. Москва: Мир, 1974. 406 с.

References (transliterated)

1. Maklakov S. V. *BPwin i Erwin. CASE-zasobi rozrobki informatsiynih system* [BPwin and Erwin. CASE-tools for developing information systems]. Moscow, Nauka Publ., 1999. 163 p.
2. Marka D. A., McGowan K. *Metodologiya strukturnogo analizu i proektuvannya SADT* [Methodology of Structural Analysis and Designing of SADT]. Moscow, Nauka Publ., 1993. 231 p.
3. Khomyak T. V., Suima I. O. *Rozrobka sistemi pidtrimki i priynyattya rishen pri vidkritti zakladu restorannogo gospodarstva* [Development of the system of support and decision-making at the opening of the restaurant establishment]. *Vestnik Nats. tekhn. un-ta "KhPI": sb. nauch. tr. Temat. vyp.: Sistemnyy analiz, upravlenie i informatsionnye tekhnologii* [Bulletin of the National Technical University "KhPI": a collection of scientific papers. Thematic issue: System analysis, management and information technology]. Kharkov, NTU "KhPI" Publ., 2017, no. 51 (1272). pp. 49–71.
4. Zgurovskiy M. Z., Pankratova N. D. *Sistemnyy analiz: problemi, metodologiya, prilozheniya* [System analysis: problems, methodology, applications]. Kiev, Nauk. dumka Publ., 2005. 742 p.
5. Stavitsky A. V. *Navchalno-metodichniy kompleks z kursiv «Prognozuvannya» ta «Finsove prognovuvannya»* [Educational and methodological complex of courses «Forecasting» and «Financial forecasting»]. Kiev, 2006. 107 p.
6. Enders W. *Applied econometric time series*. New York: John Wiley & Sons, Inc., 1995. 434 p.
7. Olskevich M. O. *Osnovi ekonometriyi chasovih ryadiv. Navchalnyi posibnik* [Fundamentals of econometrics of time series. Tutorial]. Publishing Center of LNU them. I. Franko, 2008. 321 p.
8. Davnis V. V. *Adaptivnye modeli: analiz i prognos ekonomicheskikh sistemah* [Adaptive models: analysis and forecast in economic systems]. Voronezh: Voronezh. State univ Publ., 2006. 380 p.
9. Lukashyn U. P. *Adaptivnye metody kratkosrochnogo prognovirovaniya vremennih ryadov: uchebnoe posobie* [Adaptive methods for short-term time series forecasting: study guide]. Moscow: Finance and Statistics Publ., 2003. 416 p.
10. Gmurman V. E. *Teoriya ymovirnosti i matematichna statistika: navchalnyi posibnik* [Probability Theory and Mathematical Statistics: Tutorial]. 12th edition. Moscow: Higher education Publ., 2008. 479 p.
11. Chernyak O. I., Stavitsky A. V. *Dinamichna ekonometrika: navch. posibnyk* [Dynamic ekonometrika: study guide]. Kiev: Znannya Publ., 2000. 199 p.
12. Boks J., Dgenkins G. *Analiz vremennih ryadov. Prognoz i upravlenie* [Time series analysis. Forecast and management]. Moscow: Mir Publ., 1974. 406 p.

Надійшла (received) 16.05.2019

Відомості про авторів / Сведения об авторах / About the Authors

Хом'як Тетяна Валеріївна (Хом'як Татьяна Валериевна, Khomyak Tatiana Valeriivna) – кандидат фізико-математичних наук, доцент, Національний технічний університет «Дніпровська політехніка», доцент кафедри системного аналізу та управління, м. Дніпро, Україна; ORCID: <https://orcid.org/0000-0002-6177-2827>; e-mail: khomyak.t.v@gmail.com

Малиєнко Андрій Вікторович (Малиенко Андрей Викторович, Malienko Andrey Viktorovich) – кандидат технічних наук, Національний технічний університет «Дніпровська політехніка», доцент кафедри системного аналізу та управління, м. Дніпро, Україна; ORCID: <https://orcid.org/0000-0002-3165-9233>; e-mail: andrei.malienko@gmail.com

Симонець Галина Василівна (Симонець Галина Васильевна, Simonets Halyna Vasyliivna) – Національний технічний університет «Дніпровська політехніка», студент, м. Дніпро, Україна; ORCID: <https://orcid.org/0000-0002-1322-7189>; e-mail: galya.golovko.2014@gmail.com

В. В. СМАЧИЛО**НОРМАТИВНО-ПРАВОВЕ ЗАБЕЗПЕЧЕННЯ ДІЯЛЬНОСТІ У СФЕРІ ТРАНСПОРТУ**

Економічна категорія «підприємство транспорту» є загальноживаним в економічному науковому та практичному обігу. Але, в той же час, її розуміння в правовому полі, яке визначає основи господарської діяльності в сфері транспорту, не є досить чітким. Вказане породжує проблему ідентифікації та віднесення до тієї чи іншої категорії суб'єктів господарювання транспортної галузі. Нормативно-правове забезпечення діяльності транспорту України базується на сукупності загальних законів та кодексів, які регулюють господарську діяльність будь-яких сфер національної економіки країни, а також на спеціальних нормативно-правових актах, які враховують специфіку транспортної сфери. Незважаючи на широке правове поле забезпечення діяльності транспорту, сучасні трансформації та виклики, що стоять перед галуззю, вимагають вирішення певних неузгодженостей задля коректного сприйняття понятійно-категоріального апарату, що формується в правовому полі, та його розвитку у відповідності з потребами часу. Метою дослідження є ідентифікація суб'єктів в сфері транспорту на основі законодавчо-правового підходу. Рекомендовано ввести в нормативно-правове поле визначення транспортного підприємства (підприємства транспорту), яке б корелювало з його дефініцією в Господарському Кодексі України. Пропонуємо розкрити термін «транспортне підприємство» як самостійний суб'єкт господарювання, створений компетентним органом державної влади або органом місцевого самоврядування, або іншими суб'єктами для задоволення суспільних та особистих потреб шляхом систематичного здійснення господарської діяльності в сфері транспорту, відповідно до Закону України «Про транспорт», в порядку, передбаченому Господарським Кодексом та іншими законами України. Чітке окреслення понятійно-категоріального апарату відносно сутності підприємств транспорту в законодавчому полі та його поєднання із класифікацією видів економічної діяльності дозволить в майбутньому коректно формувати сукупність для аналітичної вибірки та здійснювати подальшу класифікацію.

Ключові слова: підприємство, транспорт, транспортні підприємства, вид економічної діяльності, види транспорту, транспортна система.

В. В. СМАЧИЛО**НОРМАТИВНО-ПРАВОВОЕ ОБЕСПЕЧЕНИЕ ДЕЯТЕЛЬНОСТИ В СФЕРЕ ТРАНСПОРТА**

Экономическая категория «предприятие транспорта» является общепотребительной в экономическом научном и практическом обращении. Но, в то же время, ее понимание в правовом поле, которое определяет основы хозяйственной деятельности в сфере транспорта, не является достаточно четким. Указанное порождает проблему идентификации и отнесения к той или иной категории субъектов хозяйствования транспортной отрасли. Нормативно-правовое обеспечение деятельности транспорта Украины базируется на совокупности общих законов и кодексов, регулирующих хозяйственную деятельность любых сфер национальной экономики страны, а также на специальных нормативно-правовых актах, которые учитывают специфику транспортной сферы. Несмотря на широкое правовое поле обеспечения деятельности транспорта, современные трансформации и вызовы, стоящие перед отраслью, требуют решения определенных неточностей для корректного восприятия понятийно-категориального аппарата, который формируется в правовом поле, и его развития в соответствии с потребностями времени. Целью исследования является идентификация субъектов в сфере транспорта на основе законодательно-правового подхода. Рекомендуется ввести в нормативно-правовое поле определения транспортного предприятия (предприятия транспорта), которое бы коррелировало с его дефиницией в Хозяйственном кодексе Украины. Предлагаем раскрыть термин «транспортное предприятие» как самостоятельный субъект хозяйствования, созданный компетентным органом государственной власти или органом местного самоуправления, или другими субъектами для удовлетворения общественных и личных потребностей путем систематического осуществления хозяйственной деятельности в сфере транспорта, в соответствии с Законом Украины «О транспорте», в порядке, предусмотренном Хозяйственным Кодексом и другими законами Украины. Четкое определение понятийно-категориального аппарата относительно сущности предприятий транспорта в законодательном поле и его сочетание с классификацией видов экономической деятельности позволит в будущем корректно формировать совокупность для аналитической выборки и осуществлять дальнейшую классификацию.

Ключевые слова: предприятие, транспорт, транспортные предприятия, вид экономической деятельности, виды транспорта, транспортная система.

V. V. SMACHYLO**NORMATIVE-LEGAL FRAMEWORK OF TRANSPORT**

The economic category «transport enterprise» is commonly used in economic scientific and practical circulation. However, at the same time, its understanding in the legal field, which defines the basis of economic activity in the field of transport, is not very clear. The above-mentioned raises the problem of identification and attribution to one or another category of business entities in the transport sector. Regulatory support of the transport of Ukraine is based on a set of general laws and codes that regulate the economic activity of any areas of the national economy of the country, as well as on special regulatory acts that take into account the specifics of the transport sphere. Despite the broad legal framework for ensuring transport activity, modern transformations and challenges facing the industry require the resolution of certain inconsistencies in order to properly perceive the conceptual-categorical apparatus that is formed in the legal field and its development in accordance with the needs of the time. The purpose of the study is to identify entities in the field of transport on the basis of legislative and legal approach. It is recommended to introduce in the legal and regulatory framework the definition of a transport undertaking (transport company), which would correlate with its definition in the Commercial Code of Ukraine. We propose to disclose the term «transport enterprise» as an independent economic entity created by the competent state authority or local self-government body or other entities to meet social and personal needs through the systematic implementation of economic activities in the field of transport in accordance with the Law of Ukraine «On Transport» in accordance with the procedure provided by the Commercial Code and other laws of Ukraine. A clear definition of the conceptual-categorical apparatus relative to the essence of transport enterprises in the legislative field and its combination with the classification of economic activities will allow in the future to correctly form the population for the analytical sample and carry out further classification.

Keywords: enterprise, transport, transport enterprises, type of economic activity, types of transport, transport system.

Вступ. Економічна категорія «підприємство транспорту» є загальноживаним в економічному науковому та практичному обігу. Але, в той же час, її розуміння в правовому полі, яке визначає основи господарської діяльності в сфері транспорту, не є досить чітким.

Вказане породжує проблему ідентифікації та віднесення до тієї чи іншої категорії суб'єктів господарювання транспортної галузі.

Аналіз основних досягнень і літератури. Нормативно-правове забезпечення діяльності

транспорт України базується на сукупності загальних законів та кодексів, які регулюють господарську діяльність будь-яких сфер національної економіки країни: Господарський Кодекс України [1], Податковий Кодекс України, Цивільний Кодекс України, Кодексу законів про працю в Україні, Закон України «Про оплату праці», Закон України «Про колективні договори та угоди», тощо. Спеціалізована правова база представлена Законом України «Про транспорт» [2], Законом України «Про залізничний транспорт» [3], Законом України «Про автомобільний транспорт» [4] та іншими. Саме вказані документи стали основою вивчення нормативно-правового забезпечення діяльності суб'єктів господарювання транспортної галузі. Окрім того, досліджувалися публікації вітчизняних вчених [5,6], де представлено аналіз існуючих законодавчих положень та неузгоджень в сфері економіки транспорту.

Постановка задачі (проблеми). Незважаючи на широке правове поле забезпечення діяльності транспорту, сучасні трансформації та виклики, що стоять перед галуззю, вимагають вирішення певних неузгодженостей задля коректного сприйняття понятійно-категоріального апарату, що формується в правовому полі, та його розвитку у відповідності з потребами часу.

Метою статті є ідентифікація суб'єктів в сфері транспорту на основі законодавчо-правового підходу.

Результати. Транспорт визначається однією з найважливіших галузей суспільного виробництва, призначенням якої є задоволення потреби населення та суспільного виробництва в перевезеннях [2].

Транспортна галузь, займає важливе місце в економіці України, так, виходячи із структури валової доданої вартості за I кв. 2018р., її частка становила 8,2% (для порівняння: будівництво – лише 2,2%,

сільське господарство – 4,0%, хоча нижче, ніж в промисловості – 27,7% та торгівлі – 17,5%) [7]. Але важливість транспортного комплексу, на думку автора, визначається не лише його прямим позитивним економічним впливом на ВВП, але й опосередкованим, адже розвинена транспортна інфраструктура спричиняє мультиплікативний ефект у вигляді підвищення доступності інвесторів, туристів та власних громадян до будь-якого місця України. А це, в свою чергу, обумовлює створення логістичних хабів, нових робочих місць, поживає виробництво та торгівлю, туризм, тим самим підвищуючи загальне ВВП та ВРП. Державне управління в галузі транспорту здійснюють центральний орган виконавчої влади, що забезпечує формування та реалізує державну політику у сфері транспорту, дорожнього господарства, туризму та інфраструктури (Міністерство інфраструктури України), місцеві Ради народних депутатів та інші спеціально уповноважені на те органи відповідно до їх компетенції [2]. Склад єдиної транспортної системи визначається Законом України «Про транспорт» (рис. 1). Окрім нього, діяльність в сфері транспорту регулює цілий ряд нормативно-правових актів (Закони України «Про міський електротранспорт» [8], «Про трубопровідний транспорт» [9], «Про автомобільний транспорт» [4], «Про залізничний транспорт» [3], Повітряний та Водний Кодекси [10,11] та інші), а також передбачено ліцензування цього виду діяльності. Так, Ліцензійні умови провадження господарської діяльності з перевезення пасажирів, небезпечних вантажів та небезпечних відходів автомобільним транспортом, міжнародних перевезень пасажирів та вантажів автомобільним транспортом, затверджено постановою Кабінету Міністрів України від 02.12.2015 № 1001 [13]; Порядок здійснення державного контролю на автомобільному транспорті, затверджено



Рис. 1. Склад Єдиної транспортної системи України (побудовано на основі [2–4, 2–12, 15–19])

постановою Кабінету Міністрів України від 08.11.2006 № 1567 [14].

Основою господарської діяльності підприємств транспорту визначено перевезення пасажирів, вантажів, багажу та пошти, надання інших транспортних послуг, експлуатація і ремонт шляхів сполучення здійснюються залізницями, пароплавствами, суб'єктами господарювання у морських портах, автомобільними, авіаційними, дорожніми підприємствами та організаціями, якщо це передбачено їх статутами [2].

В той же час, безпосередньо надання транспортних послуг, як економічної сутнісної характеристики транспортних підприємств, відбувається в межах КВЕД 2010 «Транспорт, складське господарство, поштова та кур'єрська діяльність» секція Н та включає [19]: забезпечення пасажирських і вантажних перевезень (за розкладом або ні), залізничним, трубопровідним, автомобільним, водним або повітряним транспортом, а також допоміжну діяльність, таку як діяльність терміналів і автостоянок, вантажно-розвантажувальні роботи, зберігання тощо. Ця секція також включає надання в оренду транспортних засобів разом із водієм або екіпажем. Також до секції віднесена поштова та кур'єрська діяльність. В межах цієї секції КВЕД здійснюється розподіл за розділами [19]: 49 – наземний і трубопровідний транспорт; 50 – водний транспорт; 51 – авіаційний транспорт; 52 – складське господарство та допоміжна діяльність у сфері транспорту; 53 – поштова та кур'єрська діяльність. В межах кожного розділу виділяються окремі підвиди. На рис. 1 відображено основну законодавчу базу діяльності транспорту за рівнями транспортної системи та в розрізі видів транспорту із врахуванням шифрів основної економічної діяльності згідно КВЕД:2010.

Розкриємо визначення кожного виду транспорту загального користування, керуючись нормативно-правовою базою України (табл. 1).

Можемо відмітити, що діяльність багатьох видів транспорту загального користування регулюється окремими законодавчими актами. В той же час, не всі існуючі визначення видів транспорту в різних документах корелюють між собою, що викликає певну неусталеність нормативно-правової бази сфери транспорту та її понятійно-категоріального апарату. Якщо в ЗУ «Про транспорт», до того чи іншого виду транспорту віднесено не лише підприємства, що забезпечують транспортні послуги, а й обслуговуючі, то в спеціалізованих законах – передбачено лише перевезення.

Крім того, складові єдиної транспортної системи, що не входять до транспорту загального користування, не знайшли свого визначення в ЗУ «Про транспорт» [2], але представлені в інших законодавчих актах (діяльність трубопровідного транспорту регулюється окремим законом «Про трубопровідний транспорт» [9]). А визначення промислового залізничного транспорту наведено в Законі України «Про залізничний транспорт», та трактується як транспортно-технологічний комплекс, який забезпечує системне переміщення вантажів у процесі виробництва (між виробництвами, виробничими циклами, окремими операціями або підприємств-

вами в цілому) та взаємодію із залізничним транспортом загального користування і не належить до нього [3].

В Законі [2] відсутні визначення таких складових єдиної транспортної системи як «відомчий транспорт» – лише вказано, що до його складу входять транспортні засоби підприємств, установ та організацій (відповідно, діяльність відомчого транспорту буде здійснюватися в межах всієї секції Н за КВЕД 2010) та «шляхи сполучення загального користування».

Також слід відмітити наявність Проекту Закону України «Про внутрішній водний транспорт» [21], де визначено внутрішній водний транспорт як «вид транспорту, до якого належать судна, суб'єкти господарювання, що здійснюють та забезпечують перевезення переважно річковими водними шляхами вантажів, пасажирів, багажу і пошти, надають інші послуги з використанням суден внутрішнього плавання, забезпечують судноплавство по річковим водним шляхам, утримують об'єкти інфраструктури внутрішніх водних шляхів, здійснюють навігаційно-гідрографічне забезпечення судноплавства».

Зауважимо, що словосполучення водний транспорт не означене в ЗУ «Про транспорт», що знову ж таки викликає певну неусталеність у дефініціях. У той же час, таке трактування відповідає КВЕД 2010 (рис. 1). Отже, законодавча дефініція видів транспорту є неусталеною та неоднозначною в різних законодавчих актах, що потребує її упорядкування.

В ЗУ «Про транспорт» вказано за видами транспорту, що це підприємства, які не тільки надають транспортні послуги, а й виконують інші види діяльності для забезпечення транспортування (навчання, будівництво, ремонт, тощо).

Здійснимо співвіднесення складових транспортної системи та видів економічної діяльності більш детально, розкриваючи окремі розділи (рис. 2) поклавши в основу економічну сутність транспортних підприємств – ті, які надають транспортні послуги.

Відмічаємо, що в основу поділу за видами економічної діяльності частково покладено види транспорту та типи перевезень. Слід наголосити на чіткій відповідності видів економічної діяльності та класифікації діяльності за авіаційним транспортом згідно [2]; морський та річковий види транспорту представлені окремим розділом – водний транспорт, де виділені морський та річковий види транспорту за типами перевезень – пасажирські та вантажні (аналогічно й для авіаційного). Інші види транспорту загального користування, промисловий залізничний та трубопровідний транспорт узагальнено в розділі 49, де відсутній чіткий поділ за видами транспорту та типами перевезень.

Можемо відмітити, що відсутній чіткий код для такого виду діяльності як «пасажирські перевезення автомобільним транспортом», в той час, як виділяється «вантажний транспорт». Саме група 49.3 розподіляється на пасажирський наземний транспорт міського та приміського сполучення, надання послуг таксі та інший пасажирський наземний транспорт, н.в.і.у. [19].

Таблиця 1 – Ідентифікація видів транспорту в правовому полі України (узагальнено автором)

| Вид транспорту | Закон України «Про транспорт» [2] | Інші законодавчі акти |
|--------------------------|---|---|
| Залізничний транспорт | підприємства, що здійснюють перевезення пасажирів, вантажів, багажу, пошти, рухомий склад залізничного транспорту, залізничні шляхи сполучення, а також промислові, будівельні, торговельні та постачальницькі підприємства, навчальні заклади, технічні школи, дитячі дошкільні заклади, заклади охорони здоров'я, фізичної культури та спорту, культури, науково-дослідні, проектно-конструкторські організації, підприємства промислового залізничного транспорту та інші підприємства, установи та організації незалежно від форм власності, що забезпечують його діяльність і розвиток [2]. | виробничо-технологічний комплекс підприємств залізничного транспорту, призначений для забезпечення потреб суспільного виробництва і населення країни в перевезеннях у внутрішньому і міжнародному сполученнях та надання інших транспортних послуг усім споживачам без обмежень за ознаками форми власності та видів діяльності тощо [3]. |
| Морський транспорт | підприємства, що здійснюють перевезення пасажирів, вантажів, багажу, пошти, судна, судноремонтні заводи, морські шляхи сполучення, а також підприємства зв'язку, промислові, торговельні, будівельні і постачальницькі підприємства, навчальні заклади, заклади охорони здоров'я, фізичної культури, науково-дослідні, проектно-конструкторські організації та інші підприємства, установи та організації незалежно від форм власності, що забезпечують роботу морського транспорту [2]. | |
| Річковий транспорт | підприємства, що здійснюють перевезення пасажирів, вантажів, багажу, пошти, порти і пристані, судна, суднобудівно-судноремонтні заводи, ремонтно-експлуатаційні бази, підприємства шляхового господарства, а також підприємства зв'язку, промислові, торговельні, будівельні та постачальницькі підприємства, навчальні заклади, заклади охорони здоров'я, фізичної культури та спорту, культури, проектно-конструкторські організації та інші підприємства, установи і організації незалежно від форм власності, що забезпечують роботу річкового транспорту [2]. | |
| Автомобільний транспорт | підприємства, що здійснюють перевезення пасажирів, вантажів, багажу, пошти, авторемонтні і шиноремонтні підприємства, рухомий склад автомобільного транспорту, транспортно-експедиційні підприємства, а також автовокзали і автостанції, навчальні заклади, ремонтно-будівельні організації та соціально-побутові заклади, інші підприємства, установи та організації незалежно від форм власності, що забезпечують роботу автомобільного транспорту [2]. | галузь транспорту, яка забезпечує задоволення потреб населення та суспільного виробництва у перевезеннях пасажирів та вантажів автомобільними транспортними засобами [4]. |
| Авіаційний транспорт | підприємства, що здійснюють перевезення пасажирів, вантажів, багажу, пошти, аерофотозйомки, сільськогосподарські роботи, а також аеропорти, аеродроми, аероклуби, транспортні засоби, системи управління повітряним рухом, навчальні заклади, ремонтні заводи цивільної авіації та інші підприємства, установи та організації незалежно від форм власності, що забезпечують роботу авіаційного транспорту [2]. | |
| Міський електротранспорт | підприємства, що здійснюють перевезення пасажирів, вантажів, багажу, пошти, рухомий склад, трамвайні і тролейбусні лінії, ремонтно-експлуатаційні депо, службові приміщення, фунікулери, канатні дороги, ескалатори, заводи по ремонту рухомого складу і виготовленню запасних частин, споруди енергетичного господарства та зв'язку, промислові, ремонтно-будівельні, торговельні та постачальницькі організації, навчальні заклади, науково-дослідні та проектно-конструкторські установи, заклади охорони здоров'я, відпочинку, фізичної культури і спорту та інші культурно-побутові заклади і підприємства, установи та організації незалежно від форм власності, що забезпечують роботу міського електротранспорту [2]. | складова частина єдиної транспортної системи, призначена для перевезення громадян трамваями тролейбусами, поїздами метрополітену на маршрутах (лініях) відповідно до вимог життєзабезпечення населених пунктів [8]. |

Отже, діяльність в сфері пасажирських перевезень автомобільним транспортом буде відповідати КВЕД 49.3 (49.31 та 49.39).

Господарська діяльність в цій сфері, як і в будь-якій іншій, відбувається на основі українського

законодавства, відповідно до якого її можуть здійснювати суб'єкти господарювання – юридичні та фізичні особи, зареєстровані у встановленому порядку у відповідності із видами економічної діяльності [1].

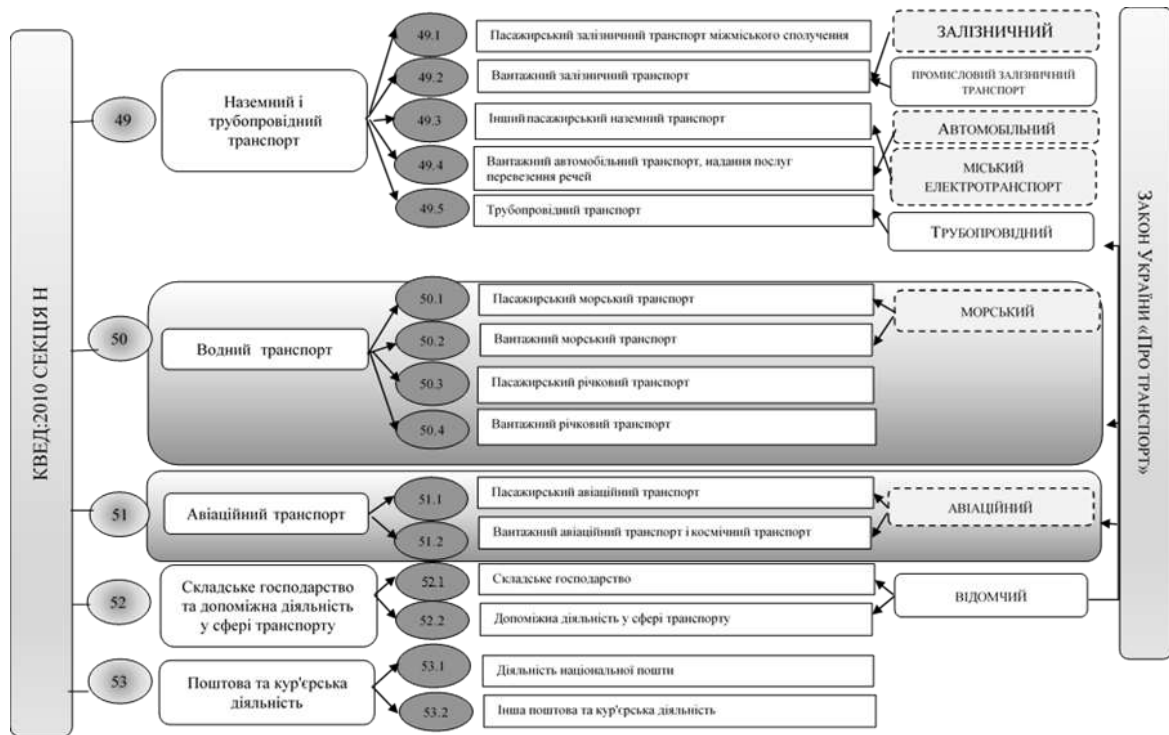


Рис. 2. Узгодження видів транспорту та видів економічної діяльності (побудовано автором на основі [2,19])

Господарський кодекс України, як одного із суб'єктів господарювання визначає, юридичну особу - підприємство.

Аналіз інших законодавчих актів в сфері транспорту демонструє певні неузгодженості як щодо суб'єктів господарювання, так і щодо розуміння підприємства транспорту того чи іншого виду. Порівняльна таблиця 2, презентує декілька понять, які характеризують суб'єктів господарювання в сфері того чи іншого виду транспорту. Серед них слід відмітити терміни: «суб'єкт господарювання», «перевізник», «підприємство (а)».

При цьому перевізник (за видами транспорту) визначається як суб'єкт господарювання (може бути як фізичною, так і юридичною особою), яка надає послуги з перевезення, транспортування. Підприємство ж є більш вужчим поняттям, аніж суб'єкт господарювання (перевізник), бо передбачає утворення юридичної особи. В той же час, знову слід відмітити неусталеність у визначеннях в законодавчому полі щодо розуміння терміну «перевізник» – фізична і юридична чи лише юридична. Така неусталеність обумовлена певними специфічними рисами в сфері діяльності транспорту: міський електротранспорт є монополізованим та не передбачає діяльність приватних структур, в той час як автомобільний транспорт є найменш монополізованим та характеризується високим рівнем конкуренції як серед юридичних, так і серед фізичних осіб, що провадять господарську діяльність в галузі перевезень пасажирів та вантажів.

Зауважимо, термін «перевізник» передбачає надання транспортних послуг, перевезення вантажів чи пасажирів тим чи іншим видом транспорту, що не

узгоджується з переліком підприємств за ЗУ «Про транспорт», де передбачені не лише перевезення, а й інші підприємства (обслуговуючі, ремонтні, тощо).

Тому вищевказані терміни не є узгодженими між собою в правовому полі транспорту України. На наш погляд, найбільш широкою характеристикою є суб'єкт господарювання, який може бути як фізичною, так і юридичною особою та здійснювати будь-які види діяльності згідно КВЕД; підприємство може бути лише юридичною особою; перевізник же обмежений лише сферою надання послуг з транспортування. Крім того, в науковому та повсякденному обігу вживаються терміни «транспортне підприємство» та «підприємство транспорту», які пропонуємо вживати як синоніми. Співвідношення вказаних термінів наведено на рис. 3.



Рис. 3. Співвідношення термінів стосовно суб'єктів, що провадять діяльність в сфері транспорту (розроблено автором)

Таким чином, законодавчі акти України не містять визначення поняття «транспортне підприємство»; зустрічається лише його ідентифікація за видами транспорту (автотранспортне підприємства, підприємства залізничного транспорту).

Виходячи з вищевказаного, доцільно ввести в нормативно-правове поле визначення транспортного

Таблиця 2 – Узагальнення визначень суб'єктів, які здійснюють в діяльність в сфері транспорту за законодавчо-правовим підходом (узагальнено автором)

| Визначення | Джерело |
|--|---------|
| Суб'єкт | |
| Суб'єкт авіаційної діяльності – фізичні та юридичні особи незалежно від форми власності, відомчої підпорядкованості, які провадять діяльність у галузі цивільної авіації; | [10] |
| Перевізник | |
| Автомобільний перевізник – фізична або юридична особа, яка здійснює на комерційній основі чи за власний кошт перевезення пасажирів чи (та) вантажів транспортними засобами; | [4] |
| Автомобільний самозайнятий перевізник – це фізична особа - суб'єкт господарювання, яка здійснює на комерційній основі чи за власний кошт перевезення пасажирів на таксі без застосування праці найманих водіїв; | |
| Автомобільним перевізником та автомобільним самозайнятим перевізником, які здійснюють перевезення пасажирів на договірних умовах, є суб'єкти господарювання, які відповідно до законодавства та одержаної ліцензії надають послуги за договором перевезення пасажирів транспортним засобом, що використовується ними на законних підставах. | |
| Перевізник – юридична або фізична особа, яка взяла на себе зобов'язання і відповідальність за договором перевезення вантажу за доставку до місця призначення довіреного їй вантажу, перевезення вантажів та їх видачу (передачу) вантажоодержувачу або іншій особі, зазначеній у документі, що регулює відносини між експедитором та перевізником; | [15] |
| Перевізник – суб'єкт господарювання, який здійснює перевезення вантажів, пасажирів та їх майна, багажу, пошти внутрішнім водним транспортом; | [20] |
| Перевізник – юридична особа, яка в установленому законодавством порядку надає транспортні послуги, здійснюючи експлуатацію та утримання об'єктів міського електричного транспорту; | [8] |
| Підприємство | |
| Підприємства залізничного транспорту – суб'єкти господарювання, які провадять діяльність у сфері залізничного транспорту | [3] |
| Автотранспортне підприємство надає послуги з перевезення пасажирів міським та приміським автомобільним транспортом, у тому числі й пільгових категорій пасажирів. | [21] |
| Підприємство з обслуговування річкових водних шляхів – юридична особа державної форми власності, утворена з метою обслуговування річкових водних шляхів, утримання та використання об'єктів інфраструктури внутрішнього водного транспорту державної форми власності, забезпечення функціонування річкової інформаційної служби, виконання інших покладених на неї відповідно до закону завдань. | [20] |

підприємства (підприємства транспорту), яке б корелювало з його дефініцією в Господарському Кодексі України. Пропонуємо розкрити термін «транспортне підприємство» як самостійний суб'єкт господарювання, створений компетентним органом ржавної влади або органом місцевого самоврядування, або іншими суб'єктами для задоволення суспільних та особистих потреб шляхом систематичного здійснення господарської діяльності в сфері транспорту, відповідно до Закону України «Про транспорт», в порядку, передбаченому Господарським Кодексом та іншими законами України.

Результати та подальші напрямки досліджень.

Чітке окреслення понятійно-категоріального апарату відносно сутності підприємств транспорту в законодавчому полі та його поєднання із класифікацією видів економічної діяльності дозволить в майбутньому коректно формувати сукупність для аналітичної вибірки та здійснювати подальшу класифікацію.

Список літератури

1. *Господарський кодекс України*. URL: <https://zakon.rada.gov.ua/laws/show/436-15> (дата звернення 20.04.2019).
2. *Закон України «Про транспорт»*. URL: <https://zakon.rada.gov.ua/laws/show/232/94-%D0%B2%D1%80> (дата звернення 20.04.2019).
3. *Закон України «Про залізничний транспорт»*. URL: <https://zakon.rada.gov.ua/laws/show/273/96-%D0%B2%D1%80> (дата звернення 20.04.2019).
4. *Закон України «Про автомобільний транспорт»*. URL: <https://zakon1.rada.gov.ua/laws/show/2344-14> (дата звернення 20.04.2019).
5. Рябова К. О. *Особливості правового статусу підприємств транспорту*. *Юридичний вісник. Повітряне і космічне право*. 2018. № 1. С. 144-149.
6. Скоромний Д.А. *Система транспорту та система перевезень за законодавством України*. URL: http://apir.org.ua/wp-content/uploads/2014/11/skoromnyu_ua.pdf (дата звернення 20.04.2019).
7. *Офіційний сайт Державної служби статистики України*. URL: <http://www.ukrstat.gov.ua/> (дата звернення 20.04.2019).
8. *Закон України «Про міський електричний транспорт»*. URL: <https://zakon.rada.gov.ua/laws/show/1914-15> (дата звернення 20.04.2019).
9. *Закон України «Про трубопровідний транспорт»*. URL: <https://zakon.rada.gov.ua/laws/show/192/96-%D0%B2%D1%80> (дата звернення 20.04.2019).
10. *Повітряний кодекс України*. URL: <https://zakon.rada.gov.ua/laws/show/3393-17> (дата звернення 20.04.2019).
11. *Водний кодекс України*. URL: <https://zakon.rada.gov.ua/laws/show/213/95-%D0%B2%D1%80> (дата звернення 20.04.2019/).
12. *Закон України «Про морські порти України»*. URL: <https://zakon.rada.gov.ua/laws/show/4709-17> (дата звернення 20.04.2019).
13. *Ліцензійні умови провадження господарської діяльності з перевезення пасажирів, небезпечних вантажів та небезпечних відходів автомобільним транспортом, міжнародних перевезень*

- пасажирів та вантажів автомобільним транспортом. URL: <https://zakon.rada.gov.ua/laws/show/1001-2015-%D0%BF> (дата звернення 20.04.2019).
14. *Порядок здійснення державного контролю на автомобільному транспорті*. URL: <https://zakon.rada.gov.ua/laws/show/1567-2006-%D0%BF> (дата звернення 20.04.2019).
 15. *Закон України «Про транспортно-експедиторську діяльність»*. URL: <https://zakon.rada.gov.ua/laws/show/1955-15> (дата звернення 20.04.2019).
 16. *Закон України «Про дорожній рух»*. URL: <https://zakon.rada.gov.ua/laws/show/3353-12> (дата звернення 20.04.2019).
 17. *Закон України «Про поштовий зв'язок»*. URL: <https://zakon.rada.gov.ua/laws/show/2759-14> (дата звернення 20.04.2019).
 18. *Закон України «Про особливості утворення акціонерного товариства залізничного транспорту загального користування»*. URL: <https://zakon.rada.gov.ua/laws/show/4442-17> (дата звернення 20.04.2019).
 19. *Класифікатор видів економічної діяльності 2010*. URL: ved.ukrstat.gov.ua/KVED2010/kv10_i.html (дата звернення 20.04.2019).
 20. *Проект Закону України «Про внутрішній водний транспорт»*. URL: http://w1.c1.rada.gov.ua/pls/zweb2/webproc4_1?pf3511=64397 (дата звернення 20.04.2019).
 21. *Офіційні відповіді*. URL: <https://zakon.rada.gov.ua/rada/show/p1180697-03> (дата звернення 20.04.2019).
- References (transliterated)**
1. *Hospodarskyi kodeks Ukrainy* [Economic Code of Ukraine]. URL: <https://zakon.rada.gov.ua/laws/show/436-15> (accessed 20.04.2019).
 2. *Zakon Ukrainy Pro transport* [Law of Ukraine On Transport]. URL: <https://zakon.rada.gov.ua/laws/show/232/94-%D0%B2%D1%80> (accessed 20.04.2019).
 3. *Zakon Ukrainy Pro zaliznychnyi transport* [Law of Ukraine on Railway Transport]. URL: <https://zakon.rada.gov.ua/laws/show/273/96-%D0%B2%D1%80> (accessed 20.04.2019).
 4. *Zakon Ukrainy Pro avtomobilnyi transport* [Law of Ukraine on Road Transport]. URL: <https://zakon1.rada.gov.ua/laws/show/2344-14> (accessed 20.04.2019).
 5. Riabova K. O. *Osoblyvosti pravovoho statusu pidpriemstv transportu* [Features of the legal status of transport enterprises]. *Yurydychnyi visnyk. Povitriane i kosmichne pravo*. 2018, no. 1, pp. 144-149.
 6. Skoromnyi D.A. *Systema transportu ta systema perevezen za zakonodavstvom Ukrainy* [Transport system that system is transported by the legislation of Ukraine]. URL: http://apir.org.ua/wp-content/uploads/2014/11/skoromnyy_ua.pdf (accessed 20.04.2019).
 7. *Ofitsiyni sait Derzhavnoi sluzhby statystyky Ukrainy*. URL: <http://www.ukrstat.gov.ua/> (accessed 20.04.2019).
 8. *Zakon Ukrainy Pro miskiyi elektrychnyi transport* [Law of Ukraine on urban electric transport]. Available at: <https://zakon.rada.gov.ua/laws/show/1914-15> (accessed 20.04.2019).
 9. *Zakon Ukrainy Pro truboprovodnyi transport* [Law of Ukraine on pipeline transport]. URL: <https://zakon.rada.gov.ua/laws/show/192/96-%D0%B2%D1%80> (accessed 20.04.2019).
 10. *Povitrianyi kodeks Ukrainy* [Air Code of Ukraine]. URL: <https://zakon.rada.gov.ua/laws/show/3393-17> (accessed 20.04.2019).
 11. *Vodnyi kodeks Ukrainy* [Water Code of Ukraine]. URL: <https://zakon.rada.gov.ua/laws/show/213/95-%D0%B2%D1%80> (accessed 20.04.2019).
 12. *Zakon Ukrainy Pro morskii porty Ukrainy* [Law of Ukraine on Sea ports of Ukraine]. URL: <https://zakon.rada.gov.ua/laws/show/4709-17> (accessed 20.04.2019).
 13. *Litsenzii umovy provadzhennia hospodarskoi diialnosti z perevezennia pasazhyriv, nebezpechnykh vantazhiv ta nebezpechnykh vidkhodiv avtomobilnym transportom, mizhnarodnykh perevezen pasazhyriv ta vantazhiv avtomobilnym transportom* [Licensing conditions for conducting business activities for the carriage of passengers, dangerous goods and hazardous wastes by road, international transportation of passengers and goods by road] URL: <https://zakon.rada.gov.ua/laws/show/1001-2015-%D0%BF> (accessed 20.04.2019).
 14. *Poriadok zdiisnennia derzhavnogo kontroliu na avtomobilnomu transporti* [The order of state control in road transport] URL: <https://zakon.rada.gov.ua/laws/show/1567-2006-%D0%BF> (accessed 20.04.2019).
 15. *Zakon Ukrainy Pro transportno-ekspedytorsku diialnist* [Law of Ukraine On Freight Forwarding Activities]. URL: <https://zakon.rada.gov.ua/laws/show/1955-15> (accessed 20.04.2019).
 16. *Zakon Ukrainy Pro dorozhnyi rukh* [Law of Ukraine on Road Traffic]. URL: <https://zakon.rada.gov.ua/laws/show/3353-12> (accessed 20.04.2019).
 17. *Zakon Ukrainy Pro poshtovyi sviazok* [Law of Ukraine on postal services]. URL: <https://zakon.rada.gov.ua/laws/show/2759-14> (accessed 20.04.2019).
 18. *Zakon Ukrainy Pro osoblyvosti utvorennia aktsionerhoho tovarystva zaliznychnoho transportu zahalnoho korystuvannia* [Law of Ukraine On peculiarities of railway transport company public]. URL: <https://zakon.rada.gov.ua/laws/show/4442-17> (accessed 20.04.2019).
 19. *Klasyfikator vydiv ekonomichnoi diialnosti 2010*. URL: ved.ukrstat.gov.ua/KVED2010/kv10_i.html (accessed 20.04.2019).
 20. *Proekt Zakonu Ukrainy «Pro vnutrishnii vodnyi transport»* [Draft Law of Ukraine On Inland Water Transport]. URL: http://w1.c1.rada.gov.ua/pls/zweb2/webproc4_1?pf3511=64397 (accessed 20.04.2019).
 21. *Ofitsiini vidpovidi* [Official answers] URL: <https://zakon.rada.gov.ua/rada/show/p1180697-03> (accessed 20.04.2019).

