

ISSN 2713-3192  
DOI 10.15622/ia.2022.21.3  
<http://ia.spcras.ru>

ТОМ 21 № 3

# ИНФОРМАТИКА И АВТОМАТИЗАЦИЯ

# INFORMATICS AND AUTOMATION



СПб ФИЦ РАН

Санкт-Петербург  
2022



# INFORMATICS AND AUTOMATION

Volume 21 № 3, 2022

Scientific and educational journal primarily specialized in computer science, automation, robotics, applied mathematics, interdisciplinary research

Founded in 2002

---

## Founder and Publisher

St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS)

---

## Editor-in-Chief

**R. M. Yusupov**, Prof., Dr. Sci., Corr. Member of RAS, St. Petersburg, Russia

---

## Editorial Council

<b>A. A. Ashimov</b>	Prof., Dr. Sci., Academician of the National Academy of Sciences of the Republic of Kazakhstan, Almaty, Kazakhstan
<b>N. P. Veselkin</b>	Prof., Dr. Sci., Academician of RAS, St. Petersburg, Russia
<b>I. A. Kalyaev</b>	Prof., Dr. Sci., Academician of RAS, Taganrog, Russia
<b>Yu. A. Merkur'yev</b>	Prof., Dr. Sci., Academician of the Latvian Academy of Sciences, Riga, Latvia
<b>A. I. Rudskoi</b>	Prof., Dr. Sci., Academician of RAS, St. Petersburg, Russia
<b>V. Sgurev</b>	Prof., Dr. Sci., Academician of the Bulgarian Academy of Sciences, Sofia, Bulgaria
<b>B. Ya. Sovetov</b>	Prof., Dr. Sci., Academician of RAE, St. Petersburg, Russia
<b>V. A. Soyfer</b>	Prof., Dr. Sci., Academician of RAS, Samara, Russia

---

## Editorial Board

<b>O. Yu. Gusikhin</b>	Ph. D., Dearborn, USA
<b>V. Delic</b>	Prof., Dr. Sci., Novi Sad, Serbia
<b>A. Dolgui</b>	Prof., Dr. Sci., St. Etienne, France
<b>M. N. Favorskaya</b>	Prof., Dr. Sci., Krasnoyarsk, Russia
<b>M. Zelezny</b>	Assoc. Prof., Ph.D., Plzen, Czech Republic
<b>H. Kaya</b>	Assoc. Prof., Ph.D., Utrecht, Netherlands
<b>A. A. Karpov</b>	Assoc. Prof., Dr. Sci., St. Petersburg, Russia
<b>S. V. Kuleshov</b>	Dr. Sci., St. Petersburg, Russia
<b>A. D. Khomonenko</b>	Prof., Dr. Sci., St. Petersburg, Russia
<b>D. A. Ivanov</b>	Prof., Dr. Habil., Berlin, Germany
<b>K. P. Markov</b>	Assoc. Prof., Ph.D., Aizu, Japan
<b>R. V. Meshcheryakov</b>	Prof., Dr. Sci., Moscow, Russia
<b>N. A. Moldovian</b>	Prof., Dr. Sci., St. Petersburg, Russia
<b>V. V. Nikulin</b>	Prof., Ph.D., New York, United States
<b>V. Yu. Osipov</b>	Prof., Dr. Sci., St. Petersburg, Russia
<b>V. K. Pshikhopov</b>	Prof., Dr. Sci., Taganrog, Russia
<b>A. L. Ronzhin</b>	Prof., Dr. Sci., Deputy Editor-in-Chief, St. Petersburg, Russia
<b>H. Samani</b>	Assoc. Prof., Ph.D., Plymouth, UK
<b>A. V. Smirnov</b>	Prof., Dr. Sci., St. Petersburg, Russia
<b>B. V. Sokolov</b>	Prof., Dr. Sci., St. Petersburg, Russia
<b>L. V. Utkin</b>	Prof., Dr. Sci., St. Petersburg, Russia

---

**Editor:** A.S. Lopotova

**Interpreter:** Ya.N. Berezina

**Art editor:** N.A. Dormidontova

---

## Editorial office address

SPC RAS, 14-th line V.O., 39 litera A , St. Petersburg, Russia, 199178

e-mail: [ia@spcras.ru](mailto:ia@spcras.ru), web: <http://ia.spcras.ru>

## The journal is indexed in Scopus

The journal is published under the scientific-methodological supervision of Department for Nanotechnologies and Information Technologies of the Russian Academy of Sciences

© St. Petersburg Federal Research Center of the Russian Academy of Sciences, 2022

# ИНФОРМАТИКА И АВТОМАТИЗАЦИЯ

Том 21 № 3, 2022

Научный, научно-образовательный журнал с базовой специализацией  
в области информатики, автоматизации, робототехники, прикладной математики  
и междисциплинарных исследований.

Журнал основан в 2002 году

---

## Учредитель и издатель

Федеральное государственное бюджетное учреждение науки  
«Санкт-Петербургский Федеральный исследовательский центр Российской академии наук»  
(СПб ФИЦ РАН)

---

## Главный редактор

**Р. М. Юсупов**, чл.-корр. РАН, д-р техн. наук, проф., Санкт-Петербург, РФ

---

## Редакционный совет

<b>А. А. Ашимов</b>	академик Национальной академии наук Республики Казахстан, д-р техн. наук, проф., Алматы, Казахстан
<b>Н. П. Веселкин</b>	академик РАН, д-р мед. наук, проф., Санкт-Петербург, РФ
<b>И. А. Каляев</b>	академик РАН, д-р техн. наук, проф., Таганрог, РФ
<b>Ю. А. Меркурьев</b>	академик Латвийской академии наук, д-р, проф., Рига, Латвия
<b>А. И. Рудской</b>	академик РАН, д-р техн. наук, проф., Санкт-Петербург, РФ
<b>В. Сгурев</b>	академик Болгарской академии наук, д-р техн. наук, проф., София, Болгария
<b>Б. Я. Советов</b>	академик РАН, д-р техн. наук, проф., Санкт-Петербург, РФ
<b>В. А. Сойфер</b>	академик РАН, д-р техн. наук, проф., Самара, РФ

---

## Редакционная коллегия

<b>О. Ю. Гусихин</b>	д-р наук, Диаборн, США
<b>В. Делич</b>	д-р техн. наук, проф., Нови-Сад, Сербия
<b>А. Б. Долгий</b>	д-р наук, проф. Сент-Этьен, Франция
<b>М. Железны</b>	д-р наук, доцент, Пльзень, Чешская республика
<b>Д. А. Иванов</b>	д-р экон. наук, проф., Берлин, Германия
<b>Х. Кайя</b>	д-р наук, доцент, Утрехт, Нидерланды
<b>А. А. Карпов</b>	д-р техн. наук, доцент, Санкт-Петербург, РФ
<b>С. В. Кулешов</b>	д-р техн. наук, Санкт-Петербург, РФ
<b>К. П. Марков</b>	д-р наук, доцент, Аизу, Япония
<b>Р. В. Мещеряков</b>	д-р техн. наук, проф., Москва, РФ
<b>Н. А. Молдовян</b>	д-р техн. наук, проф., Санкт-Петербург, РФ
<b>В.В. Никулин</b>	д-р наук, проф., Нью-Йорк, США
<b>В.Ю. Осипов</b>	д-р техн. наук, проф., Санкт-Петербург, РФ
<b>В. Х. Пшихопов</b>	д-р техн. наук, проф., Таганрог, РФ
<b>А. Л. Ронжин</b>	д-р техн. наук, проф., зам. главного редактора, Санкт-Петербург, РФ
<b>Х. Самани</b>	д-р наук, доцент, Плимут, Соединённое Королевство
<b>А. В. Смирнов</b>	д-р техн. наук, проф., Санкт-Петербург, РФ
<b>Б. В. Соколов</b>	д-р техн. наук, проф., Санкт-Петербург, РФ
<b>Л. В. Уткин</b>	д-р техн. наук, проф., Санкт-Петербург, РФ
<b>М. Н. Фаворская</b>	д-р техн. наук, проф., Красноярск, РФ
<b>А. Д. Хомоненко</b>	д-р техн. наук, проф., Санкт-Петербург, РФ
<b>Л. Б. Шереметов</b>	д-р техн. наук, Мехико, Мексика

---

**Выпускающий редактор:** А.С. Лопотова

**Переводчик:** Я.Н. Березина

**Художественный редактор:** Н.А. Дормидонтова

---

## Адрес редакции

199178, г. Санкт-Петербург, 14-я линия В.О., д. 39, литер А

e-mail: [ia@spcras.ru](mailto:ia@spcras.ru), сайт: <http://ia.spcras.ru>

## Журнал индексируется в международной базе данных Scopus

Журнал входит в «Перечень ведущих рецензируемых научных журналов и изданий, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени доктора и кандидата наук»

Журнал выпускается при научно-методическом руководстве Отделения нанотехнологий и информационных технологий Российской академии наук

© Федеральное государственное бюджетное учреждение науки

«Санкт-Петербургский Федеральный исследовательский центр Российской академии наук», 2022  
Разрешается воспроизведение в прессе, а также сообщение в эфир или по кабелю опубликованных  
в составе печатного периодического издания - журнала «ИНФОРМАТИКА И АВТОМАТИЗАЦИЯ»  
статей по текущим экономическим, политическим, социальным и религиозным вопросам  
с обязательным указанием имени автора статьи и печатного периодического издания  
журнала «ИНФОРМАТИКА И АВТОМАТИЗАЦИЯ»

## CONTENTS

### **Robotics, Automation and Control Systems**

- V. Pshikhopov, M. Medvedev, V. Kostjukov, F. Houssein, A. Kadhim  
TRAJECTORY PLANNING ALGORITHMS IN TWO-DIMENSIONAL  
ENVIRONMENT WITH OBSTACLES 459

### **Artificial Intelligence, Knowledge and Data Engineering**

- A. Korsakov, L. Astapova, A. Bakhshiev  
APPLICATION OF A COMPARTMENTAL SPIKING NEURON MODEL  
WITH STRUCTURAL ADAPTATION FOR SOLVING CLASSIFICATION  
PROBLEMS 493

- D. Galeev, V. Panishchev  
EXPERIMENTAL STUDY OF LANGUAGE MODELS OF "TRANSFORMER"  
IN THE PROBLEM OF FINDING THE ANSWER TO A QUESTION IN A  
RUSSIAN-LANGUAGE TEXT 521

- V. Pimenov, I. Pimenov  
DATA ANALYSIS AND VISUALIZATION IN THE TASKS OF THE  
PROJECT SOLUTIONS MULTICRITERIA OPTIMIZATION 543

- A. Borodinov, V. Manoilov, I. Zarutsky, A. Petrov, V. Kurochkin, A. Saraev  
MACHINE LEARNING IN BASE-CALLING FOR NEXT-GENERATION  
SEQUENCING METHODS 572

### **Mathematical Modeling, Numerical Methods**

- A. Perevaryukha  
DYNAMIC MODEL OF POPULATION INVASION WITH DEPRESSION  
EFFECT 604

- K. Loginov  
NUMERICAL SOLUTION OF THE PROBLEM OF FILTERING ESTIMATES  
INFORMATION IMPACT ON THE ELECTORATE 624

## СОДЕРЖАНИЕ

### **Робототехника, автоматизация и системы управления**

В.Х. Пшихопов, М.Ю. Медведев, В.А. Костюков, Ф. Хуссейн, А. Кадим  
АЛГОРИТМЫ ПЛАНИРОВАНИЯ ТРАЕКТОРИЙ В ДВУМЕРНОЙ СРЕДЕ  
С ПРЕПЯТСТВИЯМИ 459

### **Искусственный интеллект, инженерия данных и знаний**

А.М. Корсаков, Л.А. Астапова, А.В. Бахшиев  
ПРИМЕНЕНИЕ СЕГМЕНТНОЙ СПАЙКОВОЙ МОДЕЛИ НЕЙРОНА СО  
СТРУКТУРНОЙ АДАПТАЦИЕЙ ДЛЯ РЕШЕНИЯ ЗАДАЧ  
КЛАССИФИКАЦИИ 493

Д.Т. Галеев, В.С. Панищев  
ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ ЯЗЫКОВЫХ МОДЕЛЕЙ  
"ТРАНСФОРМЕР" В ЗАДАЧЕ НАХОЖДЕНИЯ ОТВЕТА НА ВОПРОС В  
РУССКОЯЗЫЧНОМ ТЕКСТЕ 521

В.И. Пименов, И.В. Пименов  
АНАЛИЗ И ВИЗУАЛИЗАЦИЯ ДАННЫХ В ЗАДАЧАХ  
МНОГОКРИТЕРИАЛЬНОЙ ОПТИМИЗАЦИИ ПРОЕКТНЫХ РЕШЕНИЙ 543

А.Г. Бородинов, В.В. Манойлов, И.В. Заруцкий, А.И. Петров, В.Е. Курочкин,  
А.С. Сараев  
МАШИННОЕ ОБУЧЕНИЕ В ЗАДАЧАХ BASE-CALLING ДЛЯ МЕТОДОВ  
СЕКВЕНИРОВАНИЯ НОВОГО ПОКОЛЕНИЯ 572

### **Математическое моделирование и прикладная математика**

А.Ю. Переварюха  
ДИНАМИЧЕСКАЯ МОДЕЛЬ ПОПУЛЯЦИОННОЙ ИНВАЗИИ С  
ЭФФЕКТОМ ДЕПРЕССИИ 604

К.О. Логинов  
ЧИСЛЕННОЕ РЕШЕНИЕ ЗАДАЧИ ФИЛЬТРАЦИИ ОЦЕНОК  
ИНФОРМАЦИОННОГО ВОЗДЕЙСТВИЯ НА ЭЛЕКТОРАТ 624

V. PSHIKHOPOV, M. MEDVEDEV, V. KOSTJUKOV, F. HOUSSEIN, A. KADHIM  
**TRAJECTORY PLANNING ALGORITHMS IN TWO-  
DIMENSIONAL ENVIRONMENT WITH OBSTACLES**

*Pshikhopov V., Medvedev M., Kostjukov V., Houssein F., Kadhim A.* **Trajectory Planning Algorithms in Two-Dimensional Environment with Obstacles.**

**Abstract.** This article proposes algorithms for planning and controlling the movement of a mobile robot in a two-dimensional stationary environment with obstacles. The task is to reduce the length of the planned path, take into account the dynamic constraints of the robot and obtain a smooth trajectory. To take into account the dynamic constraints of the mobile robot, virtual obstacles are added to the map to cover the unfeasible sectors of the movement. This way of accounting for dynamic constraints allows the use of map-oriented methods without increasing their complexity. An improved version of the rapidly exploring random tree algorithm (multi-parent nodes RRT – MPN-RRT) is proposed as a global planning algorithm. Several parent nodes decrease the length of the planned path in comprise with the original one-node version of RRT. The shortest path on the constructed graph is found using the ant colony optimization algorithm. It is shown that the use of two-parent nodes can reduce the average path length for an urban environment with a low building density. To solve the problem of slow convergence of algorithms based on random search and path smoothing, the RRT algorithm is supplemented with a local optimization algorithm. The RRT algorithm searches for a global path, which is smoothed and optimized by an iterative local algorithm. The lower-level control algorithms developed in this article automatically decrease the robot's velocity when approaching obstacles or turning. The overall efficiency of the developed algorithms is demonstrated by numerical simulation methods using a large number of experiments.

**Keywords:** mobile robots, motion planning, path planning, motion control, robot motion.

**1. Introduction.** Path planning is a key problem in designing motion planning systems. The problem is to determine the trajectory that connects the initial and final position of the robot and ensures that there are no collisions with obstacles. In this case, the problem can be solved by taking into account the dynamics of the robot, the uncertainty and non-stationarity of the environment, the time for calculating the path, and the physical feasibility of the trajectory. The planning problem is often formulated as the problem of optimizing the state of the robot's current position relative to the target position. Most often, this problem is solved in the configuration space [1 – 3], which consists of a set of obstacles, kinematic and dynamic constraints, and a set of robot points. In this case, a configuration is understood as a set of variables that uniquely determine the robot's position in space. Planning methods are divided into global and local. Global methods build a route based on a known map, while local methods correct the path when unmapped obstacles are detected. Planning can be carried out in discrete and continuous spaces.

Cell decomposition methods [2] are the basis of a significant number of path-planning algorithms, which implies decomposing the original space

into discrete cells, for example, according to the procedure [4]. After decomposition, the search for a path is already carried out in a discrete space that describes the original space with certain accuracy. In general, the division of space into cells that is optimal in terms of accuracy requires different resolutions for different regions of this space. In this regard, the methods of adaptive cell decomposition are widely used today [5]. The main problem of this method is the increase in computational complexity with the growth of requirements for the accuracy of planning the robot's path, which makes this problem remaining relevant [6].

Cell decomposition allows applying discrete search methods based on graph theory. The most popular methods are the A\* and D\* algorithms [7, 8].

A\* algorithm is a development of Dijkstra's algorithm, in which the computational complexity is reduced due to the heuristic function for estimating the path cost. At the same time, the algorithm demands on memory, and its computational complexity increases significantly with the number of cells on the map. In this regard, there are modifications of A\* aimed at eliminating these disadvantages. Such modifications include A\* algorithms with iterative deepening [9], A\* with memory constraints [10], hierarchical A\* [11], and A\* with dynamic change in edge weights [12, 13].

D\* algorithm [8] provides for path planning under conditions when information about the environment can be updated while the robot is moving, for example, when new obstacles are detected.

It should be noted that the considered path planning algorithms give non-smooth paths. In this regard, these methods are supplemented by additional methods that are used to smooth the path [14]. Taking into account the dynamic constraints imposed by the inertia of the robot is also a separate laborious task.

Smooth paths can be obtained using the potential field method [15], which has been studied in detail in the works of various authors [16 - 19]. The potential field method is computationally efficient and allows path planning in uncertain environments. However, taking into account the dynamic properties of the target and obstacles is a separate problem since it is required to form the functions of attraction and repulsion depending on the positions and velocities. This problem is considered in [20].

The main disadvantage that limits the application of the potential field method in practice is the possible presence of local minima. The proposed approaches to solving this drawback solve the problem in particular cases. There is currently no good solution to this problem. In addition, the coefficients of attractive and repulsive forces are usually chosen heuristically, and the potential field method experiences difficulties

when planning in a complex environment with narrow corridors, passages, and non-convex obstacles. Despite the partial solution of problems, for example, using vector histograms [19], it is required to involve other planning algorithms.

One way to solve the problem of local minima is to use random forces in them that push the robot out of the indicated minima. This idea has been developed in sampling methods, which include probabilistic planning methods.

At present, probabilistic path-planning algorithms in discrete space are widely used [21]. Algorithms of this type include Probabilistic Roadmap (PRM) [3, 22, 23] and rapidly-exploring random trees (RRT) [2, 24 - 26]. The PRM algorithm usually does not take into account the dynamics of the mobile robot. To consider it, significant computational costs are required [22, 27]. The RRT algorithm can take into account the dynamics of a mobile robot [24, 25] since it builds a feasible path in free space.

The main limitations of these methods are slow convergence to the optimal solution, high requirements for the memory used [22, 27], and limitations in planning under conditions of incomplete certainty of the environment and robot model [22]. In solving these problems, algorithms based on a probabilistic approach can be effectively used as global path planners.

In this regard, let us consider in more detail the work aimed at improving the efficiency of algorithms based on a probabilistic approach.

**1.1. Related work.** The problem of using the RRT and RPM algorithms under conditions of partial uncertainty is considered in [22], in which it is proposed to divide the entire area of operation into separate regions. Links between regions are searched using the RRT or RPM methods. Planning a local path within a particular region is carried out using a simple straight-line planning algorithm. In this case, the uncertainty of the robot model is taken into account by introducing an additive disturbing input such as white noise. The uncertainty of the map is taken into account by the probability that the cell is occupied by an obstacle. The main result of the method is a path that guarantees the successful achievement of the target point with a probability not less than a given  $p_{min}$ . The work [22] presents the results of analyzing and modeling the motion of the robot that is represented by the kinematic equations of a material point in a static environment. The main disadvantage of the method is a significant increase in computational costs for large values of the probability  $p_{min}$ . The problem of reducing computational costs and memory occupied is considered in [26, 27], which combines the RRT and potential field (P-RRT\*) methods. In this method, a random tree is built in the direction of decreasing the



potential field value, which makes it possible to increase the rate of convergence to the optimal trajectory. In [27], an improved version of the algorithm RRT\* [28] is used, in which the initial path is calculated similarly to the basic RRT algorithm. Then the number of nodes in the configuration space is increased to optimize the initial path. In [28], the analysis of the convergence of the proposed algorithm is carried out and the results of numerical simulation are presented. The disadvantages of the method presented in [27, 28] are the non-smoothness of the trajectories and the still high computational complexity.

Note that the idea of using potential fields to improve the convergence and computational efficiency is not limited only to the RRT algorithms family, but it is also used with the PRM algorithms. Thus, in [23], a hybrid method is used that includes the PRM algorithm and artificial potential fields. The node distribution uses a segmented map to create areas of low and high potentials. To eliminate the problem of local minima, a reactive approach to planning local paths in the presence of static and dynamic obstacles is used. The proposed algorithms have been tested using numerical and field experiments.

In [26], the P-RRT\* method is developed for the case of bidirectional search. In this method, nodes are generated from both the starting and the target points. Both trees create attractive fields for each other, which improves the convergence of the P-RRT\* method, especially for finding narrow passages. The work [29] also investigates the problem of increasing the rate of convergence of the path planning process using the bidirectional search RRT algorithm. The outcomes of [26, 29] indicate that the bidirectional search in a combination with potential fields makes it possible to reduce the number of iterations for finding the shortest path. Also, such algorithms allow moving in narrow corridors without hitting local minima. To reduce the memory used, the calculation of two random trees is carried out sequentially in time, which can increase the calculation time.

The paper [30] also uses a two-step planning procedure for a dynamic environment. The global planner plans the original first approximation path using the RRT method. Next, a dynamic planner is used, which changes the nodes by using repulsive forces from obstacles and inertial forces. To account for moving obstacles, this method uses a forecasting procedure. The effectiveness of this method is confirmed by the results of numerical experiments. In this work, a heuristic optimization function is used, the choice of which depends on various requirements and constraints. The correct formation of such function is the central problem of the presented method.

In [31], the problem of constructing an optimal path with a visit to the set of given target points is considered [32]. To solve this problem in an environment with obstacles, a forest of trees is built [33], where the trees start from the given target points. A forest of trees is built before they are connected to each other. This work is a development of [34], in which the construction of a forest of trees is optimized due to the priorities of the queue of nodes. The proposed method is studied in a two-dimensional and three-dimensional environment using numerical simulation methods. The reduction in computational costs and memory requirements is due to the fact that each tree is built separately from each other. In this regard, the nodes of individual trees are stored in separate arrays.

In [35], a modification of the RRT algorithm is considered, taking into account the limitations imposed by the vision system and the speed limitations. An algorithm based on computer vision that functions under conditions of an unknown image depth is proposed. A model of a robot in the form of kinematics of a material point is considered. The main contribution of the method proposed in [31] is the presence of a virtual target, which is located so that the robot is on a line that provides its specified orientation angle at the target point. The advantages of the proposed algorithm are confirmed by the results of the experiments. This article shows the efficient operation of the RRT algorithm in real conditions. However, the experiments were carried out in a simple environment.

**1.2. Contribution of the article.** The performed analysis allows us to conclude that currently the typical issues of path planning using the RRT algorithm are:

- 1) Consideration of restrictions in terms of the formation of feasible paths that take into account the dynamic properties of robots.
- 2) Optimization of computational complexity and convergence of algorithms to optimal trajectories.
- 3) Consideration of the uncertainty associated with the lack of a complete map of the environment and optimization of the processes of replanning paths in dynamic environments.

The purpose of this article is to develop planning algorithms that reduce the length of the planned trajectory, reduce its calculation time and take into account the restrictions imposed by the robots' properties. We propose a two-stage procedure, during which the MPN-RRT method is used to find the first approximation path which is optimized using the developed iteration algorithm. To take into account the dynamic restrictions on the turns of a mobile robot, it is proposed to add virtual obstacles to the map, which block unfeasible sectors of the movement. This makes it possible to

take into account these limitations without changing the developed MPN-RRT algorithm.

In summary, the contributions of this paper are listed as follows.

1) In this work, virtual obstacles are introduced on the map, located in such a way that they overlap the unfeasible sectors of the mobile robot's movement. This makes it possible to take into account the restrictions on maneuvers that depend on the linear speed of the robot and its maximum permissible angular accelerations and velocity. At the same time, the virtual obstacle method does not change the planning algorithm and allows the use of any map-oriented planning algorithm.

2) This article proposes the use of several parent nodes when building a random tree using the RRT algorithm. On the one hand, this complicates the structure of the constructed graph. On the other hand, it gives more possibilities to optimize the resulting motion trajectory. As will be shown below, the introduction of additional parent nodes makes it possible to reduce the length of the planned path with increasing computational complexity.

3) The proposed method includes global and local planning stages. The use of a two-stage procedure makes it possible to reduce the computational costs for finding an optimal solution compared to a one-stage global search. A proposed MPN-RRT algorithm with two-parent nodes is used as a global planner. An optimization algorithm is used as a local planner, taking into account the path length and deviation from the points planned by the global planner.

4) An algorithm of local optimization and path smoothing described by segments connecting the points of the first approximation path is proposed. A quality criterion is proposed, which is a convolution of two indicators, which are the path length and the deviation from the global path points. This procedure allows optimizing the path's length while limiting the deviation from the original trajectory of the first approximation. This method is based on the outcomes of [36, 37]. Its difference lies in the application of a new optimization functional. The function that was used in these papers included the probability of successful passage of a section of the trajectory and the deviation from the trajectory of the first approximation.

**2. Problem statement.** Let us consider a rectangular stationary two-dimensional environment of size  $L_x \times L_y$ , shown in Figure 1, which simulates a low-density urban development.

The lower-left corner of the map coincides with the origin of the fixed coordinate system  $Ox_g Y_g$ . The length of the environment along the  $Ox_g$  axis is  $L_x$ , and along the  $OY_g$  axis is  $L_y$ . The position of the robot in the

environment is determined by the coordinate vector  $(x, y)$  and the orientation angle  $\varphi$ . The obstacles are the rectangles which are  $r_{pi}$  wide and  $b_{pi}$  high. The position of an obstacle is determined by the coordinates of its lower-left corner  $(x_{pi}, y_{pi})$ . The target position is specified by a vector of coordinates  $(x_t, y_t)$ .

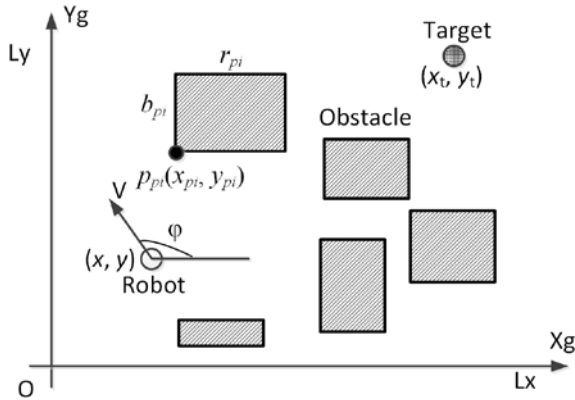


Fig. 1. 2-D obstructed environment

The mobile robot is described by the following equations [16, 38]:

$$\dot{x} = V \cos \varphi, \quad \dot{y} = V \sin \varphi, \quad \dot{\varphi} = \omega. \quad (1)$$

$$\dot{V} = -\alpha_v(V - V^*), \quad \dot{\omega} = -\alpha_\omega(\omega - \omega^*), \quad (2)$$

where  $x, y$  are the robot's coordinates in the fixed frame  $OX_gY_g$ ;  $V$  is the robot's linear velocity;  $\omega$  is the robot's angular velocity;  $\varphi$  is the robot's yaw angle;  $\alpha_v, \alpha_\omega$  are constants;  $V^*, \varphi^*$  are references for  $V$  and  $\varphi$ , respectively.

The mobile robot's design features, inertial properties, and requirements for stability and controllability do not allow it to move to an arbitrary point in space [39]. In this regard, the following restriction is introduced on the maximum turning angle depending on:

$$|\varphi^* - \varphi|_{max} = \Delta\varphi_{min} + (\pi - \Delta\varphi_{min})e^{-\frac{V - V_{min}}{\alpha(V_{max} - V_{min})}}, \quad (3)$$

where  $\Delta\varphi_{min}$  is the minimum rotation angle, independent of the robot's velocity, determined by the rolling resistance;  $V_{min}$ ,  $V_{max}$  are the minimum and maximum values of the robot's velocity, respectively;  $\alpha$  is the coefficient depending on critical velocity, robot length, friction, and rolling resistance. The function (3) is shown in Figure 2.

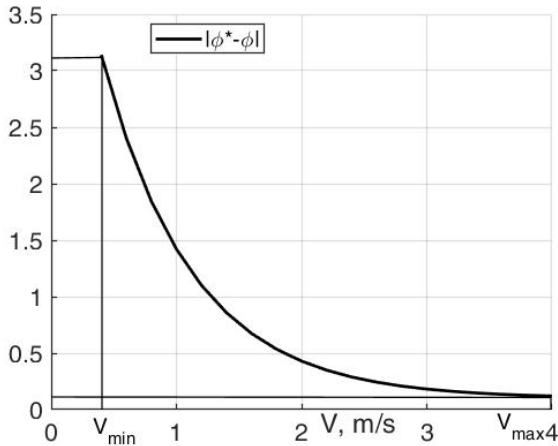


Fig. 2. Restriction on the maximum angle of rotation of the robot

The reference influence on the velocity  $V^*$  changes in accordance with the expression:

$$\begin{aligned}
 V^* &= V_{max} \frac{r}{r_s} k_\varphi k_{obs}, \quad \forall V^* \leq V_{max}, \\
 V^* &= V_{max}, \quad \forall V^* > V_{max},
 \end{aligned}
 \tag{4}$$

where  $r$  is the straight distance to the target;  $r_s$  is the distance that determines the moment when the robot starts braking;  $k_\varphi$  is the function of velocity reduction when turning around;  $k_{obs}$  is the function of decreasing velocity when approaching obstacles.

The function  $k_\varphi$  is shown in Figure 3, according to which the robot's velocity decreases linearly with an increase in the difference between the current  $\varphi$  and the desired  $\varphi^*$  angle of the movement of the robot.

The  $k_{obs}$  function is shown in Figure 4, according to which, when the robot approaches an obstacle at distance smaller than  $r_s$ , its velocity begins to decrease in proportion to the specified distance:

$$\begin{aligned}
 k_{obs} &= \frac{r_{obs}}{r_s}, \quad \forall r_{obs} < r_s, \\
 k_{obs} &= 1, \quad \forall r_{obs} \geq r_s.
 \end{aligned}
 \tag{5}$$

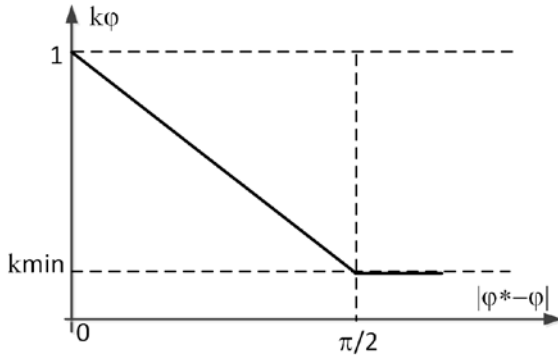


Fig. 3. Graph of the function that determines the decrease in the velocity of the robot when turning

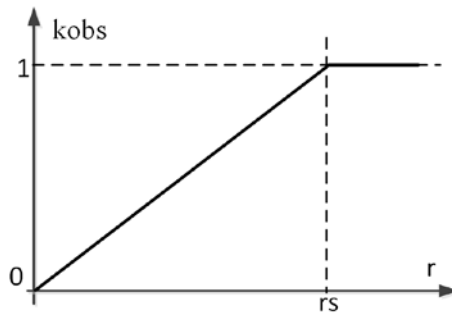


Fig. 4. Graph of the function that determines the decrease in the speed of the robot when approaching obstacles

It is required to develop an algorithm for planning the movement of a mobile robot from the current point  $(x, y)$  to a given target point  $(x_t, y_t)$ , taking into account the restrictions on the rotation angle (3) and ensuring the absence of collisions. There are no requirements for the orientation of the mobile robot at the target point. The velocity of the robot at the target point should be zero, but at intermediate points of the trajectory, the velocity is determined by the expression (4). The planned trajectory must be smooth,

i.e. have no breaks. The quality criterion of the planned trajectory is its length, i.e.:

$$D_{path} = \sum_{i=1}^N \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2} \rightarrow \min.$$

**3. Planning system algorithms.** This section considers the planning procedure, which includes the following:

- 1) Introduction of a virtual obstacle into the map to take into account the constraints (3).
- 2) Construction of a random tree for planning the first approximation path using the modified RRT method.
- 3) Finding the shortest first approximation path using the ant colony algorithm.
- 4) Local optimization and smoothing of the first approximation path.

**3.1. Adding virtual obstacles to the map.** As follows from the expression (3), if the velocity is  $V = V_{min}$ , then the robot can turn at any angle from  $-\pi$  to  $+\pi$ . As the robot's velocity increases, the sector available for its maneuver begins to decrease in accordance with (3). To consider these limits, a virtual obstacle is added to the environment map as shown in Figure 5.

Figure 5 shows the evolution of a virtual obstacle in the case when  $V_3 > V_2 > V_1$ . The virtual obstacle is V-shaped. At low speeds, this obstacle is located at a convex angle to the robot. As the velocity increases, this angle decreases to unfold and becomes sharp. The presence of a virtual obstacle makes it possible to increase the safety of the robot's movement without using additional algorithms to avoid collisions. As shown in Figure 6, if the robot moves in a narrow passage between two obstacles and receives a target point located behind it, then a collision may occur if the robot's velocity is restricted by inequalities  $0 < V_{min} \leq V \leq V_{max}$ .

The advantage of virtual obstacles for eliminating such collisions is that they do not increase the computational complexity of the used planning algorithm. The procedure for introducing a virtual obstacle is based on the expression (3), which does not present additional computational complexity. At the same time, it should be noted that the parameters  $\Delta\varphi_{min}$  and  $\alpha$ , which determine the turning radius in the expression (3), depend on the type of the robot and the conditions of its movement. In this regard, these parameters must be set, and the robot must evaluate the movement resistance coefficients.

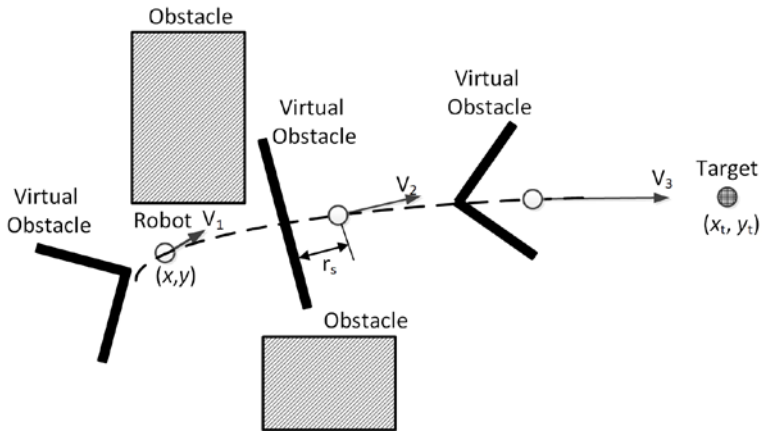


Fig. 5. Graph of the function that determines the decrease in the robot's velocity when approaching obstacles

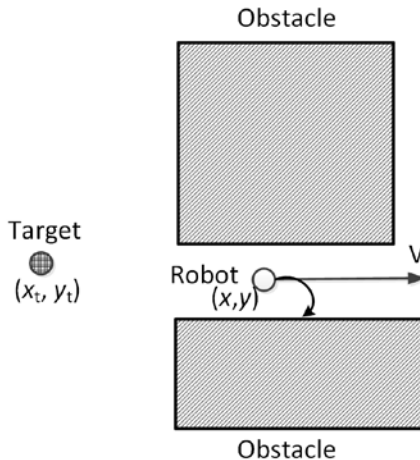


Fig. 6. An example of a collision due to the inertial properties of a mobile robot

**3.2. Two-parent nodes RRT algorithm.** The pseudocode of Algorithm 1 for constructing a graph using the RRT method is distinguished by the search for two-parent nodes for each new node.



**Algorithm 1** for constructing a graph for planning the trajectory of a mobile robot:

**Input:** Robot's position  $(x, y)$ ; environment map Map; target point  $(x_b, y_t)$ ; robot's velocity  $V$ ; yaw angle  $\varphi$ ; parameter  $L_s$ ; trajectory optimization parameter  $N_{add}$ .