Надійшла (received) 29.04.2019

Відомості про авторів / Сведения об авторах / About the Authors

Смачило Валентина Володимирівна (Смачило Валентина Владимировна, Smachylo Valentyna) – кандидат економічних наук, доцент, Харківський національний університет будівництва та архітектури, професор кафедри економіки, м. Харків, Україна; ORCID: <https://orcid.org/0000-0002-6153-1564>; e-mail: miroslava.valya@ukr.net

УПРАВЛІННЯ В ОРГАНІЗАЦІЙНИХ СИСТЕМАХ**УПРАВЛЕНИЕ В ОРГАНИЗАЦИОННЫХ СИСТЕМАХ****MANAGEMENT IN ORGANIZATIONAL SYSTEMS**

UDC 519.854.2

DOI: 10.20998/2079-0023.2019.01.04

*A. A. PAVLOV, E. B. MISURA, O. V. MELNIKOV***TOTAL WEIGHTED TARDINESS MINIMIZATION FOR TASKS WITH A COMMON DUE DATE ON PARALLEL MACHINES IN CASE OF AGREEABLE WEIGHTS AND PROCESSING TIMES**

We consider n tasks scheduling problem on m identical parallel machines by the criterion of minimizing the total weighted tardiness of tasks. All tasks arrive for processing at the same time. Weights and processing times are agreeable, that is, a greater weight of a task corresponds to a shorter processing time. In addition, we have arbitrary start times of machines for tasks processing. The times may be less or greater than the due date or to coincide with it. The problem in this formulation is addressed for the first time. It can be used to provide planning and decision making in systems with a network representation of technological processes and limited resources. We give efficient PSC-algorithm with $O(mn \log n)$ complexity that includes the polynomial component and the approximation algorithm based on permutations of tasks. The polynomial component contains sufficient signs of optimality of the obtained solutions and allows to obtain an exact solution by polynomial subalgorithm. In the case when the sufficient signs of optimality do not fulfill, we obtain approximate solution with an estimate of deviation from the optimum for each individual problem instance of any practical dimension. We show that a schedule obtained as a result of the problem solving can be split into two schedules: the schedule on machines which start time is less than or equal to the due date, and the schedule on machines which start after the due date. Optimization is only done in the first schedule. The second schedule is optimal by construction. Statistical studies of the PSC-algorithm showed its high efficiency. We solved problems with dimensions up to 40,000 tasks and up to 30 machines. The average time to solve the problem by the algorithm using the most efficient types of permutations was 27.3 ms for this dimension. The average frequency of an optimal solution obtaining amounted to 90.3 %. The average deviation from an optimum was no more than 0.000251.

Keywords: scheduling theory, parallel machines, total weighted tardiness, common due date, agreeable weights, PSC-algorithm

*O. A. ПАВЛОВ, О. Б. МИСЮРА, О. В. МЕЛЬНИКОВ***МІНІМІЗАЦІЯ СУМАРНОГО ЗВАЖЕНОГО ЗАПІЗНЕННЯ ЗАВДАНЬ ІЗ СПІЛЬНИМ ДИРЕКТИВНИМ СТРОКОМ НА ПАРАЛЕЛЬНИХ ПРИБАДАХ ДЛЯ ВИПАДКУ УЗГОДЖЕНИХ ВАГ І ТРИВАЛОСТЕЙ**

Розглядається задача складання розкладів виконання n завдань на m ідентичних паралельних приладах за критерієм мінімізації сумарного зваженого запізнення завдань. Всі завдання надходять на обслуговування одночасно. Ваги і тривалості узгоджені, тобто завданню з меншою тривалістю відповідає більша вага. Додатково, задані довільні моменти початку роботи приладів на виконання завдань, які можуть бути як менше директивного строку, так і більше, або співпадати з ним. У такій постановці задача розв'язується вперше. Вона може використовуватися для забезпечення планування та прийняття рішень в системах з мережним представленням технологічних процесів та обмеженими ресурсами. Наведено ефективний ПДС-алгоритм її розв'язання із трудомісткістю $O(mn \log n)$, який включає поліноміальну складову з достатніми ознаками оптимальності одержуваних розв'язків, яка дозволяє отримувати точний розв'язок поліноміальним підалгоритмом. У разі невиконання достатніх ознак оптимальності ми отримуємо наближений розв'язок з оцінкою відхилення отриманого розв'язку від оптимального для кожної індивідуальної задачі будь-якої практичної розмірності. Показано, що розклад, отриманий в результаті розв'язання задачі, можна умовно розбити на два розклади – розклад на приладах, момент початку роботи яких менше або дорівнює директивному строку, та розклад на приладах, що починають роботу після директивного строку. Оптимізація виконується тільки у першому розкладі. Другий розклад оптимальний за побудовою. Статистичні дослідження ПДС-алгоритму показали його високу ефективність. Розв'язувались задачі з розмірністю до 40 000 завдань з числом приладів до 30. Середній час розв'язання задачі алгоритмом, що використовує найбільш ефективні типи перестановок, склав 27,3 мс при цій розмірності. Середня частота отримання оптимального розв'язку склала до 90,3 %. Середнє відхилення від оптимуму – не більш, ніж 0,000251.

Ключові слова: комбінаторна оптимізація, теорія розкладів, паралельні прилади, сумарне зважене запізнення, спільний директивний строк, узгоджені ваги, ПДС-алгоритм

*A. A. PAVLOV, E. B. MISURA, O. V. MELNIKOV***МИНИМИЗАЦИЯ СУММАРНОГО ВЗВЕШЕННОГО ЗАПАЗДЫВАНИЯ ЗАДАНИЙ С ОБЩИМ ДИРЕКТИВНЫМ СРОКОМ НА ПАРАЛЛЕЛЬНЫХ ПРИБОРАХ ДЛЯ СЛУЧАЯ СОГЛАСОВАННЫХ ВЕСОВ И ДЛИТЕЛЬНОСТЕЙ**

Рассматривается задача составления расписаний выполнения n заданий на m идентичных параллельных приборах по критерию минимизации суммарного взвешенного опоздания заданий. Все задания поступают на обслуживание одновременно. Веса и длительности согласованы, то есть заданию с меньшей длительностью соответствует больший вес. Дополнительно, заданы произвольные моменты начала работы приборов

© A. A. Pavlov, E. B. Misura, O. V. Melnikov, 2019

на выполнение заданий, которые могут быть как меньше директивного срока, так и больше, либо совпадать с ним. В такой постановке задача решается впервые. Она может использоваться для обеспечения планирования и принятия решений в системах с сетевым представлением технологических процессов и ограниченными ресурсами. Приведен эффективный ПДС-алгоритм ее решения с трудоемкостью $O(mn \log n)$, включающий полиномиальную составляющую с достаточными признаками оптимальности получаемых решений, которая позволяет получать точное решение полиномиальным подалгоритмом. В случае невыполнения достаточных признаков оптимальности мы получаем приближенное решение с оценкой отклонения полученного решения от оптимального для каждой индивидуальной задачи любой практической размерности. Показано, что расписание, полученное в результате решения задачи, можно условно разбить на два расписания – расписание на приборах, момент начала работы которых меньше или равен директивного срока, и расписание на приборах, начинающих работу после директивного срока. Оптимизация выполняется только в первом расписании. Второе расписание оптимально по построению. Статистические исследования ПДС-алгоритма показали его высокую эффективность. Решались задачи с размерностью до 40 000 заданий с числом приборов до 30. Среднее время решения задачи алгоритмом, использующим наиболее эффективные типы перестановок, составило 27,3 мс при этой размерности. Средняя частота получения оптимального решения составила до 90,3 %. Среднее отклонение от оптимума – не более 0,000251.

Ключевые слова: комбинаторная оптимизация, теория расписаний, параллельные приборы, суммарное взвешенное запаздывание, общий директивный срок, согласованные веса, ПДС-алгоритм

Introduction. Scheduling theory problems play an important role in calendar and operational planning of discrete type manufactures. In particular, those manufactures include small-series productions, aircraft and ship-building enterprises. Almost all known scheduling problems are NP-hard problems of combinatorial optimization. Therefore, creating for them efficient approximation algorithms is very important to solve practical problems of large dimensions.

The interest for tardiness criterion is due to its practical effect in the real life [1]. This criterion is among the most interesting criteria for production systems, especially in the current situation where competitiveness is becoming more and more intensive. Suppliers do ensure their markets and customers. For that, they must have a high service quality while focusing on delivery dates.

We solve in this paper the problem of the total weighted tardiness minimization on identical parallel machines with a common due date and agreeable weights (WTPA). It can be used to provide planning and decision making in systems with a network representation of technological processes and limited resources. In particular, it is used in the algorithmic ware of the four-level model of planning (including operative planning) and decision making [2].

WTPA problem statement. Given a set of tasks $J = \{j_1, j_2, \dots, j_n\}$, the number of machines is m . For each task $j \in J$, we know its processing time l_j and the weight ω_j . The weights are agreeable: if $l_i < l_j$, then $\omega_i \geq \omega_j$. All tasks have the same due date d . We need to build a schedule σ of tasks $j \in J$ on m machines that minimizes the function:

$$F(\sigma) = \sum_{j \in J} \omega_j \max[0, C_j(\sigma) - d], \quad (1)$$

where $C_j(\sigma)$ is the completion time of a task j in a schedule σ .

We assume that all tasks of the set J arrive simultaneously. The start times of the machines' operation for the processing of tasks $r_i \geq d$, $i = \overline{1, m}$, may be different. Here, $\overline{a, b}$ denotes the interval of integer numbers from a to b . Machines' idle times and interruptions in a task's processing are not allowed.

Literature review. This problem belongs to the class of NP-hard problems since it is NP-hard already for $m = 1$ and $r_1 = 0$ [3, 4].

Many modern methods of scheduling are described in the book [5]. Unified heuristics and annotated bibliography for a large class of scheduling problems with tardiness criteria are presented in [6]. The case of the problem with equal start times of machines and equal weights of tasks (TTP problem) was investigated in [2, 7–9]. A literature review for TTP problem in different formulations is given in [10]; this problem is also considered there for the case of different release dates of tasks. The class of schedules that contain an optimal solution is defined in [7], also an estimate of the deviation of the obtained solution from the optimum for each individual problem instance was formulated there. Kovalyov et al. [8] show that there is no polynomial approximation algorithm with a guaranteed relative error of a solution for TTP problem unless $P = NP$. Lawler et al. [11] indicated that the total tardiness problem on a single machine with a common due date is solved in $O(n^2)$ time. Its version with job weights is NP-hard in the ordinary sense [12].

Books [2, 9] present efficient PSC-algorithm to solve TTP problem, prove sufficient signs of optimality of the obtained solutions, and clarify the estimate of deviation of an obtained solution from the optimum. In contrast to the well-known estimate [7] which can be arbitrarily large even at an optimal solution, the estimate from [2, 9] is adequate: it is limited from above and from the bottom and shows the maximum possible reduction of the functional during an optimal solution construction. It was shown in [2, 9] that the PSC-algorithm for the problem solving, sufficient signs of optimality of the solutions obtained, and the estimate of deviation of a solution from the optimum for each individual problem instance obtained for TTP problem, are also valid for WTPA problem in the case of equal weights.

The above review shows that the problem under consideration is not presented in the scheduling literature in original formulation. This explains why we address it here. The problem in above formulation is addressed for the first time. The purpose of this paper is to develop the PSC-algorithm for the WTPA problem solving with an estimate of deviation of an obtained solution from the optimum for each individual problem instance.

Theoretical foundations. We now present basics of PSC-algorithm for TTP problem [2, 9, 10] which underlies the PSC-algorithm for WTPA problem.

Algorithm for initial schedule construction.

1. Renumber the tasks $j \in J$ in non-decreasing order of their processing times.

2. Set *release time* (completion time of all assigned jobs) of all machines to zero: $C_j = 0 \quad \forall i = \overline{1, m}$.
3. Select the next unassigned task j with the minimum processing time l_j . Assign it to a machine i with the minimum release time C_i .
4. Determine new release time of the machine i : $C_i = C_i + l_j$.
5. If all tasks were assigned, the algorithm terminates. Otherwise, go to step 3.

As a result of the above algorithm's execution, tasks $j \in J$ are split into three subsets [2]:

$P_i(\sigma)$ is the set of non-tardy tasks in the schedule of machine i ;

$S_i(\sigma)$ is the set of "straddling" tardy tasks in the schedule of machine i for which

$$C_j - l_j < d, C_j > d, \quad \forall j \in S_i(\sigma),$$

where $C_j - l_j$ is the start time of a task j ;

$Q_i(\sigma)$ is the set of "fully" tardy tasks in the schedule of machine i for which $C_j - l_j \geq d, \forall j \in Q_i(\sigma)$.

We use the notation from [2]:

$$P = \bigcup_{i=\overline{1, m}} P_i; S = \bigcup_{i=\overline{1, m}} S_i; Q = \bigcup_{i=\overline{1, m}} Q_i;$$

$R_i(\sigma)$ the time reserve of machine i in a schedule σ : $R_i(\sigma) = d - \sum_{j \in P_i(\sigma)} l_j$; $R_\Sigma(\sigma) = \sum_{i=1}^m R_i(\sigma)$;

$\Delta_i(\sigma)$ the tardiness of straddling task $j \in S_i(\sigma)$ in regard to the due date: $\Delta_i(\sigma) = \sum_{j \in P_i(\sigma) \cup S_i(\sigma)} l_j - d$;

$$\Delta_\Sigma(\sigma) = \sum_{i=1}^m \Delta_i(\sigma).$$

Theorem 1 [7]. There is an optimal schedule that satisfies the conditions:

$$1) P \cup S = \{1, 2, \dots, |P \cup S|\};$$

2) if $|P \cup S| < n$, then $\sum_{j \in P_i \cup S_i} l_j \geq d$ and Q_i contains those and only those elements which differ from $|P \cup S| + i$ by an amount which is a multiple of $m, i = \overline{1, m}$.

Corollary 1 [7]. Suppose that the tasks processing on L -th machine cannot be started before the time point $r_L \geq 0, L = \overline{1, m}$. Schedule σ where each successive task $k = 1, 2, \dots, n$ is assigned to be processed on the machine released before the others, corresponds to the smallest sum of completion times for all tasks.

Statement 1 [7]. When constructing an optimal schedule as a result of directed permutations, tasks can be moved only between sets P and S .

Let ψ_{PS} denote a class of schedules that correspond to the conditions of Theorem 1. $\psi_P \subset \psi_{PS}$ is a class of schedules satisfying the following additional conditions:

$$1) P = \{1, 2, \dots, |P|\};$$

$$2) \min_{j \in S_i(\sigma)} l_j > \max_{i=\overline{1, m}} R_i(\sigma);$$

$$3) \text{ if } l_{j_k} \leq l_{j_l}, \text{ then } C_{j_k} - l_{j_k} \leq C_{j_l} - l_{j_l} \quad \forall j_k, j_l \in S(\sigma).$$

The number of tardy tasks on machines differs by a maximum of one in the class ψ_P [2]. We determine $R_i(\sigma)$ on machines with a smaller number of tardy tasks and $\Delta_i(\sigma)$ on those with a larger number of tardy tasks.

Two *sufficient signs of optimality of a feasible solution* were proved in [2]:

1) a schedule $\sigma \in \psi_P$ with the same number of tardy tasks on all machines (an even schedule) is optimal;

2) if $\Omega_\Sigma(\sigma) = \min(R_\Sigma(\sigma), \Delta_\Sigma(\sigma)) = 0$ in a schedule $\sigma \in \psi_P$, then the schedule σ is optimal.

The PSC-algorithm from [2] is based on directed permutations that decrease $\Delta_\Sigma(\sigma)$ on machines with a larger number of tardy tasks due to reserves $R_\Sigma(\sigma)$ on machines with a smaller number of tardy tasks or to build an even schedule. Let $\Psi(\sigma_P) \subset \psi_{PS}$ denote the class of such schedules.

Theorem 2 [2]. The following estimate of deviation of the functional value from the optimum is valid for any schedule $\sigma \in \Psi(\sigma_P)$: $f(\sigma) - f(\sigma^*) \leq \Omega_\Sigma(\sigma)$.

The case of the TTP problem, in which the start times of machines are less than the common due date, was considered in [2], and a PSC-algorithm was given. Let us consider the general case.

Study of WTPA problem properties. Let us call ω_j/l_j the priority of a task j . Since we have different start times of machines in the WTPA problem, in contrast to the TTP problem, let us give a new algorithm for construction of initial schedule $\sigma^{\text{ord}} \in \psi_P$.

Algorithm A0.

1. Renumber the tasks $j \in J$ in non-increasing order of priorities ω_j/l_j .
2. Renumber the machines in non-decreasing order of start times r_i .
3. Set the initial release times of machines: $C_i = r_i \quad \forall i = \overline{1, m}$.
4. Select an unassigned task j with the maximum priority. Assign it to a machine i with the minimum release time C_i .
5. Determine the new release time of the machine i : $C_i = C_i + l_j$.
6. If all tasks were assigned, the algorithm terminates. Otherwise, go to step 4.

Denote the obtained schedule as σ^{ord} (sigma ordered). Let us split σ^{ord} conditionally into σ^1 and σ^2 where σ^1 is the schedule of tasks on machines with $r_i < d$ and σ^2 is the schedule of tasks on machines with $r_i \geq d$.

Statement 2. The number of tardy tasks on machines in the schedule σ^1 differs by a maximum of one.

Statement 3. The number of tardy tasks on each of the machines with $r_i < d$ is greater than or equal to the number of tardy tasks on each of the machines with $r_i \geq d$.

Validity of Statements 2 and 3 is based on the algorithm for the schedule σ^{ord} construction.

The schedule σ^1 meets the requirements to the class ψ_P . The following theorem is true.

Theorem 3. The PSC-algorithm for the problem solving, sufficient signs of optimality of obtained solutions, and the estimate of deviation of a solution from the optimum for each individual problem instance which were obtained for TTP problem, are also valid for WTPA problem.

Proof. Initial schedule of the class ψ_P is constructed in TTP problem by distributing the list of tasks in non-decreasing order of their processing times, each to the machine with a minimum release time. Similarly, initial schedule in WTPA problem is built by distributing the

priority-ordered sequence of tasks, each to the machine with the minimum release time. But the priority-ordered sequence of tasks coincides with the sequence of tasks ordered by non-decreasing of their processing times, due to the agreeability of weights and processing times. Hence, the initial schedules in TTP and WTPA problems coincide. The PSC-algorithm is based on task permutations from machines with a greater number of tardy tasks to machines with a smaller number of tardy tasks. Thus, the task weights do not affect the implementation of the algorithm and are taken into account only when determining the functional value. This proves the theorem.

Theorem 4. There are no permutations of tasks in the schedule σ^{ord} between the machines $i \in \sigma^1$ and $i \in \sigma^2$ which lead to a decrease in the functional value.

Proof is based on Statement 1 and Corollary 1. Consider the schedule on machines $i_r \in \sigma^1$ and $i_s \in \sigma^2$. Suppose that a task j_k is processed on the machine i_r and a task j_p on the machine i_s . We have $l_{j_k} \leq l_{j_p}$ for these tasks, in accordance with the algorithm for the schedule σ^{ord} construction and due to the agreeability of weights and processing times. Let us swap these tasks, that is, let the task j_k be processed on the machine i_s and the task j_p on the machine i_r . As a result of such a permutation, in accordance with Corollary 1, the functional value increases. The processing times of tasks $j \in \sigma^2$ is greater than that of tasks in the sets P or S , according to Theorem 1. Therefore, according to Statement 1, tasks $j \in \sigma^2$ cannot be moved into the set P or the set S . Consequently, there are no permutations of tasks in the schedule σ^{ord} between schedules σ^1 and σ^2 which decrease the functional value. This proves the theorem.

Corollary 2. Optimization is only done in the schedule σ^1 . The schedule σ^2 is optimal by construction in accordance with Corollary 1.

Corollary 3. We check the sufficient signs of optimality on the schedule σ^1 and determine the estimate of deviation of the functional value from the optimum $\Omega_\Sigma(\sigma)$, which is equal to $\min(R_\Sigma(\sigma), \Delta_\Sigma(\sigma))$, on the schedule σ^1 , since the schedule σ^2 is optimal by construction.

Theorem 5. The functional value for WTPA problem is equal to the sum of the functional values of the schedules σ^1 and σ^2 .

Proof of Theorem 5 is obvious.

PSC-algorithm for WTPA problem solving. It has the following 8 steps.

1. Initial schedule construction by Algorithm A0.
2. Split the obtained schedule σ^{ord} into σ^1 and σ^2 where σ^1 is the schedule of tasks on the machines with $r_i < d$ and σ^2 is the schedule of tasks on the machines with $r_i \geq d$.
3. Execute the PSC-algorithm A1 or A2 [2] on the schedule σ^1 .
4. Analyze the obtained solution. If the polynomial component of the algorithm has been fulfilled, then the schedule σ^1 is optimal, go to step 5. Otherwise, go to step 6.
5. Determine the functional value for the schedule σ^1 . Go to step 7.

6. Determine the functional value and the estimate of deviation of the functional value from the optimum for the schedule σ^1 .
7. Determine the functional value for the schedule σ^2 which is optimal by construction.
8. Determine the functional value for WTPA problem in accordance with Theorem 5. Terminate.

Justification of the algorithm. As a result of the initial schedule construction for WTPA problem which is implemented by Algorithm A0, we obtain schedules σ^1 (the schedule on the machines with $r_i < d$) and σ^2 (the schedule on the machines with $r_i \geq d$). The functional value for the entire schedule σ is equal to the sum of the functional values for the schedules σ^1 and σ^2 . The schedule σ^2 is optimal by construction. But building and optimization of the schedule σ^1 is reduced to solving the TTP problem. Thus, PSC-algorithm for WTPA problem solving coincides with the PSC-algorithm for TTP problem solving on the schedule σ^1 . The polynomial component of the PSC-algorithm coincides with the polynomial approximation of the exact algorithm. As a result of the problem solving, we obtain either an exact optimal solution by the polynomial component of the algorithm (if at least one of the sufficient signs of optimality has fulfilled during its execution), or an approximate solution with an upper estimate of deviation of the functional value from the optimum which is $\Omega_\Sigma(\sigma)$.

The complexity of the polynomial component of the PSC-algorithm is determined by the complexity of the algorithm (A1 or A2) used to solve the problem. Their complexities are $O(n^2m)$ and $O(mn \log n)$, respectively.

Computational studies. To research the algorithm's efficiency, we used instances generator and solver written in C# in Microsoft Visual Studio 2010 environment. We randomly generated the task sets with uniform distribution of parameters. We chose processing times and weights of tasks from interval $[1, 200]$, then we assigned to each next task with the minimal processing time the next maximal of unassigned weights. The due date d was calculated as $0.7L/m$ where L is the sum of all processing times. The start times of machines were chosen from uniform distribution within $[0, 2d]$. We carried out 2,000 runs for each (n, m) pair. We tested problems with up to 40,000 tasks and 30 machines on a PC with 2 GBytes of RAM and a Pentium Core 2 Duo processor with 3.0 GHz frequency.

Average (for 100 runs) solving time appeared: from 1.53 ms at $n = 3,000$ and $m = 5$ to 2,025.14 ms at $n = 40,000$ and $m = 15$ for Algorithm A1; from 1.78 ms at $n = 3,000$ and $m = 30$ to 27.3 ms at $n = 40,000$ and $m = 20$ for Algorithm A2. The average frequency of obtaining an optimal solution was 90.3% for Algorithm A1 and 73.5% for Algorithm A2. The average deviation from an optimal solution was 0.000251 for Algorithm A1 and 0.000114 for Algorithm A2. After introducing additional types of permutations, Algorithm A2 becomes more efficient than Algorithm A1. In this case, we achieve an optimal functional value by the polynomial component of the algorithm for 92 % of instances.

Conclusions. We considered the problem of scheduling tasks on identical parallel machines. Weights and pro-

cessing times are agreeable, that is, a greater weight corresponds to a task with a shorter processing time. Additionally, arbitrary start times of the machines are given, they can be less than or greater than the due date, or to coincide with it. The problem in this formulation was solved for the first time. We presented an efficient PSC-algorithm for its solving, sufficient signs of optimality of obtained solutions, the estimate of deviation of an obtained solution from the optimum for each individual problem instance, and statistical studies of the PSC-algorithm which showed its high efficiency. We solved instances with dimensions up to 40,000 tasks with the number of machines up to 30. The average solving time by the algorithm that uses the most efficient types of permutations was 27.3 ms at this dimension. The average frequency of obtaining an optimal solution amounted to 90.3 %. The average deviation from an optimum was no more than 0.000251.

References

1. Yalaoui F. Minimizing total tardiness in parallel-machine scheduling with release dates. *Applied evolutionary computation*. 2012. Vol. 3, Iss. 1. P. 21–46. doi: 10.4018/jaec.2012010102
2. Zgurovsky M. Z., Pavlov A. A. *Combinatorial optimization problems in planning and decision making: theory and applications*. Cham: Springer, 2019. 526 p. doi: 10.1007/978-3-319-98977-8
3. Garey M. R., Johnson D. S. *Computers and intractability: a guide to the theory of NP-completeness*. San Francisco: W. H. Freeman and Co., 1979. 348 p.
4. Lawler E. L. A “pseudopolynomial” algorithm for sequencing jobs to minimize total tardiness. *Annals of Discrete Mathematics*. 1977. Vol. 1. P. 331–342. doi: 10.1016/S0167-5060(08)70742-8
5. Pinedo M. L. *Scheduling: theory, algorithms, and systems*. 5-th ed. Cham: Springer, 2016. 690 p. doi: 10.1007/978-3-319-26580-3
6. Kramer A., Subramanian A. A unified heuristic and an annotated bibliography for a large class of earliness-tardiness scheduling problems. URL: <https://arxiv.org/abs/1509.02384> (accessed 05.05.2019).
7. Танаев В. С., Шкурба В. В. *Введение в теорию расписаний*. Москва: Наука, 1975. 256 с.
8. Kovalyov M. Y., Werner F. Approximation schemes for scheduling jobs with common due date on parallel machines to minimize total tardiness. *Journal of Heuristics*. 2002, Vol. 8, Iss. 4. P. 415–428. doi: 10.1023/A:1015487829051
9. Згуровский М. З., Павлов А. А. *Труднорешаемые задачи комбинаторной оптимизации в планировании и принятии решений: монография*. Киев: Наукова думка, 2016. 716 с.
10. Pavlov A. A., Misura E. B., Melnikov O. V., Mukha I. P., Lishchuk K. I. Approximation algorithm for parallel machines total tardiness minimization problem for planning processes automation. *The Second International Conference on Computer Science, Engineering and Education Applications ICCSEEA 2019 (26–27 January 2019, Kiev)*. Cham, Springer Publ., 2020, pp. 459–467. doi: 10.1007/978-3-030-16621-2_43

11. Lawler E. L., Moore J. M. A functional equation and its application to resource allocation and sequencing problems. *Management Science*. 1969. Vol. 16, no. 1. P. 77–84. doi: 10.1287/mnsc.16.1.77
12. Yuan J. The NP-hardness of the single machine common due date weighted tardiness problem. *Journal of Systems Science and Complexity*. 1992. Vol. 5, no. 4. P. 328–333.

References (transliterated)

1. Yalaoui F. Minimizing total tardiness in parallel-machine scheduling with release dates. *Applied evolutionary computation*. 2012, vol. 3, iss. 1, pp. 21–46. doi: 10.4018/jaec.2012010102
2. Zgurovsky M. Z., Pavlov A. A. *Combinatorial optimization problems in planning and decision making: theory and applications*. Cham, Springer Publ., 2019. 526 p. doi: 10.1007/978-3-319-98977-8
3. Garey M. R., Johnson D. S. *Computers and intractability: a guide to the theory of NP-completeness*. San Francisco, W. H. Freeman and Co. Publ., 1979. 348 p.
4. Lawler E. L. A “pseudopolynomial” algorithm for sequencing jobs to minimize total tardiness. *Annals of Discrete Mathematics*. 1977, vol. 1, pp. 331–342. doi: 10.1016/S0167-5060(08)70742-8
5. Pinedo M. L. *Scheduling: theory, algorithms, and systems*. 5-th ed. Cham, Springer Publ., 2016. 690 p. doi: 10.1007/978-3-319-26580-3
6. Kramer A., Subramanian A. *A unified heuristic and an annotated bibliography for a large class of earliness-tardiness scheduling problems*. Available at: <https://arxiv.org/abs/1509.02384> (accessed 05.05.2019).
7. Танаев В. С., Шкурба В. В.: *Vvedenie v teoriyu raspisaniy* [Introduction to scheduling theory]. Moscow, Nauka Publ., 1975. 256 p.
8. Kovalyov M. Y., Werner F. Approximation schemes for scheduling jobs with common due date on parallel machines to minimize total tardiness. *Journal of Heuristics*. 2002, vol. 8, iss. 4, pp. 415–428. doi: 10.1023/A:1015487829051
9. Zgurovsky M. Z., Pavlov A. A. *Trudnoreshaemye zadachi kombinatornoy optimizatsii v planirovani i prinyatii resheniy: monografiya* [Intractable problems of combinatorial optimization in planning and decision-making: monograph]. Kiev, Naukova dumka Publ., 2016. 716 p.
10. Pavlov A. A., Misura E. B., Melnikov O. V., Mukha I. P., Lishchuk K. I. Approximation algorithm for parallel machines total tardiness minimization problem for planning processes automation. *The Second International Conference on Computer Science, Engineering and Education Applications ICCSEEA 2019 (26–27 January 2019, Kiev)*. Cham, Springer Publ., 2020, pp. 459–467. doi: 10.1007/978-3-030-16621-2_43
11. Lawler E. L., Moore J. M. A functional equation and its application to resource allocation and sequencing problems. *Management Science*. 1969, vol. 16, no. 1, pp. 77–84. doi: 10.1287/mnsc.16.1.77
12. Yuan J. The NP-hardness of the single machine common due date weighted tardiness problem. *Journal of Systems Science and Complexity*. 1992, vol. 5, no. 4, pp. 328–333.

Received 07.04.2019

Відомості про авторів / Сведения об авторах / About the Authors

Павлов Олександр Анатолійович (Павлов Александр Анатольевич, Pavlov Alexander Anatolievich) – доктор технічних наук, професор, завідувач кафедри автоматизованих систем обробки інформації та управління, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»; м. Київ, Україна; ORCID: <http://orcid.org/0000-0002-6524-6410>; e-mail: pavlov.fiot@gmail.com

Місюра Олена Борисівна (Мисюра Елена Борисовна, Misura Elena Borisovna) – кандидат технічних наук, старший науковий співробітник НДІ інформаційних процесів Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»; м. Київ, Україна; ORCID: <http://orcid.org/0000-0002-4210-0996>; e-mail: elena_misura@ukr.net

Мельников Олег Валентинович (Мельников Олег Валентинович, Melnikov Oleg Valentinovich) – кандидат технічних наук, провідний інженер кафедри автоматизованих систем обробки інформації та управління Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського»; м. Київ, Україна; ORCID: <http://orcid.org/0000-0002-2651-2739>; e-mail: oleg.v.melnikov72@gmail.com

T. V. KOZULIA, A. S. SVIRIDOVA, M. M. KOZULIA

PROBLEMS OF ELICITATION AND ANALYSIS OF REQUIREMENTS TO THE PROGRAM MODULE OF MONITORING ON THE BASIS OF THE CONDITIONS OF COGNITIVE ANALYSIS

There are risks of obtaining result that does not answer a work purpose statement when developing the software product for new subject area. It is necessary to pay attention on requirements to the software for elimination of such risks. This article is considering questions about collection and analysis requirements to a program module of sea water areas environmental monitoring. Program system «НДС Эколог» for ensuring continuous observation of sea water areas condition is noted by complex structure, high labor intensity of processing, storage and manipulations of data due to their significant amount. The software of this system makes a large-scale program complex for monitoring researches automation. However, according to the preliminary analysis of functionality «НДС Эколог» on quality control of dolphins living environment is not sufficient and need additional module development. Thus, the further development of this monitoring system is associated with the problem of elicitation and analyzing the requirements for the program module for calculating the coastal zones water resources quality. Elicitation requirements to additional program module is carried out on the basis of the standard documentation analysis, business processes of monitoring researches according to program engineering provisions. Identified Requirements are need check on integrity and lack of contradictions in their pithiness for further documenting according to the international recommendations standard about development of requirements specifications to the software IEEE STD 830 1993. To solve the problems of information support and software for integrated monitoring, a matrix of requirements was formed and cognitive modeling was applied. The requirements dependency matrix represents a simple and effective method of contradictions and overlappings identification, in case of simple systems consideration. Cognitive modeling promotes better understanding a problem situation, identification of contradictions and qualitative system analysis. The purpose of modeling consists in formation and specification of a studied object functioning hypothesis which consists of separate subsystems and elements. In this case, the system of requirements should reflect the causal relationships of object elements under study.

Keywords: analysis, requirements, software module, dependency matrix, cognitive modeling, environmental monitoring.

T. V. КОЗУЛЯ, А. С. СВІРІДОВА, М. М. КОЗУЛЯ

ЗАДАЧИ ЗБОРУ ТА АНАЛІЗУ ВИМОГ ДО ПРОГРАМНОГО МОДУЛЮ З МОНІТОРИНГУ НА ОСНОВІ ПОЛОЖЕНЬ КОГНІТИВНОГО АНАЛІЗУ

При розробці програмного продукту для нової предметної області існують ризики отримання результату, що не відповідає меті технічного завдання. Для усунення таких ризиків необхідно звернути увагу на вимоги до програмного забезпечення (ПЗ). У даній роботі розглянуті питання збору та аналізу вимог до програмного модулю з екологічного моніторингу морських акваторій. Програмна система «НДС Еколог» для забезпечення безперервного спостереження за станом морських акваторій відзначається складною структурою, високою трудомісткістю обробки, зберігання і маніпулювання даних у зв'язку з їх значним обсягом. Програмне забезпечення цієї системи становить масштабний комплекс програм з автоматизації моніторингових досліджень. Однак, за попереднім аналізом функціональних можливостей ПЗ «НДС Еколог» з контролю якості середовища життєдіяльності дельфінів не є достатнім і потребує розробки додаткового модулю. Таким чином, подальший розвиток цієї моніторингової системи пов'язаний з задачею збору та аналізу вимог до програмного модулю з розрахунку якості водних ресурсів прибережних зон. Збір вимог до додаткового програмного модулю проводиться на основі аналізу нормативної документації, бізнес-процесів моніторингових досліджень у відповідності до положень програмної інженерії. Визначені вимоги потребують перевірки на цілісність і відсутність протиріч у їх змістовності для подальшого документування відповідно до міжнародного стандарту рекомендацій з розробки специфікацій вимог до програмного забезпечення IEEE Std 830-1993. Для розв'язання задач інформаційного та програмного забезпечення у комплексному моніторингу було сформовано матрицю залежності вимог та застосовано когнітивне моделювання. Матриця залежності вимог являє собою простий та ефективний метод виявлення протиріч і перекриття, у випадку розгляду нескладних систем. Когнітивне моделювання сприяє кращому розумінню проблемної ситуації, виявленню суперечностей та якісному аналізу системи. Мета моделювання полягає в формуванні та уточненні гіпотези про функціонування досліджуваного об'єкта, який складається з окремих підсистем та елементів. У такому разі система вимог повинна відображати причинно-наслідкові зв'язки елементів досліджуваного об'єкта.

Ключові слова: аналіз, вимоги, програмний модуль, матриця залежності, когнітивне моделювання, екологічний моніторинг.

T. V. КОЗУЛЯ, А. С. СВІРІДОВА, М. М. КОЗУЛЯ

ЗАДАЧИ СБОРА И АНАЛИЗА ТРЕБОВАНИЙ К ПРОГРАММНОМУ МОДУЛЮ ПО МОНИТОРИНГУ НА ОСНОВЕ ПОЛОЖЕНИЙ КОГНИТИВНОГО АНАЛИЗА

При разработке программного продукта для новой предметной области существуют риски получения результата, не соответствующего цели технического задания. Для устранения таких рисков необходимо обратить внимание на требования к программному обеспечению (ПО). В данной работе рассмотрены вопросы сбора и анализа требований к программному модулю для экологического мониторинга морских акваторий. Программная система «НДС Эколог» для обеспечения непрерывного наблюдения за состоянием морских акваторий отмечается сложной структурой, высокой трудоемкостью обработки, хранения и манипулирования данных в связи с их значительным объемом. Программное обеспечение этой системы составляет масштабный комплекс программ по автоматизации мониторинговых исследований. Однако, по предварительному анализу функциональных возможностей ПО «НДС Эколог» по контролю качества среды жизнедеятельности дельфинов не является достаточным и требует разработки дополнительного модуля. Таким образом, дальнейшее развитие этой мониторинговой системы связано с задачей сбора и анализа требований к программному модулю по расчету качества водных ресурсов прибрежных зон. Сбор требований к дополнительному программному модулю проводится на основе анализа нормативной документации, бизнес-процесов мониторинговых исследований в соответствии с положениями программной инженерии. Определенные требования нуждаются в проверке на целостность и отсутствие противоречий в их содержании для дальнейшего документирования соответствии с международным стандартом рекомендаций по разработке спецификаций требований к программному обеспечению IEEE Std 830-1993. Для решения задач информационного и программного обеспечения для комплексного мониторинга было сформировано матрицу зависимости требований и применено когнитивное моделирование. Матрица зависимости требований представляет собой простой и эффективный метод выявления противоречий и перекрестий, в случае рассмотрения несложных систем. Когнитивное моделирование способствует лучшему пониманию проблемной ситуации, выявлению противоречий и качественному анализу системы. Цель моделирования заключается в

формирования и уточнения гипотезы о функционировании исследуемого объекта, состоящего из отдельных подсистем и элементов. В таком случае система требований должна отражать причинно-следственные связи элементов исследуемого объекта.

Ключевые слова: анализ, требования, программный модуль, матрица зависимости, когнитивное моделирование, экологический мониторинг.

Introduction.

The urgency of problem tasks regarding the choice of methods and means for developing requirements for an additional program module for monitoring marine waters is directly related to the development of information technology in software engineering.

The process of setting requirements for software, the procedure for their formulation and formation, as well as the existing methods and means of ensuring their quality, as well as the process of developing the software itself, remain unsecured by the fundamental theory and effective methodology. A significant number of studies on the software requirements development, their formulation quality has a chaotic, unsystematic nature. The poor software requirements formulation leads to the appearance from 35% to 55% of all defects and errors of the future software product [1]. The absence of a universal, completed and tested theory and methodology for developing clear and qualitative requirements for software, as well as methods and tools for constructing template requirements that could be applied to both the preparation of user and system requirements formulate the problem of developing software requirements [2].

Comprehensive monitoring of marine areas is characterized by the high systems dimensionality to be studied and quality control, the high handling complexity, storage and manipulation of a large amount of data. Software in this case is a large-scale complex of software modules for the control systems and incoming experimental data processing, observation data, data processing results, etc.

In software engineering field is suggested to obtain the input information according to the requirements developed for this software system. Thus, in the article is set the actual task of collecting and analyzing the program module requirements for monitoring water resources of coastal zones in case of existing software development for marine waters monitoring.

It is proposed to introduce elements of cognitive modeling theory for develop requirements system.

In accordance with the stated purpose of creating a requirements system in the article following tasks are solved:

1) process components description of the of requirements elicitation and analyzing for the additional program module development for the program system «НДС Эколог»;

2) requirements composition are determined for the additional program module development for calculating the coastal zones water resources quality;

3) requirements significance for the software module are established according to cognitive analysis results.

Characteristics of elicitation bases and requirements analysis.

Software Requirements – a set of characteristics regarding the properties, quality and features of a future software product that needs to be developed or it is at

upgrading process. Software requirements are fixed in the requirements specification, use case diagrams and other artifacts of subject analysis area distinguishing 3 basic classification components of requirements (Fig. 1) [1].

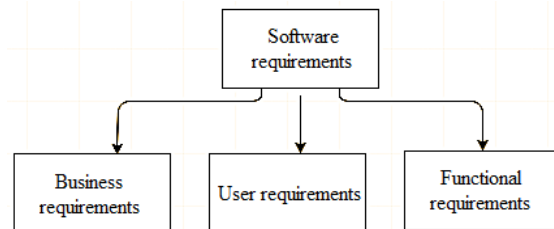


Fig. 1. Basic classification components of requirements

Business requirements determine software purpose, they are detailed in the document about project boundaries.

User requirements establish a set of user tasks that must implement a finished software product, as well as scenarios for their solution in the software hardware system. These requirements have the form of allegations, usage options, user stories, and interaction scenarios.

Functional requirements establish "what" should make a finished software product in the software and hardware system, they are detail described in the software requirements specification [3].

Functional requirements are documented in the software requirements specification, which describes the expected system behavior. Requirements types of by character are distinguished by functional and non-functional ones.

Functional requirements describe the internal system functionality, its behavior: data computation, data exchange and data management, data examination and other specific functions that system must do.

Non-functional requirements regulate the internal and external conditions or attributes of system functioning quality.

In general, functional requirements determine what the system should do, and non-functional ones - how the system should look.

Functional character – requirements for system behavior:

- business requirements;
- user requirements;
- functional requirements.

Non-functional character - requirements to the nature of the system's behavior:

- business rules;
- system requirements;
- quality attributes;
- external systems and interfaces;
- limitations.

Software requirements development is divided into four stages:

- requirements identification;
- requirement analysis;

- requirements documenting;
- requirements verification.

Requirements identification is a process of extracting information from different sources (contracts, materials, analysts, tasks and functions of the system, etc.), conducting technical interviews for formulate separate product and development process. Executor must approve requirements with customer [4].

The main source of requirements for the information system is the considerations expressed by the representatives of the Customer. The problem lies in the fact that requirements are formed for a system that is not yet created and existing, that is the initial subtask solves the problem of designing the program system but Customer representatives are not always competent in this matter. Therefore, together with the requirements witch are expressed by the Customer, it is expedient to collect requirements from other co-owners of the system: analytical executor group employees, external experts, etc.

Requirements elicitation methods:

- interviews, questionnaires;
- brainstorming seminar;
- normative documentation and legislation analysis;
- analysis of business processes.

The requirement detection stage is a divergent process that aims to gather a large amount of information. After identifying the requirements, the analyst has a large amount of various information received in the interview during the survey from questionnaires and other sources.

The obtained information should be clarified, structured, eliminated, duplicated, formulated in the form of requirements and defined their priorities. These tasks solve requirements analysis stage [3].

Requirements analysis is the process of studying users the needs and purposes, classifying and transforming them into system requirements, hardware and software requirements, establishing and resolving conflicts between requirements, defining priorities, system boundaries, and principles for interacting with functioning environment.

The purpose of requirements analysis is obtain clear and consistent requirements on the basis which it is possible to design and develop a software application. At the initial analysis stage, model requirements are created, missing information is determined, requirements are fixed and for each requirement the attributes values are set.