**Output:** List of nodes Tree; Relationship matrix Matrix\_tr.

```

1: [Tree, Matrix_tr] ← Init_Tree([x,y], [xt,yt]);
2: Map ← Save_Vir_Obs(V, φ);
3: FL_0 ← Str_Line(Tree(last), [xt, yt], Map, Ls);
4: IF FL_0=1
5:   WHILE FL_1 < Nadd
6:     X_rand ← Rand_Node(Map, Tree, Ls)
7:     X_nod1 ← Parent_Node(Tree, Map, rs, []);
8:     X_nod2 ← Parent_Node(Tree, Map, rs, X_nod1);
9:     Tree ← Expand_Tree(X_rand);
10:    Matrix_tr ← Expand_Matr(X_nod1, X_rand);
11:    Matrix_tr ← Expand_Matr(X_nod2, X_rand);
12:    FL_2 ← Str_Line(X_rand, [xt,yt], Map, Ls);
13:    IF FL_2=1
14:      FL_1 = FL_1+1;
15:      Matrix_tr ← Expand_Matr(X_rand, [ xt,yt]);
16:    END IF
17:  END WHILE
18: ELSE
19:  Matrix_tr ← Expand_Matr([x,y],[xt,yt]);
20: END IF
21: Path_opt = Aco(Matrix_tr, Tree);

```

The Init\_Tree () function initializes the tree and the matrix of paths' weights Matrix\_tr. During the initialization phase, the starting point of the path  $(x, y)$  and the target point  $(x_b, y_t)$  are added to the list of nodes. The first tree node in the list is always the starting position of the robot  $(x, y)$ . The last node in the tree list is always the target point  $(x_b, y_t)$ . The path matrix Matrix\_tr is initialized as a zero  $2 \times 2$  matrix.

The Save\_Vir\_Obs function adds a virtual obstacle to the map based on the current velocity and the robot's movement direction. In this case, the environment map is a matrix of dimension  $N_x \times N_y$ , where the number of matrix cells is determined by the expressions:

$$N_x = [Lx/Sx], \quad N_y = [Ly/Sy], \quad (6)$$

where  $S_x, S_y$  are the dimensions of the space cell along the axes  $OX_g$  and  $OY_g$ , respectively;  $\lceil * \rceil$  is the operation of rounding to the nearest integer from the top.

The value of a cell of the Map matrix equals 0 if it is free, 1 if it contains the current position of the robot, 2 if it contains the target point, and 3 if it contains an obstacle. Obstacles are recorded in the Map in accordance with Figure 7. According to this figure, if the cell with coordinates  $(x_c, y_c)$  is occupied by an obstacle, then the number 3 is written into the cell of the matrix Map  $(y_c + 1, x_c + 1)$ . Thus, the  $x_c$  coordinates are the column numbers of the Map matrix, and the  $y_c$  coordinates are the row numbers. Adding 1 to the coordinates is due to the fact that the rows and columns are numbered starting from one, and the coordinates of the cells are numbered from zero.

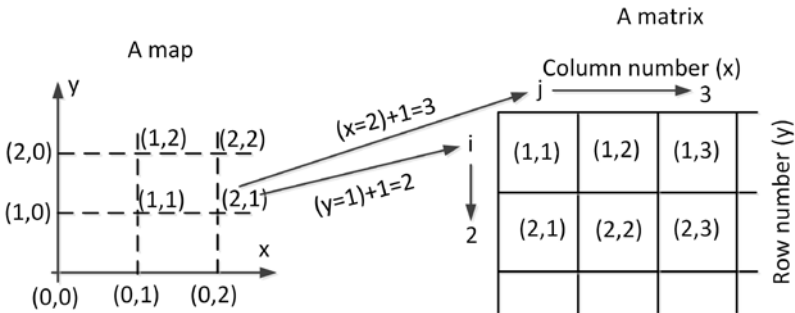


Fig. 7. Recording the environment map to the Map matrix

Line 3 of Algorithm 1 checks for the presence of a straight line connecting points  $(x, y)$  and  $(x_b, y_b)$  without collisions. Such a check is carried out using the `Str_Line ()` function. It returns the `FL_0` flag, which is equal to 1 in the presence of a collision, which is determined when the robot's trajectory approaches any obstacle closer than the distance  $L_s$ .

The main tree building loop begins with the generation of a random node `X_rand` for which the `Rand_Node ()` function is used. This function generates a node so that it is located at a distance of at least  $L_s$  from the obstacles and the other nodes of the tree. Next, the search for parent nodes is carried out using the `Parent_Node ()` function. The presence of two-parent nodes is a key difference between the used algorithm and the traditional version of RRT. The main planned effect of this introduction of two-parent nodes is a decrease in the length of the planned path, which occurs at cost of increasing the tree construction time. This effect is due to the fact that the presence of two-parent nodes increases the number of links in the graph

under construction, which makes it possible to optimize the path length more effectively.

If the parent nodes are found successfully, then a new node  $X_{rand}$  is added to the tree and the corresponding links are added to the Matrix\_tr matrix using the functions Expand\_Tree () and Expand\_Matr ().

Next, using the function, Str\_Line () is checked for a direct collision-free path from the new  $X_{rand}$  node to the target point  $(x_b, y_t)$ . If such a path exists, then the FL\_1 counter is incremented.

The search cycle continues until it finds  $N_{add}$  nodes that are added to the tree and gives a new path to the target point.

The search for a path on the constructed graph is carried out using the ant colony optimization algorithm [40]. The choice of the evolutionary algorithm is due to the need of optimizing the search time. The search for the shortest path (line 21) using the ant colony optimization algorithm is implemented by Aco () function that takes Tree and the matrix of coupling Matrix\_tr as an input. The result of the operation of the Aco () function is a list of points of the found path Path\_opt. This list defines a piecewise broken path that needs to be smoothed. It should also be noted that the RRT algorithm converges rather slowly, so it is necessary to further optimize its length.

**3.3. Path optimization and smoothing.** Path\_opt is the first approximation path. The task is to reduce the length of the curve Path\_opt, assuming that the optimized curve is found in a certain bounded neighborhood, or in a certain corridor relative to the original trajectory. To solve this problem, the following functional can be generated:

$$G(Tr) = \delta_1 \sum_{i=2}^{N-1} [(x_i - x'_i)^2 + (y_i - y'_i)^2] + \delta_2 \sum_{i=2}^{N-1} [(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2], \quad (7)$$

where the first term is the sum of the squares of the deviations of the points  $(x_i, y_i)$  of the constructed trajectory from the points of the original trajectory  $(x'_i, y'_i)$ ; the second term is the trajectory length  $T$ ; the coefficients  $0 \leq \delta_1, \delta_2 \leq 1$  are the weights of each term.

The stationary point of the functional (6) with respect to the variables  $\{x_i, y_i\}, i = 2, 3, \dots, N-1$  is found from the system of equations:

$$\begin{cases} \partial G(Tr)/\partial x_i = 0, \\ \frac{\partial G(Tr)}{\partial} y = 0, \quad i = 2, \dots, N-2 \end{cases} \quad (8)$$

System (8) is reduced to two independent systems of linear equations with respect to vectors  $X = [x_2 \dots x_{N-1}]^T$  and  $Y = [y_2 \dots y_{N-1}]^T$ . The solution of these two systems has the form:

$$\begin{aligned} X_{zt} &= [x_{st2} \dots x_{stN-1}]^T = A^{-1}B_X, \\ Y_{zt} &= [y_{st2} \dots y_{stN-1}]^T = A^{-1}B_Y, \end{aligned} \quad (9)$$

$$B_X = [\delta_1 x'_2 + \delta_2 x'_1, \delta_1 x'_3, \dots, \delta_1 x'_{N-2}, \delta_1 x'_{N-1} + \delta_2 x'_N]^T, \quad (10)$$

$$B_Y = [\delta_1 y'_2 + \delta_2 y'_1, \delta_1 y'_3, \dots, \delta_1 y'_{N-2}, \delta_1 y'_{N-1} + \delta_2 y'_N]^T, \quad (11)$$

$$A = \begin{bmatrix} \delta_1 + 2\delta_2 & -\delta_2 & 0 & 0 & \dots & 0 & 0 & 0 \\ -\delta_2 & \delta_1 + 2\delta_2 & -\delta_2 & 0 & \dots & . & 0 & 0 \\ 0 & -\delta_2 & \delta_1 + 2\delta_2 & -\delta_2 & \dots & 0 & . & 0 \\ 0 & 0 & -\delta_2 & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & . & . & 0 & 0 \\ 0 & . & 0 & \dots & . & \delta_1 + 2\delta_2 & -\delta_2 & 0 \\ 0 & 0 & . & \dots & 0 & -\delta_2 & \delta_1 + 2\delta_2 & -\delta_2 \\ 0 & 0 & 0 & \dots & 0 & 0 & -\delta_2 & \delta_1 + 2\delta_2 \end{bmatrix}. \quad (12)$$

Let us show that at the found stationary point (9) the sufficient condition for a local maximum with respect to minors of the Hessian matrix is satisfied for any values of  $\delta_1, \delta_2$  such that,  $0 \leq \delta_1, \delta_2 \leq 1, \delta_G + \delta_L = 1$ . Indeed, it is easy to observe that the Hessian matrix for functional (6) is the block-diagonal matrix:

$$Gess(G) = \begin{bmatrix} A & 0_{N-2 \times N-2} \\ 0_{N-2 \times N-2} & A \end{bmatrix}, \quad (13)$$

where  $0_{N-2 \times N-2}$  is a zero matrix of dimension  $N-2 \times N-2$ .

It is known that if each of the blocks of such a matrix is a positive definite matrix, then the resulting block diagonal matrix will also be positive definite [41]. Also, according to the corollary to the Gershgorin

theorem [41], a sufficient condition for any symmetric matrix with positive entries on the main diagonal to be positive definite is the condition of strict dominance of diagonal entries. In this case of symmetric tridiagonal Toeplitz matrix of type (9), this condition takes the form:

$$|\delta_1 + 2\delta_2| > 2|\delta_2|, \quad (14)$$

Since  $0 \leq \delta_1, \delta_2 \leq 1$ , then the expression (14) takes the form:

$$\delta_1 + 2\delta_2 > 2\delta_2 \Rightarrow \delta_1 > 0. \quad (15)$$

Thus, an additional constraint is a requirement that the coefficient  $\delta_1$  does not vanish.

Note that the global minimum of functional (6) is simultaneously attained at point (8) due to the uniqueness of the stationary point for this functional.

For practically important cases of minimization of the resulting curve, two conditions can be used. The first condition is the natural requirement of not colliding with obstacles. The second is not exceeding a certain maximum deviation from the original curve, expressed, for example, in the form of the limiting root-mean-square discrepancy  $\sigma_{kr}$  between the corresponding nodal points. This condition is written as:

$$\sqrt{\sum_{i=2}^{N-1} [(x_i - x'_i)^2 + (y_i - y'_i)^2] / (N - 2)} \leq \sigma_{kr}. \quad (16)$$

The following procedure is used to find the offset path of the minimum length, taking into account the first and second constraints. First, the values of the coefficient  $\delta_1$  are sorted out in the range  $[\delta_{1\min}, \delta_{1\max}]$  with a step  $\Delta\delta_1 \ll \delta_{1\max}$ . The value  $\delta_1 = \delta_{1\min}$  corresponds to the case when the optimized trajectory is the closest in terms of the standard deviation to the straight line connecting the starting and target points. If there are no collisions and the deviation from the initial trajectory does not exceed a given one, then the optimal trajectory for  $\delta_1 = \delta_{1\min}$  is the best. If at least one of these conditions is not met, it is necessary to move to the next larger value of the coefficient  $\delta_1$  and re-check the specified conditions.

This procedure is repeated iteratively until there is such a value  $\delta_1$  from the corresponding finite sampling sequence of the interval  $[\delta_{1\min}, \delta_{1\max}]$ , for which both of the above constraints are satisfied, or until the search is completed. To exclude the last case, it is necessary to choose the value of the step  $\Delta\delta_1$  in accordance with the value of the upper admissible boundary  $\sigma_{kr}$  of the limiting root-mean-square discrepancy between the corresponding nodal points, and the less is the  $\Delta\delta_1$  the less should be the step.

After finding the optimized curve, it has to be smoothed. For this, the method of approximation by straight-line segments, smoothly conjugate to arcs of the circles, is applied. Moreover, these arcs are constructed for each of the nodal points of the resulting curve.

The resulting trajectory in the form of a list of points that the robot must pass is fed to the motion control function, which calculates the desired direction of motion and setpoints for the linear and angular velocities of the robot. The motion control system is implemented in the form of an adaptive multi-loop system presented in [38].

**4. Numerical modeling results.** In this section, a case study of the developed algorithms is carried out. The simulation was carried out in MATLAB. During the simulation, the following parameters of the mobile robot were set:  $a_v = 1$ ;  $a_w = 5$ ;  $V_{max} = 4$  m/s;  $w_{max} = 2$  rad/s. The safe distance at which the robot starts braking is  $r_s = 6$ . The minimum distance to the obstacles is  $L_s = 1$  m. The dimensions of the environment are  $L_x = 100$  m,  $L_y = 100$  m. The maximum number of obstacles is 24. The maximum width and length of each obstacle is 11 m. An urban environment is simulated with a maximum building rate of 30 %.

**4.1. Investigation of the influence of the number of parent nodes on the trajectory length and the time of the trajectory calculation.** The study of the influence of the number of nodes on the trajectory length and the time for calculating the trajectory was carried out for various values of the parameter  $N_{add} = 20, 40, 80$ . For each number of nodes and the parameter  $N_{add}$ , 1000 tests were carried out, the results of which are summarized in Table 1 and presented in Figure 8 and Figure 9. The confidence probability is accepted as 95%. The confidence interval for the path length is 0.026, and the confidence interval for the path calculation time is 0.011.

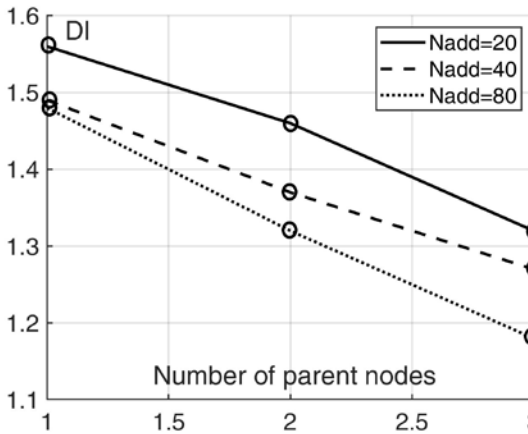


Fig. 8. Dependence of the average relative path length on the number of parent nodes

The  $N_{add}$  parameter determines the degree of optimality of the first approximation path. The larger this parameter, the higher the probability of the first approximation path approaching the optimal trajectory.

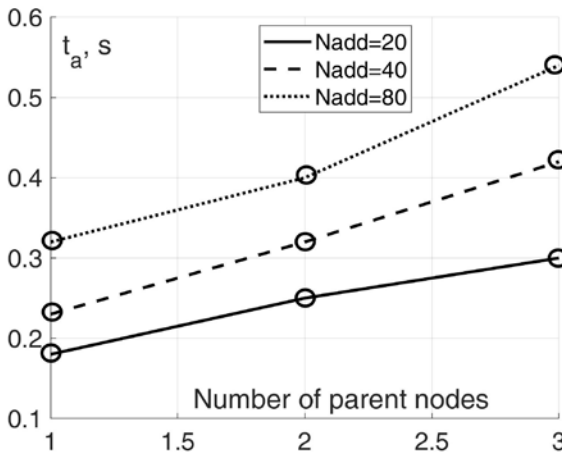


Fig. 9. Dependence of the average time for calculating the path on the number of parent nodes

Table 1. Influence of the number of parent nodes on the length of the path and the time of calculation of the trajectory

Number of nodes	Parameter $N_{add}$	Average relative path length	Average calculation time
1 node	80	1.48	0.32
	40	1.49	0.23
	20	1.56	0.18
2 nodes	80	1.32	0.4
	40	1.37	0.32
	20	1.46	0.25
3 nodes	80	1.18	0.54
	40	1.27	0.42
	20	1.32	0.3

The average relative path length is calculated according to the expression:

$$Dl = \frac{1}{1000} \sum_{i=1}^{1000} (Dl_i / r_{rti}), \quad (17)$$

where  $Dl_i$  is the length of the  $l$ -th path in the  $i$ -th experiment;  $r_{rti}$  is the distance from the starting point of the robot's position to the target point in the  $i$ -th experiment.

The calculation time was measured using the tic and toc MATLAB functions. The table shows the average calculation time calculated in accordance with the expression:

$$t_a = \frac{1}{1000} \sum_{i=1}^{1000} t_i, \quad (18)$$

where  $t_i$  is the time of calculating the path in the  $i$ -th experiment.

As can be seen from Figure 8 and Table 1, the path length decreases with an increase in the number of parent nodes in the RRT algorithm. This is explained by the fact that with an increase in the number of parent nodes, the number of links in the graph under construction increases. This entails increasing the different path options from point  $(x, y)$  to the target point  $(x_b, y_t)$ .

It follows from the simulation results that the introduction of an additional parent node results in 6.4 to 10.8% reduction of the planned path length, depending on the  $N_{add}$  parameter. The introduction of two additional



nodes makes it possible to reduce the average path length from 14.8 to 20.3%, depending on the  $N_{add}$  parameter.

Reducing the length of the planned path increases the computational cost of the algorithm. Thus, Figure 9 shows the dependence of the average time for calculating the trajectory from a different number of nodes and different values of the parameter  $N_{add}$ . From the graphs shown in Figure 9 and Table 1, it can be seen that the increase in calculation time is from 28 to 40% for one additional parent node and from 40 to 45% for two additional parent nodes.

At the same time, we note that an increase in calculation time is associated with both the procedure for constructing a random graph and further searching for a path. When using one parent node, the path from a potential new node only to the nearest graph node is checked. When using additional parent nodes, the path to several nearest nodes of the graph is checked, which leads to additional computational costs.

At the same time, it should be noted that using the ant colony algorithm for finding the shortest path allows parallelizing the computational process [42], which reduces computational costs.

**4.2. Study of the optimization and smoothing algorithm.** The used method of local optimization and trajectory smoothing is based on the optimization of functional (6). In this case, the main issue in the formation of this functional is the rational choice of the admissible average deviation  $\sigma_{kr}$  of the optimal trajectory from the first approximation path. This conclusion follows from the fact that when the condition  $\delta_1 + \delta_2 = 1$  is satisfied, there is only one independent weighting coefficient. The coefficient  $\delta_1$  was chosen as such a coefficient, which is sought using the procedure described in Section 3. Thus, the weighting coefficients of functional (6) are chosen automatically; therefore, it is necessary to choose only the parameter  $\sigma_{kr}$ . The influence of this parameter on the trajectory length is shown in Figure 10 and Figure 11.