At the requirements analysis, a situation is likely when the requirements collected are incomplete or one requirement is contrary to the other. In this case, the analyst must re-interact with the customer, refine unclear moments, fix everything that is needed by developers and designers to implement these requirements [5].

At the stage of requirements documenting, it is provided a software requirements specification – complete software behavior description that is needed to be developed. This is the base that contains of functional and non-functional (additional) requirements set for the future software product.

Functional requirements describe all user interaction with software, and non-functional requirements impose

restrictions on the software project implementation or software functionality [1].

Module requirements analysis for calculating water quality in coastal zones.

The following functional and non-functional requirements are proposed to establish for an additional software module, which are in a certain order:

- relational type of DB (R1);
- to build of an ecological balance chart (R2);
- the ecological balance chart should be columnar (R3);
- to enter the seawater collection data (R4);
- the data are stored in the database (R5);
- processing requests to the module for no more than 10 seconds (R6);
- to predict changes in water quality for a month (R7);
- forecast accuracy is not less than 85% (R8);
- forecast chart of water quality changes for a month should be linear (R9);
- to calculate the water quality assessment by a differential method (R10).

The requirements dependence matrix (tab. 1) is formed because all requirements clearly identified and numbered.

Table 1 – Requirements dependence matrix

| | R1 | R2 | R3 | R4 | R5 | R6 | R7 | R8 | R9 | R10 |
|-----|----|----|----|----|----|----|----|----|----|-----|
| R1 | | | | | X | | | | | |
| R2 | | | | | | | | | | |
| R3 | | X | | | | | | | | |
| R4 | | | | | | | | | | |
| R5 | | | | X | | | | | | |
| R6 | | X | | X | | | X | | | X |
| R7 | | | | | | | | | | |
| R8 | | | | | | | X | | | |
| R9 | | | | | | | X | | | |
| R10 | | | | | | | | | | |

Requirements dependence matrix is a simple but effective method for detecting contradictions and overlaps when the number of requirements is small. Ordered IDs are listed in the column and header row. The remaining cells are indicating that there overlapping or two requirements conflicting or independent of one another (empty cells).

Contradictory requirements need to be discussed with customers and if it possible reformulated to mitigate contradictions. Overlapping requirements should be reworded to exclude coincidences [6].

According to the formed matrix, the requirements for the new program module of the monitoring system do not have contradictions and overlaps (see tab. 1).

Requirements significance for the program module for calculating the coastal zones water resources quality according to cognitive analysis results.

Cognitive analysis helps to better understand the problem situation, to identify contradictions and to determine the systems qualitative state. The purpose of creating a cognitive model is formulated and refines the

hypothesis about the functioning of the object under study, which consists of relations between subsystems and elements.

In order for understand and analyze the relations between requirements for the software module proposed to build a structural scheme of causal relationships between them [7-9].

The determined interaction of requirements is given in the form of a cognitive map - an oriented graph with weighted arcs of the graph (interaction evaluation or factors influence).

The requirements system behavior analysis takes place in accordance with the structural scheme of the cause-effect system elements relationships (situation factors) [10].

The comparisons matrix was constructed for further requirements advantages analysis (tab. 2).

Table 2 – The comparisons matrix

| | R1 | R2 | R3 | R4 | R5 | R6 | R7 | R8 | R9 | R10 |
|-----|----|-----|----|-----|----|----|-----|----|----|-----|
| R1 | | | | | 1 | | | | | |
| R2 | | | | | | | | | | |
| R3 | | 1/2 | | | | | | | | |
| R4 | | | | | | | | | | |
| R5 | | | | 1 | | | | | | |
| R6 | | 1/2 | | 1/2 | | | 1/2 | | | 1/2 |
| R7 | | | | | | | | | | |
| R8 | | | | | | | 2 | | | |
| R9 | | | | | | | 1 | | | |
| R10 | | | | | | | | | | |

The matrix uses the following scale of the relative requirements importance:

- 0- variants cannot be compared;
- 1- equaled importance;
- 2- moderate advantage one over another;
- 3- strong advantage one over another.

If the significance of one requirement exceeds the significance of another requirement, then a less significant requirement is assigned a reverse estimate of 1/2 or 1/3. Empty table cells mean zero effect.

A quantitative estimate of the factors influence was obtained as a result of comparisons matrix construction. The requirements system cognitive model with factors is shown on figure. 2.

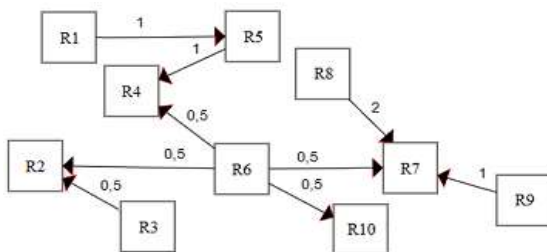


Figure 2 – Cognitive model of the requirements system

There are a number of numerical cognitive models characteristics, which are calculated on the basis of arc weights and used for the subject areas models analysis. To the number of such characteristics are included:

- the influence of one factor on another;
- influence of one factor on a system;
- impact of a system on a factor [11].

The total influence values of the factors on the map and the influence values of the map on factors is calculated for the cognitive model presented in figure. 2. The results of total calculations influences of requirements are given in table. 3.

Table 3 – Influence values

| | Influence of factors on the card | Influence of cards on factors |
|-----|----------------------------------|-------------------------------|
| R1 | 1 | |
| R2 | | 1 |
| R3 | 1,5 | |
| R4 | | 1,5 |
| R5 | 0,5 | 1 |
| R6 | 2 | |
| R7 | | 3,5 |
| R8 | 2 | |
| R9 | 1 | |
| R10 | | 0,5 |

The scales value of the requirements is shows that the greatest influence on the map are have the factors R6 and R8, and the card is have most influences on the factors R4 and R7. Based on this information, it is a priori determined that at the beginning of the development it is appropriate to take into account the requirements of R6 and R8, and finally is focus attention on requirements R4 and R7.

Conclusions.

The article is described the fundamentals characteristics of software requirements collection and analysis, provided a requirements classification, provided an overview of the main methods for requirements elicitation, identifies functional and non-functional requirements for an additional program module for monitoring water resources of coastal zones.

According to the provided information about the requirements for the module, their analysis was conducted for the absence of contradictions and overlaps (see table 1). The problem of requirement analysis for the additional module according to results of cognitive analysis regarding their significance is describes (see figure 2). A description of the requirements influence on the map and the map influence on the requirements is proposed and the priority of attention to them during the software module development is proposed (see table 3).

References

1. Грицюк І. Ю., Лешкевич І. Ф. Особливості визначення вимог до програмного забезпечення та проблеми їх аналізу. *Науковий вісник НЛТУ України «Львівська політехніка»*. 2017. № 4, т. 27. С. 148–158.
2. Говорущенко Т. О. *Теоретичні та прикладні засади інформаційної технології оцінювання достатності інформації, щодо якості у специфікаціях вимог до програмного забезпечення*. Хмельницький: ХНУ, 2018. 421 с.
3. Грицюк І. Ю., Немова О. А. Особливості формулювання вимог до програмного забезпечення. *Науковий вісник НЛТУ України «Львівська політехніка»*. 2018. № 7, т. 28. С. 135–148.

4. Дюваль, Поль М. *Непрерывная интеграция. Улучшение качества программного обеспечения и снижение риска: уч. пособие* / ред. В. А. Коваленко. Москва: Вильямс, 2017. 240 с.
5. Геци К., Мандриоли Д., Джазаири М. *Основы инженерии программного обеспечения*. Петербург: БХВ, 2016. 832 с.
6. Брукс Ф. *Мифический человек-месяц или как создаются программные системы*. Петербург: Символ-Плюс, 2015. 304 с.
7. Буров С. В., Пасичник В. В. Побудова програмних систем з використанням онтологічних моделей задач. *Системний аналіз та інформаційні технології. Матеріали доп. учасн. XVI міжнар. наук.-практ. конф.* Київ: ПІСА, 2014. С. 196–198.
8. Сиговцев Г. С., Семенов И. О. Разработка электронного учебного курса с использованием когнитивной карты как модели содержания. *Теория и практика системной динамики. Материалы междунар. науч.-практ. конф. Апатиты, 12 нояб. 2011.* Апатиты, 2012. С. 97–106.
9. Семенов И. О., Сиговцев Г. С. Математическое моделирование в планировании содержания электронного учебного курса. *Современные информационные технологии и ИТ-образование. Материалы междунар. науч.-практ. конф. Петрозаводск, 26 фев. 2012.* Петрозаводск, 2012. С. 134–139.
10. Горелова Г. В. Сложные системы. Когнитивное моделирование. *Когнитивная наука в Москве: новые исследования. Материалы междунар. науч.-практ. конф. Москва, 13 дек. 2010.* Москва, 2011. С. 80–85.
11. Горелова Г.В. Когнитивный подход к имитационному моделированию сложных систем. *Известия ЮФУ. Технические науки*. 2013. № 3. С. 239–250.
4. Duval Paul M. *Nepreryvnaya integratsiya. Uluchsheniye kachestva programmnogo obespecheniya i snizheniye riska* [Continuous integration. Improving software quality and reducing risk], uch. posobiye. Moscow, Vilyams Publ., 2017. 240 p.
5. Gaetzi K., Mandrioli D., Jazaieri M. *Osnovy inzhenerii programmnogo obespecheniya* [Basics of software engineering]. St. Petersburg, BHV Publ., 2016. 832 p.
6. Bruks F. *Mificheskij cheloveko-mesyats ili kak sozdaiutsya programmyye sistemy* [Mythical man-month or how software systems are created]. St. Petersburg, Symvol Plyus Publ., 2015. 304 p.
7. Burov Y. V., Pasichnyk V. V. Pobudova prohrannykh system z vykorystanniam ontolohichnykh modeley zadach [Construction of software systems using ontological models of tasks]. *Sistemnyy analiz ta informatsiyni tekhnolohiyi. Materialy dop. uchasn. XVI mizhnyar. nauk.-prakt. conf.* [System Analysis and Information Technology. Proceedings of the XVI International Scientific and Practical Conference]. Kyiv, IPSA Publ., 2014, pp. 196–198.
8. Sigovets G. S., Semenov I. O. Razrabotka elektronnoho uchebnogo kursa s ispol'zovaniem kognitivnoy karty kak modeli soderzhaniya [Development of an electronic training course using a cognitive map as a content mode]. *Teoriya i praktika sistemnoy dinamiki. Materiali mezhdunar. nuch.-prakt. conf. Apatity, 12 nojab. 2011* [Theory and practice of system dynamics. Materials of the international scientific-practical conference, Apatity, 12 nov. 2011]. Apatity, 2012, pp. 97–106.
9. Semenov I. O., Sygovtsev G. S. Matematicheskoye modelirovaniye v planirovaniy soderzhaniya elektronnoho uchebnogo kursa [Mathematical modeling in planning the content of the e-learning course]. *Sovremennyye informatsionnyye tekhnologii i IT-obrazovaniye. Materiali mezhdunar. nuch.-prakt. conf. Petrozavodsk, 26 Feb. 2012* [Modern information technology and IT education. Proceedings of the international scientific-practical conference. Petrozavodsk, February 26, 2012]. Petrozavodsk, 2012, pp. 134–139.
10. Gorelova G. V. Slozhnyye sistemy. Kognitivnoye modelirovaniye [Complex systems. Cognitive modeling]. *Kognitivnaya nauka v Moskve, novyye issledovaniya. Materialy mezhdunar. nuch.-prakt. conf. Moskva, 13 Des. 2010* [Cognitive science in Moscow: new research. Proceedings of the international scientific-practical conference. Moscow, December 13, 2010]. Moscow, 2011, pp. 80–85.
11. Gorelova G. V. Kognitivnyy podkhod k imitatsionnomu modelirovaniyu slozhnykh system [Cognitive approach to the simulation of complex systems] *Izvestiya YUFU. Tekhnicheskiye nauki* [News SFU. Technical science]. 2013, no. 3, pp. 239–250.

References (transliterated)

1. Gritsyuk I. Yu., Leshkevich I. F. Osoblyvosti vizualizatsiyi vymoza do programnoho zabezpechennya problem bezpeky ta yikh analizu [Features of defining requirements for software and problems of their analysis]. *Naukovyy visnyk NLTU Ukrainy "Lviv'ska politekhnika"* [Scientific bulletin of Ukraine NLTU "Lviv Polytechnic"]. 2017, no. 4, vol. 27, pp. 148–158.
2. Govoruschenko T. O. *Teoretychni ta prykladni zasady informatsiynoyi tekhnolohiyi otsynuyannya dostatnosti informatsiyi, shchodo yakosti u spetsyfikatsiyakh vymoza do programnoho zabezpechennya* [Theoretical and Applied Fundamentals of Information Technology for Assessing the Sufficiency of Information, on Quality in the Specifications of Software Requirements]. Khmelnytsky, KHNU Publ., 2018. 421 p.
3. Grycyuk I. Y., Nemova O. A. Osoblyvosti formulyuvannya vymoza do programnoho zabezpechennya [Features of software requirements wording]. *Naukovyy visnyk NLTU Ukrainy "Lviv'ska politekhnika"* [Scientific bulletin of Ukraine NLTU "Lviv Polytechnic"]. 2018, no. 7, vol. 28, pp. 135–148.

Received 17.05.2019

Відомості про авторів / Сведения об авторах / About the Authors

Козуля Тетяна Володимирівна (Козуля Татьяна Владимировна, Kozulia Tatyana Vladimirovna) – доктор технічних наук, професор, Національний технічний університет «Харківський політехнічний інститут», професор кафедри програмної інженерії та інформаційних технологій управління; м. Харків, Україна; ORCID: <https://orcid.org/0000-0001-5107-9140>; e-mail: mariya.kozulya7@gmail.com

Козуля Марія Михайлівна (Козуля Мария Михайловна, Kozulia Maria Mikhailovna) – кандидат технічних наук, доцент, Національний технічний університет «Харківський політехнічний інститут», доцент кафедри програмної інженерії та інформаційних технологій управління; м. Харків, Україна; ORCID: <https://orcid.org/0000-0002-4090-8481>; e-mail: mariya.kozulya7@gmail.com

Свіридова Альона Сергіївна (Свиридова Алёна Сергеевна, Sviridova Alyona Sergeevna) – Національний технічний університет «Харківський політехнічний інститут», студентка кафедри програмної інженерії та інформаційних технологій управління; м. Харків, Україна; ORCID: <https://orcid.org/0000-0003-1924-2520>; e-mail: http.pug@gmail.com

A. E. GOLOSKOKOV, D.V. TKACHENKO

MODELS AND SOFTWARE SOLUTIONS FOR THE PROBLEM OF DIAGNOSING THE FINANCIAL STATE OF IT-ENTERPRISE

Today, the economy of Ukraine is in a relatively unstable position; therefore, Ukrainian enterprises require effective management. But in order to effectively manage the enterprise, you need to know what state it is in. Solving the problem of diagnosing the financial state of an enterprise in the future will allow developing an apparatus of effective management decisions that will help maintain the enterprise at the proper level of functioning and ensure further development of both the enterprises and the economy as a whole. The relevance of research is manifested in the application of the results for operational and effective management. The problem is in the need to obtain a more accurate solution for the problem of diagnosing the financial state of the enterprise with the parameters that characterize the financial situation best of all. The main objective of the research was to solve the problem of diagnosing the financial state of an IT company, using a model that implements a certain approach in order to obtain a qualitative conclusion about the state of a company. A method based on the use of a fuzzy logic apparatus, namely, production models with a Mamdani fuzzy inference algorithm is proposed for solving the problem. There are 10 input parameters were allocated to determine the financial state. The criteria according to which the state was assessed were quantitative and qualitative indicators of the company's activity over the selected period. The resulting mathematical model allows to take into consideration both quantitative and qualitative indicators. The results of the research give an understanding of what indicators and how affect the financial condition of the company, and can also be used in the future, for example, to solve the forecasting problem. The implementation of research results can help speed up the diagnosis of the financial state of the enterprise and make a right management decision based on the results of diagnosis in time.

Keywords: diagnosing, financial state, financial indicator, fuzzy logic, production model, Mamdani algorithm, rule base.

О. Є. ГОЛОСКОКОВ, Д. В. ТКАЧЕНКО

МОДЕЛІ І ПРОГРАМНІ РІШЕННЯ ЗАДАЧІ ДІАГНОСТУВАННЯ ФІНАНСОВИХ СТАНІВ ІТ-ПІДПРИЄМСТВА

На сьогоднішній день, економіка України знаходиться у відносно нестабільному положенні, тому підприємства України вимагають ефективного управління. Але для того щоб ефективно управляти підприємством, необхідно знати в якому стані воно знаходиться. Рішення задачі діагностування фінансового стану підприємства в подальшому дозволить розробляти апарат ефективних управлінських рішень, які допоможуть підтримувати підприємство на належному рівні функціонування і забезпечувати подальший розвиток, як самих підприємств, так і економіки в цілому. Актуальність досліджень проявляється в застосуванні отриманих результатів для оперативного і ефективного управління. Проблема полягає в необхідності отримання більш точного рішення задачі діагностування фінансових станів підприємства, з урахуванням параметрів, що характеризують фінансовий стан найкраще. Основним завданням дослідження було рішення задачі діагностування фінансових станів ІТ-компанії, за допомогою моделі, що реалізує певний підхід з метою отримання якісного висновку про стан компанії. Для вирішення завдання пропонується метод, заснований на застосуванні апарату нечіткої логіки, а саме продукційних моделей з алгоритмом нечіткого висновку Мамдані. Для визначення фінансового стану були виділені 10 вхідних параметрів. Критеріями, за якими оцінювався стан, були кількісні та якісні показники діяльності компанії за обраний період. Отримана математична модель дозволяє враховувати як кількісні, так і якісні показники. Отримані результати дослідження дають розуміння про те, які показники і як саме впливають на фінансовий стан компанії, а також в подальшому можуть використовуватися, наприклад, для вирішення задачі прогнозування. Впровадження результатів дослідження може допомогти прискорити проведення діагностування фінансового стану на підприємстві, а також вчасно прийняти те чи інше управлінське рішення, засноване на результатах діагностування.

Ключові слова: діагностування, фінансовий стан, фінансовий показник, нечітка логіка, продукційна модель, алгоритм Мамдані, база правил.

A. E. GOLOSKOKOV, D.V. TKACHENKO

МОДЕЛИ И ПРОГРАММНЫЕ РЕШЕНИЯ ЗАДАЧИ ДИАГНОСТИРОВАНИЯ ФИНАНСОВЫХ СОСТОЯНИЙ ИТ-ПРЕДПРИЯТИЯ

На сегодняшний день, экономика Украины находится в относительно нестабильном положении, поэтому предприятия Украины требуют эффективного управления. Но для того чтобы эффективно управлять предприятием, необходимо знать в каком состоянии оно находится. Решение задачи диагностирования финансового состояния предприятия в дальнейшем позволит разрабатывать апарат эффективных управленческих решений, которые помогут поддерживать предприятие на должном уровне функционирования и обеспечивать дальнейшее развитие, как самих предприятий, так и экономики в целом. Актуальность исследований проявляется в применении полученных результатов для оперативного и эффективного управления. Проблема заключается в необходимости получения более точного решения задачи диагностирования финансовых состояний предприятия, с учетом параметров, характеризующих финансовое положение лучше всего. Основной задачей исследования являлось решение задачи диагностирования финансовых состояний ИТ-компании, с помощью модели, реализующей определенный подход с целью получения качественного заключения о состоянии компании. Для решения задачи предлагается метод, основанный на применении аппарата нечеткой логики, а именно продукционных моделей с алгоритмом нечеткого вывода Мамдани. Для определения финансового состояния были выделены 10 входных параметров. Критериями, по которым оценивалось состояние, были количественные и качественные показатели деятельности компании за выбранный период. Полученная математическая модель позволяет учитывать как количественные, так и качественные показатели. Полученные результаты исследования дают понимание о том, какие показатели и как именно влияют на финансовое состояние компании, а также в дальнейшем могут использоваться, например, для решения задачи прогнозирования. Внедрение результатов исследования может помочь ускорить проведение диагностирования финансового состояния на предприятии, а также вовремя принять то или иное управленческое решение, основанное на результатах диагностирования.

Ключевые слова: диагностирование, финансовое состояние, финансовый показатель, нечеткая логика, продукционная модель, алгоритм Мамдани, база правил.

Introduction. In the conditions that have emerged situation; therefore, Ukrainian enterprises require effective management in order to prevent the bankruptcy of

enterprises, which in turn can lead to a sharp decline of the economy. Among a number of tasks that are solving in the field of effective management, an important role takes the task of diagnosing the financial state of an enterprise, because it is impossible to manage effectively without knowing the condition of the control object.

Solving the problem of diagnosing the financial state of an enterprise in the future will allow developing an apparatus of effective management decisions that will help maintain the enterprise at the proper level of functioning and ensure further development of both the enterprises and the economy as a whole. In modern conditions of development of information technology, the automation of the diagnostic process can also significantly improve the efficiency of the enterprise. Some IT-company is considered as an enterprise.

The relevance of diagnosing the financial condition of an enterprise over time is obvious, because obtaining financial states is a necessary tool for strategic planning and management in the economic sector. Accurate determination of financial status provides an adequate understanding of the situation and the direction in which you need to act. The relevance of research is manifested in the application of the results for operational and effective management.

Timely and accurate determination of the financial state of the company is one of the main conditions for its successful development and functioning. Accurate determination of the financial condition in modern conditions is the reason for the need to have approaches that help determine the financial state of an enterprise and bring enterprises out of their unsatisfactory condition, if they are in it, by managing and subsequently ensuring their proper financial and economic position.

Formulation of the problem. First of all, the management of the enterprise needs to know the state of the object of management. To do this it is necessary to solve the problem of diagnosing the financial state of the company and to develop actions to ensure its further effective functioning and development.

Diagnosing is the process of determining and studying indicators, which characterize the state of an object. It consists in certain research methods, analysis of the obtained results and their generalization in the form of a conclusion (diagnosis) to determine possible deviations and prevent disturbances in the normal functioning of an object.

The task of diagnosing the financial state of an enterprise is characterized by the difficulty of obtaining information, a large volume of analyzed data, and multicriteria, because a large number of indicators are calculated and investigated.

Existing models and methods of diagnostic of financial state do not take into account the parameters that are quite important in modern conditions, since they do not bring them into consideration; therefore, the financial situation can be determined insufficiently accurately. Some mathematical models take into account a series of key parameters, but do not have the optimal criteria, which are necessary for determining the financial state. Also, uncertain or unclear information may be contained in the initial data and knowledge about the controlled object and

that information cannot be processed by traditional quantitative methods [1].

Therefore, in the current economic conditions, the issues of expanding and improving the models for diagnosing financial conditions are relevant.

The problem is in the need to obtain a more accurate solution for the problem of diagnosing the financial state of the enterprise with the parameters that characterize the financial situation best of all.

Thus, the main objective of the research was to solve the problem of diagnosing the financial state of an IT company, using a model that implements a certain approach in order to obtain a qualitative conclusion about the state of a company. The solution way, based on the obtained model, can be automated in the future, which is necessary for the quick and accurate determination of the financial condition of a company in order to improve the efficiency of the financial department and the company as a whole.

The subject of the research is the process of determining the financial state of the enterprise, and the object of the research is some IT-company.

The task and purpose of the research is to analyze the company's financial indicators and determine the financial condition with their help on the basis of the chosen diagnostic method.

The financial state in this case can be evaluated as good, normal or bad.

This article is aimed at implementing the practical aspects of diagnosis of the financial condition of the company. The approach to the diagnosis of the financial state of the company, realized in the construction of diagnosis model, differs from the well-known models of the mathematical apparatus, which allows to carry out diagnostic of the financial state of the company.

It is planned to build a model of diagnosing the financial state of the company such a way that it allows to take into account the optimal number of basic indicators that have a significant impact on the determination of the financial condition of the company with the greatest accuracy.

Thus, the need for accurate diagnosis of the financial condition of the enterprise is due to economic reasons and the reason for the need to make correct and effective management decisions. Solving this problem allows us to accurately assess the current financial position and activity of the enterprise.

Solution Method. The financial state of the company (which is an IT-company) depends on the results of its industrial, commercial and financial activities. It is determined on the basis of a series of indicators that most objectively reflect the trends of change in financial state, as a rule, it consists of four groups: liquidity indicators, financial stability indicators, profitability indicators (profitability), business activity indicators [2].

Analysis of Ukrainian and foreign publications and literary sources showed that the task of diagnosing the financial conditions of an IT-company can be solved using various mathematical approaches, such as: the classical coefficient method with calculating rating score, neural network technology, fuzzy logic apparatus and others [3–6].

For illustrate, the authors propose for the solving task of diagnosing financial state of company method based on the using of a fuzzy logic apparatus, namely, production models with the Mamdani fuzzy inference algorithm.

Complex dynamic economic systems (which are enterprises) may be characterized by incomplete information, instability, evolutionary changes, etc. This information is usually characterized by inaccuracy; it comes not in the form of quantitative, but in the form of linguistic variables [7].

In assessing the activity of the enterprise often have to deal with non-numeric data (qualitative characteristics). Many financial indicators do not have a precise rationing and are highly dependent on the scope of the enterprise. In such cases, often resorting to expert estimates. The classical model for assessing the financial condition cannot work with such data, so we can use fuzzy sets in this case.

The methods of fuzzy logic can significantly simplify the description of the model of objects of control and diagnosis, and are easier for hardware implementation.

Fuzzy production models are the most common type of fuzzy models used to describe, analyze and model complex systems and processes that are badly formalized [8].

The fuzzy production system is understood as the agreement of the sets of individual fuzzy production rules of the form "if A, then B", where A and B are the precondition and the conclusion of this rule in the form of fuzzy statements, designed to determine the truth level of the conclusions of fuzzy production rules, based on preconditions with a certain degree of truth of the relevant rules [8].

A fuzzy logical conclusion for a model that reflects the functioning of a company is called approximation of dependence $Y = f(x_1, x_2, \dots, x_{10})$ using a fuzzy knowledge base and fuzzy operations.

As input linguistic variables using variables, defined on the basis of selected primary financial indicators of the company (obtained from the financial accounting of the company). They are combined into two large groups [9].

The first is a group of quantitative indicators of financial condition

$$X_0 = f(X_1, X_2, X_3, X_4),$$

where X_1 – group of liquidity ratios, which including x_1 – current liquidity ratio, x_2 – absolute liquidity ratio;

X_2 – group of indicators of financial stability, which including, x_3 – availability of current assets ratio, x_4 – debt to equity ratio;

X_3 – group of profitability indicators, which including x_5 – indicator of profitability of assets, x_6 – indicator of profitability of equity;

X_4 – group of indicators of business activity, which including x_7 – asset turnover ratio, x_8 – equity turnover ratio.

Second group X_5 is quality parameters, which including x_9 – professional abilities of managers (point), x_{10} – the level of wages and social protection of employees (point) [9].

In order to be able to evaluate and process indicators x_i ($i = \overline{1,10}$) that can characterize a company in terms of financial condition, we define a single scale of three quality terms: L – a low level indicator x_i , M – a middle level indicator x_i , H – a high level indicator x_i [9].

Y will be used as the output linguistic variable. It is the financial state of the company.

The obtained output parameter Y allows us to characterize the financial state of the company as: G – good, N – normal, B – bad.

Solving the assigned problem also requires solving two subtasks, namely, the problem of parametric and structural identification, i.e. the possible range of variation of the selected parameters x_i and the output variable Y must be determined and the type of membership functions of fuzzy terms should be set for the input parameters which are used and for the output parameter too.

The membership function reflects elements from the set x on the set of numbers in the interval $[0, 1]$, which indicate the degree of belonging of each element to different qualitative terms [9].

Input indicators may correspond or do not match to recommend or standard values, which are presented in the form of some established intervals. Therefore, the trapezoidal membership function will be used to solve the problem, because it allows us to specify the basis of a fuzzy set as an interval and is simple to set [10].

The levels of all terms of each of the indicators x_i ($i = \overline{1,10}$), of a particular enterprise are set in accordance with standard values for classical criteria. If standards for the indicator do not exist, then the levels of the terms are divided on the basis of expert judgments or by comparing the values of the desired indicator for similar enterprises in different periods of time.

We reflect the ranges of change of parameters x_i , to a single universal set x in order for constructing the membership functions of three fuzzy terms of the input variable {L, M, H} [11].

Three fuzzy subsets are set whose membership functions are shown in Figure 1.

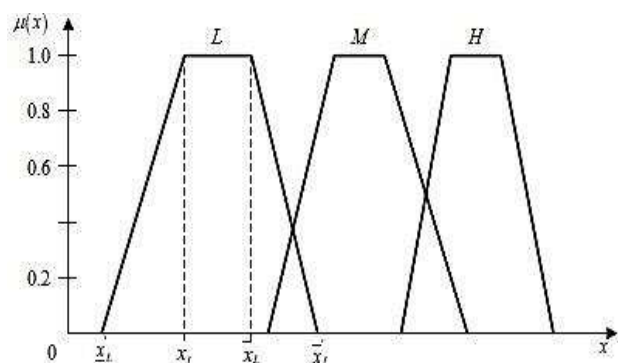


Figure 1 – Fuzzy variable x with a trapezoidal membership function

As an example, we can take the analytical form of writing a trapezoidal membership function of one fuzzy term L of the input variable x , shown in Figure 1:

$$\mu^L(x) = \begin{cases} 0, & x < \underline{x}'_L, \\ \frac{\bar{x} - \underline{x}'_L}{\underline{x}_L - \underline{x}'_L}, & \underline{x}'_L \leq x \leq \underline{x}_L, \\ 1, & \underline{x}_L \leq x \leq \bar{x}_L, \\ \frac{\bar{x}'_L - \bar{x}}{\bar{x}'_L - \bar{x}_L}, & \bar{x}_L \leq x \leq \bar{x}'_L, \\ 0, & x > \bar{x}'_L. \end{cases}$$

Membership functions of other fuzzy terms of the input variables x_i and terms of the output variable Y are built by a similar way.

A system based on fuzzy knowledge should contain a mechanism for fuzzy inference, such as to be possible to judge about the financial state of an enterprise. Therefore, the necessary stage of analysis is the creating of a system of rules [11].

A simple way to build a rule base is a full enumeration method, which will be used to solve the assigned problem.

Rules are given for each level of the system. An example of a decision rule would take the next expression: IF quantitative indicators of FSE are high AND qualitative parameters of FSE are middle OR quantitative indicators of FSE are high AND qualitative parameters of FSE are high, THEN the company's financial state is good.

The mathematical form of writing of a given decision rule using the membership functions is represented as:

$$\mu^Y = \mu^H(X_0) \cdot \mu^M(X_5) \vee \mu^H(X_0) \cdot \mu^H(X_5),$$

where $\mu^Y(X_0, X_5)$ – membership function of input variables vector (X_0, X_5) to the output variable Y ;

$\mu^{a_i}(X_i)$ – parameter membership function X to fuzzy term a_i .

Both criteria X_0 and X_5 , which represent the complex values of the specified groups of indicators, are presented in the form of mathematical dependencies on input variables [3].

The mathematical form of writing of the decision rule for determining the level H of the quantitative assessment of the FSE X_0 will take the form:

$$\begin{aligned} \mu^H(X_1, \dots, X_4) &= \mu^H(X_1) \cdot \mu^H(X_2) \cdot \mu^M(X_3) \times \\ &\times \mu^H(X_4) \vee \mu^H(X_1) \cdot \mu^M(X_2) \cdot \mu^H(X_3) \times \\ &\times \mu^H(X_4) \vee \mu^H(X_1) \cdot \mu^H(X_2) \cdot \mu^H(X_3) \times \\ &\times \mu^M(X_4). \end{aligned}$$

Next, we need to represent a relation for calculation of each of the generalized indicators of groups of influence factors X_1, \dots, X_4 .

On the example of a group of factors affecting the liquidity of an enterprise X_1 , we can present the rule for determining the generalized indicator of liquidity at level H.

This rule on natural language will be next: IF current liquidity ratio is high AND absolute liquidity ratio is high OR current liquidity ratio is middle AND absolute liquidity ratio is high, THEN the generalized liquidity ratio is high.

Next fuzzy logic equation will correspond to previous linguistic statement:

$$\mu^H(x_1, x_2) = \mu^H(x_1) \cdot \mu^H(x_2) \vee \mu^M(x_1) \cdot \mu^H(x_2).$$

The entire knowledge base is formed using the available data in this way, and a system of fuzzy logic equations will derive from it.

Algorithm for solving the problem. Fuzzy inference is the obtaining of a conclusion in the form of a fuzzy set corresponding to the current values of the inputs, using a fuzzy knowledge base and fuzzy operations [12]. The final result of the fuzzy inference is the precise value of the variable Y , obtained from the initially defined precise values x_i ($i = \overline{1, 10}$).

It is proposed to use one of the most common logical inference algorithms in fuzzy systems – the Mamdani algorithm for solving the problem.

The Mamdani fuzzy inference algorithm is implemented by the following steps [12].

At the first step, the predicate rule base is formed in the subject area, for example, using the full enumeration method; next action is checking the input variables: if the variables are precise sets, then go to the next step, which is called the fuzziness introduction stage – fuzzification [12].

In the second step, the fuzzification procedure is performed: each value of a separate input variable is associated with the value of the membership function of the corresponding term of the input linguistic variable [12]:

$$\mu_1(x), \mu_2(x), \dots, \mu_n(x),$$

where $\mu_1(x), \dots, \mu_n(x)$ – membership functions for variable x ;

The third step is the aggregation of prerequisites in fuzzy production rules. Paired fuzzy logic operations are used to find the degree of truth of the conditions of each of the rules of fuzzy products [12]:

$$\alpha_1 = \min\{\mu_{A_{11}}(x'_1), \mu_{A_{12}}(x'_2), \dots, \mu_{A_{1n}}(x'_n)\},$$

$$\alpha_2 = \min\{\mu_{A_{21}}(x'_1), \mu_{A_{22}}(x'_2), \dots, \mu_{A_{2n}}(x'_n)\},$$

.....

$$\alpha_m = \min\{\mu_{A_{m1}}(x'_1), \mu_{A_{m2}}(x'_2), \dots, \mu_{A_{mn}}(x'_n)\},$$

where m – number of variables;

n – number of rules in the base.

The fourth step is the activation procedure – finding the truncated membership function for the output variable, which is made according to the formulas [12]:

$$\mu_{B'_1} = \min\{\alpha_1, \mu_{B_1}(Y)\},$$

$$\mu_{B'_2} = \min\{\alpha_2, \mu_{B_2}(Y)\},$$

.....

$$\mu_{B'_m} = \min\{\alpha_m, \mu_{B_m}(Y)\}.$$

The fifth step is the procedure of accumulating or combining the found truncated functions in order to obtain the final fuzzy set for the output variable and the resulting

membership function, which is performed according to the formula [12]:

$$\mu_{B'}(Y) = \max\{\mu_{B'_1}(Y), \mu_{B'_2}(Y), \dots, \mu_{B'_m}(Y)\}.$$

The sixth and final step is defuzzification, or bringing to precision. Most often, the Mamdani model uses defuzzification by the centroid method, when a precise value of the output variable is defined as the center of gravity for the curve:

$$Y' = \frac{\sum_{i=1}^n Y_i \mu_{B'}(Y_i)}{\sum_{i=1}^n \mu_{B'}(Y_i)},$$

where n – the number of single-point fuzzy sets, each of which characterizes a single value of the considered output linguistic variable;

Y' – financial condition of the company [12].

Thus, having a rule base and using the algorithm above, we can get the result of a fuzzy inference, represented by the output parameter of a given model, namely, the financial condition of the company.

Numerical research. The process of solving the problem of diagnosing the financial state of IT-company can be illustrated using the MATLAB package.

Table 1 shows the values of the financial indicators of the studied company, obtained on the basis of the reporting forms of the company's financial activities (balance sheet) over the past few years.

Table 2 shows the trapezoidal numbers characterizing the corresponding membership functions for each indicator x_i .

Membership functions were built for the linguistic variable "Indicator" with the help of the MATLAB environment. An example of building the membership functions of the linguistic variable "Indicator" for the parameter x_1 is shown in Figure 2.

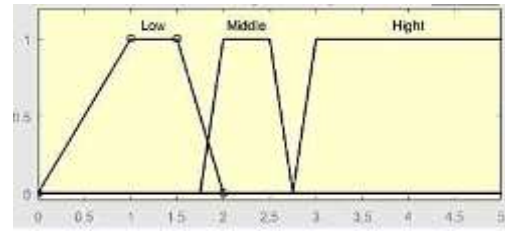


Figure 2 – Graphs of the membership functions of the linguistic variable "Indicator" for the parameter x_1

The membership functions of the linguistic variable "Indicator" is constructed for the remaining parameters x_i by the similar way.

Then, a rule base is created for each level of the system using data and MATLAB tools: first for determining the generalized indicators X_1, \dots, X_4, X_5 , then for obtaining the complex criterion X_0 and finally for determining the output parameter Y – the financial state of the company. As a result, the rule base will contain 135 rules.

Examples of the first two and the last two rules, which were formed in MATLAB, are given below.

Rules formed in MATLAB:

1. If (x1 is Low) and (x2 is Low) then (X1 is Low);
2. If (x1 is Low) and (x2 is Middle) then (X1 is Middle);
-
134. If (X0 is High) and (X5 is Low) then (Y is Normal);
135. If (X0 is Low) and (X5 is High) then (Y is Bad).

We can illustrate the obtaining of a generalized liquidity index X_1 using the Mamdani algorithm, implemented in MATLAB (Fig. 3).

Table 1 – Values of the selected financial indicators of the company

| Indicator / Year | x_1 | x_2 | x_3 | x_4 | x_5 | x_6 | x_7 | x_8 | x_9 | x_{10} |
|------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|----------|
| 2018 | 12.6 | 8.82 | 0.4 | 0.31 | 0.26 | 0.13 | 0.7 | 3.8 | 8.5 | 9 |
| 2017 | 3.92 | 0.89 | 0.29 | 0.58 | 0.48 | 0.57 | 3.23 | 4.95 | 7.5 | 7.9 |
| 2016 | 2.6 | 0.52 | 0.33 | 0.48 | 0.17 | 0.46 | 2.73 | 4.38 | 5.5 | 7.6 |
| 2015 | 1.7 | 0.03 | 0.53 | 0.67 | 0.19 | 0.15 | 1.67 | 4.25 | 4.9 | 3.6 |

Table 2 – T-numbers for the values of the linguistic variable "Indicator"

| Value / Indicator | Low | Middle | High |
|-------------------|-------------------|---------------------|-----------------|
| x_1 | [0 1 1.5 2] | [1.75 2 2.5 2.75] | [2.75 3 4 ∞] |
| x_2 | [0 0.1 0.15 0.2] | [0.2 0.25 0.5 0.75] | [0.7 0.9 ∞ ∞] |
| x_3 | [0 0.1 0.25 0.3] | [0.25 0.3 0.5 0.6] | [0.5 0.75 1 1] |
| x_4 | [0 0.2 0.4 0.5] | [0.45 0.5 0.6 0.7] | [0.65 0.75 1 1] |
| x_5 | [0 0.1 0.15 0.2] | [0.2 0.3 0.4 0.5] | [0.4 0.5 1 ∞] |
| x_6 | [0 0.1 0.15 0.25] | [0.2 0.3 0.4 0.5] | [0.45 0.5 1 ∞] |
| x_7 | [0 0.5 1.5 2] | [2 3.5 5 7.5] | [7 8 10 ∞] |
| x_8 | [0 2 3.5 4.5] | [4 4.6 7 8.5] | [8.5 9 10 ∞] |
| x_9 | [0 1 3 5] | [4.5 5 7 8.5] | [8 9 10 10] |
| x_{10} | [0 1 3 4] | [3.75 5 7 8] | [7.5 8 10 10] |
| $X_i, i = 0,5$ | [0 1 2 2.5] | [2.5 3 3.5 4] | [4 4.25 4.75 5] |

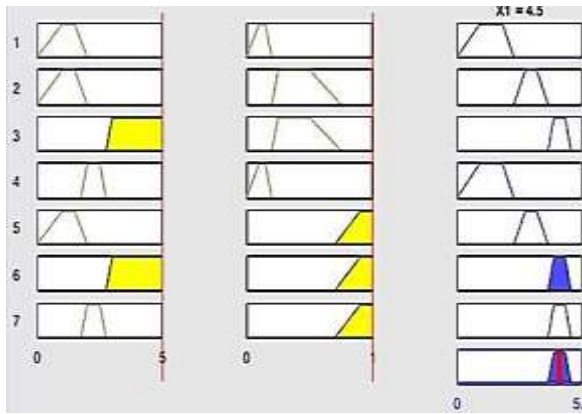


Figure 3 – Obtaining a generalized liquidity ratio X_1 (for 2018)

Figure 3 shows the result of accumulation, i.e. obtaining the final fuzzy set for the output variable X_1 and also shows the result of defuzzification using the centroid method.

Thus, it is clear that the generalized liquidity ratio is high. We obtain the remaining generalized indicators X_1, \dots, X_4, X_5 and X_0 by the similar way (Table 3).

Table 3 – Generalized indicators of the company and X_0

| Year | Generalized indicator | Value | Description |
|------|-----------------------|-------|-------------|
| 2018 | X_1 | 4.5 | High |
| | X_2 | 1.36 | Low |
| | X_3 | 3.25 | Middle |
| | X_4 | 1.33 | Low |
| | X_5 | 4.5 | High |
| | X_0 | 3.25 | Middle |
| 2017 | X_1 | 4.5 | High |
| | X_2 | 2.56 | Middle |
| | X_3 | 4.12 | High |
| | X_4 | 3.25 | Middle |
| | X_5 | 4.24 | High |
| | X_0 | 4.5 | High |
| 2016 | X_1 | 2.75 | Middle |
| | X_2 | 2.47 | Middle |
| | X_3 | 1.77 | Low |
| | X_4 | 2.61 | Middle |
| | X_5 | 3.58 | Middle |
| | X_0 | 3.25 | Middle |
| 2015 | X_1 | 1.29 | Low |
| | X_2 | 3.54 | Middle |
| | X_3 | 1.27 | Low |
| | X_4 | 1.3 | Low |
| | X_5 | 2.87 | Middle |
| | X_0 | 1.35 | Low |

Next, we need to build membership functions for the linguistic variable "The financial condition of the company." The graphs of the membership functions for the linguistic variable "The financial condition of the company" are shown in Figure 4.

We can get the output parameter Y – the financial condition of the company using the Mamdani algorithm and having data of input parameters (quantitative – a complex indicator X_0 and qualitative – a generalized indicator X_5) (Fig. 5).

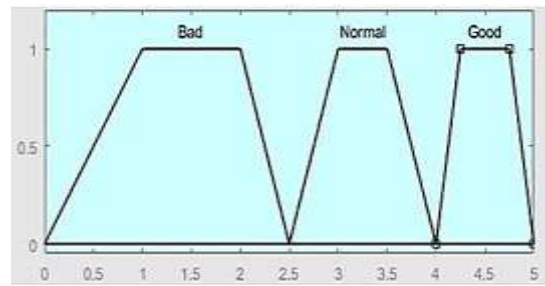


Figure 4 – Graphs of the membership functions of the linguistic variable "The financial condition of the company"

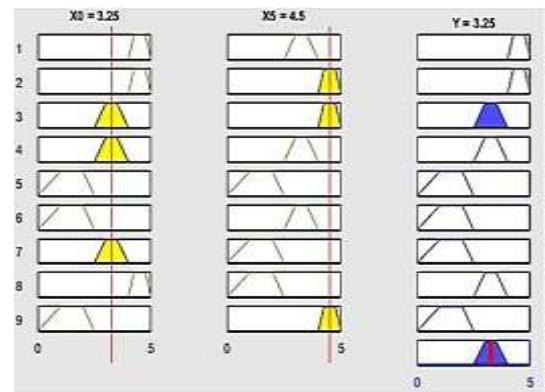


Figure 5 – Determination of the financial state Y (for 2018)

Figure 5 shows the result of the accumulation process – obtaining the final fuzzy set for the output variable Y , and also shows the result of defuzzification by the centroid method.

After completing all the above steps, we got the conclusion that the company's financial state is normal.

Table 4 shows the results of the research, which reflect the state of the company over the past few years.

Table 4 – The financial state of the company, obtained as a the result of research

| Year | Value | State |
|------|-------|--------|
| 2018 | 3.25 | Normal |
| 2017 | 4.5 | Good |
| 2016 | 3.25 | Normal |
| 2015 | 1.33 | Bad |

Investigate the change of membership functions depending on changes in the values of input parameters.

To do this, we change the parameters x_1, x_2 (leaving the remaining values of the financial indicators for 2018) affecting the generalized liquidity ratio, investigate the changes in the value of the generalized indicator X_1 , the complex indicator X_0 and the financial condition Y . The results are presented in table 5.

Similarly, we change the parameters x_5, x_6 , affecting the generalized profitability index, investigate the changes in the value of the generalized indicator X_3 , the complex indicator X_0 and the financial condition Y . The results are presented in table 6.

Next, we change the parameters x_9, x_{10} , investigate the changes in the value of the generalized quality indicator X_5 , and the financial state Y . The results are presented in table 7.

Table 5 – Changes in values X_1, X_0, Y

| | Indicator | | Generalized indicator X_1 | Description | Complex indicator X_0 | Description | Financial state Y | Description |
|------|-----------|-------|-----------------------------|-------------|-------------------------|-------------|---------------------|-------------|
| | x_1 | x_2 | | | | | | |
| 2018 | 12.6 | 8.82 | 4.5 | High | 3.25 | Middle | 3.25 | Normal |
| 2017 | 3.92 | 0.89 | 4.5 | High | 3.25 | Middle | 3.25 | Normal |
| 2016 | 2.6 | 0.52 | 2.75 | Middle | 1.31 | Low | 1.36 | Bad |
| 2015 | 1.7 | 0.03 | 1.29 | Low | 1.36 | Low | 1.36 | Bad |

Table 6 – Changes in values X_3, X_0, Y

| | Indicator | | Generalized indicator X_3 | Description | Complex indicator X_0 | Description | Financial state Y | Description |
|------|-----------|-------|-----------------------------|-------------|-------------------------|-------------|---------------------|-------------|
| | x_5 | x_6 | | | | | | |
| 2018 | 0.26 | 0.13 | 3.25 | Middle | 3.25 | Middle | 3.25 | Normal |
| 2017 | 0.48 | 0.57 | 4.12 | High | 3.25 | Middle | 3.25 | Normal |
| 2016 | 0.17 | 0.46 | 1.77 | Low | 1.36 | Low | 1.36 | Bad |
| 2015 | 0.19 | 0.15 | 1.27 | Low | 1.36 | Low | 1.36 | Bad |

Table 7 – Changes in values X_5, Y

| | Indicator | | Generalized indicator X_5 | Description | Financial state Y | Description |
|------|-----------|----------|-----------------------------|-------------|---------------------|-------------|
| | x_9 | x_{10} | | | | |
| 2018 | 8.5 | 9 | 4.5 | High | 3.25 | Normal |
| 2017 | 7.5 | 7.9 | 4.24 | High | 3.25 | Normal |
| 2016 | 5.5 | 7.6 | 3.58 | Low | 3.25 | Normal |
| 2015 | 4.9 | 3.6 | 2.87 | Low | 3.25 | Normal |

Conclusions. After analyzing the problem and the relevance of the research problem, the statement of the problem was formulated. The existing approaches for solving the problem of diagnosing financial states were considered during the research. As a result of the review, and to illustrate the example, the authors suggested using an approach based on the use of a fuzzy logic apparatus, namely, production models with the Mamdani fuzzy inference algorithm.

To determine the financial state were selected 10 input parameters. The criteria by which the state is assessed were quantitative and qualitative indicators of company's activity. The presented mathematical model allows to take into consideration both quantitative and qualitative indicators and also allow to analyze financial groups of indicators (liquidity, financial stability, profitability and business activity), and assess the level of managerial skills.

The implementation of research results can help speed up the diagnosis of the financial state of the enterprise and make a right management decision based on the results of diagnosis in time.

References

1. *Методы робастного, нейро-нечёткого и адаптивного управления: Учебник / ред. Н.Д. Егупов. Изд. 2-е. Москва: Изд-во МГТУ им Н. Э. Баумана, 2002. 744 с.*
2. Русак Н. А., Русак В. А. Диагностика финансового состояния предприятия. *Труд и социальные отношения*. Москва: АТИСО. 2012. № 9. С. 120–126.

3. Стоянов Е. А., Стоянова Е. С. *Экспертная диагностика и аудит финансово-хозяйственного положения предприятия*. Москва: Перспектива, 1993. 89 с.
4. Ковалев А. П. *Диагностика банкротства*. Москва: АО "Фин-статинформ", 1995. 96 с.
5. Плышевский Б. П. Финансовое положение: новые тенденции. *Финансы*. Москва: Изд-во "Книжная редакция финансы". 2003. № 2. С. 18–20.
6. Ротштейн А. П. *Интеллектуальные технологии идентификации: нечеткие множества, генетические алгоритмы, нейронные сети*. Винница: Універсум-Вінниця, 1999. 320 с.
7. Орлов А. И. *Нечисловая статистика*. Москва: МЗ-Пресс, 2004. 513 с.
8. Самарский А. А., Михайлов А. П. *Математическое моделирование. Идеи. Методы. Примеры*. 2-е изд. Москва: Физматлит, 2001. 320 с.
9. Матвийчук А. В., Сметанюк О. А. Диагностирование финансового состояния предприятия с применением инструментария нечеткой логики. *Фінанси України*. Київ: ДННУ "Академія фінансового управління". 2007. № 12. С. 115–128.
10. Штовба С. Д. *Введение в теорию нечетких множеств и нечеткую логику*. URL: <http://matlab.exponenta.ru/fuzzylogic/book1/index.php> (дата обращения: 25.02.2019).
11. Матвийчук А. В. *Аналіз та прогнозування розвитку фінансово-економічних систем із використанням теорії нечіткої логіки*. Київ: Центр навчальної літератури, 2005. 183 с.
12. Борисов В. В., Круглов В. В., Федулов А. С. *Нечеткие модели и сети*. 2-е изд. Москва: Горячая линия–Телеком, 2012. 284 с.

References (transliterated)

1. N. D. Egorov, ed. *Metody robastnogo, neuro-nechetkogo i adaptivnogo upravleniya. Uchebnik* [Methods of robust, neuro-fuzzy

- and adaptive control: Tutorial]. 2-e izd. Moscow, MGTU im N. E. Bauman Publ., 2002. 744 p.
- Rusak N. A., Rusak V. A. Diagnostika finansovogo sostoyaniya predpriyatiya [Diagnostics of the financial condition of the company]. *Trud i sotsial'nye otnosheniya [Labor and social relations]*. Moscow, ATISO Publ., 2012, no 9, pp. 120–126.
 - Stoyanov E. A., Stoyanova E. S. *Ekspertnaya diagnostika i audit finansovo-khozyaystvennogo polozheniya predpriyatiya* [Expert diagnostics and audit of the financial and economic situation of the enterprise]. Moscow, Perspektiva Publ., 1993. 89 p.
 - Kovalev A. P. *Diagnostika bankrotstva* [Диагностика банкротства]. Moscow, AO "Finstatinform" Publ., 1995. 96 p.
 - Plyshevskiy B. P. Finansovoe polozhenie: novye tendentsii [Financial situation: new trends]. *Finansy [Finances]*. Moscow, "Knizhnaya redaktsiya finansy" Publ., 2003, no 2, pp. 18–20.
 - Rotshteyn A. P. *Intellektual'nye tekhnologii identifikatsii: nechetkie mnozhestva, geneticheskie algoritmy, neyronnye seti* [Intellectual identification technologies: fuzzy sets, genetic algorithms, neural networks]. Vinnitsa, Universum-Vinnitsya Publ., 1999. 320 p.
 - Orlov A. I. *Nechislovaya statistika* [Non-numeric statistics]. Moscow, MZ-Press Publ., 2004. 513 p.
 - Samarskiy A. A., Mikhaylov A. P. *Matematicheskoe modelirovanie. Idei. Metody. Primery* [Math modeling. Ideas. Methods. Examples]. 2-e izd. Moscow, Fizmatlit Publ., 2001. 320 p.
 - Matviychuk A. V., Smetanyuk O. A. Diagnostirovanie finansovogo sostoyaniya predpriyatiya s primeneniem instrumentariya nechetkoy logiki [Diagnosing the financial condition of an enterprise using fuzzy logic tools]. *Finansy Ukrainy [Finance of Ukraine]*. Kiev, SAI "Academy of Financial Management" Publ., 2007, no 12, pp. 115–128.
 - Shtovba S. D. *Vvedenie v teoriyu nechetkikh mnozhestv i nechetkuyu logiku* [Introduction to the theory of fuzzy sets and fuzzy logic]. Available at: <http://matlab.exponenta.ru/fuzzylogic/book1/index.php> (accessed 20.01.2013).
 - Matviychuk A. V. *Analiz ta prohnouzuvannya rozvytku finansovo-ekonomichnykh system iz vykorystanniam teoriiy nechitkoyi lohiky* [Analysis and forecasting of the development of financial and economic systems using the theory of fuzzy logic]. Kiev, Tsentr navchal'noyi literatury Publ., 2005. 183 p.
 - Borisov V. V., Kruglov V. V., Fedulov A. S. *Nechetkie modeli i seti* [Fuzzy models and networks]. 2-e izd. Moscow, Goryachaya liniya–Telekom Publ., 2012. 284 p.

Received 26.03.2019

Відомості про авторів / Сведения об авторах / About the Authors

Голоскоков Олександр Євгенович (Голоскоков Александр Евгеньевич, Goloskokov Alexander Evgenievich)

– кандидат технічних наук, доцент, Національний технічний університет «Харківський політехнічний інститут», професор кафедри Програмної інженерії та інформаційних технологій управління, м. Харків, Україна; ORCID: <https://orcid.org/0000-0003-1824-6255>; e-mail: prof.goloskokov@gmail.com

Ткаченко Дарія Вадимівна (Ткаченко Дария Вадимовна, Tkachenko Daria Vadimovna)

– Національний технічний університет «Харківський політехнічний інститут», студент; м. Харків, Україна; ORCID: <https://orcid.org/0000-0002-5234-1013>; e-mail: tka4enko.daha@gmail.com

МАТЕМАТИЧНЕ І КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ

МАТЕМАТИЧЕСКОЕ И КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ

MATHEMATICAL AND COMPUTER MODELING

УДК 519.612

DOI: 10.20998/2079-0023.2019.01.07

*Н. А. МАРЧЕНКО, Р. О. РУДЕНКО***ПОЛІПШЕНИЙ МЕТОД ДОСЛІДЖЕННЯ СТІЙКОСТІ РОЗВ'ЯЗУВАННЯ СИСТЕМ ЛІНІЙНИХ АЛГЕБРАІЧНИХ РІВНЯНЬ**

Проведений огляд існуючих методів дослідження стійкості розв'язків систем лінійних алгебраїчних рівнянь (СЛАР), що залежать від вхідних даних, тобто варіацій параметрів. Розглянуто методи оцінки стійкості розв'язування систем лінійних алгебраїчних рівнянь, такі як числа зумовленості, модульні визначники та побудова таблиці знаків за оригінальним та поліпшеним методом побудови. Реалізоване програмне забезпечення для оцінки стійкості систем лінійних алгебраїчних рівнянь за допомогою чисел обумовленості, модульних визначників та побудови таблиці знаків для знаходження точних оцінок варіацій розв'язків залежних від варіацій параметрів СЛАР.

В роботі показано, що дослідження стійкості за числами обумовленості дають дуже грубу оцінку можливих похибок розв'язків, але вони є простими в реалізації, та для СЛАР можуть одразу показати, що деякі системи є погано зумовленими, що значно економить час дослідження, особливо якщо СЛАР мають дуже велику розмірність. Дослідження стійкості за модульними визначниками потребують великих розрахунків, але дають досить надійну оцінку зверху щодо можливих варіацій окремих компонент розв'язків систем лінійних алгебраїчних рівнянь. Це є дуже важливою особливістю метода тому, що окремі компоненти розв'язку можуть зазнавати значних варіацій, що не враховуються при дослідженні за числами обумовленості. Дослідження стійкості побудовою таблиці знаків надають можливість знайти максимальні варіації окремих компонент розв'язків системи лінійних алгебраїчних рівнянь, що насправді можуть бути значно меншими, ніж верхня оцінка можливих варіацій за методом модульних визначників. В роботі запропоновано поліпшений метод побудови таблиці знаків, що знаходить більш точний діапазон можливих варіацій розв'язків системи лінійних алгебраїчних рівнянь.

Був проведений порівняльний аналіз між традиційним методом побудови таблиці знаків за окремими визначниками та поліпшеним методом побудови таблиці знаків за похідними від ділення визначників за формулою Крамера. Згідно аналізу, поліпшений метод у 30% випадків знаходить варіації, що в 1.3 рази більші ніж варіації, що знаходить попередній метод, та у 5% випадків ці варіації перевищують попередні у 2 або більше разів. Це говорить про те, що традиційний метод у деяких випадках недооцінював можливі відхилення розв'язків, що залежать від варіацій вхідних даних.

Ключові слова: лінійна алгебра, чисельні методи, стійкість, матриця, система лінійних алгебраїчних рівнянь, визначник, модульний визначник, таблиця знаків, число обумовленості

*Н. А. МАРЧЕНКО, Р. А. РУДЕНКО***УЛУЧШЕННЫЙ МЕТОД ИССЛЕДОВАНИЯ УСТОЙЧИВОСТИ РЕШЕНИЯ СИСТЕМ ЛИНЕЙНЫХ АЛГЕБРАИЧЕСКИХ УРАВНЕНИЙ**

Проведен обзор существующих методов исследования устойчивости решений систем линейных алгебраических уравнений (СЛАУ), зависящих от входных данных, то есть вариаций параметров. Рассмотрены методы оценки устойчивости решения систем линейных алгебраических уравнений, такие как числа обусловленности, модульные определители и построение таблицы знаков по оригинальному и улучшенному методам построения. Реализовано программное обеспечение для оценки устойчивости систем линейных алгебраических уравнений с помощью чисел обусловленности, модульных определителей и построения таблицы знаков для нахождения точных оценок вариаций решений зависящих от вариаций параметров СЛАУ.

В работе показано, что исследование устойчивости по числам обусловленности дают очень грубую оценку возможных погрешностей решений, но они просты в реализации, и для СЛАУ могут сразу показать, что некоторые системы плохо обусловленными, что значительно экономит время исследования, особенно если СЛАУ имеют очень большую размерность. Исследование устойчивости по модульным определителям требуют больших расчетов, но дают достаточно надежную оценку сверху относительно возможных вариаций отдельных компонент решений систем линейных алгебраических уравнений. Это является очень важной особенностью метода потому, что отдельные компоненты решений могут испытывать значительные вариации, которые не учитываются при исследовании при помощи чисел обусловленности. Исследование устойчивости при помощи построения таблицы знаков дает возможность найти максимальные вариации отдельных компонент решений систем линейных алгебраических уравнений, которые на самом деле могут быть значительно меньше, чем верхняя оценка возможных вариаций методом модульных определителей. В работе предложен улучшенный метод построения таблицы знаков, который находит более точный диапазон возможных вариаций решений системы линейных алгебраических уравнений.

Был проведен сравнительный анализ между традиционным методом построения таблицы знаков по отдельным определителям и улучшенным методом построения таблицы знаков по производным от деления определителей по формуле Крамера. Согласно анализу, улучшенный метод в 30% случаев находит вариации, которые в 1.3 раза больше чем вариации, которые находит предыдущий метод, и в 5% случаев эти вариации превышают предыдущие в 2 или более раз. Это говорит о том, что традиционный метод в некоторых случаях недооценивал возможные отклонения решений, зависящих от вариаций входных данных.

Ключевые слова: линейная алгебра, численные методы, надежность, матрица, система линейных алгебраических уравнений, определитель, модульный определитель, таблица знаков, число обусловленности.

© Н. А. Марченко, Р. О. Руденко, 2019

N. A. MARCHENKO, R. O. RUDENKO

IMPROVED METHOD FOR STABILITY RESEARCH OF SOLUTIONS OF SYSTEMS OF LINEAR ALGEBRAIC EQUATIONS

A review of existing methods for stability research of solutions of systems of linear algebraic equations (SLAE) depending on the input data, that is, parameter variations, have been carried out. Methods for stability research of solving systems of linear algebraic equations, such as condition numbers, modular determinants and the construction of a table of signs using the original and improved construction methods, were considered. Software for stability research of systems of linear algebraic equations have been developed. The software is using condition numbers, modular determinants, and construction of table of signs to find accurate estimates of the variations of solutions of SLAEs that depend on parameters variations.

It is shown that stability research using condition numbers gives a very rough estimate of possible errors in solutions, but that research is simple to implement, and for SLAEs it can immediately show that some systems are ill-conditioned, which saves research time, especially if SLAEs have a very large dimensionality. The stability research using modular determinants requires large calculations, but they give a fairly reliable upper estimate with respect to possible variations of individual components of solutions of systems of linear algebraic equations. This is a very important feature of the method because the individual components of solutions may experience significant variations that are not taken into account in the research using condition numbers. The study of stability by constructing a table of signs makes it possible to find the maximum variations of individual components of solutions of a system of linear algebraic equations, which in fact can be significantly less than the upper estimate of possible variations found by method of modular determinants. The paper proposes an improved method for constructing a table of signs, which finds a more accurate range of possible variations of solutions of a system of linear algebraic equations.

A comparative analysis was conducted between the traditional method of constructing a table of signs using individual determinants and an improved method of constructing a table of signs based on the derivatives of the division of determinants using the Cramer formula. According to the analysis, the improved method in 30% of cases finds variations that are 1.3 times greater than the variations that the previous method finds, and in 5% of cases these variations are 2 or more times greater than the previous ones. This suggests that the traditional method in some cases underestimated the possible deviations of solutions that depend on variations of the input data.

Keywords: linear algebra, numerical methods, stability, reliability, matrix, system of linear algebraic equations, determinant, modular determinant, signs table, condition number.

Вступ. За оцінками вчених чисельні методи розв'язування СЛАР застосовуються більш ніж у половині розрахункових математичних задач [1–9]. Традиційні методи розрахунку стійкості розв'язування систем лінійних алгебраїчних рівнянь не дають надійних результатів [10, 11]. Однією з причин ненадійності розрахунків є неточності в завданні коефіцієнтів рівнянь. Для реальних об'єктів ці коефіцієнти не можуть точно відповідати значенням, заданим при проектуванні, крім того, в ході експлуатації параметри об'єкта можуть зазнавати змін, що впливають на розв'язки. При деяких умовах навіть невеликі зміни в параметрах можуть привести до значних змін у розв'язках, і відповідно до значних змін у поведінці об'єктів.

У даній роботі увага приділяється погрішностям розв'язків, що виникають через варіації вхідних даних розрахунків.

Норма матриці, числа зумовленості. Класичним методом оцінки стійкості розв'язку системи лінійних алгебраїчних рівнянь є числа зумовленості [1, 2, 10–13]. Якщо розглядається система лінійних рівнянь, яка записується в матричному вигляді як

$$Ax = b, \quad (1)$$

де A – матриця СЛАР;
 x – невідомий вектор, що є розв'язком системи;
 b – заданий вектор правих частин.

Тоді числом зумовленості називають добуток норми матриці A на норму оберненої матриці. При цьому число зумовленості системи лінійних рівнянь пов'язує властивості матриці системи рівнянь з похибкою розв'язку системи лінійних рівнянь. В найгіршому випадку відносна невизначеність розв'язку x буде перевищувати відносну невизначеність вхідних даних у кількість разів.

Навіть малі числа зумовленості матриці не є надійним показником стійкості отриманих розв'язків. Це пов'язано з тим, що числа зумовленості лише

гарантують деякий діапазон значень міри вектору розв'язку x , наприклад, його довжини, але не мають відношення до окремих складових розв'язку x_i , що в деяких випадках можуть навіть змінювати знак при достатньо невеликих значеннях числа зумовленості.

Перевагою використання чисел зумовленості є простота використання та мала кількість необхідних обчислювань.

Оцінка похибки за модульними визначниками. В роботах [10, 11] як методи розрахунку стійкості додатково пропонується використовувати оцінки похибки за модульними визначниками та за таблицею їх знаків.

Модульні визначники дозволяють дати оцінку максимально можливої невизначеності визначників системи лінійних алгебраїчних рівнянь, якщо відома невизначеність вхідних даних. Модульний визначник розраховується аналогічно до розрахунку визначника матриці, але усі складові беруться за модулем та усі знаки «мінус» у формулі замінюються на знаки «плюс».

Модульний визначник чисельно пов'язує властивості вхідних даних з мірою відносної невизначеності визначників матриці.

Справжнє значення визначника $\overline{\det}$ можна оцінити за допомогою формули:

$$\det_{\text{НОМ}} - n \varepsilon \det_{\text{МОД}} \leq \overline{\det} \leq \det_{\text{НОМ}} + n \varepsilon \det_{\text{МОД}}, \quad (2)$$

де $\det_{\text{НОМ}}$ – визначник матриці, обчислений за стандартною формулою;

$\det_{\text{МОД}}$ – модульний визначник матриці;

$\overline{\det}$ – справжнє значення визначника матриці;

n – порядок матриці;

ε – можлива відносна похибка в визначенні параметрів матриці.

Виходячи з того, що формула відома Крамера дозволяє отримати значення розв'язків x_i системи лінійних алгебраїчних рівнянь як ділення визначників, та

відомостей про невизначеність визначників системи лінійних алгебраїчних рівнянь можемо дати верхню оцінку невизначеності окремих складових розв'язку x_i . Використовуючи формулу Крамера

$$x_i = \frac{D^{<i>}}{D}, \quad (3)$$

де x_i – i -та компонента розв'язку СЛАР;
 D – визначник СЛАР, що досліджується;
 $D^{<i>}$ – визначник, отриманий із D заміною i -го стовбця на стовбець вільних членів СЛАР, можливо оцінити варіації розв'язку x_i при відносних варіаціях коефіцієнтів $\pm \varepsilon$:

$$\frac{D^{<i>\min}}{D_{\max}} < x_i < \frac{D^{<i>\max}}{D_{\min}}, \quad (4)$$

де x_i – i -та компонента розв'язку СЛАР;
 D_{\min} і D_{\max} – відповідно мінімально і максимально можливе значення визначника СЛАР, отримане за формулою (3);

$D^{<i>\min}$ і $D^{<i>\max}$ – відповідно мінімально і максимально можливе значення визначника $D^{<i>}$, отримане за формулою (3).

Перевагою використання модульних визначників є відносна простота використання (порівняно до методу побудови таблиці знаків) та можливість отримання верхньої границі невизначеності розв'язків.

Головними недоліками використання модульних визначників є значна кількість необхідних обчислювань для матриць великої розмірності, та те, що модульні визначники зазвичай дають завищені значення невизначеності, які не можуть бути досягненні [10, 11]. Для більш точної оцінки необхідно користуватися методом побудови таблиці знаків та розраховувати максимальні і мінімальні значення розв'язків x_i .

Оцінка похибки за таблицею знаків для визначників. Метод оцінки за допомогою «таблиці знаків» ґрунтується на побудові такого прикладу можливого відхилення, що призводить до максимально можливих змін значення розв'язків x_i . Побудова прикладу відхилення, як і у попередньому методі, використовує формулу Крамера (3) як формулу прямого знаходження розв'язків x_i

Розглянемо визначник n -го порядку

$$D = \begin{vmatrix} a_{11} \pm \varepsilon_{11} a_{11} & \cdots & a_{1n} \pm \varepsilon_{1n} a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} \pm \varepsilon_{n1} a_{n1} & \cdots & a_{nn} \pm \varepsilon_{nn} a_{nn} \end{vmatrix}, \quad (5)$$

як функцію n^2 змінних величин ε_{ij} . Випишемо головну лінійну частину приросту цієї функції:

$$\Delta_{\text{лін}} = \sum_{i=1}^n \sum_{j=1}^n a_{ij} A_{ij} \Delta \varepsilon_{ij}, \quad (6)$$

де n – порядок матриці,
 a_{ij} – елемент визначника з координатами (i, j) ;
 A_{ij} – алгебраїчне доповнення елемента a_{ij} ;
 $\Delta \varepsilon_{ij}$ – можлива відносна похибка в визначенні елемента a_{ij} , $i = 1, 2, \dots, n$, $j = 1, 2, \dots, n$.

Кожен елемент таблиці знаків («+» або «-») співпадає зі знаком добутку $a_{ij} A_{ij}$.

Аналогічно попередньому методу за формулами (3) і (4) оцінюються варіації розв'язку x_i , але завдяки знаходженню конкретного прикладу варіації визначників ця оцінка буде точнішою за попередню.

Додаткові складності вносить наявність спільних коефіцієнтів визначників у чисельнику та знаменнику, а саме, $n - 1$ стовпців у визначниках D та $D^{<i>}$, збігаються. Це призводить до суперечливих змін розв'язків x_i , що не дозволяє знайти максимально можливе відхилення розв'язків при зміні коефіцієнтів ε_{ij} , що задовольняють нерівності $|\varepsilon_{ij}| \leq \varepsilon_0$, де ε_0 – задана максимально можлива відносна похибка в визначенні елементів a_{ij} .

Поліпшений метод побудови таблиці знаків.

Розглянемо розв'язок x_k :

$$x_k = \frac{D^{<k>}}{D} = \frac{\begin{vmatrix} a_{11} \pm \varepsilon_{11} a_{11} & \cdots & b_1 \pm \mu_1 b_1 & \cdots & a_{1n} \pm \varepsilon_{1n} a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} \pm \varepsilon_{n1} a_{n1} & \cdots & b_n \pm \mu_n b_n & \cdots & a_{nn} \pm \varepsilon_{nn} a_{nn} \end{vmatrix}}{\begin{vmatrix} a_{11} \pm \varepsilon_{11} a_{11} & \cdots & a_{1k} \pm \varepsilon_{1k} a_{1k} & \cdots & a_{1n} \pm \varepsilon_{1n} a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} \pm \varepsilon_{n1} a_{n1} & \cdots & a_{nk} \pm \varepsilon_{nk} a_{nk} & \cdots & a_{nn} \pm \varepsilon_{nn} a_{nn} \end{vmatrix}},$$

де $k = 1, 2, \dots, n$;

μ_1, \dots, μ_n – можливі відносні похибки в визначенні елементів b_1, \dots, b_n .

Запишемо формулу для функції x_k , що залежить від ε_{ij} :

$$x_k(\varepsilon_{ij}) = \frac{D^{<k>}(\varepsilon_{ij})}{D(\varepsilon_{ij})} = \frac{\varepsilon_{ij} a_{ij} A^{<k>} + D^{<k>}}{\varepsilon_{ij} a_{ij} A_{ij} + D}, \quad (7)$$

де $i = 1, 2, \dots, n$, $j = 1, 2, \dots, n$, та знайдемо часткову похідну розв'язку x_k за змінною ε_{ij} :

$$x_k(\varepsilon_{ij})' = \frac{a_{ij} A^{<k>}_{ij} D - a_{ij} A_{ij} D^{<k>}}{(\varepsilon_{ij} a_{ij} A_{ij} + D)^2}. \quad (8)$$

Головна лінійна частина приросту розв'язку x_k в точці $\varepsilon_{ij} = 0$ (його повний диференціал) без урахування стовпця $j = k$ буде дорівнювати:

$$\Delta_{\text{лін}}^{<k>} = \sum_{i=1}^n \sum_{j=1, j \neq k}^n \frac{a_{ij} A^{<k>}_{ij} D - a_{ij} A_{ij} D^{<k>}}{D^2} \cdot \Delta \varepsilon_{ij}, \quad (9)$$

де $\Delta \varepsilon_{ij}$ – приріст змінної ε_{ij} . Вважаючи, як і раніше, що для всіх i та j , $j \neq k$ дотримується нерівність $|\varepsilon_{ij}| \leq \varepsilon_0$, отримаємо, що найбільше збільшення розв'язку в лінійному наближенні буде при $|\varepsilon_{ij}| = \varepsilon_0$, і воно дорівнює

$$\Delta_{\text{лін макс}}^{<k>} = \sum_{i=1}^n \sum_{j=1, j \neq k}^n \frac{|a_{ij} A^{<k>}_{ij} D - a_{ij} A_{ij} D^{<k>}|}{D^2} \varepsilon_0. \quad (10)$$

Кожен елемент (окрім випадку $j = k$) повної «таблиці знаків» співпадає зі знаком виразу $a_{ij} A^{<k>}_{ij} D - a_{ij} A_{ij} D^{<k>}$.

За умови $j = k$ визначники $D^{<k>}$ та D не мають спільних коефіцієнтів, тому максимізація (мінімізація)

розв'язку x_k за коефіцієнтами μ_k та ε_{ik} може проводитися незалежно для визначників $D^{<k>}$ та D , але слід враховувати знаки визначників.

Запишемо формулу для функції x_k , що залежить від змінних μ_k та ε_{ik} :

$$x_k(\mu_k, \varepsilon_{ik}) = \frac{\mu_k b_i A^{<k>}_{ik} + D^{<k>}}{\varepsilon_{ik} a_{ik} A_{ik} + D}. \quad (11)$$

Знайдемо часткову похідну розв'язку x_k за змінною μ_k при $\varepsilon_{ik} = 0$, вона буде дорівнювати:

$$\frac{\partial x_k(\mu_k, \varepsilon_{ik})}{\partial \mu_k} = \frac{b_i A^{<k>}_{ik}}{D}. \quad (12)$$

Головна лінійна частина приросту розв'язку x_k в точці $\varepsilon_{ik} = 0$, $\mu_k = 0$ (його повний диференціал) з урахування лише стовпця $j = k$ буде дорівнювати:

$$\Delta_{\text{лін}}^{<k>} = \sum_{i=1}^n \left(\frac{b_i A^{<k>}_{ik}}{D} \Delta \mu_k - \frac{a_{ik} A_{ik} D^{<k>}}{D^2} \Delta \varepsilon_{ik} \right), \quad (13)$$

де $\Delta \mu_k$ – приріст змінної μ_k ;

$\Delta \varepsilon_{ik}$ – приріст змінної ε_{ik} .

Вважаючи, як і раніше, що для всіх i та j ($j = k$) дотримується нерівність $|\varepsilon_{ij}| \leq \varepsilon_0$, отримаємо, що найбільше збільшення розв'язку в лінійному наближенні буде при $|\varepsilon_{ij}| = \varepsilon_0$, і воно дорівнює

$$\Delta_{\text{лін макс}}^{<k>} = \sum_{i=1}^n \left(D + \frac{|a_{ik} A_{ik} D^{<k>}|}{D^2} \right) \cdot \varepsilon_0. \quad (14)$$

Звідси отримаємо, що кожен елемент ε_{ij} у випадку $j = k$ повної таблиці знаків («+» або «-») має протилежний знак від знаку добутку $a_{ik} A_{ik} D^{<k>}$, та кожен елемент $\Delta \mu_k$ має знак, співпадаючий зі знаком виразу $\frac{b_i A^{<k>}_{ik}}{D}$, та знаки дорівнюють нулю, якщо відповідні вирази дорівнюють нулю.

Програмне забезпечення дослідження стійкості розв'язків СЛАР. Для можливості порівняння результатів було створено програмне забезпечення, що дозволяє користувачеві досліджувати стійкість СЛАР, що представлені в матричному вигляді (1), інтерфейс якого зображено на рис. 1.

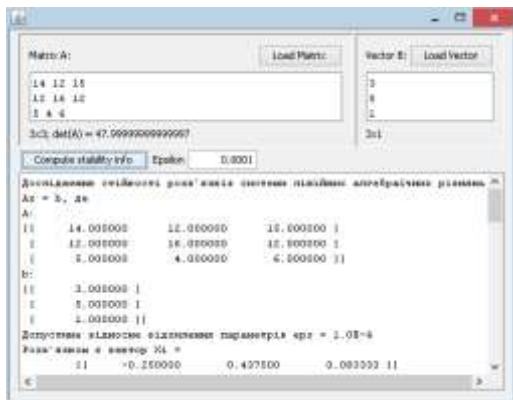


Рис. 1. Вікно програми

Користувач може ввести матрицю системи A та вектор правих частин b . При цьому введення можливо як з клавіатури, так и шляхом їх завантаження за допомогою кнопок «Load Matrix» та «Load Vector». В нижній частині вікна передбачено виведення звіту з результатами дослідженням стійкості.

Звіт з дослідженням стійкості містить:

- розв'язки системи лінійних алгебраїчних рівнянь;
- можливі варіації розв'язків СЛАР;
- можливі варіації значень визначників, які використовуються у розрахунках;
- детальну інформацію щодо умов, при яких розв'язки або визначники змінюють знак.

Чисельний експеримент з дослідженням поліпшеного методу побудови таблиці знаків. Для порівняння результатів дослідження стійкості систем лінійних алгебраїчних рівнянь методом побудови таблиці знаків за визначниками та побудови таблиці знаків за похідними ділення двох визначників згідно формули Крамера були проаналізовані випадково створені системи лінійних алгебраїчних рівнянь.

Графік з порівнянням розподілу відносних відхилень розв'язків x_i , що залежать від відносних відхилень параметрів ε за двома методами наведений на рис. 2, та графік розподілу уточнень оцінок за похідними ділення двох визначників згідно формули Крамера для кожного розв'язку x_i для СЛАР розмірності $n = 5$ наведений на рис. 3.

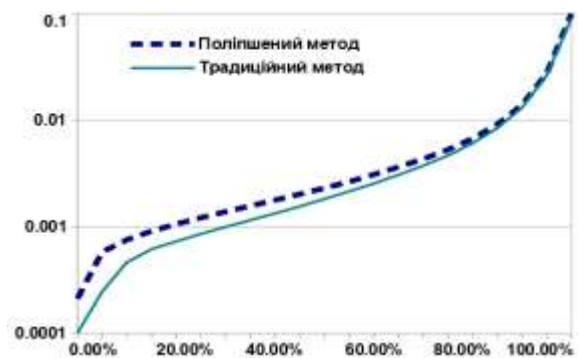


Рис. 2. Розподіл відносних відхилень розв'язків x_i

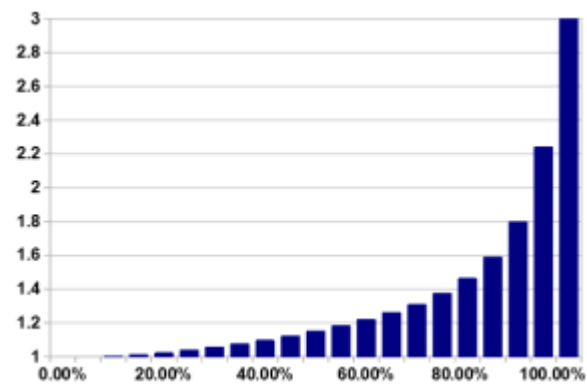


Рис. 3. Розподіл уточнень оцінок

Стосовно уточнень оцінок кожного окремого розв'язку можемо бачити, що уточнення для 50 % розв'язків не перевищують 1.15 разів, уточнення для

30 % розв'язків знаходяться у діапазоні між 1.15 та 1.46, уточнення для 15 % розв'язків знаходяться у діапазоні між 1.46 та 2.24, та уточнення для 5 % розв'язків перевищують 2.24 разів. Максимальне уточнення дорівнює 127.43, що може бути критичним у деяких випадках.

Висновки. В роботі був проведений огляд існуючих методів дослідження стійкості розв'язків систем лінійних алгебраїчних рівнянь, що залежать від вхідних даних, тобто варіацій параметрів СЛАР. Були реалізовані методи оцінки стійкості систем лінійних алгебраїчних рівнянь за допомогою чисел обумовленості, модульних визначників та побудови таблиці знаків для знаходження точних оцінок варіацій розв'язків, залежних від варіацій параметрів СЛАР.

За допомогою розробленого програмного забезпечення був проведений порівняльний аналіз стійкості розв'язків систем лінійних рівнянь за чотирма розглянутими методами. Показано, що дослідження стійкості за числами обумовленості дають дуже грубу оцінку можливих похибок розв'язків.

В ході дослідження також отримано, що запропонований поліпшений метод побудови таблиці знаків у 30 % випадків знаходить варіації, що в 1.3 рази більші ніж варіації, що знаходить класичний метод, та у 5 % випадків ці варіації перевищують попередні у 2 або більше разів. Результати роботи також доповідались на конференції [14].

Список літератури

1. Кутнів М. В. *Чисельні методи*. Львів: НУ «Львівська політехніка», 2008. 200 с.
2. Шахно С. М. *Чисельні методи лінійної алгебри*. Львів: ВЦ ЛНУ ім. І. Франка, 2007. 245 с.
3. Красс М. С., Чупрынов Б. П. *Математика для экономистов*. Санкт-Петербург: Питер, 2005. 464 с.
4. *Высшая математика для экономистов* / ред. проф. Н. Ш. Кремера. Москва: ЮНИТИ-ДАНА, 2007. 471 с.
5. Красс М. С. *Математика для экономических специальностей*. Москва: ИНФРА-М, 1998. 464 с.
6. Ильина В. А., Силаев П. К. *Численные методы для физиков-теоретиков*. Москва-Ижевск: Институт компьютерных исследований, 2003. 132 с.
7. Кудрявцев Л. Д. О некоторых математических вопросах теории электрических цепей // *Успехи математических наук*. 1948. Т. 3, № 4(26). С. 80–118.
8. Коваленко А. А. *Основы линейной алгебры*. Барнаул: Изд-во АлтГПА, 2010. 118 с.
9. Гун Г. Я. *Математическое моделирование процессов обработки металлов давлением*. Москва: Металлургия, 1983. 352 с.
10. Петров Ю. П. *Как получить надежные решения систем уравнений*. Санкт-Петербург: БХВ-Петербург, 2012. 176 с.
11. Петров Ю. П. *Обеспечение достоверности и надежности компьютерных расчетов*. Санкт-Петербург: БХВ-Петербург, 2008. 160 с.
12. Уоткинс Д. С. *Основы матричных вычислений*. Москва: БИНОМ. Лаборатория знаний, 2006. 664 с.
13. Деммель Дж. *Вычислительная линейная алгебра. Теория и приложения*. Москва: Мир, 2001. 430 с.
14. Руденко Р. О. Розробка програмного забезпечення для дослідження стійкості чисельних методів лінійної алгебри. *Труди 23-го міжнародного молодіжного форуму «Радіоелектроніка та молодь у XXI столітті»*. Т. 9. Харків: ХНУРЕ, 2019. С. 130–131.