Figure 10 shows the result of modeling the procedure for finding a displaced trajectory of the minimized length, taking into account the limiting conditions for the following input data:  $\sigma_{kr} = 10$  m, the length of the original curve is 148.5 m. As a result of applying the procedure of local optimization and smoothing described in Section 3, a trajectory with a length of 71.63 m is obtained. The parameter change interval is  $\delta_1 [0.02 \ 0.98]$ , and step is  $\Delta\delta_1 = 0,02$ . The optimal value of the parameter is  $\delta_1 = 0.1$ . In this case, the standard deviation at the nodes of the original and displaced curves is 9.85 m, which is less than the maximum allowable value. Similar simulation results for  $\sigma_{kr} = 5$  m are shown in Figure 11, taking into account the limiting conditions and for the same initial trajectory

of the first approximation as in the previous case. The interval for changing the parameter  $\delta_1$  and the step are the same. As a result, the length of the optimized curve is 90.8 m with the optimal value  $\delta_1 = 0.6$ .

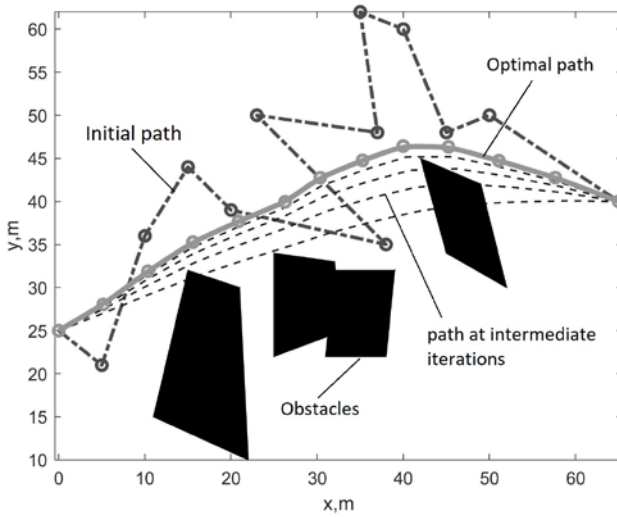


Fig. 10. First approximation path and optimal trajectory for  $\sigma\kappa=10$  m

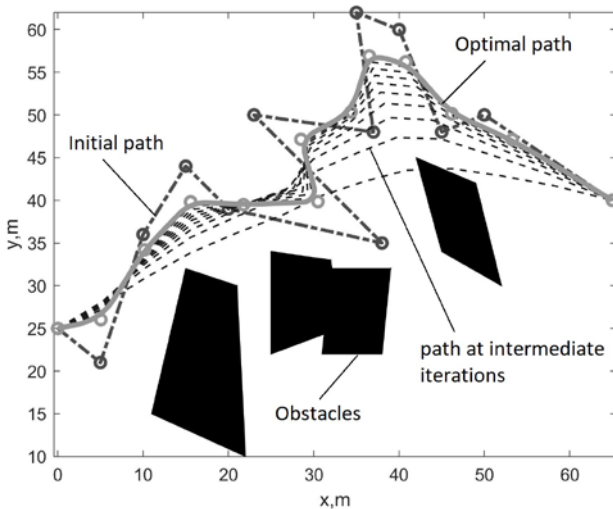


Fig. 11. First approximation path and optimal trajectory for  $\sigma\kappa=5$  m

Thus, in the first case with  $\sigma_{kr} = 10$  m, the length of the first approximation path is reduced by approximately 51.8%. In the second case, with  $\sigma_{kr} = 5$  m, the length of the first approximation path is reduced by about 38.8%.

Figure 10 and Figure 11 show the first approximation path obtained using the RRT algorithm and the ant colony algorithm, the trajectories of intermediate iterations, and the optimal trajectory of the movement. The nodes of the optimized trajectory according to the considered algorithm are shown separately.

Based on the simulation, it can be concluded that with the tightening of restrictions on deviations from an initial curve, i.e. as the value of  $\sigma_{kr}$  decreases, the length of the optimized curve increases. Thus, solving the problem of optimizing the original trajectory, taking into account the restrictions on avoiding collisions and not exceeding the maximum permissible standard deviation from the original curve, requires prioritization when answering the question of whether it is more important to significantly reduce the length of the curve, or to prevent deviation from the original curve by more than a given value  $\sigma_{kr}$ .

#### 4.3. Study of a trajectory planning and motion control system.

The results of the simulation of the mobile robot control system are shown in Figure 12 - 14. Figure 12 shows the obstacles, the virtual obstacle at the initial moment of time, the initial position of the robot with the symbol "\*", the target position of the robot with the symbol "+", the path obtained using the RRT algorithm with two-parent nodes, and the smoothed path obtained as a result of smoothing the initial trajectories. To plan the path in Figure 12, the RRT algorithm generated a graph that includes 341 nodes. As a result of the search for the shortest path using the ant colony algorithm, a path was found containing the following 8 nodes, including the starting and target nodes:

$$Tr = \begin{bmatrix} 18 & 26 & 33 & 41 & 50 & 50 & 48 & 93 \\ 84 & 90 & 96 & 91 & 98 & 94 & 93 & 52 \end{bmatrix}. \quad (19)$$

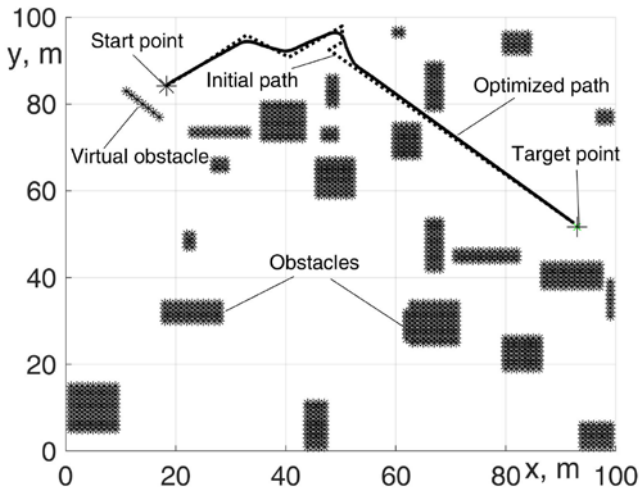


Fig. 12. The trajectory of a mobile robot in an environment with obstacles

The path length obtained as a result of the application of the RRT algorithm is approximately 107 m. The length of the optimized trajectory is 99.9 m. Thus, the reduction in the path is about 6.6%. The results of simulation on a sample of 1000 random situations showed that the smoothing and optimization procedure gives an effect of approximately 6-8% on the path's length. In this case, the greater the number of intermediate points between the starting and target points, the more decrease in the length of the path. For the first approximation paths consisting of two and three points, the use of the local optimization procedure does not make sense.

Figure 13 shows the change in the velocity of the mobile robot in time and the reference velocity. In the time interval from 0 to 5 s, the robot accelerates to  $V_{max}$ . Then, in the interval from 5 to 13 seconds, the motion control system reduces the velocity four times. The first three velocity maneuvers are due to turns, and the last speed reduction is due to passing close to an obstacle. When approaching the target point, the robot also slows down.

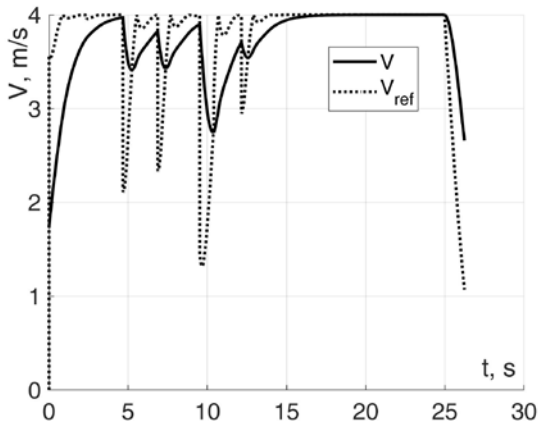


Fig. 13. The velocity of a mobile robot in an environment with obstacles

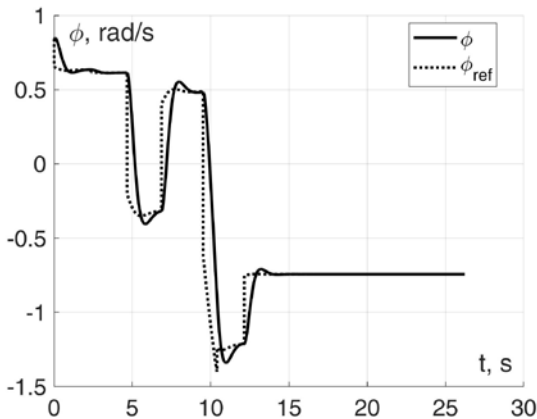


Fig. 14. Yaw angle of a mobile robot in an environment with obstacles

Figure 14 shows the change in the yaw angle of the robot when moving to the target point. It can be seen from this graph that the robot makes turns at the initial moment, and then three turns when moving between the nodes of the planned path.

Figure 15 shows the simulation results demonstrating the movement of the robot in an environment with obstacles, including the movement in a narrow corridor formed by the obstacle and the boundary of the area.

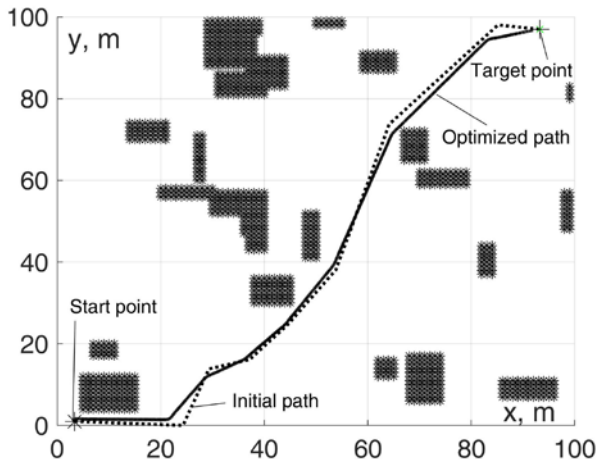


Fig. 15. The robot's trajectory in the presence of a narrow corridor

Studies have shown that the presence of narrow corridors can significantly increase the planning time of the robot's trajectory. In this regard, the RRT algorithm uses a variant of generating a new potential node in a certain limited area relative to the last node added to the graph. This way of building a tree gives an additional effect in terms of reducing the length of the planned trajectory. So, for the MPN-RRT algorithm with two-parent nodes, an additional 1000 experiments were carried out with the generation of a new potential node within a radius of 20 meters from the last node added to the tree. An additional reduction in the trajectory length is obtained, ranging from 5 to 10%, depending on the value of the parameter  $N_{add}$ . However, in the MPN-RRT algorithm, there is a significant increase in the trajectory search time, which is on average 30% compared to the variant of random generation of a new node over the entire area of operation. This increase in the computation time is observed only on average. For situations with narrow corridors, the algorithm with a local area for generating a new potential node does not lead to an increase in the computation time. In this regard, it is possible to recommend the use of a unit for evaluating the functioning environment in the robot control system. Such a block can be built using intelligent technologies. Thus, [43] presents a neural network that estimates the geometric complexity of the environment. This estimate is used to determine the clearest areas for planning a route for a group of robots.

Note that generating a new potential node in a certain bounded area relative to the constructed tree leads to trajectories consisting of a larger number of points. However, the application of the above-described iterative local optimization and smoothing procedure allows obtaining a smooth trajectory with a minimized length. In this case, the effect of reducing the path length becomes higher and amounts to approximately 10% for the simulated movement conditions. A more significant decrease in the trajectory length is due to an increase in the optimization efficiency with an increase in the number of nodal points of the first approximation trajectory. Figure 16 shows the results of one of the numerical experiments in which a new potential node is generated within a radius of 20 meters relative to the last added node.

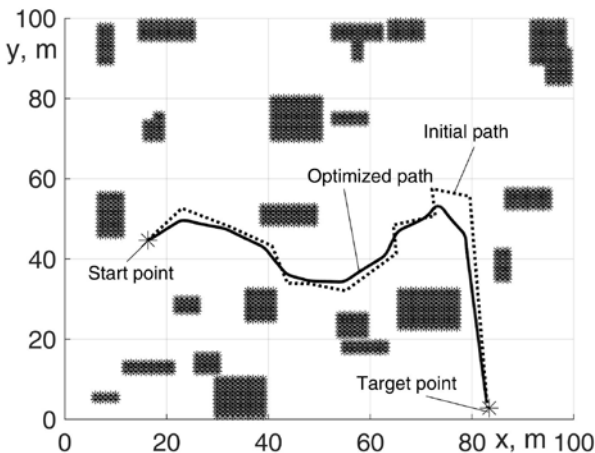


Fig. 16. The robot's trajectory when a new potential node is generated within a radius of 20 meters relative to the last added node

**7. Discussion and conclusion.** Some discussions about the highlights of the approach proposed in this article are summarized as follows.

1) Taking into account the dynamic and design constraints of mobile robots when planning a path usually leads to the complication of planning algorithms. In this article, it is proposed to take into account the restrictions on the angle of rotation of the robot by introducing virtual obstacles. Such obstacles are located in such a way that they cover sectors into which the mobile robot cannot turn at a given time. As a result, the planned trajectory is physically realizable. Thus, compared to the existing

methods of considering dynamic constraints, the proposed approach does not affect the planning method if it uses a map to calculate the trajectory.

2) The potential of using probabilistic methods for path planning is limited by a large number of iterations to obtain the optimal solution. This article proposes an improved MPN-RRT algorithm that uses multiple parent nodes. This algorithm allows creating a graph with a large number of connections, resulting in an increase in the optimality of the planned path.

3) To reduce the number of iterations in the path search, a two-stage path planning method is proposed. At the first stage, the first approximation path is searched using the proposed MPN-RRT algorithm. At the second stage, the resulting trajectory is optimized and smoothed by an iterative local search algorithm. Compared to the one-stage MPN-RRT algorithm, the two-stage method, which includes MPN-RRT and an iterative optimization and smoothing procedure, reduces the time it takes to find a trajectory of the same length by 7–13%. This effect is due to the slow convergence of algorithms based on the probabilistic method.

4) The numerical experiments performed have shown that for a simulated urban environment, the MPN-RRT algorithm with two-parent nodes finds a trajectory of the same length as the original RRT algorithm approximately 30–40% faster. Adding a third parent node reduces the path search time by roughly 25%. With a fixed trajectory calculation time, the introduction of an additional parent node reduces the trajectory length by 6–11%. The introduction of two additional parent nodes reduces the length of the trajectory by 15–20% compared to the original RRT algorithm.

As a further development of the work, it is planned to study the proposed algorithms in a dynamic environment. It is planned to re-plan the rest of the path preservation of those nodes that the robot has not yet passed. Another important issue not covered in this article is the consideration of the uncertainties of the mobile robot model and their influence on the procedure of introducing virtual obstacles. In addition, it can be noted that it is possible to use a large number of the local optimization algorithms at the second stage of planning. Studying their effectiveness and finding the optimal moment for switching from global planning to local optimization and smoothing is also the subject of a separate study.

## References

1. K.A. Kazakov, V.A. Semenov, "Obzor sovremennykh metodov planirovaniya puti," TRUDY ISP RAN, T. 28(4), 2016, s. 241-294.
2. Sánchez-Ibáñez, J.R.; Pérez-del-Pulgar, C.J.; García-Cerezo, A. Path Planning for Autonomous Mobile Robots: A Review. *Sensors* 2021, 21, 7898. <https://doi.org/10.3390/s21237898>



3. L.E. Kaviraki; P. Svestka; J.-C. Latombe; M.H. Overmars, "Probabilistic roadmaps for path planning in high-dimensional configuration spaces," *IEEE Transactions on Robotics and Automation*, v. 12 (4), pp. 566-580, 1996.
4. А.Г. Сухарев, "Оптимальные стратегии поиска экстремума," СССР. Вычислительная математика и математическая Computational физика, Т. 11(4), с. 910-924, 1971.
5. F. Samaniego, J. Sanchis, S. García-Nieto, R. Simarro. Recursive Rewarding Modified Adaptive Cell Decomposition (RR-MACD): A Dynamic Path Planning Algorithm for UAVs. *Electronics*, v. 8 (3), 306, 2019.
6. R. Toodesh, and S. Verhagen. "Adaptive, variable resolution grids for bathymetric applications using a quadtree approach," *Journal of Applied Geodesy*, v. 12(4), 2018, pp. 311-322.
7. P.E. Hart, N.J. Nilsson, B.A Raphael, "Formal Basis for the Heuristic Determination of Minimum Cost Paths," *IEEE Transactions on Systems Science and Cybernetics*, v. 2, pp. 100 – 107, 1968.
8. A. Stentz, "Optimal and efficient path planning for partially known environments," In *Intelligent Unmanned Ground Vehicles*, Springer, Boston, MA, USA, 1997, pp. 203–220.
9. Q. Wang, Y. Hao, F. Chen. "Deepening the IDA\* algorithm for knowledge graph reasoning through neural network architecture," *Neurocomputing*, v. 429, 2021, pp. 101-109.
10. R. Zhou, E.A. Hansen, "Memory-Bounded {A\*} Graph Search," *The Florida AI Research Society Conference - FLAIRS*, pp. 203–209, 2002.
11. R. Holte, M. Perez, R. Zimmer, A. MacDonald, "Hierarchical A\*: Searching abstraction hierarchies efficiently," *Proceedings of the thirteenth national conference on Artificial intelligence*, v. 1, pp. 530–535, 1996.
12. B. Liu, X. Xiao and P. Stone, "A Lifelong Learning Approach to Mobile Robot Navigation," in *IEEE Robotics and Automation Letters*, v. 6(2), 2021, pp. 1090-1096.
13. B.Y. Chen, X.-W. Chen, H.-P. Chen, W.H.K. Lam, "Efficient algorithm for finding k shortest paths based on re-optimization technique," *Transportation Research Part E: Logistics and Transportation Review*, v. 133, 2020, Article number 101819.
14. X. Zhang, B. Wylie, C. Oscar, C.A. Moore, "Time-Optimal and Collision-Free Path Planning for Dual-Manipulator 3D Printer," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2020-October, 9283493, pp. 2389-2396.
15. O. Khatib, "Real-Time Obstacles Avoidance for Manipulators and Mobile Robots," *International Journal of Robotics Research*, vol. 5(2), pp. 90–98, 1986.
16. V.Kh. Pshikhopov (Ed.), D. Beloglazov, V. Finaev, V. Guzik, E. Kosenko, V. Krukhmalev, M. Medvedev, V. Pereverzev, A. Pyavchenko, R. Saprykin, I. Shapovalov, V. Soloviev. *Path Planning for Vehicles Operating in Uncertain 2D Environments*, Elsevier, Butterworth-Heinemann, 2017. 312 p, ISBN: 9780128123058.
17. S.S. Ge, Y.J. Cui, "New potential functions for mobile robot path planning," *IEEE Transactions on Robotics and Automation*. v. 16 (5), pp. 615 – 620, 2000.
18. A.C. Woods, H.M. La, "A Novel Potential Field Controller for Use on Aerial Robots," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 49 (4), 7932539, pp. 665-676, 2019.
19. Y. Koren, J. Borenstein, "Potential field methods and their inherent limitations for mobile robot navigation," *International Conference on Robotics and Automation* v. 2, pp. 1398 – 1404, 1991.
20. В.Х. Пшихопов, М.Ю. Медведев. "Групповое управление движением мобильных роботов в неопределенной среде с использованием неустойчивых режимов", *Труды СПИИРАН*. 2018. Вып. 60. С. 39-63.

21. Zhou, C., Huang, B. & Fránti, P. A review of motion planning algorithms for intelligent robots. *Journal of Intelligent Manufacturing*, 2021. <https://doi.org/10.1007/s10845-021-01867-z>.
22. S. Chakravorty, S. Kumar, "Generalized Sampling-Based Motion Planners," *IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics*, vol. 41(3), 2011.
23. A. Ravankar, Ab. Ravankar, T. Emaru, Y. Kobayashi, "HPPRM: Hybrid Potential Based Probabilistic Roadmap Algorithm for Improved Dynamic Path Planning of Mobile Robots," *IEEE Access*, vol. 8, pp. 221743 – 221766, 2020.
24. S.M. LaValle, J. Kuffner, "Randomized kinodynamic planning," *Int. Journal of Robotics Research*, vol. 20(5), pp. 378–400, 2001.
25. S.M. LaValle, J.J. Kuffner, "Rapidly-exploring random trees: Progress and prospects," 2000 Workshop on the Algorithmic Foundations of Robotics, pp. 293–308, 2000.
26. X. Wang, X. Li, Y. Guan, J. Song, R. Wang, "Bidirectional Potential Guided RRT\* for Motion Planning," *IEEE Access*, vol. 7, pp. 95046 – 95057, 2019.
27. A. Qureshi, Y. Ayaz, "Potential functions based sampling heuristic for optimal path planning," *Autonomous Robot*, vol. 40, pp 1079–1093, 2016.
28. S. Karaman, E. Frazzoli, "Sampling-based algorithms for optimal motion planning," *The International Journal of Robotics Research*, 30(7), pp. 846–894, 2011.
29. L. Chen, Y. Shan, W. Tian, B. Li, and D. Cao, "A Fast and Efficient Double-Tree RRT \*Like Sampling-Based Planner Applying on Mobile Robotic Systems," *IEEE/ASME Transactions on Mechatronics*, vol. 23(6), pp. 2568 – 2578, 2018.
30. J. Wang, M.Q.-H. Meng, O. Khatib, "EB-RRT: Optimal Motion Planning for Mobile Robots," *IEEE Transactions on Automation Science and Engineering*, v. 17(4), pp. 2063-2073, 2020.
31. J. Janos, V. Vonasek, R. Penicka, "Multi-Goal Path Planning Using Multiple Random Trees," *IEEE Robotics and Automation Letter*. v. 6(2), 2021, pp. 4201-4208.
32. S.N. Spitz and A.A.G. Requicha, "Multiple-goals path planning for coordinate measuring machines," in *Proc. IEEE Int. Conf. Robot. Automat.*, vol. 3, 2000, pp. 2322–2327.
33. D. Devaurs, T. Siméon, and J. Cortés, "A multi-tree extension of the transition-based RRT: Application to ordering-and-pathfinding problems in continuous cost spaces," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2014, pp. 2991–2996.
34. V. Vonásek and R. Pěnička, "Space-filling forest for multi-goal path planning," in *Proc. 24th IEEE Int. Conf. Emerg. Technol. Factory Automat.*, 2019, pp. 1587–1590.
35. R. Wang, X. Zhang, Y. Fang, B. Li, "Virtual-Goal-Guided RRT for Visual Servoing of Mobile Robots With FOV Constraint," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2021.
36. V. Kostjukov, V. Pshikhopov, M. Medvedev, "Optimization of mobile robot movement on a plane with finite number of repeller sources," *SPIIRAS Proceedings*. v. 19(1), pp. 43-78, 2020.
37. V. Kostjukov, M. Medvedev, V. Pshikhopov, "Method for Optimizing of Mobile Robot Trajectory in Repeller Sources Field," *Informatics and Automation*. v. 20(3), pp. 690-726, 2021.
38. V. Pshikhopov, M. Medvedev, "Multi-Loop Adaptive Control of Mobile Objects in Solving Trajectory Tracking Tasks," *Automation and Remote Control*, v. 81(11), pp. 2078–2093, 2020.
39. B.A. Güvenç, L. Güvenç, S. Karaman, "Robust Yaw Stability Controller Design and Hardware-in-the-Loop Testing for a Road Vehicle," *IEEE Transactions on Vehicular Technology*, v. 58(2), pp. 555-571, 2009.

40. M. Dorigo, L.M. Gambardella, "Ant Colony System: A Cooperative Learning Approach to the Traveling Salesman Problem," *IEEE Transactions on Evolutionary Computation*, v. 1(1), pp. 53-66, 1997.
41. J.H. Wilkinson, "The algebraic Eigenvalue Problem," Oxford, Clarendon Press, 1965.
42. W. Chen and J. Zhang, "An Ant Colony Optimization Approach to a Grid Workflow Scheduling Problem With Various QoS Requirements," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, v. 39(1), pp. 29-43, 2009.
43. A.R. Gaiduk, O.V. Martjanov, M.Yu. Medvedev, V.Kh. Pshikhopov, N. Hamdan, A. Farhood, "Neural network based control system for robots group operating in 2-d uncertain environment," *Mekhatronika, Avtomatizatsiya, Upravlenie*. v. 21(8), pp. 470 – 479, 2020.

**Pshikhopov Viacheslav** — Ph.D., Dr.Sci., Professor, Director, R&d department of robotics and control systems, Southern Federal University (SFedU). Research interests: planning and control of autonomous robots and manipulators with a focus on time-optimal control, control in obstructed unmapped environments, and intelligent and group control. The number of publications — 300. pshichop@rambler.ru; 2, Shevchenko St., 347922, Taganrog, Russia; office phone: +7(863)437-1694.

**Medvedev Mikhail** — Ph.D., Dr.Sci., Associate Professor, Leading researcher, Southern Federal University (SFedU). Research interests: planning and control of autonomous robots with a focus on adaptive and robust control, path planning methods, and neural network planning. The number of publications — 260. medvmihal@gmail.com; 2, Shevchenko St., 347922, Taganrog, Russia; office phone: +7(8634)371-694.

**Kostjukov Vladimir** — Ph.D., Associate professor, Southern Federal University (SFedU). Research interests: applied electrodynamics, dynamical systems modeling, path planning methods, aerodynamical modeling. The number of publications — 65. wkost-einheit@yandex.ru; 2, Shevchenko St., 347922, Taganrog, Russia; office phone: +7(8634)371-694.

**Houssein Firas** — Junior researcher, Southern Federal University (SFedU). Research interests: path planning algorithms, optimization algorithms, machine learning algorithms, and dynamical systems modeling. The number of publications — 0. info@sfedu.ru; 2, Shevchenko St., 347922, Taganrog, Russia; office phone: +7(8634)681-894.

**Kadhim Azhar** — Ph.D., Assistant, Department of mechanical techniques, Technical Institute of Nasiriyah. Research interests: path planning algorithms, machine learning methods. The number of publications — 0. info@nust.edu.iq; 26, Baghdad St., 64001, Nasiriyah, Iraq; office phone: +9647800778836.

**Acknowledgements.** The study is supported by the Russian Science Foundation, grant 22-29-00533, executed at Joint-stock Company "Scientific-Design bureau of Robotics and Control Systems".

В.Х. ПШИХОПОВ, М.Ю. МЕДВЕДЕВ, В.А. КОСТЮКОВ, Ф. ХУССЕЙН,  
А. КАДИМ

## АЛГОРИТМЫ ПЛАНИРОВАНИЯ ТРАЕКТОРИЙ В ДВУМЕРНОЙ СРЕДЕ С ПРЕПЯТСТВИЯМИ

*Пшихопов В.Х., Медведев М.Ю., Костюков В.А., Хуссейн Ф., Кадим А.* **Алгоритмы планирования траекторий в двумерной среде с препятствиями.**

**Аннотация.** В данной статье предложены алгоритмы планирования и управления движением мобильного робота в двухмерной стационарной среде с препятствиями. Задача состоит в том, чтобы сократить длину запланированного пути, учесть динамические ограничения робота и получить плавную траекторию. Для учета динамических ограничений мобильного робота на карту добавляются виртуальные препятствия, перекрывающие невыполнимые участки движения. Такой способ учета динамических ограничений позволяет использовать картографические методы без увеличения их сложности. В качестве алгоритма глобального планирования используется модифицированная версия алгоритма быстрого исследования случайных деревьев (Multi parent nodes RRT – MPN-RRT). В этом алгоритме, в отличие от оригинальной версии, используется несколько родительских узлов, что уменьшает длину запланированной траектории по сравнению с исходной версией RRT с одним узлом. Кратчайший путь на построенном графе находится с помощью алгоритма оптимизации муравьиной колонии. Методами численного моделирования показано, что использование двух родительских узлов позволяет уменьшить среднюю длину пути для городской среды с низкой плотностью застройки. Для решения проблемы медленной сходимости алгоритмов, основанных на случайном поиске и сглаживании путей, алгоритм RRT дополнен алгоритмом локальной оптимизации. Алгоритм RRT ищет глобальный путь, который сглаживается и оптимизируется итеративным локальным алгоритмом. Алгоритмы управления нижнего уровня, разработанные в этой статье, автоматически уменьшают скорость робота при приближении к препятствиям или повороте. Общая эффективность разработанных алгоритмов продемонстрирована методами численного моделирования с использованием большого количества экспериментов.

**Ключевые слова:** мобильный робот, планирование движения, управление движением, движение робота, планирование пути.

**Пшихопов Вячеслав Хасанович** — д-р техн. наук, профессор, директор, НИИ робототехники и процессов управления, Южный федеральный университет (ЮФУ). Область научных интересов: управление мобильными роботами в неопределенных средах, оптимальное управление роботами, анализ и синтез систем группового управления, интеллектуальное управление и планирование в робототехнике. Число научных публикаций — 300. pshichop@rambler.ru; ул. Шевченко, 2, 347922, Таганрог, Россия; р.т.: +7(863)437-1694.

**Медведев Михаил Юрьевич** — д-р техн. наук, доцент, ведущий научный сотрудник, Южный федеральный университет (ЮФУ). Область научных интересов: адаптивное и робастное управление мобильными роботами, оценивание возмущений, методы анализа и синтеза систем автоматического управления. Число научных публикаций — 260. medvmihal@gmail.com; ул. Шевченко, 2, 347922, Таганрог, Россия; р.т.: +7(8634)371-694.

**Костюков Владимир Александрович** — канд. техн. наук, доцент, Южный федеральный университет (ЮФУ). Область научных интересов: математическое моделирование, мобильные роботы, алгоритмы планирования пути, ветроэнергетические установки. Число научных публикаций — 65. [wkost-einheit@yandex.ru](mailto:wkost-einheit@yandex.ru); ул. Шевченко, 2, 347922, Таганрог, Россия; р.т.: +7(8634)371-694.

**Хусейн Фирас** — младший научный сотрудник, Южный федеральный университет (ЮФУ). Область научных интересов: алгоритмы планирования пути, алгоритмы оптимизации, машинное обучение, моделирование динамических системы. Число научных публикаций — 0. [info@sfnedu.ru](mailto:info@sfnedu.ru); ул. Шевченко, 2, 347922, Таганрог, Россия; р.т.: +7(8634)681-894.

**Кадим Азар** — Ph.D., ассистент, кафедра механических технологий, Технический институт Насирии. Область научных интересов: алгоритмы планирования пути, методы машинного обучения. Число научных публикаций — 0. [info@nust.edu.iq](mailto:info@nust.edu.iq); ул. Багдадская, 26, 64001, Эн-Насирия, Ирак; р.т.: +9647800778836.

**Поддержка исследований.** Работа выполнена при финансировании Российского научного фонда (грант 22-29-00533, выполняемый в ОАО «НКБ робототехники и систем управления»).

### Литература

1. K.A. Kazakov, V.A. Semenov, "Reviews of modern path planning methods," Proceedings of ISP RAS, v. 28(4), 2016, pp. 241-294.
2. Sánchez-Ibáñez, J.R.; Pérez-del-Pulgar, C.J.; García-Cerezo, A. Path Planning for Autonomous Mobile Robots: A Review. *Sensors* 2021, 21, 7898. <https://doi.org/10.3390/s21237898>
3. L.E. Kavraki; P. Svestka; J.-C. Latombe; M.H. Overmars, "Probabilistic roadmaps for path planning in high-dimensional configuration spaces," *IEEE Transactions on Robotics and Automation*, v. 12 (4), pp. 566-580, 1996.
4. A.G. Sukharev, "Optimal strategies of the search for an extremum," *U.S.S.R. Computational Mathematics and Mathematical Physics*, v. 11(4), pp. 910-924, 1971. Translated from Russian, *Journal Vychisl. Mat. i Mat. Fiz.*
5. F. Samaniego, J. Sanchis, S. García-Nieto, R. Simarro. "Recursive Rewarding Modified Adaptive Cell Decomposition (RR-MACD): A Dynamic Path Planning Algorithm for UAVs," *Electronics*, v. 8 (3), 306, 2019.
6. R. Toodesh, and S. Verhagen. "Adaptive, variable resolution grids for bathymetric applications using a quadtree approach," *Journal of Applied Geodesy*, v. 12(4), 2018, pp. 311-322.
7. P.E. Hart, N.J. Nilsson, B.A. Raphael, "Formal Basis for the Heuristic Determination of Minimum Cost Paths," *IEEE Transactions on Systems Science and Cybernetics*, v. 2, pp. 100 – 107, 1968.
8. A. Stentz, "Optimal and efficient path planning for partially known environments," In *Intelligent Unmanned Ground Vehicles*, Springer, Boston, MA, USA, 1997, pp. 203–220.
9. Q. Wang, Y. Hao, F. Chen. "Deepening the IDA\* algorithm for knowledge graph reasoning through neural network architecture," *Neurocomputing*, v. 429, 2021, pp. 101-109.
10. R. Zhou, E.A. Hansen, "Memory-Bounded {A\*} Graph Search," *The Florida AI Research Society Conference - FLAIRS*, pp. 203–209, 2002.

11. R. Holte, M. Perez, R. Zimmer, A. MacDonald, "Hierarchical A\*: Searching abstraction hierarchies efficiently," Proceedings of the thirteenth national conference on Artificial intelligence, v. 1, pp. 530–535, 1996.
12. B. Liu, X. Xiao and P. Stone, "A Lifelong Learning Approach to Mobile Robot Navigation," in IEEE Robotics and Automation Letters, v. 6(2), 2021, pp. 1090-1096.
13. B.Y. Chen, X.-W. Chen, H.-P. Chen, W.H.K. Lam, "Efficient algorithm for finding k shortest paths based on re-optimization technique," Transportation Research Part E: Logistics and Transportation Review, V. 133, 2020, Article number 101819.
14. X. Zhang, B. Wylie, C. Oscar, C.A. Moore, "Time-Optimal and Collision-Free Path Planning for Dual-Manipulator 3D Printer," IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2020-October, 9283493, pp. 2389-2396.
15. O. Khatib, "Real-Time Obstacles Avoidance for Manipulators and Mobile Robots," International Journal of Robotics Research, vol. 5(2), pp. 90–98, 1986.
16. V.Kh. Pshikhopov (Ed.), D. Beloglazov, V. Finaev, V. Guzik, E. Kosenko, V. Krukhmalev, M. Medvedev, V. Pereverzev, A. Pyavchenko, R. Saprykin, I. Shapovalov, V. Soloviev. Path Planning for Vehicles Operating in Uncertain 2D Environments, Elsevier, Butterworth-Heinemann, 2017. 312 p, ISBN: 9780128123058.
17. S.S. Ge, Y.J. Cui, "New potential functions for mobile robot path planning," IEEE Transactions on Robotics and Automation. v. 16 (5), pp. 615 – 620, 2000.
18. A.C. Woods, H.M. La, "A Novel Potential Field Controller for Use on Aerial Robots," IEEE Transactions on Systems, Man, and Cybernetics: Systems, 49 (4), 7932539, pp. 665-676, 2019.
19. Y. Koren, J. Borenstein, "Potential field methods and their inherent limitations for mobile robot navigation," International Conference on Robotics and Automation v.2, pp. 1398 – 1404, 1991.
20. V. Pshikhopov, M. Medvedev, "Group control of autonomous robots motion in uncertain environment via unstable modes," SPIIRAS Proceedings. v. 60(5), pp. 39-63, 2018.
21. Zhou, C., Huang, B. & Fränti, P. A review of motion planning algorithms for intelligent robots. Journal of Intelligent Manufacturing, 2021. <https://doi.org/10.1007/s10845-021-01867-z>.
22. S. Chakravorty, S. Kumar, "Generalized Sampling-Based Motion Planners," IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics, vol. 41(3), 2011.
23. A. Ravankar, Ab. Ravankar, T. Emaru, Y. Kobayashi, "HPPRM: Hybrid Potential Based Probabilistic Roadmap Algorithm for Improved Dynamic Path Planning of Mobile Robots," IEEE Access, vol. 8, pp. 221743 – 221766, 2020.
24. S.M. LaValle, J. Kuffner, "Randomized kinodynamic planning," Int. Journal of Robotics Research, vol. 20(5), pp. 378–400, 2001.
25. S.M. LaValle, J.J. Kuffner, "Rapidly-exploring random trees: Progress and prospects," 2000 Workshop on the Algorithmic Foundations of Robotics, pp. 293–308, 2000.
26. X. Wang, X. Li, Y. Guan, J. Song, R. Wang, "Bidirectional Potential Guided RRT\* for Motion Planning," IEEE Access, vol. 7, pp. 95046 – 95057, 2019.
27. A. Qureshi, Y. Ayaz, "Potential functions based sampling heuristic for optimal path planning," Autonomous Robot, vol. 40, pp 1079–1093, 2016.
28. S. Karaman, E. Frazzoli, "Sampling-based algorithms for optimal motion planning," The International Journal of Robotics Research, 30(7), pp. 846–894, 2011.
29. L. Chen, Y. Shan, W. Tian, B. Li, and D. Cao, "A Fast and Efficient Double-Tree RRT Like Sampling-Based Planner Applying on Mobile Robotic Systems," IEEE/ASME Transactions on Mechatronics, vol. 23(6), pp. 2568 – 2578, 2018.