References (transliterated)

1. Kutniv M. V. *Chysel'ni metody* [Numerical methods]. Lviv, Lviv Polytechnic National University Publ., 2008. 200 p.
2. Shakhno S. M. *Chysel'ni metody liniynoyi algebrы* [Numerical methods of linear algebra]. Lviv, VTS LNU im. I. Franko Publ., 2007. 245 p.
3. Krass M. S., Chuprynov B. P. *Matematika dlya ekonomistov* [Mathematics for Economists]. St. Petersburg, Piter Publ., 2005. 464 p.
4. *Vyssshaya matematika dlya ekonomistov* [Higher Mathematics for Economists] N. Sh. Kremer ed. Moscow, UNITY-DANA Publ., 2007. 471 p.
5. Krass M. S. *Matematika dlya ekonomicheskikh spetsial'nostey* [Mathematics for economic specialties]. Moscow, INFRA-M Publ., 1998. 464 p.
6. P'ina V. A., Silayev P. K. *Chislennyye metody dlya fizikov-teoretikov*. [Numerical methods for theoretical physicists]. Moscow-Izhevsk, Institute of Computer Science Publ., 2003. 132 p.
7. Kudryavtsev L. D. *O nekotorykh matematicheskikh voprosakh teorii elektricheskikh tsepey* [On Some Mathematical Questions in the Theory of Electrical Circuits]. *Uspekhi Matematicheskikh Nauk*. 1948, vol. 3, no. 4 (26), pp. 80–118.
8. Kovalenko A. A. *Osnovy lineynoy algebrы* [Basics of linear algebra]. Barnaul, AltGPA Publ., 2010. 118 p.
9. Gun G. Ya. *Matematicheskoye modelirovaniye protsessov obrabotki metallov davleniyem* [Mathematical modeling of metal forming processes]. Moscow, Metallurgy Publ., 1983. 352 p.
10. Petrov Yu. P. *Kak poluchat' nadezhnyye resheniya sistem uravneniy* [How to obtain reliable solutions of systems of equations]. St. Petersburg, BHV-Petersburg Publ., 2012. 176 p.
11. Petrov Yu. P. *Obespecheniye dostovernosti i nadezhnosti komp'yuternykh raschetov* [Ensuring the reliability and reliability of computer calculations]. St. Petersburg, BHV-Petersburg Publ., 2008. 160 p.
12. Watkins David S. *Fundamentals of Matrix Computations*. 2nd Edition. New York, John Wiley & Sons, Inc., 2002 (Russ. ed.: Watkins D. S. *Osnovy matrichnykh vychisleniy*. Moscow, BINOM; Laboratory of Knowledge Publ., 2006. 664 p.
13. Demmel James W. *Applied Numerical Linear Algebra*. Philadelphia, Society for Industrial and Applied Mathematics, 1997 (Russ. ed.: Demmel' Dzh. *Vychislitel'naya lineynaya algebra. Teoriya i prilozheniya*. Moscow, Mir Publ., 2001. 430 p.
14. Rudenko R. O. *Rozrobka prohrannoho zabezpechennya dlya doslidzhennya stiykosti chysel'nykh metodiv liniynoyi algebrы* [Software development for studying the stability of numerical methods of linear algebra]. *Trudy 23-ho mizhnarodnoho molodizhnoho forumu «Radioelektronika ta molod' u KHKHI stolitti»*. Т. 9. [Proc. of the 23rd International Youth Forum "Radio Electronics and Youth in the 21st Century". Vol. 9]. Kharkiv, KNURE Publ., 2019, p. 130–131.

Надійшла (received) 13.05.2019

Відомості про авторів / Сведения об авторах / About the Authors

Марченко Наталія Андріївна (Марченко Наталья Андреевна, Marchenko Natalya Andriivna) – кандидат технічних наук, доцент, Національний технічний університет «Харківський політехнічний інститут», доцент кафедри «Системний аналіз та інформаційно-аналітичні технології»; м. Харків, Україна; ORCID: <https://orcid.org/0000-0001-9889-3713>; e-mail: mna123@ukr.ua

Руденко Роман Олександрович (Руденко Роман Александрович, Rudenko Roman Oleksandrovych) – Національний технічний університет «Харківський політехнічний інститут», магістрант; м. Харків, Україна; ORCID: <https://orcid.org/0000-0002-9424-6639>; e-mail: roman.rudenko.a@gmail.com

I. I. МАРЧЕНКО, М. М. МАЛЬКО, І. Г. МАРЧЕНКО

КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ ПРОЦЕСІВ ДИФУЗІЇ У ПОХИЛИХ ПРОСТОРОВО-ПЕРІОДИЧНИХ ПОТЕНЦІАЛАХ

Нещодавно було показано, що в істотно нерівноважних системах коефіцієнт дифузії може вести себе немонотонно з температурою. Одним із прикладів таких систем з аномальною температурною залежністю є рух броунівських часток в просторово-періодичних структурах. Метою статті було дослідження зміни температурної залежності дифузії в недодемпфованих системах з низьким коефіцієнтом тертя. В роботі методами комп'ютерного моделювання вивчено зміна коефіцієнта дифузії частинок в широкому діапазоні температур в нахилених просторово-періодичних потенціалах для різних значень коефіцієнта тертя. Показано, що дифузія досягає максимуму при певній величині зовнішньої сили. Її значення залежить від величини коефіцієнта тертя. Показано, що на відміну від звичайної залежності Аррениуса, в разі нахиленого періодичного потенціалу, максимальний коефіцієнт дифузії зростає, а не зменшується з пониженням температури експоненціальним чином. Встановлено, що така залежність характерна для всіх недодемпфованих систем. Показано, що для просторово-періодичних структур існує обмежена ділянка сил, в якій спостерігається зростання коефіцієнта дифузії з зменшенням температури. Це область так званої температурно-аномальної дифузії (ТАД). Визначено ширину і положення області ТАД в залежності від коефіцієнта тертя γ і параметрів системи. Показано, що зі зменшенням γ , ширина області ТАД зменшується пропорційно γ . При цьому коефіцієнт дифузії в області ТАД, навпаки зростає $\sim \gamma$. Отримані дані про температурно-аномальну дифузію мають важливе значення для різних областей фізики і техніки та відкривають перспективи новітніх технологій управління процесами дифузії.

Ключові слова: дифузія, комп'ютерне моделювання, періодичні структури, рівняння Ланжевена, періодичні поля

И. И. МАРЧЕНКО, М. Н. МАЛЬКО, И. Г. МАРЧЕНКО

КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ ПРОЦЕССОВ ДИФфузии В НАКЛОННЫХ ПРОСТРАНСТВЕННО-ПЕРИОДИЧЕСКИХ ПОТЕНЦИАЛАХ

Недавно было показано, что в существенно неравновесных системах коэффициент диффузии может вести себя немонотонно с температурой. Одним из примеров таких систем с аномальной температурной зависимостью является движение броуновских частиц в пространственно-периодических структурах. Целью статьи было исследование изменения температурной зависимости диффузии в недодемпфированных системах с низким коэффициентом трения. В работе методами компьютерного моделирования изучено изменение коэффициента диффузии частиц в широком диапазоне температур в наклонных пространственно-периодических потенциалах для различных значений коэффициента трения. Показано, что диффузия достигает максимума при определенной величине внешней силы. Ее значение зависит от величины коэффициента трения. Показано, что в отличие от обычной зависимости Аррениуса, в случае наклонного периодического потенциала, максимальный коэффициент диффузии возрастает, а не убывает с понижением температуры экспоненциальным образом. Установлено, что такая зависимость характерна для всех недодемпфированных систем. Показано, что для пространственно-периодических структур существует ограниченный участок сил, в котором наблюдается рост коэффициента диффузии с уменьшением температуры. Это область так называемой температурно-аномальной диффузии (ТАД). Определены ширина и положение области ТАД в зависимости от коэффициента трения γ и параметров системы. Показано, что с уменьшением γ , ширина области ТАД уменьшается $\sim \gamma$. При этом коэффициент диффузии в области ТАД, наоборот возрастает пропорционально γ . Полученные данные о температурно-аномальной диффузии имеют важное значение для различных областей физики и техники и открывают перспективы создания новейших технологий управления процессами диффузии.

Ключевые слова: диффузия, компьютерное моделирование, периодические структуры, уравнения Ланжевена, периодические поля

I. I. MARCHENKO, M. M. MALKO, I. G. MARCHENKO

COMPUTER SIMULATION OF DIFFUSION PROCESSES IN TILT SPATIO-PERIODIC POTENTIALS

It was recently shown that in essentially nonequilibrium systems, the diffusion coefficient can behave nonmonotonically with temperature. One example of such systems with anomalous temperature dependence is the motion of Brownian particles in spatially periodic structures. The aim of the article was to study the change in the temperature dependence of diffusion in underdamped systems with a low coefficient of friction. In this paper, computer simulation methods are used to study the change in the diffusion coefficient of particles in a wide range of temperatures in oblique spatially periodic potentials for different values of the friction coefficient. It is shown that diffusion reaches a maximum at a certain external force. Its value depends on the coefficient of friction. It is shown that, in contrast to the usual Arrhenius dependence, in the case of an inclined periodic potential, the maximum diffusion coefficient increases while temperature is decreasing exponentially. It is established that such a dependence is common to all underdamped systems. It is shown that for spatially periodic structures there is a limited portion of forces in which an increase in the diffusion coefficient while decreasing temperature is observed. This is the area of the so-called temperature-anomalous diffusion (TAD). The width and position of the TAD region are determined depending on the friction coefficient γ and the system parameters. It has been shown that a decrease in γ , width TAD region decreases proportionally γ . In this case, the diffusion coefficient in the TAD region, on the contrary, increases $\sim \gamma$. The data obtained on the temperature and the anomalous diffusion are important for various fields of physics and engineering, and opens new prospects for a diffusion process control technology.

Keywords: diffusion, computer simulation, periodic structures, Langevin equations, periodic fields

Вступ. У останні роки спостерігається зростаючий інтерес до експериментального вивчення прискорення дифузії часток шляхом використання зовнішнього поля [1–4]. Змінюючи характеристики поля можна ефективно впливати на процеси дифузії. Це відкриває нові технологічні можливості керування дифузією.

Перші теоретичні дослідження руху Броунівських часток у похилих періодичних потенціалах були

виконані Х. Ріскеном [5]. Було показано, що для недодемпфованого випадку важливим у поведінці ансамблю частинок є поява «локалізованих» та «бігучих» рішень. Х. Ріскеном були отримані вирази для рухливості частинок. У той же час поведінку коефіцієнту дифузії не було досліджено.

Це певною мірою пов'язано з тим що, що аналітичні методи рішення рівняння Фокера-Планка, які використовувалися авторами, погано пристосовані до

систем з малою дисипацією. Альтернативним способом вивчення процесів дифузії і транспорту частинок є пряме чисельне моделювання ансамблю частинок за допомогою стохастичних рівнянь Ланжевена. Ф. Марчезоні [6] було встановлено, що дифузія частинок в системах з малою дисипацією істотно зростає поблизу деякої критичної сили. Подальше вивчення дифузії під впливом постійної сили було пов'язано з роботами групи К. Лінденберг. У роботі [7] вони вперше показали, що в похилих періодичних потенціалах коефіцієнт дифузії може вести себе аномальним чином. При деякому значенні сили він зростає з пониженням температури T . Температурну залежність максимального коефіцієнта дифузії вони апроксимували ступеневою залежністю: $D_{\max} \sim T^{-3.5}$. Подальші дослідження [8] показали, що така апроксимація справедлива лише в дуже вузькому діапазоні температур. Було встановлено, що в недодемпфованих системах існує обмежений інтервал діючих сил, в якому коефіцієнт дифузії зростає з пониженням температури експоненціальним чином ($D \sim \exp(|\varepsilon|/kT)$). Було показано, що фізичною причиною такого зростання є експоненціальне зростання кореляційного часу τ_{cor} із зниженням температури. У [9] була побудована феноменологічна модель, що пояснює таку температурну залежність τ_{cor} . Інтервал сил, в якому дифузія зростає зі зменшенням температури, був нами названий областю температурно-аномальної дифузії (ТАД) [10]. І. Соколов і Б. Лінднер [11], провівши чисельне моделювання для ряду різних значень коефіцієнтів тертя, підтвердили наші висновки [8, 9] про існування ТАД в обмеженому інтервалі прикладених сил (кордонах ТАД).

Однак, до теперішнього часу відсутні дані щодо посилення дифузії в області низьких значень коефіцієнтів тертя і температур. Метою даної роботи є дослідження залежності коефіцієнтів дифузії від температури і тертя методами комп'ютерного моделювання у похилих просторово-періодичних потенціалах.

Методика моделювання. Рух частинок на одновимірній решітці під дією зовнішньої сили F описується рівнянням Ланжевена:

$$m\ddot{x} = -\frac{dU}{dx} - \gamma\dot{x} + F + \sqrt{2\gamma kT}\xi(t), \quad (1)$$

де t – час, x – координата частинки в одновимірній решітці, m – її маса, γ – коефіцієнт тертя і $\xi(t)$ – білий Гаусів шум з одиничною інтенсивністю. Крапка зверху означає диференціювання по часу.

Потенційна енергія U частинки в одновимірній періодичній решітці дорівнює:

$$U(x) = -\frac{U_0}{2} \cos\left(\frac{2\pi}{a}x\right), \quad (2)$$

де a – постійна решітки, а U_0 – висота потенційного бар'єру. Параметри використовуваного просторово-періодичного потенціалу були тими ж, що і в роботах [8,9]: $U_0 = 0.08$ еВ, $a = 2.0$ Å. Маса частинок відповідала масі водню і дорівнювала 1 атомній одиниці маси.

На рухому частку діє періодична сила з боку решітки: $F_{\text{lat}} = -\frac{dU}{dx} = F_0 \sin\left(\frac{2\pi}{a}x\right)$. Величина $F_0 = \frac{\pi}{a}U_0$

називається критичною силою [12, 13], відповідає мінімальній діючій силі, необхідній для подолання в в'язкому середовищі енергетичного бар'єру, що розділяє два сусідніх положення частинки на одновимірній решітці.

Стохастичні рівняння (1) – (2) для кожної частки вирішувалися чисельно методом Ейлера [14] з кроком по часу що становив менше 0.01 періоду власних малих коливань. Статистичне усереднення проводилося по ансамблю з кількістю частинок не менше $N = 10^5$. Початкові умови задавалися наступним чином: частка розміщувалася на початку координат і їй випадковим чином присвоювалася швидкість. Швидкість мала розподіл Максвелла по температурі. Для досягнення рівноваги як за швидкостями, так і за координатами, проводилася термалізація системи протягом 10^4 кроків по часу. Як показали розрахунки, після цього часу розподіл, як за координатами, так і за швидкостями ансамблю частинок не змінювалося. В процесі термалізації частки могли здійснювати стрибок в сусідні елементарні комірки одновимірної решітки. Для того щоб дифузія частинок походила з початку координат такі частинки переміщувалися в першу елементарну комірку шляхом трансляції на цілу кількість постійних решітки.

Коефіцієнт дифузії обчислювався по дисперсії в розподілі ансамблю частинок. Звичайна дифузія характеризується лінійною залежністю середньоквадратичного відхилення частинок у часі:

$$\langle \sigma^2 \rangle \propto \alpha t. \quad (3)$$

У разі реалізації особливих режимів дифузії ця залежність змінюється:

$$\langle \sigma^2 \rangle \propto \alpha t^\alpha. \quad (4)$$

При $\alpha < 1$ говорять про субдифузію, а при $\alpha > 1$ – про супердифузію. Як було показано раніше, в системах з низьким тертям особливі режими дифузії носять перехідний у часі характер [15, 16] і, після встановлення стаціонарного стану, дисперсія в розподілі часток описується стандартним виразом (1). На рис. 1 показано стандартний графік зміни дисперсії з часом ансамблю частинок під дією зовнішньої сили. З рисунку видно, що режим «звичайної» дифузії настає на великих проміжках часу. Таким чином, для коректного обчислення коефіцієнта дифузії необхідний досить великий проміжок часу протягом якого $\langle \sigma^2 \rangle \propto \alpha t$.

За визначенням в 1-мірній системі коефіцієнт дифузії дорівнює:

$$D = \lim_{t \rightarrow \infty} \frac{\langle (x - \langle x \rangle)^2 \rangle}{2t} = \lim_{t \rightarrow \infty} \frac{\sigma^2}{2t}, \quad (5)$$

де дужки $\langle \dots \rangle$ позначають усереднення по ансамблю. Для коректного обчислення коефіцієнта дифузії визначався час t_{lin} досягнення лінійної залежності дисперсії від часу. Коефіцієнт дифузії визначався лінійною апроксимацією дисперсії від часу при $t > 100t_{\text{lin}}$.

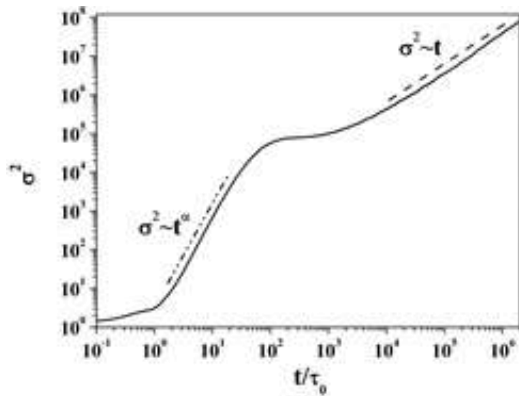


Рис. 1. Залежність дисперсії ансамблю частинок від часу.
Діюча сила $F = 0.11F_0$, $\gamma = 3 \times 10^{-2}$, $T' = 0.39$

Для зіставлення даних, отриманих в даній роботі, з результатами інших авторів, ми при поданні результатів моделювання використовували безрозмірні значення температури T' та коефіцієнтів тертя γ' [7, 12]:

$$T' = \frac{Tk}{U_0}; \gamma' = \gamma \frac{a}{\pi\sqrt{2m}U_0} \quad (6)$$

Результати і обговорення. Нами було проведено чисельне моделювання рівнянь (1) для різних значень коефіцієнтів тертя і температур. На рис. 2 представлено залежності коефіцієнтів дифузії від сили. Наведені три групи графіків, відповідають різним коефіцієнтам тертя. Літерне позначення кривих відповідає різним температурам. З груп графіків видно, що положення і ширина області ТАД ($\Delta_{ТАД}$) залежить від γ' . Як впливає з рисунка, ці величини лінійно зменшуються з γ' . Детальний аналіз усіх графіків показує, що при цьому максимальне значення D лінійно зростає зі зменшенням γ' .

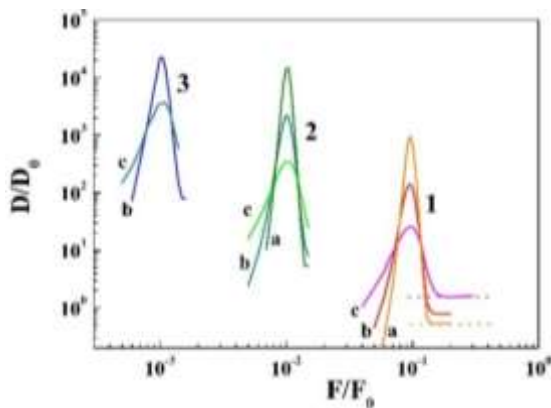


Рис. 2. Залежності коефіцієнтів дифузії від сили для різних $\gamma' = 3 \cdot 10^{-2}$, $3 \cdot 10^{-3}$ та $3 \cdot 10^{-4}$ – відповідно групи кривих 1–3.
Температури $T'_1 = 0.13$, $T'_2 = 0.19$, $T'_3 = 0.39$

Щоб зрозуміти фізичні причини такої поведінки проаналізуємо зміну функції розподілу часток за швидкостями $n(V)$ зі зміною γ' . На рис. 3 наведено приклади графіки функції розподілу часток за швидкостями для двох значень сил ($F_1 = 0.1F_0$ та $F_2 = 0.01F_0$) при різних значеннях γ' , які так само відрізняються на порядок ($\gamma' = 3 \cdot 10^{-3}$ та $\gamma' = 3 \cdot 10^{-4}$).

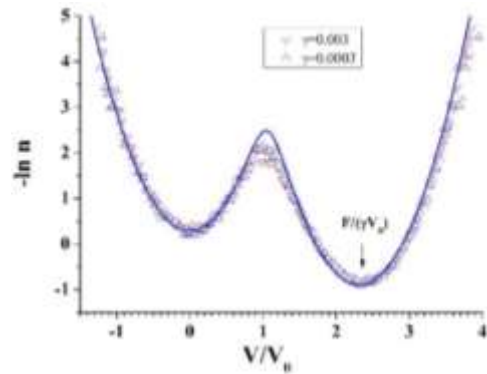


Рис. 3. Залежність функції розподілу часток за швидкостями для різних γ при значенні сили з максимальним $D(F)$. $T' = 0.19$, $\nabla - \gamma = 0.003$, $\Delta - \gamma = 0.0003$. Суцільною лінією нанесені розрахункові значення виходячи з моделі двух'ямний потенціалу [9]. Функція має мінімуми при $V = 0$ і $V = F/\gamma$

Як видно при використаних значеннях F і γ функції розподілу часток за швидкостями збігаються. На рисунку величина $V_0 = \sqrt{U_0/m}$.

З аналізу графіків, наведених на рис. 3 і інших графіків $n(V)$, слід, що при малих γ функція розподілу за швидкостями залежить тільки від співвідношення F/γ . Вперше це було відзначено в монографії Х. Рискена (див. рис. 11.22–11.22а [5]). З цього факту можна отримати скейлінгове залежності, як для рухливості, так і для коефіцієнтів дифузії. У стаціонарному випадку, якщо відома функція розподілу частинок за швидкостями $n(V)$, можна знайти рухливість частинок: $\mu(F; \gamma) = \langle V \rangle(F; \gamma)/F$, де $\langle V \rangle(F; \gamma) = \int_{-\infty}^{\infty} V n(V; F, \gamma) dV$. Оскільки $n(V; F, \gamma) = n(V, F/\gamma)$, то для рухливості має виконуватися співвідношення:

$$\gamma_1 \mu(F_1; \gamma_1) = \gamma_2 \mu(F_1 \gamma_2 / \gamma_1; \gamma_2) \quad (7)$$

На рис. 4 показані залежності $\gamma\mu$ від безрозмірної величини $f = F/(\gamma V_F)$ для різних значень γ . Величина V_F відповідає швидкості, при якій кінетична енергія частинки дорівнює половині енергетичного бар'єру $U(F)$ для переходу її з одного положення з мінімальною потенційною енергією в сусіднє на одновимірній решітці. Збіг графіків $\gamma\mu(\gamma, F/(\gamma V_F))$ для різних значень γ підтверджує те, що скейлінгове співвідношення (5) виконується для систем з низьким тертям. Аналогічно з рухливістю, скейлінгове залежності можуть бути отримані і для коефіцієнта дифузії. Відповідно до співвідношенням Кубо коефіцієнт дифузії може бути отриманий з автокореляційної функції [5]. У стаціонарному випадку коефіцієнт дифузії може бути розрахований наступним чином [17]:

$$D = \frac{1}{Q} \int_{-\infty}^{\infty} \left[\int_{-\infty}^V (u - \langle u \rangle) n(u) du \right]^2 / n(V) dV, \quad (8)$$

де Q – інтенсивність теплового шуму в просторі швидкостей. Аналогічно з рухливістю, виходячи з виду залежності $n(V; F, \gamma) = n(V; F/\gamma)$, легко показати, що

для коефіцієнтів дифузії в разі малого тертя виконується наступне співвідношення:

$$\frac{D_1(F_1; \gamma_1)}{\gamma_1} = \frac{D_2\left(\frac{F_1 \gamma_2}{\gamma_1}; \gamma_2\right)}{\gamma_2} \quad (9)$$

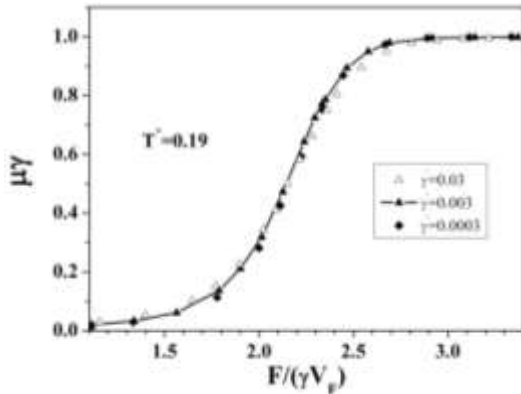


Рис. 4. Рухливість помножена на коефіцієнт тертя в залежності від величини $F/(\gamma V_F)$ для різних γ

На рис. 5 наведені графіки залежності коефіцієнтів дифузії від діючої сили для різних коефіцієнтів тертя, що відрізняються на порядок.

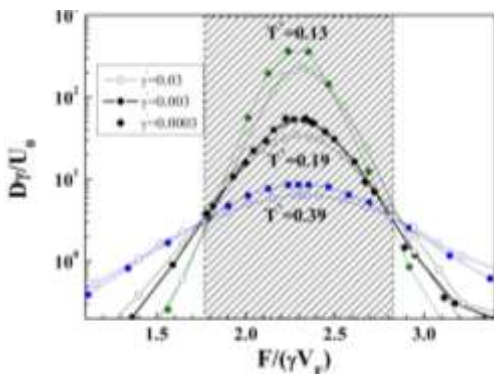


Рис. 5. Залежності коефіцієнтів дифузії від діючої сили для різних γ

Моделювання проводилося при трьох різних температурах. Під дією сили змінюється величина бар'єра $U(F)$. Однак, зі зменшенням γ , величини сил, що обмежують область ТАД так само зменшуються і, відповідно, при прагненні коефіцієнта тертя до нуля $U(F) \rightarrow U_0$.

На рис. 5 область ТАД, в якій коефіцієнт дифузії зростає з пониженням температури, виділений штрихуванням. З графіків видно, що скейлінгове співвідношення (7) виконується з хорошою точністю. Невелике розходження в даних для $\gamma' = 3 \cdot 10^{-2}$ і $\gamma' = 3 \cdot 10^{-3}$ пов'язане з згадуваним вище зміною $U(F)$. Тому відмінності в графіках нівелюються зі зменшенням γ' . Як видно з графіків криві які відповідають $\gamma' = 3 \times 10^{-3}$ і $\gamma' = 3 \cdot 10^{-4}$ дійсно практично збігаються.

На рис. 6 показано поведінку максимального значення коефіцієнта дифузії D_{\max} в залежності від температури. Видно, що при низьких температурах

коефіцієнт дифузії експоненціально зростає зі зворотною температурою. Це, як було показано раніше [9], визначається зростанням кореляційного часу. Як випливає з наведених на рис. 6 даних при $\gamma < 0.03$ графіки залежностей $D\gamma/U_0$ не залежать від γ .

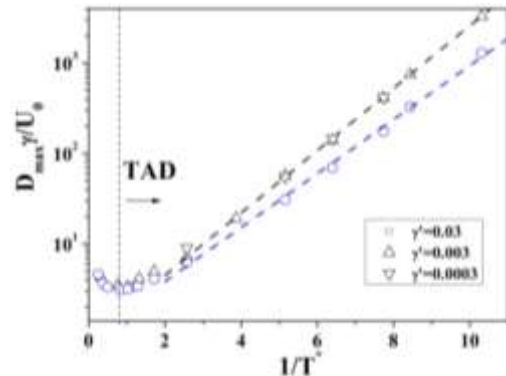


Рис. 6. Залежність коефіцієнта дифузії від зворотної температури в зоні ТАД: $\gamma = 3 \cdot 10^{-2}$, $F/F_0 = 0,1$. Стрілкою показаний інтервал температур в якому коефіцієнт дифузії зростає з пониженням температури

Таким чином, з аналізу даних комп'ютерного моделювання випливає, що при малих γ ширина області ТАД лінійно зменшується з коефіцієнтом тертя: $\Delta F_{\text{TAD}} \approx \gamma V_F$. Однак, максимальне значення D в цій області лінійно зростає зі зменшенням γ . Центр ТАД розташований поблизу сили $F \approx 2\gamma V_F$.

Висновки. У роботі досліджена дифузія частинок під дією зовнішніх сил в просторово-періодичних системах, що характеризуються малими значеннями коефіцієнта тертя γ . Показано, що у всіх недодемпфированих періодичних системах існує обмежена область температурно-аномальної дифузії (ТАД). У цій області сил коефіцієнт дифузії зростає з пониженням температури.

Визначено ширину та положення області ТАД в залежності від γ . Показано, що зі зменшенням γ ширина області зменшується $\sim \gamma$. При цьому коефіцієнти дифузії в цій області, навпаки зростають $\sim \gamma$.

Отримані залежності важливі для експериментального виявлення явища температурно-аномальної дифузії і подальшого його використання в різних галузях фізики, хімії та біології.

Список літератури

1. Lee S.-H., Grier D.G. Giant Colloidal Diffusivity on Corrugated Optical Vortices. *Phys. Rev. Let.* 2006, vol. 96. P. 190601.
2. Tierno P., Reimann P., Johansen T.H., Sagués F. Giant transversal particle diffusion in a longitudinal magnetic ratchet. *Phys. Rev. Let.* 2010, vol. 105. P. 230602.
3. Eshuis P., van der Weele K., Lohse D., van der Meer D. Experimental Realization of a Rotational Ratchet in a Granular Gas. *Phys. Rev. Let.* 2010, vol. 104. P. 248001.
4. Pagliara S., Schwall C., Keyser U.F. Optimizing Diffusive Transport Through a Synthetic Membrane Channel. *Advanc. Mat.* 2013, vol. 25. P. 844.
5. Risken H. *The Fokker-Planck Equation and Methods of Solution and Applications*. Springer, 1989. 472 p.
6. Costantini G., Marchesoni F. Threshold diffusion in a tilted washboard potential. *Europhys. Lett.* 1999, vol. 48. P. 491–497.
7. Lindenberg K., Lacasta A.M., Sancho J.M., Romero A.H. Transport and diffusion on crystalline surfaces under external forces. *New Jour. of Phys.* 2005, vol. 7. P. 29.

8. Marchenko I.G., Marchenko I.I. Diffusion in the systems with low dissipation: Exponential growth with temperature drop. *Europhysics Letters*. 2012, vol. 100. P. 5005.
9. Marchenko I.G., Marchenko I.I., Zhiglo A.V. Particle transport in space-periodic potentials in underdamped systems. *Eur. Phys. Jour.* 2014, vol. B87. P. 10.
10. Marchenko I.G., Marchenko I.I., Zhiglo A.V. Enhanced diffusion with abnormal temperature dependence in underdamped space-periodic systems subject to time-periodic driving. *Phys. Rev.* 2018, vol. E97. P. 012121 (16 pp.).
11. Lindner B., Sokolov I.M. Giant diffusion of underdamped particles in a biased periodic potential. *Phys. Rev.* 2016, vol. E93. P. 042106.
12. Lindenberg K., Sancho J.M., Lacasta A.M., Sokolov I.M. Dispersionless Transport in a Washboard Potential *Phys. Rev. Lett.* 2007, vol. 98. P. 020602.
13. Reimann, P., Van den Broeck C., Linke H., Hänggi P., Rubí J. M., Pérez-Madrid A. Diffusion in tilted periodic potentials: Enhancement, universality, and scaling. *Phys. Rev.* 2002, vol. E65. P. 031104.
14. Кузнецов Д.Ф. *Стохастические дифференциальные уравнения: теория и практика численного решения*. Санкт-Петербург: Политехнический университет, 2007. 769 с.
15. Lindenberg K., Sancho J.M., Lacasta A.M., Sokolov I.M. Dispersionless Transport in a Washboard Potential. *Phys. Rev. Lett.* 2007, Vol. 98. P. 020602.
16. Marchenko I.G., Marchenko I.I. Anomalous Temperature Dependence of Diffusion in Crystals in Time-Periodic External Fields. *JETP Letters*. 2012, vol. 95, #3. P. 137–142.
17. Lindner B., Nicola E.M. Critical Asymmetry for Giant Diffusion of Active Brownian Particles. *Phys. Rev. Lett.* 2008, vol. 101. P. 190603 (4 pp.).
4. Pagliara S., Schwall C., Keyser U.F. Optimizing Diffusive Transport Through a Synthetic Membrane Channel. *Advanc. Mat.* 2013, vol. 25. P. 844.
5. Risken H. *The Fokker-Planck Equation and Methods of Solution and Applications*. Springer, 1989. 472 p.
6. Costantini G., Marchesoni F. Threshold diffusion in a tilted washboard potential. *Europhys. Lett.* 1999, vol. 48. P. 491–497.
7. Lindenberg K, Lacasta A.M., Sancho J.M., Romero A.H. Transport and diffusion on crystalline surfaces under external forces. *New Jour. of Phys.* 2005, vol. 7. P. 29.
8. Marchenko I.G., Marchenko I.I. Diffusion in the systems with low dissipation: Exponential growth with temperature drop. *Europhysics Letters*. 2012, vol. 100. P. 5005.
9. Marchenko I.G., Marchenko I.I., Zhiglo A.V. Particle transport in space-periodic potentials in underdamped systems. *Eur. Phys. Jour.* 2014, vol. B87. P. 10.
10. Marchenko I.G., Marchenko I.I., Zhiglo A.V. Enhanced diffusion with abnormal temperature dependence in underdamped space-periodic systems subject to time-periodic driving. *Phys. Rev.* 2018, vol. E97. P. 012121 (16 pp.).
11. Lindner B., Sokolov I.M. Giant diffusion of underdamped particles in a biased periodic potential. *Phys. Rev.* 2016, vol. E93. P. 042106.
12. Lindenberg K., Sancho J.M., Lacasta A.M., Sokolov I.M. Dispersionless Transport in a Washboard Potential *Phys. Rev. Lett.* 2007, vol. 98. P. 020602.
13. Reimann, P., Van den Broeck C., Linke H., Hänggi P., Rubí J. M., Pérez-Madrid A. Diffusion in tilted periodic potentials: Enhancement, universality, and scaling. *Phys. Rev.* 2002, vol. E65. P. 031104.
14. Kuznetsov D.F. *Stokhasticheskiye differentsyalnye uravneniya: teoriya y praktika chyslenmoho resheniya* [Stochastic differential equations: theory and practice of numerical solution]. Sankt-Peterburh. Polytekhnicheskyyi unyversytet, 2007.– 769 p.
15. Lindenberg K., Sancho J.M., Lacasta A.M., Sokolov I.M. Dispersionless Transport in a Washboard Potential. *Phys. Rev. Lett.* 2007, Vol. 98. P. 020602.
16. Marchenko I.G., Marchenko I.I. Anomalous Temperature Dependence of Diffusion in Crystals in Time-Periodic External Fields. *JETP Letters*. 2012, vol. 95, #3. P. 137–142.
17. Lindner B., Nicola E.M. Critical Asymmetry for Giant Diffusion of Active Brownian Particles. *Phys. Rev. Lett.* 2008, vol. 101. P. 190603 (4 pp.).

References (transliterated)

1. Lee S.–H., Grier D.G. Giant Colloidal Diffusivity on Corrugated Optical Vortices. *Phys. Rev. Lett.* 2006, vol. 96. P. 190601.
2. Tierno P., Reimann P., Johansen T.H., Sagués F. Giant transversal particle diffusion in a longitudinal magnetic ratchet. *Phys. Rev. Lett.* 2010, vol. 105. P. 230602.
3. Eshuis P., van der Weele K., Lohse D., van der Meer D. Experimental Realization of a Rotational Ratchet in a Granular Gas. *Phys. Rev. Lett.* 2010, vol. 104. P. 248001.

Надійшло (received) 21.03.2019

Відомості про авторів / Сведения об авторах / About the Authors

Марченко Ігор Іванович (Марченко Игорь Иванович, Marchenko Igor Ivanovych) – кандидат технічних наук, доцент, Національний технічний університет «Харківський політехнічний інститут», доцент кафедри системного аналізу та інформаційно-аналітичних систем, м. Харків, Україна; ORCID: <https://orcid.org/0000-0002-3071-9169>; e-mail: igor.marchenko@hotmail.com

Малько Максим Миколайович (Малько Максим Николаевич, Malko Maksym Mykolayovych) – кандидат технічних наук, доцент, Національний технічний університет «Харківський політехнічний інститут», декан факультету комп'ютерних наук та програмної інженерії, м. Харків, Україна; ORCID: <https://orcid.org/0000-0002-0125-2141>; e-mail: maxim_malko@hotmail.com

Марченко Іван Григорович (Марченко Иван Григорьевич, Marchenko Ivan Gryhorovych) – доктор фізико-математичних наук, старший науковий співробітник, завідувач відділом Національного наукового центру «Харківський фізико-технічний інститут» м. Харків, Україна; ORCID: <https://orcid.org/0000-0003-1341-4950>; e-mail: marchkipt@gmail.com

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

INFORMATION TECHNOLOGY

UDC 004.8

DOI: 10.20998/2079-0023.2019.01.09

R. V. SHAPTALA, G. D. KYSELEV**USING GRAPH EMBEDDINGS FOR WIKIPEDIA LINK PREDICTION**

Link prediction is an important area of study in network analysis and graph theory which tries to answer the question of whether two nodes in the graph might have an association in the future. Nowadays, graphs are ubiquitously present in our lives (social networks, circuits, roads etc.), which is why the problem is crucial to the development of intelligent applications. In the past, there have been proposed methods of solving link prediction problem through algebraic formulations and heuristics, however, their expressive power and transferability fell short. Recently, graph embedding methods have risen to popularity because of their effectiveness and the ability to transfer knowledge between tasks. Inspired by the famous in machine learning and natural language processing research Word2Vec approach, these methods try to learn a distributed vector representation, called an embedding, of graph nodes. After that a binary classifier given a pair of embeddings predicts the probability of the existence of a link between the encoded nodes. In this paper, we review several graph embedding approaches for the problem of Wikipedia link prediction, namely Wikipedia2vec, Role2vec, AttentionWalk and Walklets. Wikipedia link prediction tries to find pages that should be interlinked due to some semantic relation. We evaluate prediction accuracy on a hold-out set of links and show which one proves to be better at mining associations between Wikipedia concepts. The results include qualitative (principal component analysis dimensionality reduction and visualization) and quantitative (accuracy) differences between the proposed methods. As a part of the conclusion, further research questions are provided, including new embedding architectures and the creation of a graph embedding algorithms benchmark.

Keywords: graph embeddings, link prediction, Wikipedia2vec, Role2vec, AttentionWalk, Walklets, principle component analysis.

Р. В. ШАПТАЛА, Г. Д. КИСЕЛЬОВ**ВИКОРИСТАННЯ ВЕКТОРНИХ ПРЕДСТАВЛЕНЬ ГРАФІВ ДЛЯ ПРОГНОЗУВАННЯ ЗВ'ЯЗКІВ У WIKIPEDIA**

Прогнозування зв'язків є важливою областю дослідження в аналізі мереж та теорії графів, яка намагається відповісти на питання, чи можуть два вузли у графі в майбутньому мати зв'язок. На сьогоднішній день графи повсюдно присутні у нашому житті (соціальні мережі, електротехніка, дороги і т.д.), тому проблема має вирішальне значення для розвитку інтелектуальних додатків. У минулому були запропоновані методи вирішення задачі прогнозування зв'язків за допомогою алгебраїчних формулювань і евристик, однак їхня виразність і переносимість не були задовільними. Останнім часом методи побудови векторних представлень зросли у популярності через їх ефективність і здатність передавати знання між завданнями. Натхненний знаменитим в машинному навчанні та обробці природних мов дослідницьким підходом Word2Vec, ці методи намагаються вивчити розподілене векторне представлення. Після цього бінарний класифікатор, заданий парою таких векторів, прогнозує ймовірність існування зв'язку між закодованими вузлами. У даній роботі ми розглянемо декілька підходів до вбудовування графіків для проблеми прогнозування зв'язків у Wikipedia, а саме Wikipedia2vec, Role2vec, AttentionWalk та Walklets. Прогнозування посилань у контексті Wikipedia – це знаходження сторінок, які пов'язані через певні смислові відносини. Ми оцінюємо точність прогнозування на відокремленому наборі зв'язків і показуємо, який з методів краще знаходить асоціації між сутностями у Вікіпедії. Отримані результати включають якісні (метод головних компонентів для зменшення розмірності та візуалізації) і кількісні (точність) відмінності між запропонованими методами. У рамках висновку наводяться подальші дослідницькі питання, включаючи нові архітектури побудови векторних представлень та створення загальноприйнятого тесту ефективності таких представлень.

Ключові слова: векторні представлення даних, прогноз зв'язків, Wikipedia2vec, Role2vec, AttentionWalk, Walklets, метод головних компонентів.

Р. В. ШАПТАЛА, Г. Д. КИСЕЛЕВ**ИСПОЛЬЗОВАНИЕ ВЕКТОРНЫХ ПРЕДСТАВЛЕНИЙ ГРАФОВ ДЛЯ ПРОГНОЗИРОВАНИЯ СВЯЗЕЙ В WIKIPEDIA**

Прогнозирование связей является важной областью исследования в анализе сетей и теории графов, которая пытается ответить на вопрос, могут два узла в графе в будущем иметь связь. На сегодняшний день графы повсеместно присутствуют в нашей жизни (социальные сети, электротехника, дороги и т.д.), поэтому проблема имеет решающее значение для развития интеллектуальных приложений. В прошлом были предложены методы решения задачи прогнозирования связей с помощью алгебраических формулировок и эвристик, однако их выраженность и переносимость ни были удовлетворительными. В последнее время методы построения векторных представлений выросли в популярности из-за их эффективности и способности передавать знания между задачами. Вдохновленный знаменитым в машинном обучении и обработке естественных языков исследовательским подходом Word2Vec, эти методы пытаются изучить распределенное векторное представление. После этого бинарный классификатор, заданный парой таких векторов, прогнозирует вероятность существования связи между закодированными

© R.V. Shaptala, G.D. Kyselov, 2019

узлами. В данной работе мы рассмотрим несколько подходов к построению векторных представлений для проблемы прогнозирования связей в Wikipedia, а именно Wikipedia2vec, Role2vec, AttentionWalk и Walklets. Прогнозирование связей в контексте Wikipedia – это нахождение страниц, которые связаны через определенные смысловые отношения. Мы оцениваем точность прогнозирования на отдельном наборе связей и показываем, какой из методов лучше находит ассоциации между сущностями в Википедии. Полученные результаты включают качественные (метод главных компонент для уменьшения размерности и визуализации) и количественные (точность) различия между предлагаемыми методами. В рамках заключения приводятся дальнейшие исследовательские вопросы, включая новые архитектуры построения векторных представлений и создание общепринятого теста эффективности таких представлений.

Ключевые слова: векторные представления графов, прогноз связей, Wikipedia2vec, Role2vec, AttentionWalk, Walklets, метод главных компонент.

Introduction. Networks and graphs have become ubiquitously important to model difficult systems that consist of various elements. Graph data science has a large number of applications in various fields like logistics, social networks, recommendation engines, and communication networks. There have been a lot of research in the area of the possibility to predict new links between elements in the topology of the graph based on the properties of its elements. Such a task is called link prediction and is defined as the problem of predicting new relationships in networks. Link prediction's goal is to find the initial rules of the graph link formation by inferring lost or possible relationships, given currently observed connections. The area is growing fast and is becoming more and more interesting as a research vector since it can help us predict how real-life networks will progress and evolve in time [1].

One of the applications of graphs and an example of complex networks are web-scale knowledge bases [2]. They provide a representation of world knowledge that is structured, with projects such as the Google Knowledge Vault [3], Freebase [4] and DBPedia [5]. These technologies are at the core of a wide range of applications such as question answering, recommender systems and chatbots. Unfortunately, these knowledge bases are incomplete because of the complexity of our world. That is why predicting missing entries or link prediction is one of the main problems in knowledge engineering. Knowledge bases encode data as a directed graph with edges (links, relations) between nodes (concepts, entities). The topological structure and nature among the relations present in these bases often make the tasks of filling in the missing links of a knowledge base possible. The idea behind link prediction is the automatic search for such regularities.

There are two types of approaches that are usually used to define models for graph-based problems [6]. The first one works with the initial graph adjacency matrix, while the second – with an inferred vector space. The popularity of the last approach has gradually increased lately. They try to represent the graph in a vector space that is going to preserve its properties. Having such an encoding is extremely convenient in the graph-related problems. The vectors are used as inputs (features) to a machine learning algorithm which parameters are trained based on the dataset. This helps negate the need for difficult classification algorithms which work directly with the graph.

However, the dimensions of the trained vectors become an additional hyperparameter and searching for an optimal one can be difficult. For example, higher dimensionality might increase the reconstruction precision but will have higher time and space complexities. The choice can also be domain-specific depending on the task: for example, lower number of dimensions might result in

better link prediction accuracy if the model only captures local relations between entities [6].

Preliminaries. A graph $G(V, E)$ is a collection of $V = \{v_1, \dots, v_n\}$ nodes and $E = \{e_{ij}\}_{i,j=1}^n$ edges. The adjacency matrix S of graph G contains indicators associated with each edge in the following way: $s_{ij} = 1$ if v_i and v_j are connected to each other, and $s_{ij} = 0$ otherwise. For undirected graphs, $s_{ij} = s_{ji} \forall i, j \in \{1, \dots, n\}$.

Given a graph $G(V, E)$, a graph embedding is a mapping $f: v_i \rightarrow w_i \in \mathbb{R}^d \forall i \in \{1, \dots, n\}$ such that $d \ll |V|$ and the function f retains some similarity notion defined on graph G .

Consequently, a graph embedding encodes each node in a low-dimensional feature vector that can retain the relations between nodes.

Graph embeddings. In this section we describe evaluated graph embedding approaches.

Random walks are at the core of numerous existing graph embedding methods. Since such approaches have numerous problems that arise from their exploitation of random walks (like the features that can not transfer knowledge to other nodes and networks as they are unique to each entity. Role2Vec framework tries to overcome this drawback by the use of attributed random walks. This algorithm was chosen because it is a basis for generalizing other similar methods like DeepWalk, node2vec, and many others that are based on random walks. The proposed framework helps these methods be more applicable for both transductive and inductive learning as well as for use on graphs with other features (if they exist) [7]. This is accomplished by learning functions that are applicable to unseen entities and networks. The authors show that Role2vec is more efficient in terms of predictive performance as well as requires less space than other methods on a variety of graphs. Role2Vec uses the extensible notion of attributed random walks that is not connected to a specific node but is instead based on a function that maps a node feature vector to a class, so that two nodes belong to the same class if they are topologically similar. Role2vec provides several valuable advantages to any method that is built upon it. Firstly, it is naturally inductive as the learned embeddings generalize to new entities and across networks and therefore might be used for transfer learning. Secondly, authors claim that their approach is able to capture structural similarity more efficiently. Thirdly, the Role2vec framework is way more space-efficient since representations are learned for classes (not nodes) and consequently require less space than existing methods. Fourthly, the proposed framework has an ability to work with graphs with features (if such exist or are available).

Graph embedding methods encode nodes in a continuous vector space, capturing various classes of

information present in the network. These methods have numerous hyperparameters (for example the length of a random walk) which have to be manually tuned for every graph. AttentionWalk is a method of graph embedding where previously fixed hyperparameters are replaced with trainable ones that are automatically learned via backpropagation [8]. The authors propose an attention model on the power series of the transition matrix, which decides where to take the next walk in order to optimize a long-term goal. Different to other attention models, the AttentionWalk uses attention parameters only on the training data itself (the random walk), while during model inference there are no attention layers. The authors did a series of tests on link prediction tasks, trying to produce embeddings that capture the graph structure, transferring the representation to unseen information. It is also claimed that AttentionWalk improves state-of-the-art results on a set of real-world graph datasets, for example collaboration, biological, and social networks. The final result of this approach is that automatically-learned attention parameters tend to correspond with the optimal choice of hyperparameters that are manually tuned in other methods.

Another approach which is of particular interest to us is called Walklets [9], a novel method for learning multiscale representations of nodes in a graph. These vectors explicitly encode multiscale relationships in a way that is analytically derivable compared to previous works. The proposed method creates these multiscale relationships by subsampling random walks of different length on the nodes of a graph. By skipping over steps in each random walk, Walklets generates a different training dataset than similar approaches. More specifically, it creates a corpus of node pairs which are reachable via paths of a fixed length. This corpus is then used to find a set of hidden representations, each of which encodes successively higher order relationships from the adjacency matrix. The authors demonstrate the efficiency of Walklets' hidden representations on several multi-label graph classification tasks for social applications. Their results claim that Walklets outperforms other methods based on neural matrix factorization. One of the most important benefits of Walklets is that it is an online learning algorithm, so it can scale to networks with an enormous number of nodes and links.

A different, yet powerful algorithm that can be used for Wikipedia link prediction is Wikipedia2vec – an open source tool for learning embeddings of words and entities from Wikipedia [10]. Not only does this tool enable researchers to easily obtain high-dimensional embeddings of words and entities from a Wikipedia dump, it also provides the source code, documentation, and pretrained vectors for twelve most popular languages at <http://wikipedia2vec.github.io>. The learned embeddings can easily be applied via transfer learning for natural language processing (NLP) models. The tool can be installed via Python programming language package repository PyPI. The pretrained embeddings have been learned by iterating over entire Wikipedia pages and joint optimization of three different submodels: model of Wikipedia graph, which learns entity embeddings by predicting neighboring entities in Wikipedia's page network – an undirected graph whose nodes are entities and

edges represent links between entities, based on each entity in Wikipedia (it does not matter if both pages link to each other or only one of them references another one – the link is created anyway); word-based skip-gram model, which learns word embeddings by predicting neighboring words for each word in a text contained on a Wikipedia page; anchor context model, which aims to place similar words and entities near one another in the vector space, and to create interactions between embeddings of words and those of entities. Here, we obtain referent entities and their neighboring words from links contained in a Wikipedia page, and the model learns embeddings by predicting neighboring words given each entity.

These three submodels are all inspired by the skip-gram model [11], which is a neural network model with a training objective to find embeddings that are useful for predicting context items (i.e., neighboring words or entities) given a target item.

To predict links between two nodes in a graph we use a simple one hidden layer perceptron on the concatenation of the embeddings of both nodes. The final classification task is trained using ADAM [12] optimization algorithm with the learning rate of 0.01 and 100 hidden layer units.

Evaluation and results.

We evaluated the described approaches on the SNAP Wikipedia [13] navigation paths dataset. This dataset has a set of Wikipedia links, collected through the human-computation game, called Wikispeedia. In there, users are asked to navigate from a starting Wikipedia node to a given article, through clicking Wikipedia links. A condensed version of Wikipedia (4,604 articles) is used.

For our project, 107444 links were used as positive examples and the same quantity was generated as negative examples. Thus, the dataset is balanced and we can use accuracy to measure the performance of the implemented approaches. 20 % of the data was held out for testing and the results are presented on this test set.

Quantitative results of our evaluation are summarized in Table 1. Walklets significantly outperforms every other approach that we tested due to the subsampling that is inherent in the algorithm, capturing not only first-order information, but also encoding the relations between nodes further from the start of the random walk.

Table 1 – Evaluated embeddings link prediction accuracy

| Embeddings | Accuracy |
|---------------|----------|
| Role2vec | 0.723 |
| AttentionWalk | 0.699 |
| Walklets | 0.877 |
| Wikipedia2vec | 0.734 |

To provide some qualitative results, we have also tried plotting the resulting embeddings. Since all of the tested approaches provide high-dimensional representations, the first problem that arises is to reduce these dimensions to human-readable form. For that we use principal component analysis (PCA) with the number of principle components set to 2. In our case, PCA transforms the data to a new coordinate system where the highest variance by some

projection of the data comes to lie on the first coordinate and the second highest variance on the second coordinate. From Fig. 1–4 it can be seen that AttentionWalk could not capture meaningful information, since there are no well-defined clusters on the visualization. Role2vec and Wikipedia2vec managed to group similar concepts in several clusters, however Walklets show a better space division than them. This correlates with the quantitative results that were shown previously.

Conclusions. In this paper, we reviewed several graph embedding approaches for the problem of Wikipedia link prediction, namely Wikipedia2vec, role2vec, AttentionWalk and Walklets. Qualitative and quantitative results show that Walklets due to its implicit multiscale relationship capture system have more expressive power for the given problem.

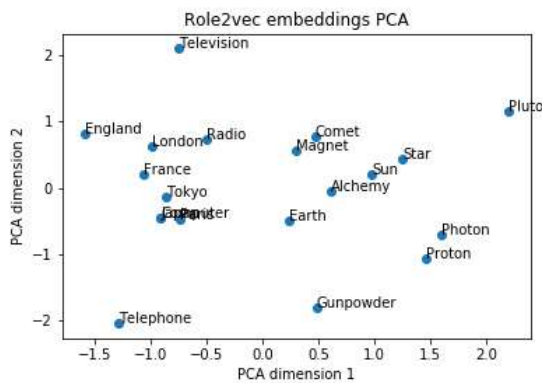


Fig. 1. Role2vec embeddings reduced to 2-D by PCA

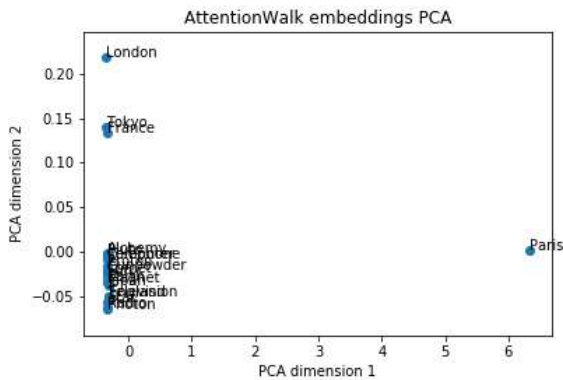


Fig. 2. AttentionWalk embeddings reduced to 2-D by PCA

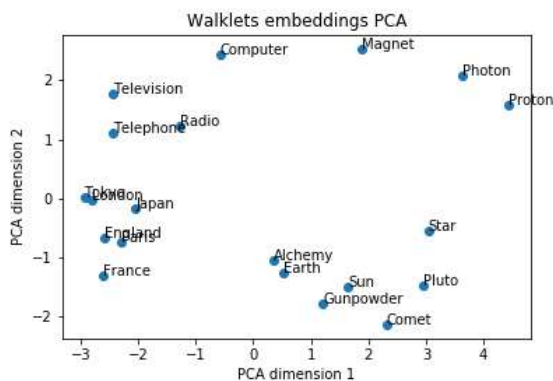


Fig. 3. Walklets embeddings reduced to 2-D by PCA

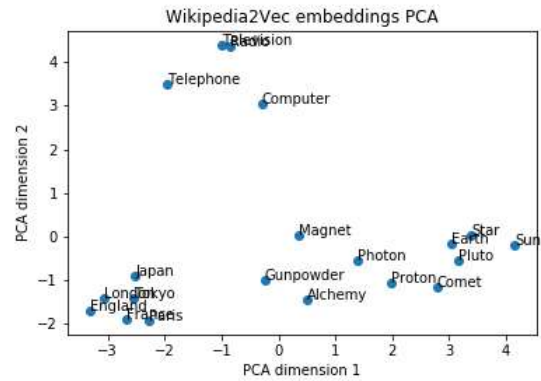


Fig. 4. Wikipedia2Vec embeddings reduced to 2-D by PCA

We consider the following research directions valid for future work: the creation of a standard benchmark dataset for link prediction of sufficient size to test accuracy, speed and scalability of graph embedding approaches; experimenting with new architectures, that would capture more information inherent to the link prediction problem, since our work did not achieve perfect prediction accuracy.

References

1. Martínez V., Berzal F., Cubero J. C. A survey of link prediction in complex networks. *ACM Computing Surveys (CSUR)*. 2017. Vol. 49, № 4. Article 69.
2. Minervini P., Costabello L., Muñoz E., Nováček V., Vandenburg P. Y. Regularizing knowledge graph embeddings via equivalence and inversion axioms. *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Vol. 10534. Cham: Springer, 2017. P. 668–683.
3. Dong X., Gabrilovich E., Heitz G., Horn W., Lao N., Murphy K., Strohmann T., Sun S., Zhang W. Knowledge vault: A web-scale approach to probabilistic knowledge fusion. *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. New York: ACM, 2014. P. 601–610.
4. Bollacker K., Evans C., Paritosh P., Sturge T., Taylor J. Freebase: a collaboratively created graph database for structuring human knowledge. *Proceedings of the 2008 ACM SIGMOD international conference on Management of data*. Vancouver: ACM, 2008. P. 1247–1250.
5. Auer S., Bizer C., Kobilarov G., Lehmann J., Cyganiak R., Ives Z. Dbpedia: A nucleus for a web of open data. *The semantic web*. Berlin, Heidelberg: Springer, 2007. P. 722–735.
6. Goyal P., Ferrara E. Graph embedding techniques, applications, and performance: A survey. *Knowledge-Based Systems*. 2018. Vol. 151. P. 78–94.
7. Ahmed N. K., Rossi R., Lee J. B., Willke T. L., Zhou R., Kong X., Eldardiry H. Learning role-based graph embeddings. *StarAI workshop, IJCAI 2018*. *arXiv preprint arXiv:1802.02896*. 2018.
8. Abu-El-Haija S., Perozzi B., Al-Rfou R., Alemi A. A. Watch your step: Learning node embeddings via graph attention. *Advances in Neural Information Processing Systems*. Vol. 31. Curran Associates, Inc. Publ., 2018. P. 9180–9190.
9. Perozzi B., Kulkarni V., Chen H., Skiena S. Don't Walk, Skip! Online Learning of Multi-scale Network Embeddings. *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. New York: ACM, 2017. P. 258–265.
10. Yamada I., Asai A., Shindo H., Takeda H., Takefuji Y. Wikipedia2Vec: An Optimized Implementation for Learning Embeddings from Wikipedia. *arXiv preprint arXiv:1812.06280*. 2018.
11. Mikolov T., Sutskever I., Chen K., Corrado G. S., Dean J. Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*. Vol. 2. Curran Associates, Inc. Publ., 2013. P. 3111–3119.
12. Kingma D. P., Ba J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*. 2014.

13. West R., Pineau J., Precup D. June. Wikispedia: An online game for inferring semantic distances between concepts. *Twenty-First International Joint Conference on Artificial Intelligence*. Pasadena, 2009. P. 1598–1603.

References (transliterated)

1. Martínez V., Berzal F., Cubero J. C. A survey of link prediction in complex networks. *ACM Computing Surveys (CSUR)*. 2017, vol. 49, no. 4, article 69.
2. Minervini P., Costabello L., Muñoz E., Nováček V., Vandenbussche P. Y. Regularizing knowledge graph embeddings via equivalence and inversion axioms. *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Cham, Springer Publ., 2017, vol. 10534, pp. 668–683.
3. Dong X., Gabrilovich E., Heitz G., Horn W., Lao N., Murphy K., Strohmann T., Sun S., Zhang W. Knowledge vault: A web-scale approach to probabilistic knowledge fusion. *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. New York, ACM Publ., 2014, pp. 601–610.
4. Bollacker K., Evans C., Paritosh P., Sturge T., Taylor J. Freebase: a collaboratively created graph database for structuring human knowledge. *Proceedings of the 2008 ACM SIGMOD international conference on Management of data*. Vancouver, ACM Publ., 2008, pp. 1247–1250.
5. Auer S., Bizer C., Kobilarov G., Lehmann J., Cyganiak R., Ives Z. Dbpedia: A nucleus for a web of open data. *The semantic web*. Berlin, Heidelberg, Springer Publ., 2007, pp. 722–735.
6. Goyal P., Ferrara E. Graph embedding techniques, applications, and performance: A survey. *Knowledge-Based Systems*. 2018, vol. 151, pp. 78–94.
7. Ahmed N. K., Rossi R., Lee J. B., Willke T. L., Zhou R., Kong X., Eldardiry H. Learning role-based graph embeddings. *StarAI workshop, IJCAI 2018*. *arXiv preprint arXiv:1802.02896*. 2018.
8. Abu-El-Haija S., Perozzi B., Al-Rfou R., Alemi A. A. Watch your step: Learning node embeddings via graph attention. *Advances in Neural Information Processing Systems*. Vol. 31. Curran Associates, Inc. Publ., 2018, pp. 9180–9190.
9. Perozzi B., Kulkarni V., Chen H., Skiena S. Don't Walk, Skip! Online Learning of Multi-scale Network Embeddings. *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. New York, ACM Publ., 2017, pp. 258–265.
10. Yamada I., Asai A., Shindo H., Takeda H., Takefuji Y. Wikipedia2Vec: An Optimized Implementation for Learning Embeddings from Wikipedia. *arXiv preprint arXiv:1812.06280*. 2018.
11. Mikolov T., Sutskever I., Chen K., Corrado G. S., Dean J. Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*. Vol. 2. Curran Associates, Inc. Publ., 2013, pp. 3111–3119.
12. Kingma D. P., Ba J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*. 2014.
13. West R., Pineau J., Precup D. June. Wikispedia: An online game for inferring semantic distances between concepts. *Twenty-First International Joint Conference on Artificial Intelligence*. Pasadena, 2009, pp. 1598–1603.

Received 13.05.2019

Відомості про авторів / Сведения об авторах / About the Authors

Шаптала Роман Віталійович (Шаптала Роман Витальевич, Shaptala Roman Vitaliyovych) – аспірант кафедри системного проектування Навчально-наукового комплексу «Інститут прикладного системного аналізу», Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»; м. Київ, Україна; ORCID: <https://orcid.org/0000-0002-4367-5775>; e-mail: r.shaptala@gmail.com

Кисельов Геннадій Дмитрович (Киселев Геннадий Дмитриевич, Kyselev Gennadiy Dmytrovych) – кандидат технічних наук, старший науковий співробітник, доцент, заступник завідувача кафедри системного проектування Навчально-наукового комплексу «Інститут прикладного системного аналізу», Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»; м. Київ, Україна; ORCID: <https://orcid.org/0000-0003-2682-3593>; e-mail: kiselev@cad.kiev.ua

D. L. ORLOVSKYI, A. M. KOPP, A. A. PYVOVAROVA

AN INFORMATION TECHNOLOGY FOR THE SUPPORT FOR SOLVING THE ANALYSIS PROBLEM OF THE CUSTOMERS' VALUE WITHIN THE ENTERPRISE CRM-STRATEGY IMPLEMENTATION

This study considers a CRM-approach and methods for analyzing customer base. Mathematical and algorithmic support was developed for the analysis of customer value in a trading enterprise. Algorithmic support is presented using the IDEF0 notation. A software solution was developed to assess customer value in a commercial enterprise. A fragment of the data model for the software solution was developed using the IDEF1x notation. The parameters for estimating the value of customers in a commercial enterprise were calculated using the processed data obtained by applying the developed software solution. The developed software solution allows to segment the customer base according to different criteria and provide marketing recommendations to each groups of customers. Therefore, customers were divided into groups according to the calculated parameters. As a result of the analysis, each group of clients receives a marketing recommendations. After the application of recommendations an increase in the purchasing power of clients, belonging to the group to which the marketing recommendation has been provided, is expected. Another used method for analyzing the customer base in the developed software solution is the sales funnel, which displays the effectiveness of the marketing unit at each of the sales stages, thereby allowing to identify weak points in the sales department. Thus, unlike existing software solutions, the developed software allows not only to segment customers by their costs, but also to provide marketing recommendations in order to increase the enterprise's profit. The generation of recommendations is based on the developed mathematical support. By using this mathematical support, the calculations are carried out. Customers are allocated into certain groups, each of which is provided with the relevant marketing recommendations, by using the results of the performed calculations. In addition, the developed software solution allows not only to design sales funnels, but also to analyze sales funnels.

Keywords: client, customer value, customer classification, CRM analysis, value analysis, sales funnel, RFM analysis.

Д. Л. ОРЛОВСЬКИЙ, А. М. КОПП, А. А. ПИВОВАРОВА

ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ПІДТРИМКИ ПРОЦЕСУ ОЦІНКИ ЦІННОСТІ КЛІЄНТА В РАМКАХ РЕАЛІЗАЦІЇ CRM-СТРАТЕГІЇ ПІДПРИЄМСТВА

Розглянуто CRM-підхід і методи аналізу клієнтської бази. Було розроблено математичне та алгоритмічне забезпечення аналізу цінності клієнта в торговому підприємстві. Алгоритмічне забезпечення представлено в нотації IDEF0. Було розроблено програмне рішення для оцінки цінності клієнта в торговому підприємстві. На його основі розроблений фрагмент моделі даних в нотації IDEF1x. На основі опрацьованих даних за допомогою розробленого програмного рішення розраховуються параметри оцінки цінності клієнтів в торговому підприємстві. Розроблене програмне рішення дозволяє сегментувати клієнтську базу за різними критеріями та формувати маркетингові рекомендації для кожної групи клієнтів. Виходячи з розрахованих параметрів, клієнти діляться на групи. В результаті аналізу, кожна група клієнтів отримує маркетингову рекомендацію. Після застосування рекомендації очікується підвищення купівельної здатності клієнтів, що відносяться до групи, до якої була надана маркетингова рекомендація. Ще одним використовуваним методом аналізу клієнтської бази в розробленому програмному рішенні є воронка продажів, яка відображає ефективність роботи маркетингового підрозділу на кожному з етапів продажу, тим самим дозволяючи виявити слабкі місця в роботі відділу продажів. Таким чином, на відміну від існуючих рішень, розроблене програмне забезпечення дозволяє не тільки сегментувати клієнтів за їх вартістю, але й надавати маркетингові рекомендації для збільшення прибутку підприємства. Формування рекомендацій здійснюється на основі розробленого математичного забезпечення, на основі якого виконуються розрахунки, за результатами яких клієнти розподіляються на певні групи, кожній з яких надаються відповідні маркетингові рекомендації. Крім того, розроблене програмне рішення дозволяє виконувати не тільки побудову воронки продажів, а й їхній аналіз.

Ключові слова: клієнт, цінність клієнта, класифікація клієнтів, CRM-підхід, аналіз цінності, воронка продажів, RFM-аналіз.

Д. Л. ОРЛОВСКИЙ, А. М. КОПП, А. А. ПИВОВАРОВА

ИНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ ПОДДЕРЖКИ ПРОЦЕССА ОЦЕНКИ ЦЕННОСТИ КЛИЕНТА В РАМКАХ РЕАЛИЗАЦИИ CRM-СТРАТЕГИИ ПРЕДПРИЯТИЯ

Рассмотрены CRM-подход и методы анализа клиентской базы. Было разработано математическое и алгоритмическое обеспечение анализа ценности клиента в торговом предприятии. Алгоритмическое обеспечение представлено в нотации IDEF0. Было разработано программное решение для оценки ценности клиента в торговом предприятии. На его основе разработан фрагмент модели данных в нотации IDEF1x. На основе обработанных данных с помощью разработанного программного решения рассчитываются параметры оценки ценности клиентов в торговом предприятии. Разработанное программное решение позволяет сегментировать клиентскую базу по различным критериям и формировать маркетинговые рекомендации для каждой группы клиентов. Исходя из рассчитанных параметров, клиенты делятся на группы. В результате анализа, каждая группа клиентов получает маркетинговую рекомендацию. После применения рекомендации ожидается повышение покупательной способности клиентов, относящихся к группе, к которой была предоставлена маркетинговая рекомендация. Еще одним используемым методом анализа клиентской базы в разработанном программном решении является воронка продаж, которая отображает эффективность работы маркетингового подразделения на каждом из этапов продажи, тем самым позволяя выявить слабые места в работе отдела продаж. Таким образом, в отличие от существующих решений, разработанное программное обеспечение позволяет не только сегментировать клиентов по их стоимости, но и представлять маркетинговые рекомендации для увеличения прибыли предприятия. Формирование рекомендаций осуществляется на основе разработанного математического обеспечения, на основе которого выполняются расчеты. По результатам расчетов клиенты делятся на определенные группы, каждой из которых предоставляются соответствующие маркетинговые рекомендации. Кроме того, разработанное программное решение позволяет выполнять не только построение воронок продаж, но и их анализ.

Ключевые слова: клиент, ценность клиента, классификация клиентов, CRM-подход, анализ ценности, воронка продаж, RFM-анализ.

Introduction. The basis of the success of any company is the correct strategy of work with customers. Business grows as the customer base grows and customers become stronger. Meeting customer needs is becoming a

key factor in ensuring and maintaining the company's competitiveness [1].

In this regard, the task of "knowing" your customer becomes more and more relevant, having a complete

© D. L. Orlovskiy, A. M. Kopp, A. A. Pyovarov, 2019

picture of what, when and to what extent he has bought and how is satisfied with cooperation with you. This information allows you to predict buying behavior and determine the prospects for further development of relations [2].

Not all customers are equally useful. The Rule of Pareto, a pattern that was taken in the middle of the last century, says, – only 20% of existing customers provides 80% of the total profit. Modern calculations often show some other results, but the original meaning is retained: working with a smaller portion of customers provides the most part of the profit. The problem is to identify those customers who will provide this profit [3].

At the moment, there is quite a lot of software on the market that deals with customer relationship management. One of the most popular are: Microsoft Dynamics CRM, vTiger CRM, Zoho CRM and others. In these systems, much attention is being paid to customer relationship management, but there is no analysis of the value of each customer [4].

The object of the research includes approaches to managing relationships with customers. The subject of the research includes methods for assessing the value of customers.

CRM-approach and its characteristics. The CRM approach is based on the use of management and information technology, through which the company collects information about its customers at all stages of their lifecycle from engaging and retaining loyalty programs, removes information from it and uses it in the interests of its business to build mutually beneficial relationships. The result of applying CRM is to increase competitiveness and profit because relationships built on the basis of a personalized approach allow you to attract new customers and keep the old ones [5].

If we consider the CRM system as a set of technologies, then it is a set of applications that are bound by single business logic and integrated into a corporate information environment based on a single database [6]. Often, in parallel with the CRM-system, an ERP-system is implemented at the enterprise, or the CRM-module is part of the ERP II system. Customer Relationship Management software enables automation of business processes related to marketing, sales and service. As a result, the development of a personalized offer to a particular customer, which is offered to him in certain, favorable for the transaction, time and transmitted to him by the most convenient communication channel for him [7].

The CRM system provides coordination of the activities of different units based on a common information platform for customer interaction. This application avoids the situation when the marketing, sales and service departments are separate from each other, to coordinate their actions and the overall vision of the customer. In addition, like any other information system, CRM can significantly accelerate the flow of information and make it credible, which in turn increases the efficiency of responses to inquiries, the speed of circulation of funds and reduces costs [8].

One of the methods for analyzing a customer base is the RFM analysis (recency, frequency, monetary – recent,

frequency, cost) is the technology used to identify customers who are most inclined to respond to a new offer. This technology is widely used in direct marketing [9].

A way to measure the effectiveness of actions to increase sales – a funnel sales. It shows at which stage of work is the loss of the largest number of customers. With its help you can visualize the entire sales cycle. That is, it reflects the total number of customers who are at different stages of relationship with the seller [10]. The problem is the need to determine at what stage of the work there are problems, namely the loss of customers, and how to overcome this problem. The funnel of sales and its analysis allows not only to find out the root of the problem, but also to choose ways to overcome it.

Research problem statement. The main purpose of any company's activity is to generate profits, which is based on the correct strategy of working with customers. The profit of the company grows with the increase of the customer base and strengthening of relations with customers. In this regard, the increasingly important task is to have a complete picture of what, when and to what extent the customer buys and how satisfied with the cooperation with you [1].

In the course of the work it is necessary to analyze the existing methods: CRM-approach, basic approaches to solving the problem of customer value analysis and to review software tools that are used to solve customer relationship management tasks.

After that, you need to develop a mathematical solution to the problem of analyzing the value of customers in a trading company and to simulate the process of analyzing the value of customers in a trading company.

Next it is necessary to develop requirements for application software solution, database structure and directly applied software solution. On the basis of test data, it is necessary to test the software solution and analyze the results.

Mathematical support for solving the analysis problem of the customers' value in a trading company. In our country, mobile operators are the first to measure customer value. Their need was not accidental – due to the high penetration of mobile communication, the cost of attracting one customer became higher and higher – it's time to get rid of unprofitable sales channels and change the distribution model.

Most marketers and analysts use a set of indicators in their work that allows you to evaluate the effectiveness of advertising campaigns: the level of failures, the number and proportion of conversions, the outflow of customers, and the cost of customer involvement. These indicators are able to give a general idea of the effect of marketing activities and the level of customer loyalty.

CLV (from Customer Lifetime Value – consumer's lifetime value, customer value for business) is the current value of the probable future net profit expected to be obtained from a particular customer during his entire lifetime, an indicator of "customer life cycle value" for the firm. This is an indicator of the potential of the consumer (buyer), taking into account 3 criteria: customer, time and net profit. Objectives of using CLV: attracting "valuable" customers, increasing the "value" of customers, increasing

“loyalty” of customers, optimizing relations with them. CLV is also a methodology and an effective tool for analyzing customer relationship management, the principle of segmentation of the customer base of the enterprise. The basis of the methodology is the message: over time, the value of the customer increases, but the total number of “valuable” customers decreases [11].

The mathematically simplified CLV model can be represented as follows:

$$CLV = Revenue \times \frac{r}{1 + d - r},$$

where *Revenue* is the gross amount of contributions for a certain period of time;

r is the percentage of customer retention;

d is the size of the discount.

This formula focuses on how much the customer spends on (monetary value), assuming a continuous valuation period for all customers [12].

To move to the next level, we will need to enter two variables of the time intervals – recent and frequency:

$$CLV = \sum_{t=0}^T \frac{(p_t - c_t)r_t}{(1 + i)^t} - AC,$$

where p_t is the price paid by the customer at time t ;

c_t is the direct costs for customer service at time t ;

i is the size of the discount or the price of capital for the firm/company;

r_t is the likelihood of a re-purchase or that the customer's life cycle will not end at the time t ;

AC is the cost of acquisition;

T is the time period selected for CLV evaluation [12].

The probability of a re-purchase or that the customer's life cycle does not end at the time t is calculated using the following formula:

$$r_t = \frac{(n_0 + n) - w}{n_0 + n},$$

where n_0 is the number of customers at the beginning of the period;

n is the number of new customers;

w is the number of customers who have gone.

This value can range from 0% to 100%:

- 100%: the organization has successfully saved all its customers (there was no departure);
- 0%: the organization has stopped its activities.

LTV (Lifetime Value) is the aggregate profit of a company that is received from a single customer for all time with him. The simple formula has the following form [13]:

$$LTV = D - V,$$

where D is the revenue from the customer;

V is the costs of attracting and retaining a customer.

There is also a basic formula. This calculation method is more precise than the simple formula [13]:

$$LTV = SP \times SN \times ST,$$

where SP is the average selling cost;

SN is the fair number of sales per month;

ST is the average customer retention time in months.

Also, a predictive formula can be used, which has the following form [12]:

$$LTV = T \times AOV \times AGM \times ALT,$$

where T is the average number of orders (sales) per month;

AOV is the average check;

ALT is the average duration of customer interaction with the company (in months);

AGM is the share of revenue in revenue.

Also, LTV can be calculated according to the following formula (for n customers) [12]:

$$LTV = \sum_{x=1}^n \frac{ARPU_x - Costs_x}{(1 + WACC)^x} - SAC,$$

where $ARPU_x$ is the average revenue per customer;

$WACC$ is the average weighted value of capital;

$Costs_x$ is the annual expenses for user support for a certain period;

SAC is the cost to the customer.

Customer referral value (CRV):

$$CRV_i = \sum_{t=1}^T \sum_{y=1}^{n_1} \frac{A_{ty} - a_{ty} + M_{ty} + ACQ1_{ty}}{(1 + r)^t} + \sum_{l=1}^T \sum_{y=n_1}^{n_2} \frac{ACQ2_{ty}}{(1 + r)^t},$$

where T is the number of time periods (years or months);

a_{ty} is the marginal profit (cost of sales minus variable costs) of the customer who would not have bought on the recommendation;

A_{ty} is the value of the customer involved for the customer;

n_1 is the number of customers who came for the recommendation;

n_2 is the number of customers who came without a recommendation;

M_{ty} is the marketing costs for customer retention;

$ACQ1_{ty}$ is the saving on attracting the customer, who came for the recommendation (did not have to spend on advertising for primary attraction);

$ACQ2_{ty}$ is the saving on attracting a customer who came independently (did not have to spend on referral payments and marketing) [12].

The first sentence is the value of the customers who came for the recommendation, and the second term is the savings on customers who came on their own.

The Value of a Loyal Customer (VLC) is a measure that quantifies the cumulative income or profit that each target market customer generates over the customer's life cycle. This indicator is calculated according to the following formula [11]:

$$VLC = P \times RF \times CM \times BLC,$$

where P is the income per unit (average revenue from the purchase);

CM is the marginal profit (difference between income and expense);

RF is the frequency of purchases;

BLC is the customer life cycle.

To calculate the customer's life cycle (BLC), you need to make the following calculations using the three formulas:

$$D\# = n_0 - n, DR = \frac{D\#}{n_0}, BLC = \frac{1}{DR},$$

where n_0 is the number of customers at the beginning of the period;

n is the final number of customers;

$D\#$ is the outflow of customers;

DR is the speed of outflow of customers.

The cost of customer acquisition (CAC, Customer Acquisition Cost) consists of aggregate marketing and sales costs. In simple words, SAS is the total amount of all marketing efforts required to attract one customer. This factor largely determines whether your company has a viable business model that allows you to maintain a low level of CAC when scaling your business. In addition, CAC has a tendency to continuous growth – as new promotional campaigns start up and the development of new techniques of promotion. The simplified formula has the following form [10]:

$$CAC = \frac{MCC}{CA},$$

where MCC is the marketing expenses;

CA is the number of attracted customers.

There is also a complex formula that will give you a more accurate result:

$$CAC = \frac{MCC + W + S + PS + O}{CA},$$

where MCC is the marketing costs;

CA is the number of customers attracted;

W is the salary related to marketing sales;

S is the software cost;

PS are the any additional professional services used in marketing or sales;

O is the other overhead costs associated with marketing and sales.

Algorithmic support for solving the problem of analyzing the value of customers in a trading company.

Based on the information above, a detailed chart of the first level of the customer value analysis process in the IDEF0 notation was developed. It includes the following processes: data preparation, RFM analysis, parameter calculation to assess customer value, analysis of results and recommendations. The diagram is shown in Fig. 1.

Development of software solution for solving the problem of analyzing the value of customers in a trading company. A fragment of the data model, shown in Fig. 2, was also developed.

One of the key entities is the counterparty – one of the parties to the contract in civil law. In this structure, contractors are suppliers and customers. Counterparties are divided into legal entities and individuals. An email address and phone number for communication is known about them, which may be several, since now almost every person has a phone number of different operators. Also, the obligatory attribute is the date of the counterparty's registration (that is, the date the company added the counterparty to its database).

- Suppliers deliver goods to the company, which are then sold to customers. The obligatory attribute of a product is a commodity group and units of measurement. A separate entity is the supplied goods, which is necessary to prevent the company from selling the unmodified product. For each item, a delivery is made, which specifies the supplier, the date and the goods that were

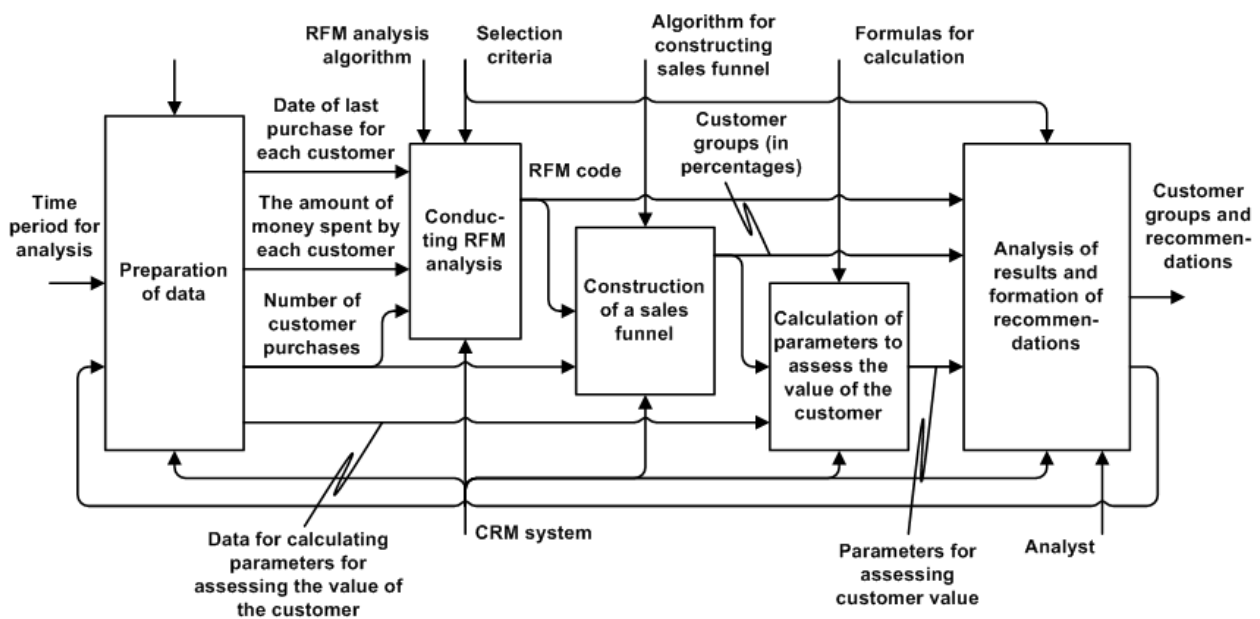


Fig. 1. Detailed diagram of the first level

delivered for this delivery at a certain price, as well as the batch. The customer can pass different stages of work with the customer. Usually, these stages are then displayed in the sales funnel. Therefore, at the stages of a particular customer, the start and end of the stage, the code of the manager involved in this stage, and the sale with which this stage is linked, is indicated. Also, the customer may have certain statuses. For each customer, the date of becoming a status is registered.

The customer may have a discount on a certain group of goods. The discount card table shows the discount product, the discount customer, the discount amount, and the start and end dates of the discount.

For each sale, be sure to fix the costs for it. The sale records the date, code of the manager who was engaged in this sale and the customer's code. There is a table of sold goods, which displays the quantity of goods and the price at which it was sold.

A sales manager (customer service) and a supply manager (work with suppliers) and an analyst will work with a software solution that implements support for solving the problem of customer value analysis in a trading company. The features which are available to the analyst include RFM analysis, construction of a funnel of sales, and calculations by the formulas for the analysis of the value of the customer.

In the customer window, the manager has the opportunity to add a new customer by adding the following information about the customer: name, tax number and certificate number, if it is a legal person and full name, first name, patronymic and tax number. In other case, an electronic address, a telephone number (one or more), and an automatic registration date (the date of the first input of the information for this counterparty) is added automatically. When editing data, the manager can edit the data that was entered earlier, as well as add the status of the customer and stages of working with him. Also, the

customer may be given a discount on a group of goods, while fixing the period and the group of goods for which it is provided.

Any customer can be ordered for sale. The manager adds the sales date, the manager who made it, the customer to whom the product was sold, the product that was sold (one or more), as well as their number. The sale price is automatically calculated by the system, depending on the value of the product already in the database.

In the window of work with the supplier, the manager has the opportunity to add a new supplier by adding the following information about the supplier: name, tax number and certificate number, if it is a legal entity and full name, name, and tax number. In other case, an electronic address, a telephone number (one or more), and an automatic registration date (the date of the first input of the information for this counterparty) is added automatically. When editing data, the manager has the ability to edit the data that was entered earlier.

Also, the window of work with the supplier should be a window of delivery of delivery. First, the manager is able to add a new item if it has been delivered, but the database does not have such a name yet. This includes the introduction of the name of the product, the addition of the product group and the unit of measurement. When making a delivery, the manager introduces the date of this delivery, adds the supplier, who carries it, the goods that are delivered their price and the size of the batch.

The manager has the opportunity to conduct RFM-analysis of the customer base, which will give him the opportunity to get a report in which customers will be divided into groups, which were obtained through RFM-analysis. Each group of customers will be given a marketing recommendation, in which the company's revenue can increase.