30. J. Wang, M.Q.-H. Meng, O. Khatib, "EB-RRT: Optimal Motion Planning for Mobile Robots," *IEEE Transactions on Automation Science and Engineering*, v. 17(4), pp. 2063-2073, 2020.
31. J. Janos, V. Vonasek, R. Penicka, "Multi-Goal Path Planning Using Multiple Random Trees," *IEEE Robotics and Automation Letter*. v. 6(2), 2021, pp. 4201-4208.
32. S.N. Spitz and A.A.G. Requicha, "Multiple-goals path planning for coordinate measuring machines," in *Proc. IEEE Int. Conf. Robot. Automat.*, vol. 3, 2000, pp. 2322-2327.
33. D. Devaurs, T. Siméon, and J. Cortés, "A multi-tree extension of the transition-based RRT: Application to ordering-and-pathfinding problems in continuous cost spaces," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2014, pp. 2991-2996.
34. V. Vonásek and R. Pěnička, "Space-filling forest for multi-goal path planning," in *Proc. 24th IEEE Int. Conf. Emerg. Technol. Factory Automat.*, 2019, pp. 1587-1590.
35. R. Wang, X. Zhang, Y. Fang, B. Li, "Virtual-Goal-Guided RRT for Visual Servoing of Mobile Robots With FOV Constraint," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2021.
36. V. Kostjukov, V. Pshikhopov, M. Medvedev, "Optimization of mobile robot movement on a plane with finite number of repeller sources," *SPIIRAS Proceedings*. v. 19(1), pp. 43-78, 2020.
37. V. Kostjukov, M. Medvedev, V. Pshikhopov, "Method for Optimizing of Mobile Robot Trajectory in Repeller Sources Field," *Informatics and Automation*. v. 20(3), pp. 690-726, 2021.
38. V. Pshikhopov, M. Medvedev, "Multi-Loop Adaptive Control of Mobile Objects in Solving Trajectory Tracking Tasks," *Automation and Remote Control*, v. 81(11), pp. 2078-2093, 2020.
39. B.A. Güvenç, L. Güvenç, S. Karaman, "Robust Yaw Stability Controller Design and Hardware-in-the-Loop Testing for a Road Vehicle," *IEEE Transactions on Vehicular Technology*, v. 58(2), pp. 555-571, 2009.
40. M. Dorigo, L.M. Gambardella, "Ant Colony System: A Cooperative Learning Approach to the Traveling Salesman Problem," *IEEE Transactions on Evolutionary Computation*, v. 1(1), pp. 53-66, 1997.
41. J. H. Wilkinson, "The algebraic Eigenvalue Problem," Oxford, Clarendon Press, 1965.
42. W. Chen and J. Zhang, "An Ant Colony Optimization Approach to a Grid Workflow Scheduling Problem With Various QoS Requirements," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, v. 39(1), pp. 29-43, 2009.
43. A.R. Gaiduk, O.V. Martjanov, M.Yu. Medvedev, V.Kh. Pshikhopov, N. Hamdan, A. Farhood, "Neural network based control system for robots group operating in 2-d uncertain environment," *Mekhatronika, Avtomatizatsiya, Upravlenie*. v. 21(8), pp. 470 - 479, 2020.

А.М. КОРСАКОВ, Л.А. АСТАПОВА, А.В. БАХШИЕВ  
**ПРИМЕНЕНИЕ СЕГМЕНТНОЙ СПАЙКОВОЙ МОДЕЛИ  
НЕЙРОНА СО СТРУКТУРНОЙ АДАПТАЦИЕЙ ДЛЯ РЕШЕНИЯ  
ЗАДАЧ КЛАССИФИКАЦИИ**

*Корсаков А.М., Астапова Л.А., Бахшиев А.В. Применение сегментной спайковой модели нейрона со структурной адаптацией для решения задач классификации.*

**Аннотация.** Рассматриваются варианты применения сегментной спайковой модели нейрона с возможностью структурной адаптации для решения задач классификации. Проводится анализ современного состояния спайковых нейронных сетей. Делается вывод о крайне низком количестве работ по исследованию сегментных моделей нейрона. В качестве модели нейрона для данной работы обосновывается выбор сегментной спайковой модели. Приводится краткое описание такой модели, отмечены её основные особенности, позволяющие производить её структурное реконфигурирование. Описывается способ структурной адаптации модели ко входному паттерну импульсов. Приводится общая схема организации сегментных спайковых нейронов в сеть для решения задачи классификации. В качестве кодирования числовой информации в паттерны импульсов выбирается временное кодирование. Приводятся краткие результаты экспериментов по решению задачи классификации на общедоступных наборах данных (Iris, MNIST). Делается вывод о сопоставимости полученных результатов с результатами, полученными классическими методами. Кроме того, приводится подробное пошаговое описание экспериментов по определению состояния телеуправляемого необитаемого подводного аппарата: определение расстояния такого аппарата до дна и определение характера его движения. Показано соответствие полученных результатов реальному состоянию телеуправляемого необитаемого подводного аппарата. Сделан вывод о перспективности применения спайковых сегментных моделей нейрона с возможностью структурной адаптации при решении задач классификации. Рассмотрены дальнейшие перспективные продолжения исследований основанных на сегментных спайковых моделях нейрона.

**Ключевые слова:** нейроморфные системы, спайковый нейрон, спайковые нейронные сети, задача классификации, телеуправляемый необитаемый подводный аппарат.

**1. Введение.** Искусственные нейронные сети (ИНС) представляют собой упрощённые модели биологических нейронных сетей [1]. В классических ИНС входы и выходы нейронов обычно имеют вид аналоговых величин, в то время как в биологических нейронах эти сигналы определяются частотой импульсов [2]. Однако экспериментальные данные нейробиологии свидетельствуют о том, что биологическая нервная система кодирует информацию не только при помощи частоты возбуждения нейронов, но и через моменты времени генерации импульсов (спайков) [3]. Используя биологически правдоподобные модели спайковых нейронов в качестве базовой единицы для построения спайковых нейронных сетей (СНС), авторы [4-6] кодируют и обрабатывают нейронную информацию при помощи



рассчитанных последовательностей спайков. СНС часто называют третьим поколением нейронных сетей [7]. Они обладают значительно большими возможностями настройки нейронных сигналов, а также аппроксимации любой непрерывной функции [8]. Кроме того, показано, что они являются подходящим инструментом для обработки пространственно-временной информации [9].

Обучение с учителем в ИНС обеспечивает механизм принудительного сопоставления входных данных с желаемым результатом [10, 11]. Экспериментальные исследования показали, что обучение с учителем существует и в биологической нервной системе [12], но нет четкого понимания, как конкретно биологические нейронные сети реализуют этот процесс. Показано, что алгоритмы обучения с учителем для традиционных ИНС, такие как алгоритм обратного распространения ошибки [13], не могут быть использованы непосредственным образом для СНС [14].

Обучение с учителем для СНС является важной областью исследований. Исследователи проделали большой объём работы по обучению с учителем для СНС и достигли определённых результатов [14]. Алгоритмы обучения с учителем для СНС, предложенные в последние годы, можно разделить на несколько категорий с разных точек зрения [15]:

- по способу сетевой архитектуры их можно разделить на алгоритмы обучения с учителем для однослойных СНС, многослойных СНС с прямым распространением сигнала и рекуррентных СНС;
- по режиму работы их можно разделить на алгоритмы онлайн обучения и алгоритмы оффлайн обучения (алгоритмы пакетного обучения) [16];
- по способу кодирования информации их можно разделить на СНС, которые в ответ на входные временные или пространственно-временные данные в качестве выходных данных генерируют один спайк, и на СНС, которые генерируют последовательность спайков;
- по структурной динамике их можно разделить на обучение при фиксированных структурах СНС (регулярные СНС) и обучение в реконфигурируемых СНС (нерегулярные СНС) [17];
- по способу представления знаний их можно разделить на обучение, не основанное на знаниях, и обучение, направленное на получение знаний [18].

Произведя анализ требований к СНС для решения задачи классификации с учётом выше сказанного, авторы пришли к следующим результатам:

- с точки зрения сетевой архитектуры предлагается использовать СНС с прямым распространением сигнала,
- с точки зрения режима работы в настоящей работе будет использоваться оффлайн обучение. Более предпочтительное онлайн обучение на данный момент является лишь перспективным направлением, по которому в настоящее время ведутся работы,
- с точки зрения кодирования информации используется временное (time-to-first-spike) кодирование, что повышает энергоэффективность системы в случае аппаратной реализации,
- с точки зрения структурной динамики предлагаемый подход относится к реконфигурируемым растущим нейронным сетям, однако, в рассмотренных в разделах 6 и 7 примерах используются фиксированные структуры, что является достаточным для рассматриваемых задач.

Следует отметить, что с точки зрения авторов приведённая выше классификация является не полной, поскольку в ней не рассматривается возможность структурного реконфигурирования самой модели нейрона. К настоящему времени работы, использующие такой подход, практически отсутствуют. Применение модели нейрона с учетом описания структуры дендритного дерева предложено в [19], однако модель не рассматривает сигналы в сети как спайковые.

Анализ существующих решений, демонстрирующих спайковые нейронные сети и методы их структурного и параметрического обучения, позволяет сделать вывод о том, что для существенного увеличения адаптивной способности таких сетей представляется необходимым рассматривать не только параметрическую, но и структурную адаптацию нейронных сетей, под которой будем понимать изменение топологии связей между узлами (нейронами) сети и усложнение модели каждого узла путем перехода к сегментным (компаратментным) моделям нейронов. По мнению авторов, это даст возможность более эффективно решать задачи классификации на основе СНС.

В качестве альтернативы традиционным стратегиям развития, регулирующим количество нейронов в процессе обучения и адаптирующим веса нейронов в динамическом режиме, в настоящей работе предлагается подход, предлагающий, помимо возможности динамического изменения структуры нейронной сети и связей между отдельными нейронами, использование структурной онлайн перестройки самой модели нейрона.

В настоящей работе рассматривается возможность решения задачи классификации с использованием СНС с сегментными

нейронами. Для этого предлагается специальный алгоритм структурной адаптации нейронов ко входным паттернам импульсов, а также общая схема организации таких нейронов в сеть для решения поставленной задачи. Работа алгоритмов демонстрируется как на общедоступных модельных наборах данных (*Iris*, MNIST), так и на реальных данных, полученных в ходе экспериментов на телеуправляемом необитаемом подводном аппарате.

Далее в разделах 2 и 3 представлена структурная схема модели нейрона, использовавшаяся при моделировании, описан принцип структурной адаптации спайковой сегментной модели нейрона. В разделе 4 представлена общая схема решения задачи классификации паттернов импульсов. Раздел 5 описывает способ перевода числовой информации в паттерны импульсов. В разделе 6 кратко приведены результаты экспериментов по решению задачи классификации на общедоступных наборах данных, тогда как в разделе 7 представлено подробное пошаговое описание экспериментов по решению задачи классификации на данных состояния телеуправляемого необитаемого подводного аппарата. В заключительном разделе 8 приведено обсуждение результатов работы.

**2. Сегментная спайковая модель нейрона с возможностью структурной адаптации.** На рисунке 1 представлена схема сегментной спайковой модели нейрона (CSNM – Compartmental Spiking Neuron Model) [20], использованная в ходе выполнения работы. Предполагается, что на вход поступают спайки, которые преобразуются в синапсах в выходную величину, отображающую влияние синаптического тока на сегмент мембраны нейрона.

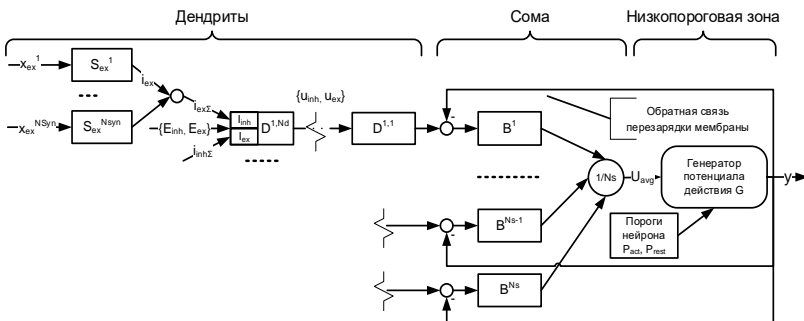


Рис. 1. Схема модели нейрона

Здесь  $B^1$ - $B^{Ns}$  – участки мембраны, описывающие тело нейрона.  $D^{ij}$  – участки мембраны, описывающие  $i$ -й дендрит.  $S_{ex}^i$  – модели

синаптической передачи. Их суммарное влияние оказывает либо возбуждающее (через ионный канал  $I_{inh}$ ), либо тормозное (через ионный канал  $I_{ex}$ ) воздействие на соответствующий участок мембраны.

Полагаем, что нейроны обмениваются информацией через события (возникновение спайков), которые можно представить по формуле:

$$x_i = \begin{cases} 1, & t \in [t_i; t_i + \Delta t] \\ 0, & t \notin [t_i; t_i + \Delta t] \end{cases} \quad (1)$$

где  $x_i$  – выход нейрона в  $i$ -й интервал времени, а  $t$  – время.

Представленная структурная организация модели нейрона позволяет проводить расчет систем дифференциальных уравнений для каждого нейрона независимо, поскольку нейроны обмениваются исключительно информацией о моменте времени возникновения спайка, а входной вектор системы уравнений модели нейрона формируется внутри модели как функция от моментов времени прихода спайков.

Система в целом характеризуется набором из следующих гиперпараметров: размер нейрона  $N_s$  (число сегментов сомы), множество длин дендритов  $N_d$  и число возбуждающих и тормозных синапсов  $N_{syn}$  для каждого сегмента мембраны. Их значения определяют структуру системы, которая формирует требуемую реакцию нейрона на входные воздействия.

Настоящая статья не ставит своей задачей подробное математическое описание использовавшейся модели нейрона. Такое описание можно найти в работе [21].

**3. Постановка задачи.** Отметим возможности рассмотренной модели нейрона, играющие существенную роль при решении задачи её структурного обучения:

- создание произвольного числа сегментов тела нейрона, что означает возможность изменять размерность входного паттерна импульсов,

- независимое наращивание длин дендритов. Увеличение длины дендрита приводит к временной задержке сигнала и к уменьшению его амплитуды,

- добавление произвольного числа синапсов на произвольном участке мембраны (соме, дендрите). Добавление синапса приводит к увеличению амплитуды сигнала,

- произвольная организация связей (как возбуждающих, так и тормозных) между отдельными элементами.

Всё вышеперечисленное означает возможность структурного реконфигурирования модели нейрона, что, в свою очередь, позволяет изменять время прихода и амплитуду сигнала с каждого из дендритов на сому нейрона.

В рамках решения задачи структурной адаптации работа будет вестись с паттернами импульсов, под которыми будем понимать совокупность сигналов, поступающих на входы нейрона (дендриты) за некоторый промежуток времени. Отдельный входной сигнал представляет собой одиночный импульс единичной амплитуды. Соответственно отдельные сигналы в паттерне имеют одинаковую (единичную) амплитуду, но, в общем случае, различное время поступления на соответствующий вход (дендрит) нейрона. Считаем, что за определённый промежуток времени поступления паттерна, на отдельном входе (дендрите) может быть зафиксировано не более одного входного сигнала.

Результатом распознавания конкретного паттерна будем считать генерацию выходного сигнала нейроном. Выходной сигнал генерируется в низкопороговой зоне нейрона (генератор потенциала действия) и представляет собой импульсы единичной амплитуды, формируемые при превышении порога величиной мембранного потенциала. Само же значение потенциала внутри низкопороговой зоны нелинейно связано с суммой сигналов, пришедших на вход нейрона. Таким образом, для преодоления порога и генерации потенциала действия, необходимо максимизировать сумму вкладов в потенциал внутри низкопороговой зоны от всех дендритов, что может быть достигнуто синхронизацией действия входных сигналов в нейроне, т. е. одновременностью прихода всех импульсов паттерна на тело нейрона.

Как было показано выше, синхронизация может быть достигнута путём добавления в модель отдельных сегментов дендритов. Однако, это приводит к уменьшению амплитуды входного сигнала, приходящего на сому, на данном дендрите, что могло бы стать препятствием для выбора фиксированного, универсального для всех входных паттернов, порога генерации. Уменьшение сигнала компенсируется добавлением синапсов на конечный сегмент дендрита, что приводит к увеличению амплитуды входного сигнала на соме.

Процедуру наращивания длины дендрита будем называть синхронизацией, тогда как процедуру добавления синапсов – нормализацией (рисунок 2). Таким образом, нашей задачей является максимизация потенциала внутри низкопороговой зоны нейрона путём синхронизации и нормализации, с целью вызвать генерацию

потенциала действия. Процессы синхронизации и нормализации в рамках структурной адаптации к конкретному паттерну соответствуют обучению нейрона данному паттерну импульсов. Генерация потенциала действия при фиксированном паттерне импульсов и соответствующей ему конфигурации обученного нейрона является результатом распознавания данного паттерна импульсов.

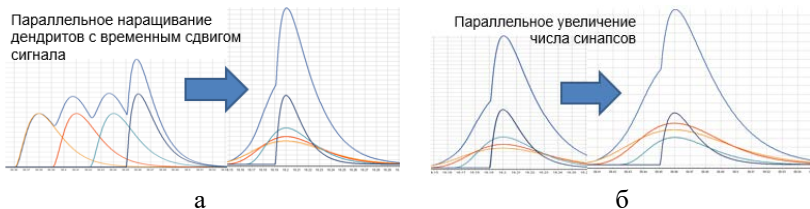


Рис. 2. Структурная адаптация модели нейрона: синхронизация (а), нормализация (б)

Алгоритм расчёта структурной адаптации модели нейрона приведён на рисунке 3.

Результаты экспериментальных исследований структурной адаптации спайковой модели нейрона представлены в работе [22].