The manager will be able to receive a report with a detailed sales funnel: there will be shown how many customers are at each of its stages in quantitative and percentage terms. And also this funnel will be displayed

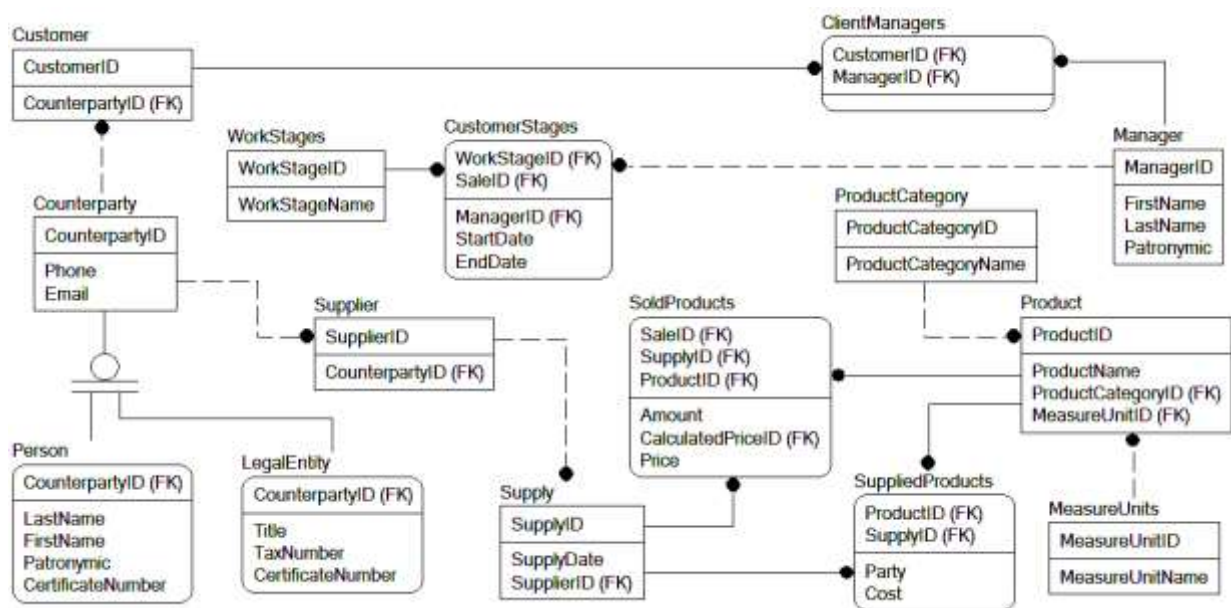


Fig. 2. Fragment of the data model

graphically: in the form of an inverted triangle, which is divided into stages in an appropriate proportion.

The manager can receive a report with calculations, which includes the calculation of the following values:

- LTV (on the life value of the consumer);
- SAS (cost of customer acquisition);
- CLV (profit from a single customer);
- VLC (cost of loyal customer);
- CRV (referral cost of the customer).

The results of these calculations will be reflected in the report, which will be divided into customers, depending on the value of these metrics. Each group of customers will be given a marketing recommendation, in which the company's revenue can increase.

Fig. 3 shows a sequence diagram for a software solution to support the solution of the problem of analyzing the value of customers in a trading company.

A software solution was developed, which included test data, which was the subject of further analysis. Analysis window contains 4 tabs: RFM analysis, CLV/CRV, other indicators, sales funnel.

Fig. 4 shows the tab "RFM-analysis". The user selects a period of time for analysis and clicks the button "Calculate". Customers who made purchases in the selected period are divided into 6 groups and are entered into the appropriate blocks. Also, the user has the ability to sort by specific product or product group. This allows you to detail the analysis performed. Each group has its name and marketing recommendation. As a result of the analysis the following groups are obtained:

- the best customers (regularly make a purchase, spend large amounts of money and recently made the last purchase), recommendation: provide a discount for regular customers;

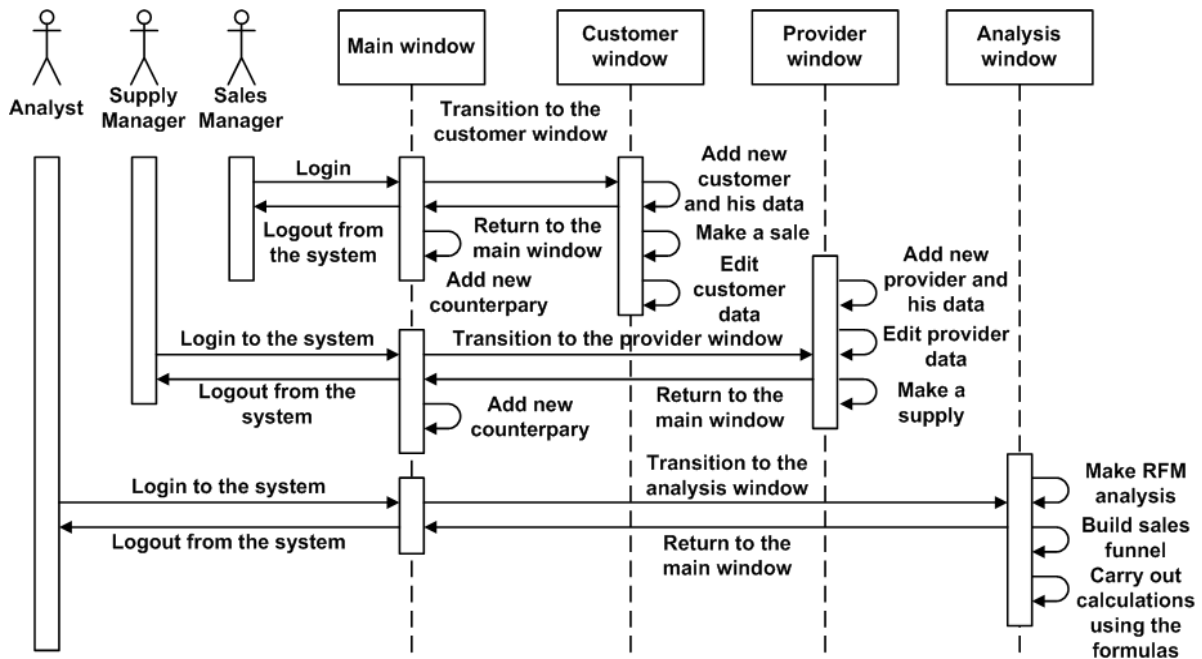


Fig. 3. Sequence diagram

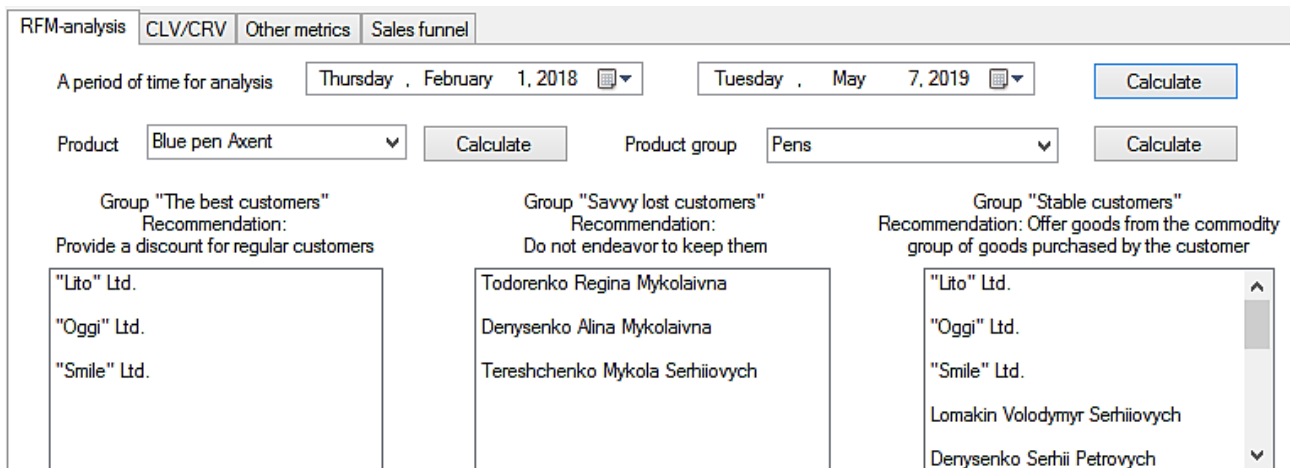


Fig. 4. "RFM-analysis" tab

- lost customers (regularly bought in the past, spent a large sum of money, but for some reason they have not bought anything for a long time); recommendation: the maximum allowable discount for a short period;
- savvy lost customers (little was bought, little wasted money and bought for a long time), a recommendation: do not endeavor to keep them;
- spenders (spend the most money from all, while not taking into account the frequency of their purchases), the recommendation: to offer goods with the most expensive price from the commodity group of goods purchased by the customer;
- stable customers (the frequency of their purchases most), recommendation: to offer goods from the commodity group of goods purchased by the customer;
- new customers (made their first purchase recently), recommendation: offer them a discount for their initial customers.

Each of these groups is provided with marketing recommendations that will allow you to maximize profits for each specific group of customers.

Fig. 5 shows the CLV/CRV tab. The user selects a period of time for analysis and clicks the button "Calculate". Customers who made purchases in the chosen period are divided into 4 groups and entered into the corresponding blocks. Each group has its name and marketing recommendation, namely:

- low CLV and high CRV: offer discounts to initial customers;
- high CLV and high CRV: informing them about referral programs through personalized communication;
- low CLV and low CRV: minimize investment in stimulating purchases using cheap communication channels;

- high CLV and low CRV: regular reminders about company offers (2 times more often than other customers).

In addition to sorting by date, the user has the opportunity to choose sorting by product or product group. In the case when the analysis is made taking into account the period for analysis, all the data obtained are stored in the database. This allows you to speed up the program in the future: no need to re-count the metrics since they will already be in the database. Real databases can have hundreds of thousands of records. And the transfer of metrics for each customer can take a significant period of time.

In the case where the user chooses the time period for analysis already in the database, the calculation is performed again, but data are not entered into the database. This prevents the accumulation and duplication of data. Such software solutions can be used for decades or more. That is why the issue of disk space saving is an important issue that cannot be overlooked. Thus, the developed software solution is suitable for use over many years as it prevents the accumulation of unnecessary data in the database.

Fig. 6 shows the "Other metrics" tab. The user selects a period of time for analysis and clicks the button "Calculate". This tab defines 2 groups of customers: low LTV and VLC. A low-LTV group has the recommendation to offer goods from the product group of goods purchased by the customer. For a low-VLC group, there is a recommendation: to minimize investment in stimulating purchases using cheap communication channels (e.g. email). In addition to sorting by date, the user has the opportunity to choose sorting by product or product group.

In the case when the analysis is made taking into account the period for analysis, all the data obtained are stored in the database. This allows you to speed up the program in the future: no need to re-count the metrics since they will already be in the database. Real databases can have hundreds of thousands of records. And the transfer of metrics for each customer can take a significant period of time.

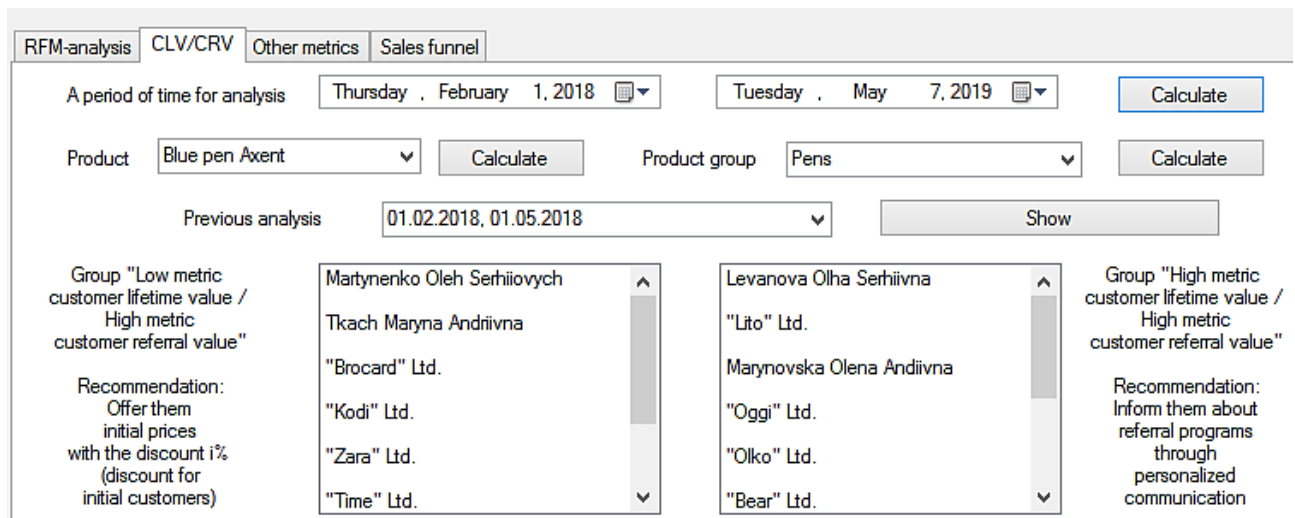


Fig. 5. "CLV/CRV" tab

In case the user chooses the time period for analysis already in the database, the calculation is performed again, but data are not entered into the database. This prevents the accumulation and duplication of data. Such software solutions can be used for decades or more. That is why the issue of disk space saving is an important issue that cannot be overlooked. Thus, the developed software solution is suitable for use over many years as it prevents the accumulation of unnecessary data in the database.

Fig. 7 shows the “Sales funnel” tab. By clicking on the “Build a sales funnel” button, a funnel for sales is under construction. This tab shows the conversion of the entire funnel and micro conversion for each stage. The user has the opportunity to select a period of time for analysis. The first phase micro conversion in any case is 100%, so it is not included in the block where the micro conversions are represented in steps. Stages can be any number, they are defined by the user (in the steps tab on the customer). The

first stage is at the top of the sales funnel. The latter is at the bottom of the sales funnel. That is, the steps are displayed from top to bottom.

Figure 8 depicts an unbalanced sales funnel. As can be seen from the figure, in the first stage there was a large number of customers, but only 16% of them went into the second stage (micro conversion of the second stage), which is a rather small indicator, compared to the micro conversions of other stages. So, we can conclude that after the first stage there is a large loss of customers. This provides the basis for further analysis of the first stage, to determine the problem situation. The conversion of the entire funnel is only 3%, which is very small compared to the funnel conversion in Figure 7 (44%). That is, only 3% of customers who were in the first stage, have reached the last stage.

The developed software solution allows to segment the customer base according to different criteria and

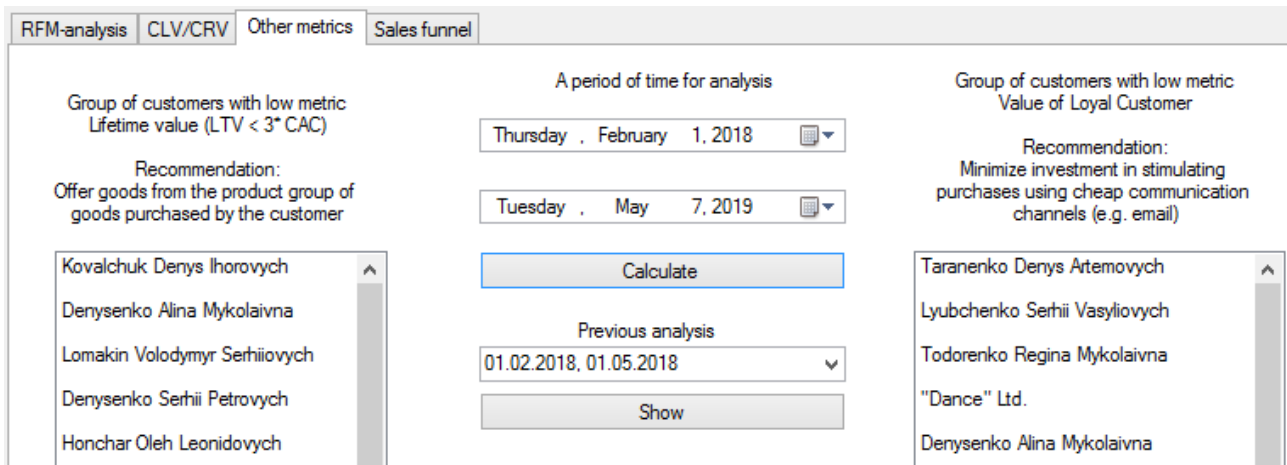


Fig. 6. “Other metrics” tab

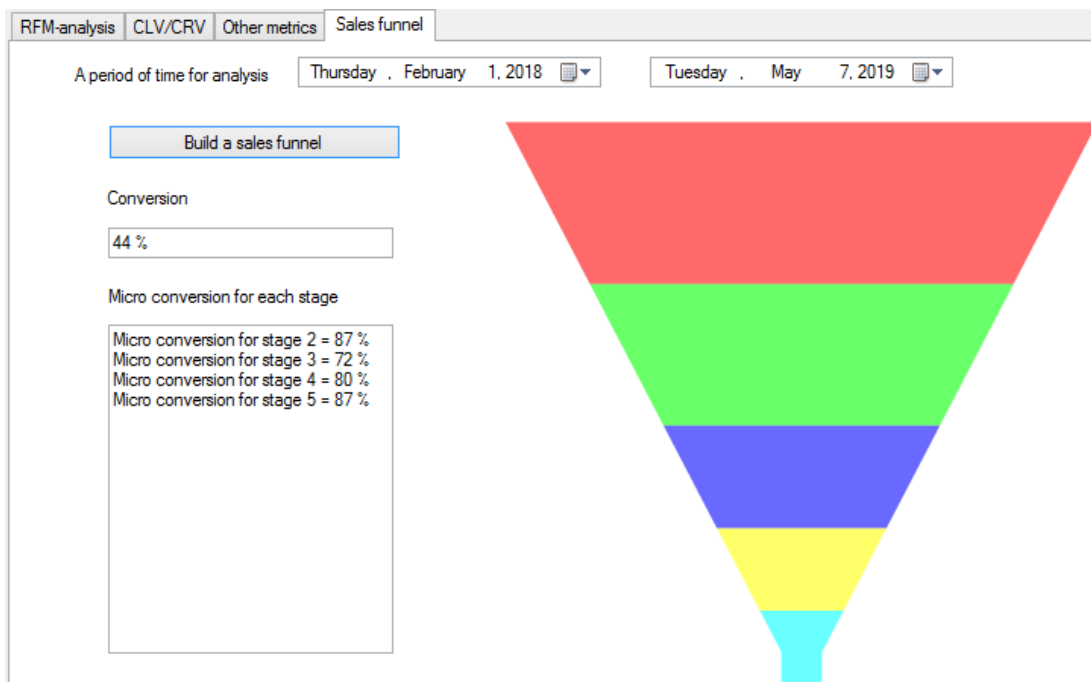


Fig. 7. “Sales funnel” tab

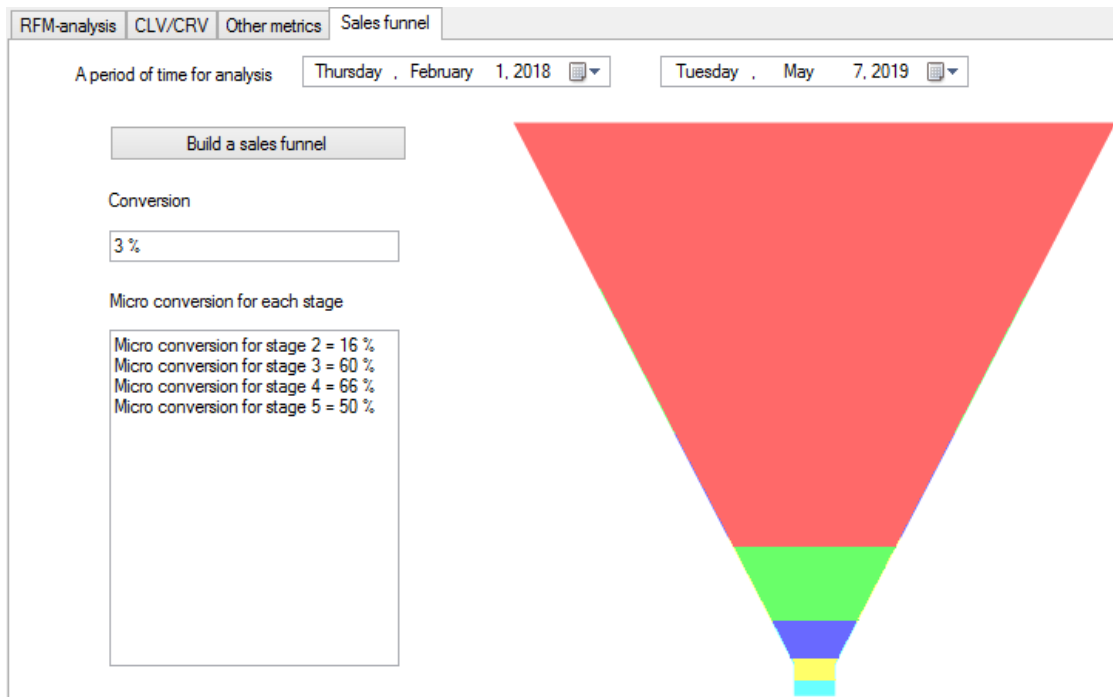


Fig. 8. Unbalanced sales funnel

provide marketing recommendations to each group of customers.

Conclusions. Whereas previously, the enterprise could become a market leader, focusing on the production and quality of goods, today it should take into account all the concepts of marketing. An enterprise should not just study the needs of the target market, but try to personally identify the needs of each customer, his individual preferences and offers the product or service he needs.

The developed software solution for solving the problem of customer value analysis is preferable to its counterparts, since it involves RFM analysis, which allows not only to segment the customers at their value, but also to issue marketing recommendations to increase company revenue. Similarly, it is done with the help of the developed mathematical support, that is, on the basis of calculations, customers are divided into groups and each group is provided with its marketing recommendations. Also, it includes not only the construction of sales funnels, but also its analysis.

References

1. Черкашин П. *Готовы ли вы к войне за клиента? Стратегия управления взаимоотношениями с клиентами (CRM)*. Москва: ООО «ИНТУИТ.ру», 2014. 384 с.
2. Гринберг П. *CRM со скоростью света: привлечение и удержание клиентов в реальном времени через Интернет*. СПб.: Символ Плюс, 2012. 528 с.
3. Гужва В. М. *Інформаційні системи і технології на підприємствах*. Київ: КНЕУ, 2011. 400 с.
4. Герасименко В. В. *Маркетинг*. Проспект, 2015. 489 с.
5. Андреева К. *Лидогенерация. Маркетинг, который продает*. СПб.: Питер, 2014. 240 с.
6. Пэйн Э. *Руководство по CRM: путь к совершенствованию менеджмента клиентов*. Гревцов-издатель, 2007. 384 с.
7. Каверина А. С. Анализ существующих методов управления клиентской базой для повышения конкурентоспособности

- аптечной организации. *Бюллетень сибирской медицины*. 2014. Т. 2, № 4. С. 172–180.
8. Кудинов А., Сорокин М., Гольшера Е. *CRM: российская практика эффективного бизнеса*. 1С-Паблишинг, 2008. 374 с.
9. Баженов Р. И. RFM-анализ клиентской базы в прикладном решении 1С: Предприятие. *Информатизация и связь*. 2014. № 2. С. 51.
10. Фатхутдинов Р. А. *Инновационный менеджмент*. СПб.: Питер, 2012. 400 с.
11. *LTV: CAC Ratio*. URL: <https://www.geckoboard.com/learn/kpi-examples/marketing-kpis/ltv-cac-ratio> (access date: 01.04.2018).
12. Kumar V. *Profitable customer engagement: concept, metrics, and strategies*. SAGE Publications India, 2013. 309 p.
13. Молино П. *Технологии CRM. Экспресс-курс*. Москва: Фаир-Пресс, 2012. 272 с.

References (transliterated)

1. Cherkashin P. *Gotovy li vy k voyne za klijenta? Strategiya upravleniya vzaimootnosheniyami s klijentami (CRM)* [Are you ready for a war for a client? Customer Relationship Management (CRM) Strategy]. Moscow, "INTUIT.ru" LLC, 2014. 384 p.
2. Grinberg P. *CRM so skorostyu sveta: privlecheniye i uderzhaniye klijentov v realnom vremeni cherez Internet* [CRM at the speed of light: attracting and retaining customers in real time via the Internet]. Saint-Petersburg, SimvolPlus, 2012. 528 p.
3. Guzhva V. M. *Informatsiini systemy i tekhnologii na pidpriyemstvakh* [Information systems and technologies at the enterprises]. Kyiv, KNEU, 2011. 400 p.
4. Gerasimenko V. V. *Marketing* [Marketing]. Prospekt, 2015. 489 p.
5. Andreyeva K. *Lidogeneratsiya. Marketing kotoryy prodayet* [Marketing that sells]. Saint-Petersburg, Piter, 2014. 240 p.
6. Peyn E. *Rukovodstvopo CRM: put k sovershenstvovaniyu menedzhmenta klijentov* [CRM Guide: The Way to Improve Client Management]. Grevtsov-publisher, 2007. 384 p.
7. Kaverina A. S. *Analiz sushchestvuyushchikh metodov upravleniya klijentskoj bazy dlya povysheniya konkurentosposobnosti aptechnoy organizatsii* [Analysis of the existing client base management methods to increase the competitiveness of the pharmacy organization]. *Byulleten sibirskoy meditsiny* [Bulletin of Siberian Medicine]. 2014, no. 4(2), pp. 172–180.
8. Kudinov A., Sorokin M., Golyshera E. *CRM: rossiyskaya praktika effektivnogo biznesa* [CRM: Russian practice of effective business]. 1С-Publishing, 2008. 374 p.

9. Bazhenov R. I. RFM-analiz kliyentskoy bazy v prikladnom reshenii 1C: Predpriyatiye [RFM analysis of the client base in the 1C: Enterprise application solution]. *Informatizatsiya i svyaz* [Informatization and communication]. 2014, no. 2. pp. 51.
10. Fatkhutdinov R. A. *Innovatsionnyy menedzhment* [Innovative management]. Saint-Petersburg, Piter, 2012. 400 p.
11. LTV:CAC Ratio. Available at: <https://www.geckoboard.com/learn/kpi-examples/marketing-kpis/ltv-cac-ratio>. (accessed 01.04.2018).
12. Kumar V. *Profitable customer engagement: concept, metrics, and strategies*. SAGE Publications India, 2013. 309 p.
13. Molino P. *Tekhnologii CRM. Ekspres-kurs* [CRM technology. Express course]. Moscow, Fair-Press, 2012. 272 p.

Received 05.04.2019

Відомості про авторів / Сведения об авторах / About the Authors

Орловський Дмитро Леонідович (Orlovskiy Dmytro Leonidovych) – кандидат технічних наук, доцент, Національний технічний університет «Харківський політехнічний інститут», доцент кафедри програмної інженерії та інформаційних технологій управління; м. Харків, Україна; ORCID: <https://orcid.org/0000-0002-8261-2988>; e-mail: ordm@kpi.kharkov.ua

Копп Андрій Михайлович (Kopp Andrii Mykhailovych) – Національний технічний університет «Харківський політехнічний інститут», асистент кафедри програмної інженерії та інформаційних технологій управління, аспірант; м. Харків, Україна; ORCID: <https://orcid.org/0000-0002-3189-5623>; e-mail: kopp93@gmail.com

Пивоварова Анастасія Андріївна (Pivovarova Anastasiia Andriivna) – Національний технічний університет «Харківський політехнічний інститут», студент; м. Харків, Україна; ORCID: <https://orcid.org/0000-0002-8255-8467>; e-mail: nysic96@gmail.com

O. O. MISHCHENKO, V. Y. VOLOVSHCHYKOV, V. F. SHAPO, M. A. GRINCHENKO

SYNTHESIS OF LOCAL AREA NETWORK STRUCTURE IN UNCERTAIN CONDITIONS OF INITIAL INFORMATION

The problem of taking into account the uncertainty of the initial information is identified in the tasks of the structure synthesis of local area networks. The analysis of the factors generating uncertainty is carried out, the sources of uncertainty are singled out and the approaches to formalization and research of uncertainty are systematized. Probabilistic-statistical, interval, fuzzy and deterministic approaches were reviewed, as well as statistical, subjective, logical interpretations of probability and stochastic description, concept of interval-probabilistic approach, continuously determined and discretely determined models. The formulation of the problem was formed and it was proposed to apply the theory of fuzzy sets to formalize and study an uncertainty in synthesizing the structure of the local computing network, the main elements of which are user points, the centers of processing of the information, switching devices and communication channels. The implementation of information and computing works which sources are subscriber points, is assigned to the centers of information processing. It is assumed the popularity of volumes of information and computing works and geographical coordinates of the location of network elements. A model for a research problem based on a cost criterion has been developed. Cost parameters are divided into well-defined and having an uncertain nature. Fuzzy parameters are proposed to describe the membership functions of the form "approximately equal" or "is approximately in the interval". Restrictions on technical capabilities of information processing centers, switching devices, traffic in communication channels are singled out. The technology of model research based on the idea of solving the problems of linear programming in fuzzy setting is proposed. The technology provides the decision maker with information about the network structure, its cost and the level of ownership of the solution. The results of the work can be used in the analysis of approaches to solving problems of structural-topological synthesis of local area networks, design, development and implementation of appropriate software solutions.

Keywords: local area network, uncertainty, probabilistic and statistical approaches, interval approach, fuzzy approach, deterministic approaches, structure synthesis, model, fuzzy cost criteria, research technology

O. O. МИЩЕНКО, В. Ю. ВОЛОВЩИКОВ, В. Ф. ШАПО, М.А. ГРИНЧЕНКО

СИНТЕЗ СТРУКТУРИ ЛОКАЛЬНОЇ ОБЧИСЛЮВАЛЬНОЇ МЕРЕЖІ В УМОВАХ НЕВИЗНАЧЕНОСТІ ВИХІДНОЇ ІНФОРМАЦІЇ

Ідентифікована проблема врахування невизначеності вихідної інформації в задачах синтезу структури локальних обчислювальних мереж. Проведено аналіз факторів, що породжують невизначеність, виділені джерела невизначеності і систематизовані підходи до формалізації і дослідження невизначеності. Розглянуті ймовірнісно-статистичний, інтервальний, нечіткий і детермінований підходи, а також статистична, суб'єктивна, логічна інтерпретація ймовірності і стохастичний опис, концепція інтервально-ймовірнісного підходу, модель безперервно-детермінована і дискретно-детермінована. Сформульовано постановку задачі і запропоновано застосовувати теорію нечітких множин до формалізації і дослідження невизначеності при синтезі структури локальної обчислювальної мережі, основними елементами якої є абонентські пункти, центри обробки інформації, комутаційні прилади та канали зв'язку. Виконання інформаційно-обчислювальних робіт, джерелами яких є абонентські пункти, покладено на центри обробки інформації. Припускається, що обсяги інформаційно-обчислювальних робіт і географічні координати розташування елементів мережі є відомими. Розроблено модель задачі дослідження на основі критерію вартості. Параметри вартості розділені на чітко задані і ті, що мають невизначену природу. Нечіткі параметри запропоновано описувати функціями на лежності виду "приблизно дорівнює" або "знаходиться приблизно в інтервалі". Виділено обмеження на технічні можливості центрів обробки інформації, комутаційних приладів, трафік в каналах зв'язку. Запропоновано технологію дослідження моделі, яка заснована на ідеї вирішення задач лінійного програмування в нечіткій постановці. Технологія забезпечує особу, яка приймає рішення, інформацією про структуру мережі, її вартість та рівень приналежності рішення. Результати роботи можуть бути використані при аналізі підходів до вирішення задач структурно-топологічного синтезу локальних обчислювальних мереж, проектуванні, розробці та впровадженні відповідних програмних рішень.

Ключові слова: локальна обчислювальна мережа, невизначеність, ймовірнісно-статистичні підходи, інтервальний підхід, нечіткий підхід, детерміновані підходи, синтез структури, модель, нечіткий критерій вартості, технологія дослідження

A. A. МИЩЕНКО, В. Ю. ВОЛОВЩИКОВ, В. Ф. ШАПО, М. А. ГРИНЧЕНКО

СИНТЕЗ СТРУКТУРЫ ЛОКАЛЬНОЙ ВЫЧИСЛИТЕЛЬНОЙ СЕТИ В УСЛОВИЯХ НЕОПРЕДЕЛЕННОСТИ ИСХОДНОЙ ИНФОРМАЦИИ

Идентифицирована проблема учета неопределенности исходной информации в задачах синтеза структуры локальных вычислительных сетей. Проведен анализ факторов, порождающих неопределенность, выделены источники неопределенности и подходы к формализации и исследованию неопределенности. Рассмотрены вероятностно-статистический, интервальный, нечеткий и детерминированный подходы, а также статистическая, субъективная, логическая интерпретация вероятности и стохастическое описание, концепция интервально-вероятностного подхода, модель непрерывно-детерминированная и дискретно-детерминированная. Сформулирована постановка задачи и предложено применять теорию нечетких множеств к формализации и исследованию неопределенности при синтезе структуры локальной вычислительной сети, основными элементами которой являются абонентские пункты, центры обработки информации, коммутационные приборы и каналы связи. Выполнение информационно-вычислительных работ, источниками которых являются абонентские пункты, возложено на центры обработки информации. Предполагается, что объемы информационно-вычислительных работ и географические координаты расположения элементов сети известны. Разработана модель задачи исследования на основе критерия стоимости. Параметры стоимости разделены на четко заданные и имеющие неопределенную природу. Нечеткие параметры предложено описывать функциями принадлежности вида "приблизительно равно" или "находится приблизительно в интервале". Выделены ограничения на технические возможности центров обработки информации, коммутационных приборов, трафик в каналах связи. Предложена технология исследования модели, основанная на идее решения задач линейного программирования в нечеткой постановке. Технология обеспечивает лица, принимающего решения, информацией о структуре сети, ее стоимости и уровне принадлежности решения. Результаты работы могут быть использованы при анализе подходов к решению задач структурно-топологического синтеза локальных вычислительных сетей, проектировании, разработке и внедрении соответствующих программных решений.

Ключевые слова: локальная вычислительная сеть, неопределенность, вероятностно-статистические подходы, интервальный подход, нечеткий подход, детерминированные подходы, синтез структуры, модель, нечеткий критерий стоимости, технология исследования

Introduction. The efficiency of using computing and peripheral equipment is increased due to the implementation of local area networks (LAN) [1].

It is necessary to go through many stages to get benefits from using the LAN: determine the purpose, choose the implementation option, analyze information flows, identify data storage and processing centers, form the structure, configure and test the LAN. One of the key stages is the LAN structure formation. It is important to take into account many different parameters, including those that may have an indefinite nature. Therefore, issues of formalization and uncertainty research are important and relevant.

Analysis of approaches to the formulation and study of uncertainty. We can identify factors causing uncertainty [2]: complexity, human and environmental factors. Then the sources of uncertainty can be divided into three groups: the randomness of the processes under consideration, the distortion and inaccessibility of information about possible events and processes, and opposition from other systems.

Systematizing [3–7], we single out approaches to formalization and the study of uncertainty.

Probabilistic-statistical approaches. This class of approaches is based on a statistical, subjective, and logical interpretation of probability with a stochastic description. The statistical interpretation of probability identifies probability with the relative frequency of a mass random event occurrence at sufficiently long tests. In conditions of limited experiment we are limited to selective estimates. The reliability of the results obtained on their basis depends on the existence of qualitative and voluminous statistical information. In a subjective interpretation of the probability, the probability is interpreted as the "degree of confidence" in the value of the uncertainty factor under consideration. The logical interpretation of probability is associated with an attempt to validate hypotheses based on logical considerations. Stochastic description is expedient if uncertainty factors can be attributed to probabilistic character and their probability density can be set.

The interval approach assumes the description of uncertainty factors in the interval form by the range of possible values of variables or dependencies. Accordingly, the length of the range is interpreted as a natural measure of uncertainty. The concept of the interval-probabilistic approach describes the available data by a set of close probability distributions, and weakens the assumption of statistical homogeneity of the observed events in the construction of probability hypotheses.

The fuzzy approach assumes that the elements of human thinking are not numbers, but sets, for which the transition from "belonging" to "non-belonging" is continuous. The main features of the approach are: the use of fuzzy and linguistic variables, the use of simple and complex relationships.

Deterministic approaches to accounting for uncertain factors play a significant role in the practice of creating various systems. Input and output variables are deterministic, internal connections in the system are known. It is assumed that there is a causal relationship between the

choice of a certain alternative and the onset of a corresponding outcome. Continuously deterministic and discretely deterministic models can be used to model systems.

Formulation of the problem. When building the structure of a LAN, it is important to take into account its constituent elements, the links between them, the presence of information flows, and the impact of external and internal environmental factors. This leads to the fact that the synthesis of the LAN structure can be influenced by uncertainty. Ignoring the accounting for uncertainty may lead to inadequate models and the adoption of unjustified decisions. Probabilistic-statistical, deterministic interval and fuzzy approaches can be used to formalize and assess uncertainties. If there is a lack of information for the application of probabilistic models, difficulties in operation with random variables, the ability to work with interval values within the limits of the fuzzy approach, it is expedient to apply the theory of fuzzy sets.

Thus, the paper proposes to consider the problem of the synthesis of the LAN structure in the conditions of uncertainty of the initial information, the generating factor of which we will consider the external environment. A fuzzy approach is applied to formalize uncertainty.

When building a LAN structure, problem statements based on the following criteria may be relevant depending on the performance indicator: cost, performance, reliability [8, 9]. Since the external environment has the greatest influence on the parameters that are independent of the internal processes of the enterprise, attention is paid to the cost criterion.

Among the variety of topological realizations of LAN [10], the paper considers radial node with an arbitrary number of levels of data switching.

Based on the ideology of the implementation and use of LAN, we will assume that the LAN has the task of serving the set I ($i \in I$) subscriber stations (SS). SS are the sources of informational computational works (ICW), the execution of which is entrusted to the set J ($j \in J$) of information processing centers (PIC). The transfer of the ICW from the SS to the PIC is performed via communication channels (CC) using the set K ($k \in K$) of switching devices (SD). It is assumed that the ICW volumes $\{h_{ij}\}$, the geographical coordinates of the SS, PIC and SD are known, and the distance vector $d = \{d_{vw}\}$, where $v, w \in IUJUK$, $v \neq w$ is defined accordingly.

Thus, the aim of the work is to develop a model for the synthesis of the LAN structure based on the cost criterion, taking into account the uncertainty of the initial data. The development of technology for the study of the synthesis model of the LAN structure is also given special attention in the work.

Model of synthesis of LAN structure. From the point of view of the presence in the LAN structure of the SS, PIC, SD, each of which can be implemented by one of the valid options $s \in S_i$, $q \in Q_j$, $m \in M_k$, we introduce the vector of Boolean variables $x = \{x_{is}\}$, $y = \{y_{jq}\}$, $z = \{z_{km}\}$, and the topology and variants of the implementation of the CC with the type l from the admissible set L will be described with $\alpha = \{\alpha_{ki}^l\}$, $\beta = \{\beta_{kj}^l\}$, $\gamma = \{\gamma_{jk}^l\}$ and $\varepsilon = \{\varepsilon_{kn}^l\}$. Moreover, let α_{ki}^l be a Boolean variable, which

determines the variant of connecting the SS i to the output of the SD k using CC l , β_{kj}^l – PIC j to the output of the SD k using CC l , γ_{jk}^l – PIC j to the output of the SD k using CC l , and ε_{kn}^l – from SD n to SD k using CC l . To build a LAN cost function, we define the vector of parameters $a = \{a_{is}\}$, $c = \{c_{jq}\}$, $r = \{r_{km}\}$ and $w = \{w_l\}$, which we associate with the costs of using SS, PIC, SD and CC models with various technical characteristics.

Analysis [11] showed the expediency of separating all cost parameters into two subsets. Let the first subset form a set of parameters that can be clearly specified $\bar{a} = \{\bar{a}_{is}\}$, $\bar{c} = \{\bar{c}_{jq}\}$, $\bar{r} = \{\bar{r}_{km}\}$ and $\bar{w} = \{\bar{w}_l\}$.

We associate the second subset with a set of parameters, the formation of which can be influenced by the external environment and which can have an indefinite nature $\tilde{a} = \{\tilde{a}_{is}\}$, $\tilde{c} = \{\tilde{c}_{jq}\}$, $\tilde{r} = \{\tilde{r}_{km}\}$ and $\tilde{w} = \{\tilde{w}_l\}$. To describe the latter, we will use the elements of the fuzzy sets theory and put in conformity with each fuzzy parameter the accessory function (AF) of the form "approximately equal to" or "is approximately in the range" $\mu(\tilde{a}_{is})$, $\mu(\tilde{c}_{jq})$, $\mu(\tilde{r}_{km})$ and $\mu(\tilde{w}_l)$.

The physical connection of the PIC j of the variant q to the SD is possible if the PIC has active g_{jq} network interfaces. The maximum number of connections to the SD SS, PIC and SD is limited by the permissible number of ports \bar{u}_{km} of the SD k with technical characteristics m . In order to avoid packet loss, the total traffic of the f_{tp} channel (t, p) connected to the SD p must not exceed the bandwidth \bar{u}_{pm} of this SD. At the same time, the total traffic of each channel (t, p) must not exceed the bandwidth of the CC b_{tp}^l when it is implemented by type l .

Therefore, the LAN structure must satisfy the following constraints.

Limitations on the technical capabilities of the PIC for connecting to them the SD:

$$N_j(\{y_{jq}\}, \{g_{jq}\}, \{y_{jk}^l\}) \geq 0, j \in J.$$

Limitations on the technical capabilities of the SD to connect to them the SS, PIC and SD:

$$P_k(\{\alpha_{ki}^l\}, \{\beta_{kj}^l\}, \{\varepsilon_{kn}\}, \{z_{km}\}, \{\bar{u}_{km}\}) \geq 0, k \in K.$$

Limitations on traffic entering to SD:

$$T_k^1(\{f_{ik}\}, \{z_{km}\}, \{\bar{u}_{km}\}) \geq 0, k \in K.$$

Restrictions on traffic in the CC between the SS and SD, PIC and SD, SD and SD:

$$T_{ki}^\alpha(\{f_{ki}\}, \{\alpha_{ki}^l\}, \{b_{ki}^l\}) > 0, k \in K, i \in I_k;$$

$$T_{kj}^\beta(\{f_{kj}\}, \{\beta_{kj}^l\}, \{b_{kj}^l\}) > 0, k \in K, j \in J_k;$$

$$T_{jk}^\gamma(\{f_{jk}\}, \{\gamma_{jk}^l\}, \{b_{jk}^l\}) > 0, j \in J, k \in K_j;$$

$$T_{kn}^\varepsilon(\{f_{kn}\}, \{\varepsilon_{kn}^l\}, \{b_{kn}^l\}) > 0, k, n \in K.$$

We define the fuzzy criterion of LAN cost, in the form of a function F , as the algebraic sum of the function \bar{F} with clearly defined parameters and fuzzy \tilde{F} :

$$\bar{F}(\bar{a}, x, \bar{c}, y, \bar{r}, z, \bar{w}, \alpha, \beta, \gamma, \varepsilon, d);$$

$$\tilde{F}(\tilde{a}, x, \tilde{c}, y, \tilde{r}, z, \tilde{w}, \alpha, \beta, \gamma, \varepsilon, d).$$

Model research technology. The technology will be based on the results of [12]. Then the AF of fuzzy parameters can be represented as exponential functions. For example, for \tilde{a}_{is} an AF, that provides a value equal to 1 in \tilde{a}_{is}^* , will take the form $\mu(\tilde{a}_{is}) = e^{-(\tilde{a}_{is} - \tilde{a}_{is}^*)^2 / 2D_{is}}$, where D_{is} is a parameter of the Gaussian fuzzy value \tilde{a}_{is} . Therefore, the AF F will have the form:

$$\mu(F) = e^{-(F - m(x, y, z, \alpha, \beta, \gamma, \varepsilon))^2 / 2D(x, y, z, \alpha, \beta, \gamma, \varepsilon)}.$$

If to suppose that it is required to get such alternative, the generalized characteristic of which as a level of belonging of decision must be not worse than the predetermined value θ , we equate $\mu(F)$ to θ . Using classical transformations, we get:

$$(F - m(x, y, z, \alpha, \beta, \gamma, \varepsilon))^2 = -2D(x, y, z, \alpha, \beta, \gamma, \varepsilon) \ln \theta.$$

Following the optimistic scenario of the synthesis of the LAN structure, the cost criterion F can be represented as:

$$m(x, y, z, \alpha, \beta, \gamma, \varepsilon) - (-2D(x, y, z, \alpha, \beta, \gamma, \varepsilon) \ln \theta)^{0.5}.$$

The pessimistic scenario should comply with the cost criterion F , defined follows:

$$m(x, y, z, \alpha, \beta, \gamma, \varepsilon) + (-2D(x, y, z, \alpha, \beta, \gamma, \varepsilon) \ln \theta)^{0.5}.$$

Then, by optimizing the obtained function F with the limitations of the model for the synthesis of the LAN structure, we can find an alternative $(x, y, z, \alpha, \beta, \gamma, \varepsilon)$, providing the decision maker with information about the network structure, its cost and the level of the decision not worse than θ .

Conclusion. The paper deals with the problem of formalization and the study of uncertainty. Probabilistic-statistical, interval, fuzzy and deterministic approaches are analyzed. The problem of the synthesis of the LAN structure in the conditions of uncertainty of the initial information is formulated. The model for the synthesis of the LAN structure formalizes the uncertainty using a fuzzy approach. The model research technology illustrates the peculiarities of solving the problem of LAN structure synthesis in fuzzy formulation.

Further studies will focus on the analysis of approaches to solving problems of structural-topological synthesis of LAN and the development of appropriate software solutions.

References

- Андерсон К., Минаси М. *Локальные сети: Полное руководство*. К.: БЕК+, СПб.: КОРОНА принт, 1999. 624 с.

2. Смирнова К.А. Понятие неопределенности экономических систем и подходы к ее оценке. *Вестник МГТУ*. Мурманск: МГТУ, 2008. Том 11. № 2. С. 241–246.
3. Воловщиков В.Ю. Подходы к решению задач системной оптимизации в условиях неопределенности исходной информации. *Авіаційно-космічна техніка і технологія*. Харків: Нац. Аерокосмічний ун-т "Харк. авіац. ін-т", 2002. Вип.29. С. 199–203.
4. Derhami S., Smith A.E. An Integer Programming Approach for Fuzzy Rule-Based Classification Systems. *European Journal of Operational Research*, 2017. Vol. 256, iss. 3. P. 924–934.
5. Кабаева Е.В. Применение методов интервального анализа. *Научный альманах*. Тамбов: ООО Консалтинговая компания Юком, 2015. № 11-3(13). С. 149–151.
6. Авраменко В.П., Ткаченко В.Ф., Серeda Л.Б. Принятие управленческих решений в условиях неопределенности и нечеткости исходной информации. *Радіоелектроніка, інформатика, управління*. Запорозжє: ЗНТУ, 2010. № 2. С. 101–106.
7. Симанков В.С., Бучацкая В.В., Теплоухов С.В. Подход к учету неопределенности исходной информации в системных исследованиях. *Вестник Адыгейского государственного университета*. Майкоп: ФГБОУВПО "Адыгейский государственный университет", 2017. № 3(206). С. 100–108.
8. Олифер В., Олифер Н. *Компьютерные сети. Принципы, технологии, протоколы*. СПб.: Питер, 2018. 992 с.
9. Зайченко Е. Ю. *Анализ и синтез структуры глобальных вычислительных сетей*. К.: ЗАО "Укрспецмонтажпроект", 1998. 107 с.
10. Таненбаум Э. *Компьютерные сети*. СПб.: Питер, 2019. 960 с.
11. Годлевский М.Д., Воловщиков В.Ю. Модель статической задачи структурного синтеза корпоративной информационно-вычислительной системы. *Східно-європейський журнал передових технологій*. Харків: ППТЦ, 2006. № 2/2 (20). С. 110–113.
12. Серая О.В., Зарубин В.С., Зинченко И.В., Лолашвили Б.Г. Решение задач линейного программирования в нечеткой постановке. *Вестник Национального технического университета "ХПИ"*. Харьков: НТУ "ХПИ", 2006. № 11. С. 160–167.
1. Anderson K., Minasi M. *Lokal'nyye seti: Polnoye rukovodstvo* [Local Area Networks: The Complete Guide]. Kiev, VEK+ Publ., St. Petersburg, KORONA print Publ., 1999. 624 p.
2. Smirnova K.A. Ponyatiye neopredelennosti ekonomicheskikh sistem i podkhody k yeye otsenke [The concept of uncertainty of economic systems and approaches to its assessment]. *Vestnik MGTU* [Bulletin of MSTU]. Murmansk, MSTU, 2008. Vol. 11, no. 2, pp. 241–246.
3. Volovshchikov V.Y. Podkhody k resheniyu zadach sistemnoy optimizatsii v usloviyakh neopredelennosti iskhodnoy informatsii [Approaches to solving problems of system optimization in the conditions of uncertainty of the initial information]. *Aviatsiynokosmichna tekhnika i tekhnologiya* [Aerospace engineering and technology]. Kharkiv, National Aerospace University "Kharkiv Aviation Institute" Publ., 2002. Vip. 29, pp. 199–203.
4. Derhami S., Smith A.E. An Integer Programming Approach for Fuzzy Rule-Based Classification Systems. *European Journal of Operational Research*. 2017, vol. 256, iss. 3, pp. 924–934.
5. Kabayeva E.V. Primeneniye metodov interval'nogo analiza [Application of interval analysis methods]. *Nauchnyy al'manakh* [Scientific almanac]. Tambov, Consulting company Ucom LLC Publ., 2015. No. 11-3(13), pp. 149–151.
6. Avramenko V.P., Tkachenko V.F., Sereda L.B. Prinyatiye upravlencheskikh resheniy v usloviyakh neopredelennosti i nechetkosti iskhodnoy informatsii [Making management decisions in the face of uncertainty and vagueness of the initial information]. *Radiyelektronika, informatika, upravlinnya* [Radio electronics, informatics, management]. Zaporozhye, ZNTU Publ., 2010. No. 2, pp. 101–106.
7. Simankov V.S., Buchatskaya V.V., Teploukhov S.V. Podkhod k uchetu neopredelennosti iskhodnoy informatsii v sistemnykh issledovaniyakh [Approach to taking into account the uncertainty of the initial information in system studies]. *Vestnik Adygeyskogo gosudarstvennogo universiteta* [Bulletin of the Adyge State University]. Maikop, FSBEI of HPE "Adyge State University" Publ., 2017. No. 3 (206), pp. 100–108.
8. Oliner V., Oliner N. *Komp'yuternyye seti. Printsipy, tekhnologii, protokoly* [Computer networks. Principles, technologies, protocols]. St. Petersburg, Peter Publ., 2018. 992 p.
9. Zaychenko Ye. Yu. *Analiz i sintez struktury global'nykh vychislitel'nykh setey* [Analysis and synthesis of the structure of global computer networks]. Kyiv, CJSC "Ukrspetsmontazhproekt" publ., 1998. 107 p.
10. Tanenbaum E. *Komp'yuternyye seti* [Computer networks]. St. Petersburg, Peter Publ., 2019. 960 p.
11. Godlevskiy M.D., Volovshchikov V.Y. Model' staticheskoy zadachi strukturnogo sinteza korporativnoy informatsionno-vychislitel'noy sistemy [The model of the static problem of structural synthesis of a corporate information-computing system]. *Skhidno-yevropeys'kyy zhurnal peredovykh tekhnolohiy* [Eastern European Journal of Advanced Technology]. Kharkiv, PPTC Publ., 2006. No. 2/2 (20), pp. 110–113.
12. Seraya O.V., Zarubin V.S., Zinchenko I.V., Lolashvili B.G. Resheniye zadach lineynogo programmirovaniya v nechetkoy postanovke [Solving linear programming problems in fuzzy setting]. *Vestnik Natsional'nogo tekhnicheskogo universiteta "KHPI"*, [Bulletin of the National Technical University "KhPI"]. Kharkiv: NTU "KhPI", 2006. No. 11, pp. 160–167.

Received 03.04.2019

Відомості про авторів / Сведения об авторах / About the Authors

Миценко Орина Олексіївна (Миценко Арина Алексеевна, Mishchenko Oryna Oleksiivna) – Національний технічний університет «Харківський політехнічний інститут», студентка; м. Харків, Україна; ORCID: <https://orcid.org/0000-0002-6783-673X>; e-mail: arinochka012@gmail.com

Воловщиков Валерій Юрійович (Воловщиков Валерий Юрьевич, Volovshchikov Valeriy Yuriyovich) – кандидат технічних наук, доцент, Національний технічний університет «Харківський політехнічний інститут», доцент кафедри програмної інженерії та інформаційних технологій управління; м. Харків, Україна; ORCID: <https://orcid.org/0000-0003-4454-2314>; e-mail: valera@kpi.kharkov.ua

Шапо Владлен Феліксович (Шапо Владлен Феликсович, Shapo Vladlen Felixovitch) – кандидат технічних наук, доцент, Національний університет «Одеська морська академія», доцент кафедри теорії автоматичного управління і обчислювальної техніки; м. Одеса, Україна; ORCID: <https://orcid.org/0000-0002-3921-4159>; e-mail: stani@te.net.ua

Гринченко Марина Анатоліївна (Гринченко Марина Анатольевна, Grinchenko Marina Anatoliivna) – кандидат технічних наук, доцент, Національний технічний університет «Харківський політехнічний інститут», доцент кафедри стратегічного управління; м. Харків, Україна; ORCID: <https://orcid.org/0000-0002-8383-2675>; e-mail: marinagrunchenko@gmail.com

Е. П. ПАВЛЕНКО, В. М. БУТЕНКО, В. А. ГУБИН

ИССЛЕДОВАНИЕ МЕТОДОВ РАЗРАБОТКИ ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ КОМПЬЮТЕРНОЙ ИНЖЕНЕРИИ НА ОСНОВЕ ТИПОВЫХ ПРОГРАММНЫХ ЭЛЕМЕНТОВ

В работе рассматриваются проблемы повышения эффективности разработки ИС, и, в частности, вопросы уменьшения сроков разработки программного комплекса ИС. Проведен анализ технологий разработки программного комплекса в жизненном цикле ИС. Структурный подход к программированию предлагал осуществлять декомпозицию программ методом пошаговой детализации. Разработка структур программы выполняется при помощи построения структур входных и выходных данных, идентификации связей обработки между этими данными, формирования структуры программы на основании структур данных и обнаруженных соответствий. Преодолеть фактор сложности можно, если отойти от прямолинейного подхода к решению поставленной задачи, состоящего в последовательном и линейном наращивании исходного текста программы оператор за оператором, в результате чего получается одна длинная и аморфная программа. Оказывается действенным принцип модульности: исходная задача разбивается на относительно независимые части; они реализуются отдельными программными модулями, которые затем связываются в единое целое на этапе компоновки. Выделены особенности технологии автоматизированного синтеза программ, а именно, технологии сборки программ из типовых программных элементов. Определены основные понятия указанной выше технологии, проведено исследование конструирования программ из блоков и выявлена проблемная область. Исходными данными для постановки и решения задач синтеза системы программных модулей являются множество информационных массивов системы, для которых определены: входные, выходные и промежуточные данные; множество альтернативных процедур обработки данных; последовательности выполнения процедур в процессе обработки; способы обмена с внешней памятью. Рассмотрен подход для выделения типовых программных элементов, удовлетворяющих определенным критериям. На основании рассмотренного подхода, а также с учетом его недостатков, был предложен усовершенствованный метод классификации типовых программных элементов и способ проектирования ПО на их основе с учетом минимизации времени и стоимости проекта.

Ключевые слова: программное обеспечение, компьютерная инженерия, информационные системы, компоненты, типовые проектные решения, затраты на разработку.

Є. П. ПАВЛЕНКО, В. М. БУТЕНКО, В. А. ГУБІН

ДОСЛІДЖЕННЯ МЕТОДІВ РОЗРОБКИ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ КОМП'ЮТЕРНОЇ ІНЖЕНЕРІЇ НА ОСНОВІ ТИПОВИХ ПРОГРАМНИХ ЕЛЕМЕНТІВ

У роботі розглядаються проблеми підвищення ефективності розробки ІС, зокрема, питання зменшення термінів розробки програмного комплексу ІС. Проведено аналіз технологій розробки програмного комплексу в життєвому циклі ІС. Структурний підхід до програмування пропонував здійснювати декомпозицію програм методом покрокової деталізації. Розробка структур програми виконується за допомогою побудови структур вхідних і вихідних даних, ідентифікації зв'язків обробки між цими даними, формування структури програми на підставі структур даних і виявлених відповідностей. Подолати фактор складності можна, якщо відійти від прямолінійного підходу до вирішення поставленого завдання: послідовному і лінійному нарощуванню вихідного тексту програми оператор за оператором, в результаті чого виходить одна довга і аморфна програма. Виявляється дієвим принцип модульності: вихідна задача розбивається на відносно незалежні частини; вони реалізуються окремими програмними модулями, які потім зв'язуються в єдине ціле на етапі компоновання. Виділено особливості технології автоматизованого синтезу програм, а саме, технології складання програм з типових програмних елементів. Визначено основні поняття зазначеної вище технології, проведено дослідження конструювання програм з блоків і виявлена проблемна область. Вихідними даними для постановки і рішення задач синтезу системи програмних модулів є множина інформаційних масивів системи, для яких визначені: вхідні, вихідні та проміжні дані; множина альтернативних процедур обробки даних; послідовності виконання процедур в процесі обробки; способи обміну із зовнішньою пам'яттю. Розглянуто підхід для виділення типових програмних елементів, які відповідають певним критеріям. На підставі розглянутого підходу, а також з урахуванням його недоліків, був запропонований вдосконалений метод класифікації типових програмних елементів і спосіб проектування ПЗ на їх основі з урахуванням мінімізації часу і вартості проекту.

Ключові слова: програмне забезпечення, комп'ютерна інженерія, інформаційні системи, компоненти, типові проектні рішення, витрати на розробку.

Y. P. PAVLENKO, V. M. BUTENKO, V. O. GUBIN

RESEARCH OF METHODS OF DEVELOPMENT OF SOFTWARE COMPUTER ENGINEERING BASED ON TYPICAL SOFTWARE ELEMENTS

The paper deals with the problems of increasing the effectiveness of the development of IS, and, in particular, the issues of reducing the development time of the software package of IS. The analysis of technology development software in the life cycle of IS. A structural programming approach suggested decomposing programs in a step-by-step manner. The development of program structures is carried out using the construction of input and output data structures, identification of processing links between these data, formation of a program structure based on data structures and detected matches. It is possible to overcome the complexity factor if we deviate from a straightforward approach to solving the problem posed, consisting in sequential and linear extension of the source code of the program operator-by-operator, resulting in one long and amorphous program. Here, the modularity principle is effective: the initial problem is divided into relatively independent parts; they are implemented by separate software modules, which are then linked into a single unit at the layout stage. The features of the technology of automated program synthesis, namely, the technology of assembling programs from typical program elements, are highlighted. The basic concepts of the above technology have been identified, a study has been conducted to design programs from blocks and a problem area has been identified. The initial data for the formulation and solution of problems for the synthesis of a system of program modules are the set of information arrays of the system, for which there are defined: input, output and intermediate data; many alternative data processing procedures; sequence of procedures in the processing; ways of sharing with external memory. The approach to the selection of typical program elements that meet certain criteria is considered. On the basis of the considered approach, and also taking into account its shortcomings, an improved method was proposed for classifying typical program elements and a method for designing software based on them, taking into account minimizing the time and cost of the project.

Keywords: software, computer engineering, information systems, components, typical design solutions, development costs.

© Е. П. Павленко, В. М. Бутенко, В. А. Губин, 2019

Введение. Создание программных систем для автоматизации предприятий является сложной и поэтапной задачей, поэтому для определения необходимых ресурсов, в том числе времени, важно регламентировать порядок разработки и использования некоторых видов типовых программных элементов по стадиям проектирования системы.

Для ускорения времени разработки и повышения качества кода программы разработано большое количество методик, технологий и подходов. Все они относятся не только к организации жизненного цикла ПО, но и к построению рациональной структуры организации, занимающейся разработкой информационных систем (ИС), и к каждому разработчику в частности.

Одним из решений проблем создания программного комплекса компьютерной инженерии является использование технологий автоматизации программирования, которые предполагают сокращение сроков разработки программных продуктов и повышение их характеристик качества.

Недостатки, изначально заложенные в требованиях на систему, невозможно компенсировать последующими версиями программных средств на различных этапах жизненного цикла. Поэтому проблема создания рациональной структуры ПО ИС является ключевой – её решение открывает новые возможности повышения эффективности ИС, а практическая значимость возрастает по мере усложнения и интеграции задач в ИС.

Постановка проблемы. Таким образом, можно прийти к выводу, что эффективность разработки ПО ИС определяется совокупностью множества факторов. Анализ показал многообразие существующих подходов к разработке ПО, их достоинства и недостатки, а также выявил общие задачи этапа реализации ПО ИС. Проанализированы структурный и объектно-ориентированный подходы к разработке ПО ИС.

Существенную помощь на этапе проектирования могут оказать разнообразные CASE-средства, но с их помощью можно построить лишь костяк ПО, их использование подразумевает дальнейшую доработку бизнес логики и отладку всего ПО.

На стадии проектирования ИС модели расширяются, уточняются и дополняются диаграммами, отражающими структуру программного обеспечения: архитектуру ПО, структурные схемы программ и диаграммы экранных форм.

Таким образом, возникает задача исследования технологий использования синтеза программ из типовых элементов, позволяющих избежать доработки выходного продукта, а также в исследовании повышения эффективности их применения.

Анализ состояния проблемы. Структурный подход к программированию в том виде, в котором он был сформулирован в 70-х годах XX в., предлагал осуществлять декомпозицию программ методом пошаговой детализации [1]. Результатом декомпозиции является структурная схема программы, которая представляет собой многоуровневую иерархическую схему взаимодействия подпрограмм по управлению. Минимально такая схема отображает два уровня иерархии, т. е. показывает общую структуру программы [2].

Однако тот же метод позволяет получить структурные схемы с большим количеством уровней.

Для анализа технологичности полученной иерархии модулей используют структурные карты Константайна [3] или Джексона [4]. На структурной карте отношения между модулями представляют в виде графа, вершинам которого соответствуют модули и общие области данных, а дугам – межмодульные вызовы и обращения к общим областям данных.

Существуют методики проектирования программного обеспечения Джексона [5] и Варнье – Орра [6], основанные на декомпозиции данных. При создании своей методики авторы исходили из того, что структуры исходных данных и результатов определяют структуру программы.

Разработка структур программы в соответствии с [5] выполняется следующим образом: строят изображение структур входных и выходных данных; выполняют идентификацию связей обработки между этими данными; формируют структуру программы на основании структур данных и обнаруженных соответствий; добавляют блоки обработки элементов, для которых не обнаружены соответствия; анализируют и обрабатывают несоответствия; добавляют необходимые операции (ввод, вывод, открытие/закрытие файлов); записывают программу в структурной нотации.

Однако методика [5, 6] может использоваться только в том случае, если данные разрабатываемых программ могут быть представлены в виде иерархии или совокупности иерархий.

Практика создания сложных программных систем говорит о том, что преодолеть фактор сложности можно, если отойти от прямолинейного подхода к решению поставленной задачи, состоящего в последовательном и линейном наращивании исходного текста программы оператор за оператором, строка за строкой, в результате чего получается одна длинная и аморфная программа. Здесь оказывается действенным принцип модульности. Он заключается в том, что исходная задача разбивается на относительно независимые части; они реализуются отдельными программными модулями, которые затем связываются в единое целое на этапе компоновки [7].

Процесс разбиения одной сложной задачи на подзадачи и распределение функций между модулями определяется принятым методом проектирования, а синтаксические и функциональные средства оформления программных модулей в виде законченных программных единиц определяются возможностями алгоритмического языка и операционной системы.

Для объектно-ориентированных программных систем применяется иной подход [8]. Поскольку экземпляры классов обмениваются сообщениями, определяют для каждого класса поступающие его экземплярам сообщения и на их основе строят диаграммы перехода (описать класс как конечный автомат). Строят также модели состояний для каждого объекта и определяют списки событий, изменяющих состояние объектов.

После выделения классов и их неформального описания строятся модели процессов, которые должны

быть реализованы в будущем ПО. В такой модели отражаются внешние события (действия пользователя) и вызываемые этими событиями действия с экземплярами классов. Если пользователь указал все характеристики данного правильно, нужно построить экземпляр объекта “данное” и включить его в множество данных [9].

Разработка объектно-ориентированных программных систем для железнодорожного транспорта имеет свою специфику, в частности, необходимость поддерживать многоуровневые структуры данных и реализовывать структуру классов с наследованием [10], [11]. Выбор технологии разработки ПО для таких систем предлагается выполнять с помощью процедур, описанных в [12]. Методы разработки программного обеспечения возможно использовать в распределенных вычислениях в различных компонентах информационных систем на железнодорожном транспорте [13].

Таким образом, налицо многообразие существующих подходов к разработке ПО, имеющих свои достоинства и недостатки. Выявились также общие задачи этапа реализации ПО ИС.

Метод решения проблемы. Причины применения технологии типового проектирования ПО:

- существенно снижаются затраты на проектирование, разработку и модернизацию ИС;
- больше возможностей обеспечивать должный научно-технический уровень разработки ИС (в отличие от технологии индивидуального проектирования).

Технология типового проектирования ПО является одной из разновидностей индустриального проектирования. Заключается в создании ПО информационной системы из готовых типовых элементов.

Область применения: автоматизация деятельности таких объектов, для которых характерны общие правила функционирования и управления. В первую очередь, сюда относятся экономические системы, для которых характерны:

- схожая структура и правила управления;
- единые стандарты отчетности;
- схожие комплексы используемых технических и программных средств;
- единая цель существования: извлечение прибыли.

Процесс проектирования ПО ИС состоит из следующих основных этапов:

- разбиение проекта информационной системы на отдельные составляющие (компоненты);
- выбор и приобретения имеющихся на рынке типовых проектных решений (тиражируемых продуктов) для каждого компонента ИС;
- настройка и доработка приобретенных типовых проектных решений в соответствии с требованиями конкретной предметной области.

Должна быть принципиальная возможность декомпозиции информационной системы на множество составляющих компонентов (комплексов, подсистем, программных модулей). В зависимости от уровня декомпозиции информационной системы различают следующие классы типовых проектных решений: элементные, подсистемные, объектные.

Типовое проектное решение (ТПР) – это представленное в виде комплекта проектной документации или набора программных модулей проектное решение, пригодное к многократному использованию.

Основные черты ТПР:

- типовые проектные решения ориентированы на автоматизацию деятельности множества однородных объектов (путем настройки под конкретные особенности каждого из них);
- основная цель применения ТПР – уменьшение трудоемкости и стоимости проектирования или разработки ПО ИС;
- создание ТПР возможно только после тщательного и всестороннего изучения предметной области и предполагает обобщение накопленного в частных случаях опыта (путем классификации, типизации, абстрагирования, унификации).

Типовые решения бывают простыми или комбинированными. Простые ТПР охватывают только какой-либо один вид обеспечения ИС, комбинированные – два и более. Примеры простых ТПР: классификаторы ИО, прикладные программы общего и специального назначения, инструктирующие руководства по управлению бизнес-процессами, рекомендации по составлению ТЗ.

Требования, выдвигаемые к типовым проектным решениям:

- возможность использования для создания новой ИС при минимальном участии разработчиков ТПР;
- соответствие требованиям положений и стандартов, распространяемых на информационную систему в целом или ее часть;
- способность удовлетворять максимально возможному числу потребностей в рамках своего функционального назначения;
- возможность адаптации к конкретным условиям проекта путем изменения параметров.

В качестве типового элемента при элементном проектировании используются простые ТПР, относящиеся к отдельной задаче ИС. В этом случае ИС комплектуется как множество ТПР по отдельным разрозненным задачам. Дополнительные элементы, для которых отсутствуют ТПР, разрабатываются вручную.

Исходными данными для постановки и решения динамических задач синтеза системы модулей являются множество информационных массивов системы, для которых определены: входные, выходные и промежуточные данные; множество альтернативных процедур обработки данных, входящих в состав алгоритмов, обеспечивающих решение задач системы; возможные последовательности выполнения процедур в процессе обработки; способы обмена с внешней памятью модульной системы обработки данных; взаимосвязи процедур с информационными массивами; временные характеристики обращения к массивам.

Выделим множество программных блоков

$$P = \{P_1, P_2, \dots, P_n\},$$

которые можно использовать при проектировании ПО ИС. Каждый блок характеризуется затратами на

разработку Z_1, Z_2, \dots, Z_n и временем на разработку T_1, T_2, \dots, T_n .

Функциональность программной системы включает в себя m различных функций по обработке информации

$$F = \{f_1, f_2, \dots, f_m\}.$$

Можно выделить два основных критерия оценки эффективности синтеза ПО ИС из функциональных блоков: затраты и время.

Постановка задачи формулируется следующим образом: найти такое подмножество программных блоков

$$\begin{aligned} P' &\subset P, \\ P' &= \{P_i, P_k, \dots, P_l\}, \end{aligned}$$

чтобы реализовать все m функций по обработке информации.

Ограничения:

$$\begin{aligned} Z &= Z_{\text{сопр}} + Z_i + Z_k + \dots + Z_l \rightarrow \min, \\ T &= T_{\text{сопр}} + T_i + T_k + \dots + T_l \rightarrow \min, \end{aligned}$$

где $Z_{\text{сопр}}$ – затраты на сопряжение программных блоков,

$T_{\text{сопр}}$ – время на сопряжение программных блоков.

Предлагается следующий метод решения задачи выбора типовых программных элементов.

Шаг 1. Рассмотрим функцию f_1 . Если программный блок P_1 необходим при реализации этой функции, то $Z = Z_1$, $T = T_1$.

Шаг 2. Если программный блок P_i необходим при реализации этой функции, то $Z = Z + Z_i$, $T = T + T_i$. Иначе переходим к шагу 3.

Шаг 3. Если функция f_1 реализована полностью, то переходим к рассмотрению функции f_2 .

Шаг 4. Если программный блок P_1 необходим при реализации этой функции и он не был задействован при реализации предыдущей функции, то $Z = Z + Z_1$, $T = T + T_1$.

Шаг 5. Если программный блок P_i необходим при реализации функции, то $Z = Z + Z_i$, $T = T + T_i$. Иначе переходим к шагу 5.

Шаг 6. Если функция реализована полностью, то переходим к рассмотрению следующей функции

Шаг 7. Если все функции рассмотрены, то оцениваем $Z_{\text{сопр}}$ и $T_{\text{сопр}}$ и фиксируем затраты и время по первому варианту рассмотрения.

$$\begin{aligned} Z^{(1)} &= Z_{\text{сопр}} + Z, \\ T^{(1)} &= T_{\text{сопр}} + T. \end{aligned}$$

Шаг 8. Начнем с рассмотрения функции f_2 и повторим шаги 1–7. Зафиксируем затраты и время по следующему варианту рассмотрения.

Шаг 9. Определим

$$Z = \min_i Z^{(i)}.$$

Таким образом, мы определили пару (Z, T) , а также подмножество программных блоков P' , являющееся решением задачи.

Выводы. В работе проведён анализ проблемы обоснования выбора программных компонентов при

проектировании программного обеспечения ИС, анализ технологий разработки программного комплекса в жизненном цикле ИС, исследован принцип модульности при проектировании ПО.

На основании проведённого исследования разработан метод решения задачи выбора типовых программных элементов, предложены критерии выбора программных элементов, разработан алгоритм решения задачи обоснования выбора программных компонентов. Полученные результаты работы опробованы на реальном примере – при проектировании ПО ИС, состоящей из 4 функциональных задач.

Разработанный метод в результате небольших изменений может быть использован для обоснования выбора программных элементов при проектировании других видов ПО, и разрабатываемое программное обеспечение может быть улучшено за счёт увеличения его функциональности и сокращения затрат на разработку.

Список литературы

1. Структурное программирование. URL: <http://www.maksakovsa.ru/TehProgram/StrProgr/index.html> (дата обращения: 16.12.2018)
2. Лингер Р., Миллс Х., Уитт Б. Теория и практика структурного программирования. URL: <https://www.livelib.ru/book/1001405237/about-teoriya-i-praktika-struktornogo-programmirovaniya-richard-linger> (дата обращения: 16.12.2018)
3. Средства структурного проектирования. URL: <http://www.interface.ru/fset.asp?Url=/case/defs71.htm> (дата обращения: 04.02.2019)
4. Структурные карты Джексона. URL: <http://mashxxl.info/article/433911/> (дата обращения: 04.02.2019)
5. Проектирование программного обеспечения. Основы программной инженерии. URL: http://swebok.sorlik.ru/2-software_engineering_designi (дата обращения: 04.02.2019)
6. Современные методологии структурного анализа и проектирования. URL: <http://swwsys.ru/index.php?page=article&id=1025> (дата обращения: 04.02.2019)
7. Проектирование программного обеспечения. URL: <http://window.edu.ru/library/pdf2txt/965/28965> (дата обращения: 04.02.2019)
8. Объектно-ориентированный анализ и проектирование. URL: <http://www.hardline.ru/1/5/1390/1789-6.htm> (дата обращения: 04.02.2019)
9. Буч Г. Объектно-ориентированный анализ и проектирование с примерами приложений. URL: <https://www.twirpx.com/file/279137> (дата обращения: 04.02.2019)
10. Listrovoy S. V., Butenko V. M., Bryksin V. O., Golovko O. V. Development of method of definition maximum clique in a non-oriented graph. *Eastern European Journal of Enterprise Technologies*. 2017. Vol. 5, № 4 (89). – P. 12–17. EID: 2-s2.0-85032585697 DOI: 10.15587/1729-4061.2017.111056.
11. Лістровий С. В., Панченко С. В., Мойсеєнко В. І., Бутенко В. М. Математичне моделювання в розподілених інформаційно-керуючих системах залізничного транспорту: Монографія. Харків: ФОП Бровін О. В., 2017. 220 с
12. Павленко Е. П., Берюх И. С., Айвазов В. А. Выбор технологии разработки программного обеспечения информационных систем. *ScienceRise*. Харьков, 2015. № 5/2. С.40–43
13. S. V. Listrovoy, O. V. Golovko, V. M. Butenko, M. V. Ushakov Formulation of the Problem of Maximum Clique Determination in Non-Oriented Graphs. *International Journal of Engineering & Technology*. Vol. 7, no. 4.3 (2018): Special Issue 3. P. 293–297.

References (transliterated)

1. *Strukturoye programmirovaniye*. [Structured programming] Available at: <http://www.maksakovsa.ru/TehProgram/StrProgr/index.html> (accessed 16.12.2018).

2. Linger R., Mills H., Witt B. *Teoriya i praktika struktornogo programmirovaniya*. [Theory and practice of structured programming] Available at: <https://www.livelib.ru/book/1001405237/about-teoriya-i-praktika-struktornogo-programmirovaniya-richard-linger> (accessed 16.12.2018).
3. *Sredstva struktornogo proyektirovaniya*. [Structural Design Tools] Available at: <http://www.interface.ru/fset.asp?Url=/case/defs71.htm> (accessed 04.02.2019).
4. *Strukturnyye karty Dzheksona*. [Jackson Structural Maps] Available at: <http://mash-xxl.info/article/433911> (accessed 04.02.2019).
5. *Proyektirovaniye programmnoy obespecheniya. Osnovy programmnoy inzhenerii*. [Software design. Software Engineering Basics] Available at: http://swebok.sorlik.ru/2-software_engineering_designi (accessed 04.02.2019).
6. *Sovremennyye metodologii struktornogo analiza i proyektirovaniya*. [Modern methodologies of structural analysis and design] Available at: <http://swsys.ru/index.php?page=article&id=1025> (accessed 04.02.2019).
7. *Proyektirovaniye programmnoy obespecheniya*. [Software design] Available at: <http://window.edu.ru/library/pdf2txt/965/28965> (accessed 04.02.2019).
8. *Ob'yektno-oriyentirovannyi analiz i proyektirovaniye*. [Object-oriented analysis and design] Available at: <http://www.hardline.ru/1/5/1390/1789-6.htm> (accessed 04.02.2019).
9. Booch G. *Ob'yektno-oriyentirovannyi analiz i proyektirovaniye s primerami prilozheniy*. [Object-oriented analysis and design with examples of applications] Available at: <https://www.twirpx.com/file/279137> (accessed 04.02.2019).
10. Listrovoy S. V., Butenko V. M., Bryksin V. O., Golovko O. V. Development of method of definition maximum clique in a non-oriented graph. *Eastern European Journal of Enterprise Technologies*. 2017, vol. 5, no. 4 (89), pp. 12–17. EID: 2-s2.0-85032585697 DOI: 10.15587/1729-4061.2017.111056).
11. Listrovoy S. V., Panchenko S. V., Mojseenko V. I., Butenko V. M. *Matematychni modeli v rozpodilennykh informatsiynokeryuyuchykh systemakh zaliznychnoho transportu* [Mathematical modeling in distributed information-control systems of railway transport]. Kharkiv, 2017. 220 p.
12. Pavlenko E. P., Beruh I. C., Ayvazov V. A. Vybory tekhnologii razrabotki programmnoy obespecheniya informatsiynnykh sistem [The choice of technology for developing software information systems]. *ScienceRise* [ScienceRise]. Kharkov, 2015, no. 5/2, pp. 40–43
13. Listrovoy S. V., Golovko O. V., Butenko V. M., Ushakov M. V. Formulation of the Problem of Maximum Clique Determination in Non-Oriented Graphs. *International Journal of Engineering & Technology*. 2018, vol. 7, no. 4.3: Special Issue 3, pp. 293–297

Поступила (received) 19.02.2019

Відомості про авторів / Сведения об авторах / About the Authors

Павленко Євген Петрович (Павленко Евгений Петрович, Pavlenko Yevhen Petrovych) – кандидат технічних наук, доцент, Український державний університет залізничного транспорту, доцент кафедри спеціалізованих комп'ютерних систем; м. Харків, Україна; ORCID: <http://orcid.org/0000-0002-7626-9933>; e-mail: evgenijpavlenko821@gmail.com

Бутенко Володимир Михайлович (Бутенко Владимир Михайлович, Butenko Vladimir Mihajlovych) – кандидат технічних наук, доцент, Український державний університет залізничного транспорту, доцент кафедри спеціалізованих комп'ютерних систем; м. Харків, Україна; ORCID: <http://orcid.org/0000-0001-9958-3960>; e-mail: docent.butenko@gmail.com

Губін Вадим Олександрович (Губин Вадим Александрович, Gubin Vadim Oleksandrovych) – Національний університет радіоелектроніки, старший викладач кафедри штучного інтелекту; м. Харків, Україна; ORCID: <http://orcid.org/0000-0003-1850-1930>; e-mail: vadim.gubin@nure.ua

ЗМІСТ

| | |
|--|----|
| СИСТЕМНИЙ АНАЛІЗ І ТЕОРІЯ ПРИЙНЯТТЯ РІШЕНЬ | 3 |
| <i>Pavlov A. A.</i> Combinatorial optimization under uncertainty and formal models of expert estimation | 3 |
| <i>Хом'як Т. В., Малієнко А. В., Симонець Г. В.</i> Застосування методів згладжування для прогнозування обсягу виробництва..... | 8 |
| <i>Смачило В. В.</i> Нормативно-правове забезпечення діяльності у сфері транспорту | 13 |
| УПРАВЛІННЯ В ОРГАНІЗАЦІЙНИХ СИСТЕМАХ..... | 20 |
| <i>Pavlov A. A., Misura E. B., Melnikov O. V.</i> Total weighted tardiness minimization for tasks with a common due date on parallel machines in case of agreeable weights and processing times..... | 20 |
| <i>Kozulia T. V., Sviridova A. S., Kozulia M. M.</i> Problems of elicitation and analysis of requirements to the program module of monitoring on the basis of the conditions of cognitive analysis | 25 |
| <i>Goloskokov A. E., Tkachenko D.V.</i> Models and software solutions for the problem of diagnosing the financial state of IT-enterprise | 30 |
| МАТЕМАТИЧНЕ І КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ..... | 38 |
| <i>Марченко Н. А., Руденко Р. О.</i> Поліпшений метод дослідження стійкості розв'язування систем лінійних алгебраїчних рівнянь..... | 38 |
| <i>Марченко І. І., Малько М. М., Марченко І. Г.</i> Комп'ютерне моделювання процесів дифузії у похилих просторово-періодичних потенціалах..... | 43 |
| ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ..... | 48 |
| <i>Shaptala R. V., Kyselev G. D.</i> Using graph embeddings for Wikipedia link prediction..... | 48 |
| <i>Orlovskiy D. L., Kopp A. M., Pyvovarova A. A.</i> An information technology for the support for solving the analysis problem of the customers' value within the enterprise CRM-strategy implementation..... | 53 |
| <i>Mishchenko O. O., Volovshchikov V. Y., Shapo V. F., Grinchenko M. A.</i> Synthesis of local area network structure in uncertain conditions of initial information..... | 63 |
| <i>Павленко Е. П., Бутенко В. М., Губин В. А.</i> Исследование методов разработки программного обеспечения компьютерной инженерии на основе типовых программных элементов | 67 |

CONTENT

| | |
|--|----|
| SYSTEM ANALYSIS AND DECISION-MAKING THEORY | 3 |
| <i>Pavlov A. A.</i> Combinatorial optimization under uncertainty and formal models of expert estimation | 3 |
| <i>Khomyak T. V., Malienko A. V., Symonets H. V.</i> Application of smoothing methods for forecasting production volume..... | 8 |
| <i>Smachylo V. V.</i> Normative-legal framework of transport | 13 |
| MANAGEMENT IN ORGANIZATIONAL SYSTEMS | 20 |
| <i>Pavlov A. A., Misura E. B., Melnikov O. V.</i> Total weighted tardiness minimization for tasks with a common due date on parallel machines in case of agreeable weights and processing times..... | 20 |
| <i>Kozulia T. V., Sviridova A. S., Kozulia M. M.</i> Problems of elicitation and analysis of requirements to the program module of monitoring on the basis of the conditions of cognitive analysis | 25 |
| <i>Goloskokov A. E., Tkachenko D.V.</i> Models and software solutions for the problem of diagnosing the financial state of IT-enterprise | 30 |
| MATHEMATICAL AND COMPUTER MODELING..... | 38 |
| <i>Marchenko N. A., Rudenko R. O.</i> Improved method for stability research of solutions of systems of linear algebraic equations..... | 38 |
| <i>Marchenko I. I., Malko M. M., Marchenko I. G.</i> Computer simulation of diffusion processes in tilt spatio-periodic potentials..... | 43 |
| INFORMATION TECHNOLOGY | 48 |
| <i>Shaptala R. V., Kyselev G. D.</i> Using graph embeddings for Wikipedia link prediction..... | 48 |
| <i>Orlovskiy D. L., Kopp A. M., Pyvovarova A. A.</i> An information technology for the support for solving the analysis problem of the customers' value within the enterprise CRM-strategy implementation..... | 53 |
| <i>Mishchenko O. O., Volovshchikov V. Y., Shapo V. F., Grinchenko M. A.</i> Synthesis of local area network structure in uncertain conditions of initial information..... | 63 |
| <i>Pavlenko Y. P., Butenko V. M., Gubin V. O.</i> Research of methods of development of software computer engineering based on typical software elements | 67 |

НАУКОВЕ ВИДАННЯ

**ВІСНИК НАЦІОНАЛЬНОГО ТЕХНІЧНОГО УНІВЕРСИТЕТУ «ХПІ».
СЕРІЯ: СИСТЕМНИЙ АНАЛІЗ, УПРАВЛІННЯ ТА ІНФОРМАЦІЙНІ
ТЕХНОЛОГІЇ**

Збірник наукових праць

№ 1'2019

Наукові редактори: М. Д. Годлевський, д-р техн. наук, професор, НТУ «ХПІ», Україна
О. С. Куценко, д-р техн. наук, професор, НТУ «ХПІ», Україна
Технічний редактор: М. І. Безменов, канд. техн. наук, професор, НТУ «ХПІ», Україна

Відповідальний за випуск М. І. Безменов, канд. техн. наук, професор

АДРЕСА РЕДКОЛЕГІЇ ТА ВИДАВЦЯ: 61002, Харків, вул. Кирпичова, 2, НТУ «ХПІ».
Кафедра системного аналізу та інформаційно-аналітичних технологій
Тел.: (057) 707-61-03, (057) 707-66-54; e-mail: bezmenov@kpi.kharkov.ua

Підп. до друку 05.07.2019 р. Формат 60×84 1/8. Папір офсетний.
Друк офсетний. Гарнітура Таймс. Умов. друк. арк. 8,0. Облік.-вид. арк. 8,75.
Тираж 100 пр. Зам. № 84. Ціна договірна.

Друкарня «ФОП Пісня О. В.»

Свідоцтво про державну реєстрацію ВО2 № 248750 від 13.09.2007 р.
61002, Харків, вул. Гіршмана, 16а, кв. 21, тел. (057) 764-20-28