**4. Общая схема решения задачи классификации с использованием сегментной спайковой модели нейрона.** Важной особенностью обученного нейрона является его способность отвечать на паттерны импульсов до некоторого предела отличные, чем паттерн импульсов, которому данный нейрон обучался. Однако, время ответа нейрона на паттерн импульсов, отличный от паттерна импульсов, которому нейрон обучался, различается со временем ответа нейрона на паттерн импульсов, которому тот был обучен. В первом случае это время оказывается больше, т. е. чем ближе входной паттерн импульсов, к паттерну импульсов, которому обучен нейрон, тем быстрее происходит генерация выходного импульса в нейроне.

На рисунке 4 представлена общая структурная схема организации в сеть обученных нейронов, решающая поставленную задачу.

Пусть задачу классификации требуется решить для  $K$  различных классов. Результатом решения является отнесение тестового паттерна импульсов к одному из этих классов, либо не отнесение ни к одному из них.

На представленной схеме каждому классу соответствуют  $M$  нейронов ( $M$  – размер обучающей выборки для класса), каждый из которых обучен своему обучающему примеру (обучающему паттерну).

В результате предъявления тестового примера (тестового паттерна) всей системе, каждый из обученных нейронов производит сравнение своего обучающего паттерна с тестовым и, при превышении суммарным значением мембранного потенциала фиксированного порога, происходит генерация выходного импульса теми из нейронов, для которых указанный порог оказался превышен.

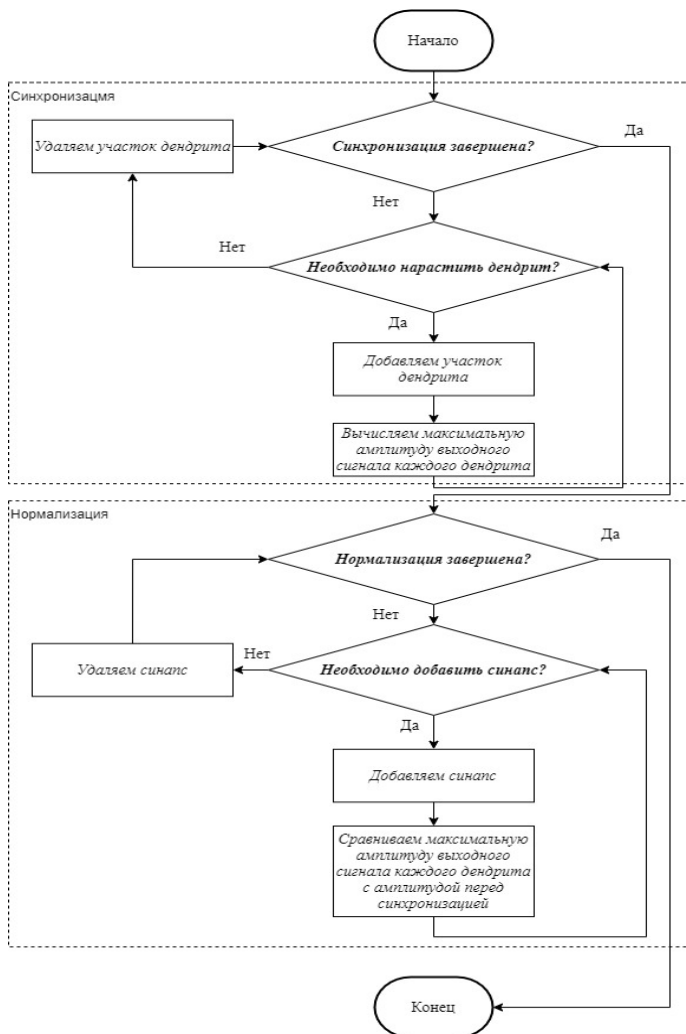


Рис. 3. Алгоритм расчёта структурной адаптации модели нейрона

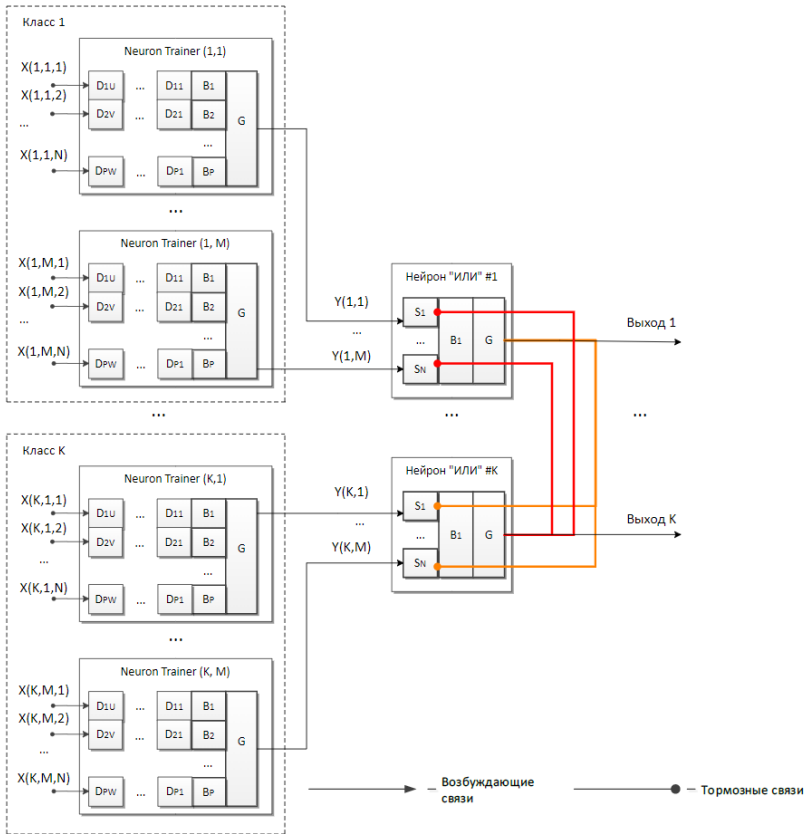


Рис. 4. Общая схема решения задачи классификации паттернов импульсов

Выходы нейронов, принадлежащих одному классу, поступают в качестве входов на соответствующий данному классу нейрон, реализующий функцию логического «ИЛИ». Логический элемент «ИЛИ» подразумевает наличие нескольких входов, при этом активность лишь одного из входов элемента приводит к активности его единственного выхода (рисунок 5).

На рисунке 5  $G$  – генератор потенциала действия,  $B^1$  – сегмент сомы нейрона,  $S^n$  –  $n$ -й синапс сомы нейрона.



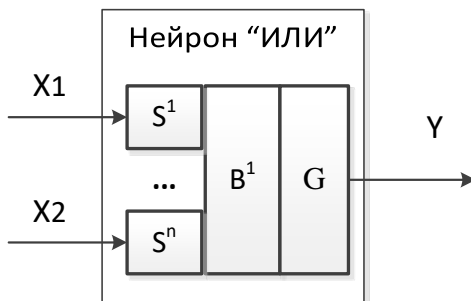


Рис. 5. Структурная схема нейрона, реализующего логическую функцию «ИЛИ»

Нейрон, реализующий функцию логического «ИЛИ» имеет лишь один элемент сомы, на котором располагаются возбуждающие синапсы в количестве, равном размерности входного паттерна. Каждый отдельный сигнал паттерна поступает на соответствующий ему возбуждающий синапс. Таким образом, активность лишь одного из входов нейрона позволяет модели генерировать выходной сигнал.

Как указывалось ранее, быстрее выходной импульс генерируют те из нейронов, для которых их обучающий паттерн оказался ближе к тестовому. Таким образом, естественным решением является организация тормозных связей от каждого из нейронов «ИЛИ» ко всем остальным, т. е. нейрон, чей обучающий паттерн оказался ближе к тестовому («нейрон-победитель»), успевает затормозить все остальные нейроны. На рисунке 4 тормозные связи от первого нейрона «ИЛИ» к  $K$ -ому обозначены оранжевым цветом, а от  $K$ -ого к первому – красным. Реакция соответствующего нейрона «ИЛИ» и означает отнесение тестового паттерна к соответствующему классу.

Стоит отметить и возможность генерации выходных спайков несколькими нейронами «ИЛИ», если сигналы поступили на их входы в пределах одного расчётного интервала времени. Таким образом, сформулируем три варианта ответа схемы на рисунке 4 на предъявление ей тестового паттерна:

- нейрон, отклик с которого пришёл первым, успевает затормозить все остальные нейроны, в результате активизируется выход лишь с одного нейрона «ИЛИ». Такое распознавание будем называть «уверенным»,

- на входной паттерн два или более нейрона реагируют одновременно, в результате активизируются выходы с двух или нескольких нейронов «ИЛИ». Такое распознавание будем называть «неуверенным»,

– в результате поступления на вход нейронов тестового паттерна активации выходов нейронов «ИЛИ» не происходит. Такую ситуацию будем называть «отсутствием распознавания».

**5. Кодирование числовых данных.** Перед подачей входных данных в сеть необходимо закодировать числовые признаки в значения временных задержек. Согласно [23], использовавшийся метод кодирования можно определить, как «time-to-first-spike», или, другими словами, временное кодирование.

Определим входной набор данных как:

$$X = \{(x_1^1; x_2^1; \dots; x_N^1; y^1), (x_1^2; x_2^2; \dots; x_N^2; y^2), \dots, (x_1^M; x_2^M; \dots; x_N^M; y^M)\}, \quad (2)$$

где  $M$  – общий размер входных данных, а  $N$  – размерность пространства признаков.

Согласно [24] задержку  $d_n^m$   $n$ -ого признака  $m$ -ого примера входных данных  $X$  можно вычислить по формуле (3).

$$d_n^m = \frac{x_n^m - x_{min}^n}{x_{max}^n - x_{min}^n} \cdot T, \quad (3)$$

где  $x_{min}^n$  – минимальное значение  $n$ -ого признака по всему набору данных,  $x_{max}^n$  – максимальное значение  $n$ -ого признака по всему набору данных, а  $T$  – ширина временного окна.

Можно заметить, что кодирование данных в данном случае является простой линейной нормализацией с заданным диапазоном.

Очевидно, что в процессе распознавания паттерна импульсов значимыми являются лишь относительные смещения отдельных спайков в паттерне, смещение всего паттерна по временной шкале никак не сказывается на результате распознавания. Поскольку в рассматриваемых ниже задачах, помимо относительных, значимы и абсолютные величины времени спайков, необходимо привязать паттерн импульсов к некоторому абсолютному значению. Для решения этой задачи в модель вводился дополнительный – калибровочный – дендрит. Значение временной задержки на таком дендрите всегда принимает максимальное значение, т. е. равное ширине временного окна. Таким образом, при добавлении калибровочного дендрита во входной вектор, размерность пространства признаков становится равной  $N+1$ .

**6. Эксперименты по решению задачи классификации на общедоступных наборах данных.** Для проведения исследований

разработанной схемы для решения задачи классификации (рисунок 4) необходимы общедоступные наборы данных, применимые в данной области. Наиболее популярными наборами реальных данных являются *Iris* и MNIST.

Набор данных *Iris* [25] представляет собой описание цветков (по 50 измерений) трех типов. Для каждого цветка измерены четыре величины, которые выступают в качестве признаков. Набор данных *Iris* широко используется в машинном обучении в связи с небольшим размером и двумя линейно неразделимыми классами. Подробно результаты работы с базой данных *Iris* с использованием предложенной на рисунке 4 схемы описаны авторами в [26]. При обучении лишь по одному примеру на класс авторам удалось достичь результатов, сопоставимых с существующими на данный момент решениями.

База данных MNIST [27] – объёмная база данных образцов рукописного написания цифр является стандартом для исследований методов распознавания изображений с помощью машинного обучения в первую очередь на основе нейронных сетей. База данных состоит из 60000 изображений в оттенках серого размером 28 x 28, состоящих из 10 цифр, а также тестового набора из 10000 изображений. Цифры нормализованы по размеру и центрированы. В данном наборе каждое изображение представлено массивом чисел от 0 до 255, что соответствует интенсивностям пикселей, расположенных построчно. Исходными признаками можно считать интенсивности каждого пикселя, т. е. размер исходного пространства признаков равен 784.

Помимо решения самой задачи классификации, работа с базой данных MNIST предполагала значительный объём исследований, связанный с понижением размерности пространства признаков, где авторами к настоящему моменту получены оригинальные решения (в результате размерность признаков удалось снизить до значения два). Тем не менее, работа в этом направлении в настоящий момент всё ещё продолжается, поэтому здесь приведём лишь предварительный результат.

В качестве обучающей выборки в экспериментах с базой данных MNIST было использовано лишь по одному обучающему примеру для каждого из десяти классов (в противовес 6000 обучающих примеров на класс, используемых при стандартных экспериментах с MNIST). Результаты сведены в таблице 1.

Из таблицы 1 видно, что результаты предложенной сети сопоставимы с результатами, полученными на стандартных классификаторах.

Таблица 1. Результаты стандартных классификаторов на наборе данных MNIST в сравнении с предложенным методом, полученные на одних и тех же исходных данных

Метод	Средняя точность классификации, %
Перцептрон	36,8 ± 7,5
Логистическая регрессия	66,9
kNN	77
Random Forest	52,4 ± 2,8
Support Vector Machine	77,3
MPL	81,1 ± 0,4
<b>Предложенная сеть</b>	<b>68,7</b>

**7. Эксперименты по решению задачи классификации на данных состояния телеуправляемого необитаемого подводного аппарата.** В разделе 6 были представлены экспериментальные результаты, полученные в ходе решения модельных задач с использованием стандартных наборов входных данных. Далее представлены результаты экспериментов, полученные на реальных входных данных в ходе натурных испытаний телеуправляемого необитаемого подводного аппарата (ТНПА) [28, 29].

Актуальной задачей при управлении ТНПА является оценка состояния такого аппарата до дна. Одним из решений данной задачи может быть использование эхолота в составе доплеровского лага. Однако, точность измерения не всегда соответствует требуемой, а ошибки измерения зачастую носят критический характер.

В рамках настоящей работы был произведён эксперимент по решению данной задачи с использованием сегментной СНС. В качестве входных данных для обучения и классификации использовались гистограммы распределения яркостей пикселей видеокadres, полученных с видеокамеры, установленной на ТНПА. При приближении ТНПА ко дну, происходит постепенное изменение гистограммы изображения дна за счёт увеличения эффекта отражения от дна света от установленного на ТНПА прожектора (рисунок 6).



Рис. 6. Кадры с изображением дна, полученные с видеокамеры, установленной на ТНПА, на расстоянии: а) 7 метров; б) 2 метра; в) 0,5 метров

Характер изменения гистограммы продемонстрирован на рисунке 7.



Рис. 7. Характер изменения гистограммы при приближении ТНПА ко дну

На рисунке 7 по оси абсцисс отложена яркость пикселей на изображениях, по оси ординат – количество пикселей. Для построения гистограммы весь диапазон яркости (0–255) разбивался на восемь равных отрезков. Из рисунка 7 легко можно заметить, что правая часть гистограммы не является значимой с точки зрения классификации положения ТНПА по глубине. Исходя из этого три правых столбца гистограммы не учитывались в дальнейшем процессе обучения и классификации.

В рассматриваемом эксперименте оставшиеся пять столбцов гистограммы использовались в качестве признаков (и один калибровочный), тогда как число классов было выбрано равным трём. Для каждого класса использовался лишь один обучающий пример. Соответствующие обучающим примерам изображения проиллюстрированы на рисунке 6.

Создание моделей СНС с возможностью структурного реконфигурирования как самой сети, так и её элементов, подразумевает необходимость обеспечить возможность описывать системы, в которых заранее не определена топология связей между элементами системы и их структурное описание. Для этой цели была разработана специальная программная среда NMSDK, описанная в [30]. В указанной программной среде была построена схема

классификации, позволявшая относить входной кадр к одному из трёх классов: «0,5», «2,0» и «7,0» (рисунок 8).

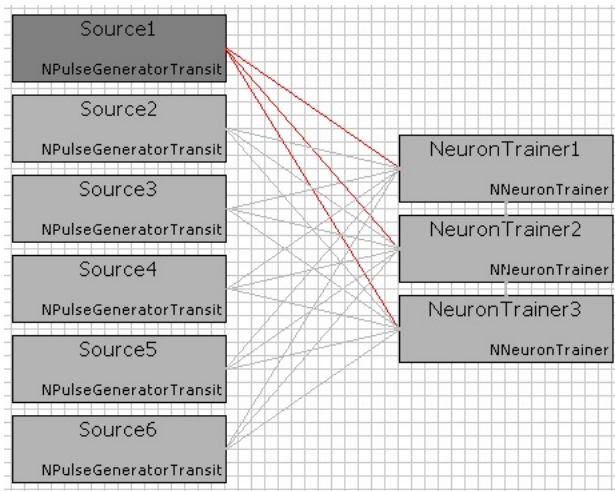


Рис. 8. Схема классификации кадров по расстоянию ТНПА от дна в среде NMSDK

На рисунке 8 справа показаны нейроны CSNM: «NeuronTrainer1» – для класса «0,5»; «NeuronTrainer2» – для класса «2,0», а «NeuronTrainer3» – для класса «7,0». «Source1» – I-й генератор импульсов, причём все блоки «Source» в совокупности задают требуемый паттерн импульсов.

В соответствии с методом кодирования, описанным в разделе 5, изображённым на рисунок 7 кривым соответствуют задержки спайков (в секундах), представленные на рисунке 9.

"0,5"	0,156	0,341	0,228	0,182	0,089	0,002	0	0
"2,0"	0,009	0,484	0,351	0,154	0	0	0	0
"7,0"	0	0,292	0,658	0,047	0,001	0	0	0

Рис. 9. Величины задержки спайков для обучающих примеров по трём классам (в секундах)

Как упоминалось ранее, три правых столбца таблицы на рисунке 9 не являются значимыми с точки зрения классификации положения ТНПА по глубине. Исходя из этого, нейроны с возможностью структурной адаптации обучались паттернам импульсов, сформированным спайками с задержками, приведёнными в

первых пяти столбцах таблицы на рисунке 9. Шестой дендрит являлся калибровочным (раздел 5), задержка на котором принимала постоянное значение, равное 0,2 секунды (принятой ширине временного окна). Структуры обученных нейронов, приведённые на рисунках 10-12, получены в среде моделирования NMSDK. Изменяемая часть в левых сторонах рисунков показывает количество дендритов, добавленных в ходе структурной адаптации.

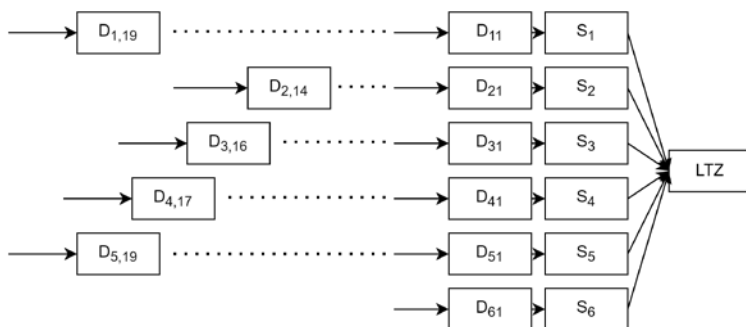


Рис. 10. Структура нейрона, отвечающего классу «0,5», после обучения

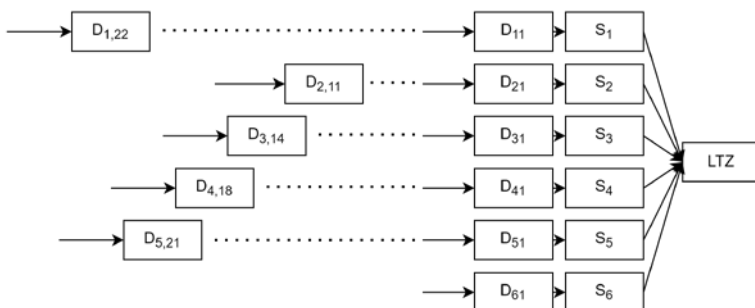


Рис. 11. Структура нейрона, отвечающего классу «2,0», после обучения

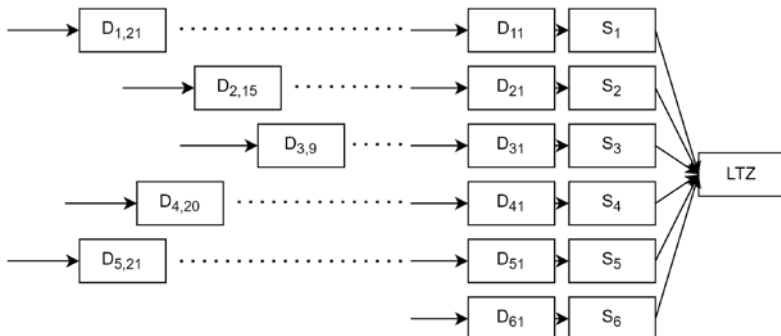


Рис. 12. Структура нейрона, отвечающего классу «7,0», после обучения

На рисунках 10-12 обозначены:  $LTZ$  – низкопороговая зона нейрона;  $S_i$  –  $i$ -й сегмент тела нейрона (сомы);  $D_{ij}$  –  $j$ -й сегмент  $i$ -го дендрита нейрона.

После процедуры обучения на вход системы подавалась последовательность из сорока восьми кадров, соответствующая процессу погружения ТНПА с высоты семи метров до дна. Результаты расчёта продемонстрированы на рисунке 13.



Рис. 13. Результаты расчёта процесса погружения ТНПА

На рисунке 13 по оси абсцисс отложены номера измерений, серым обозначены точки, соответствующие классу «0,5»; оранжевым –



классу «2,0»; синим – классу «7,0». Шкала А показывает номера отвечающих нейронов. Нулевое значение на шкале А иллюстрирует отсутствие отклика соответствующего обученного нейрона. Ответ нейрона, обученного примеру из класса «0,5» обозначен значением «1»; нейрона, обученного примеру из класса «2,0» – значением «2»; нейрона, обученного примеру из класса «7,0» – значением «3».

Из рисунка 13 видно, что характер отклика нейронов соответствует характеру движения ТНПА (погружение). На кадрах 38 и 39 можно видеть отклик сразу двух нейронов: нейрона, соответствующего классу «0,5», и нейрона, соответствующего классу «2,0». Это соответствует промежуточному положению ТНПА между точками с расстоянием 0,5 и 2 метра от дна, т. е. примерно одному метру. Такой одновременный отклик нейронов был определен в разделе 4 как «неуверенное» распознавание.

Следующий эксперимент был посвящён определению характера движения ТНПА. В ходе эксперимента ТНПА совершал движение по прямоугольной траектории на одной глубине (маневр «коробочка») (рисунок 14).

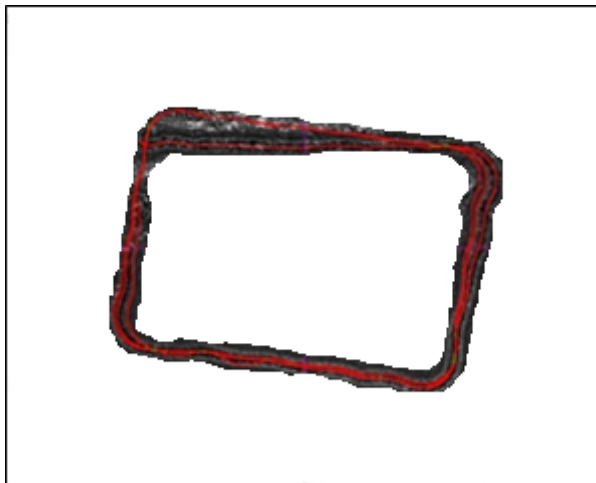


Рис. 14. Траектория движения ТНПА в ходе маневра «коробочка»

Данная траектория обеспечивалась работой двух из десяти движителей ТНПА (рисунки 15 и 16).

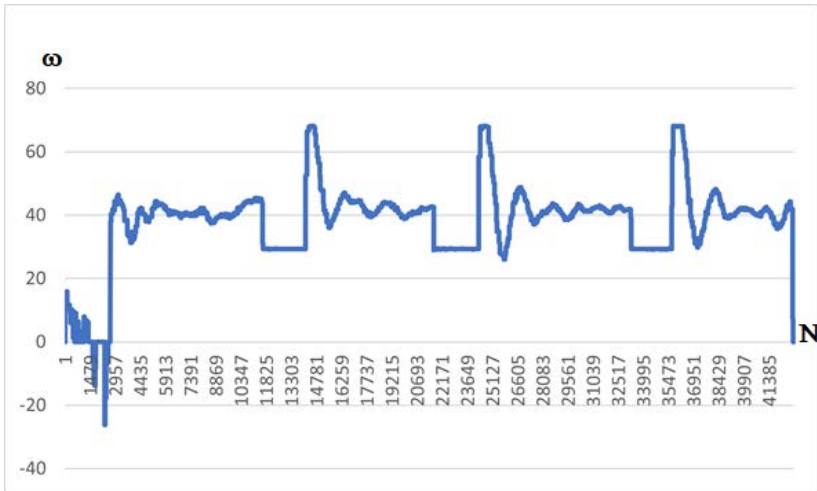


Рис. 15. Изменение скорости вращения правого двигателя ТНПА в ходе эксперимента

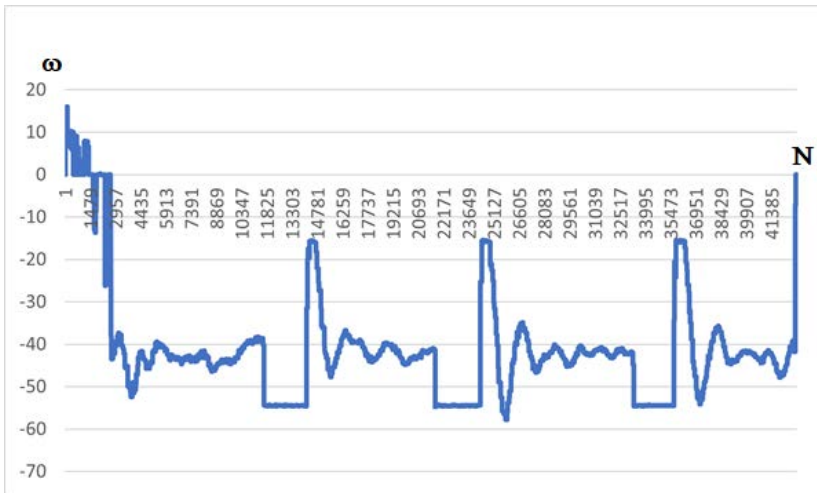


Рис. 16. Изменение скорости вращения левого двигателя ТНПА в ходе эксперимента

На графиках, на рисунках 15 и 16, показаны скорости вращения правого и левого двигателей ТНПА (в радианах в минуту с учётом направления вращения) во времени (в тактах измерения) в ходе эксперимента.

В рассматриваемом эксперименте показания датчиков левого и правого двигателей использовались в качестве признаков, тогда как число классов было выбрано равным трём. Для каждого класса использовался лишь один обучающий пример, причём примеры для обучения выбирались случайным образом.

Исходя из сказанного выше и разработанных методов классификации с использованием сегментных СНС (раздел 4), была построена схема классификации, позволяющая относить входной кадр к одному из трёх классов: «направо», «налево» и «прямо» (рисунок 17).

На рисунке 17 справа показаны нейроны CSNM: «NeuronTrainer1» – для класса «налево»; «NeuronTrainer2» – для класса «направо», а «NeuronTrainer3» – для класса «прямо». «Source1» – I-й генератор импульсов, причём все блоки «Source» в совокупности задают требуемый паттерн импульсов.

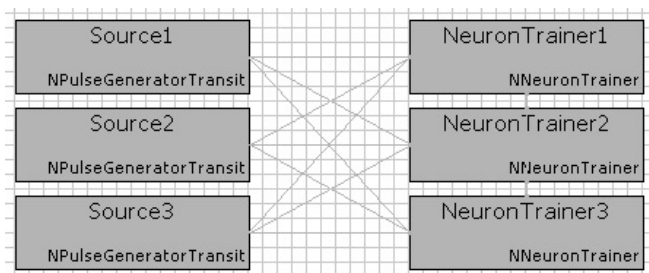


Рис. 17. Схема классификации кадров по характеру движения ТНПА в среде NMSDK

В соответствии с методом кодирования, описанным в разделе 5, выбранным в качестве обучающих примеров тактам измерения соответствуют следующие задержки спайков (в секундах):

Класс	M_Right	M_Left	1	2	3
«направо»	29,2276	-54,344	0,117406	0,009317	0,2
«налево»	63,4238	-18,2747	0,189838	0,107025	0,2
«прямо»	39,867	-43,7107	0,139942	0,038121	0,2

Рис. 18. Исходные данные (слева) и величины задержки спайков для обучающих примеров по трём классам в секундах (справа)

Нейроны (рисунок 17) обучались паттернам импульсов, сформированными спайками с задержками, приведёнными в таблице на рисунке 18. Третий дендрит являлся калибровочным (раздел 5), задержка на котором принимала постоянное значение, равное 0,2 секунды (принятой ширине временного окна). Структуры обученных нейронов, приведённые на рисунках 19–21, получены в среде моделирования NMSDK. Изменяемая часть в левых сторонах рисунков показывает количество дендритов, добавленных в ходе структурной адаптации.

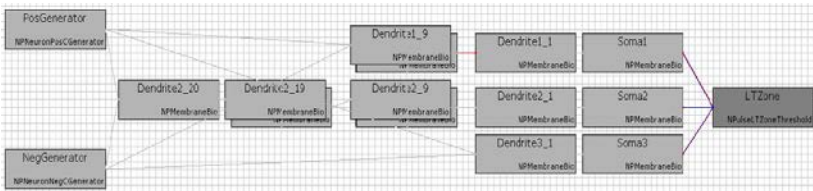


Рис. 19. Структура нейрона, отвечающего классу «направо», после обучения

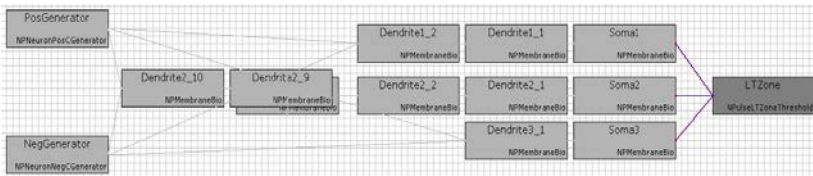


Рис. 20. Структура нейрона, отвечающего классу «налево», после обучения

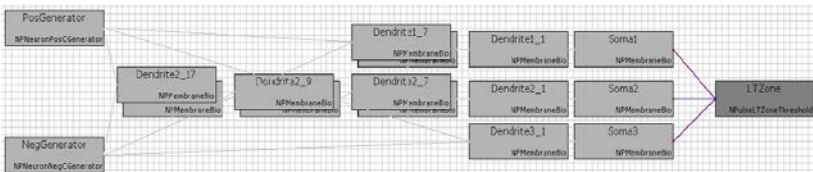


Рис. 21. Структура нейрона, отвечающего классу «прямо», после обучения

После процедуры обучения на вход системы подавалась последовательность из 42851 записей значений скоростей вращения правого и левого движителей ТНПА. Результаты расчёта продемонстрированы на рисунке 22.

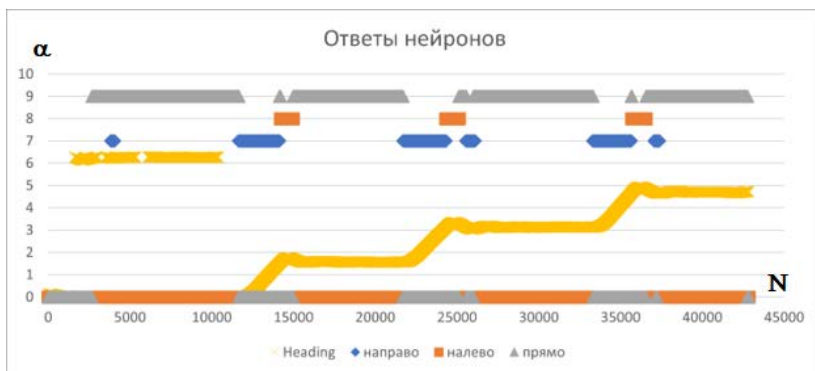


Рис. 22. Результаты расчёта процесса движения ТНПА в ходе маневра «коробочка»

На рисунке 22 по оси абсцисс отложены номера измерений, по оси ординат – угол курса движения ТНПА в радианах, синим цветом обозначены точки, соответствующие классу «направо»; оранжевым – классу «налево»; серым – классу «прямо». Жёлтые точки показывают угол курса движения ТНПА в радианах. Нулевое значение иллюстрирует отсутствие отклика соответствующего обученного нейрона. Ответ нейрона, обученного примеру из класса «направо» обозначен значением «7»; нейрона, обученного примеру из класса «налево» – значением «8»; нейрона, обученного примеру из класса «прямо» – значением «9».

Из рисунка 22 видно, что характер отклика нейронов соответствует характеру движения ТНПА.

Следует подчеркнуть, что во всех приведённых в разделах 6 и 7 экспериментах размер обучающей выборки составлял лишь один пример на класс. Эксперименты на тех же самых входных данных с использованием классических методов результатов не дали. Для получения результатов классическими методами необходимо увеличить размер обучающей выборки.

**8. Заключение.** Анализ существующих архитектур спайковых нейронных сетей показал, что для большинства спайковых нейронных сетей создана традиционная стратегия развития, которая изменяет их структуру путем регулирования количества нейронов в процессе обучения, а также в динамическом режиме адаптирует сеть под задачу, настраивая веса нейронов в реальном времени. По мнению авторов, адаптивная способность спайковой нейронной сети может быть увеличена, если использовать модели нейронов с изменяемой структурой дендритного дерева и синаптического аппарата.

Использование наряду с настройкой параметров нейронов и структуры нейронной сети, дополнительной возможности настройки внутренней структуры самих нейронов, по мнению авторов этой работы, может значительно повысить эффективность реализации поведенческих функций в нейроморфных системах управления.

Представленные в работе алгоритмы структурной адаптации для решения задач распознавания паттернов импульсов, были протестированы на решении задачи классификации. Опираясь на рассмотренные примеры, область применимости разработанной системы можно отнести к решению задач диагностики. По мнению авторов, основной областью применимости могли бы стать задачи с низкой доступностью обучающей выборки, высокой сложностью оценки состояния системы оператором, или даже с отсутствием возможности такой оценки, и с крайне высокими последствиями ошибки, например – задачи аварийной диагностики сложных систем.

По мнению авторов, полученные результаты показывают эффективность разработанных нейроморфных средств при решении задач классификации. Намечены пути продолжения этих исследований, направленные на разработку нейробиологических реализаций спайковых нейронных сетей и переход к их аппаратной реализации.

### Литература

1. Prieto A. et al. Neural networks: An overview of early research, current frameworks and new challenges // *Neurocomputing*. 2016. vol. 214. pp. 242–268.
2. Haykin S.S. Neural networks and learning machines // Harlow: Pearson Education. 2009. 934 p.
3. Walter F., Röhrbein F., Knoll A. Computation by time // *Neural Processing Letters*. 2016. vol. 44. no. 1. pp. 103–124.
4. Gerstner W., Kistler W.M. Spiking neuron models: Single neurons, populations, plasticity // Cambridge: Cambridge University Press. 2002. 496 p.
5. Izhikevich E.M. Which model to use for cortical spiking neurons // *IEEE Transactions on Neural Networks*. 2003. vol. 14. pp. 1569–1572.
6. Lin X., Zhang T. Dynamical properties of piecewise linear spiking neuron model // *Tien Tzu Hsueh Pao/Acta Electronica Sinica*. 2009. vol. 37. no. 6. pp. 1270–1276.
7. Ghosh-Dastidar S., Adeli H. Spiking neural networks // *International Journal of Neural Systems*. 2009. vol. 19. no. 4. pp. 295–308.
8. Maass W. Lower bounds for the computational power of networks of spiking neurons // *Neural computation*. 1996. vol. 8. no. 1. pp. 1–40.
9. Kulkarni S.R., Rajendran B. Spiking neural networks for handwritten digit recognition—Supervised learning and network optimization // *Neural Networks*. 2018. vol. 103. pp. 118–127.
10. Almási A.D. et al. Review of advances in neural networks: Neural design technology stack // *Neurocomputing*. 2016. vol. 174. pp. 31–41.
11. Denève S., Alemi A., Bourdoukan R. The brain as an efficient and robust adaptive learner // *Neuron*. 2017. vol. 94. no. 5. pp. 969–977.

12. Glaser J.I. et al. The roles of supervised machine learning in systems neuroscience // *Progress in neurobiology*. 2019. vol. 175. pp. 126–137.
13. Rummelhart, D.E. Learning representations by back-propagating errors // *Nature*. 1986. vol. 323. no. 9. pp. 533–536.
14. Lin X. et al. Supervised learning algorithms for spiking neural networks: a review // *Acta electronica sinica*. 2015. vol. 43. no. 3. pp. 577.
15. Wang X., Lin X., Dang X. Supervised learning in spiking neural networks: A review of algorithms and evaluations // *Neural Networks*. 2020. vol. 125. pp. 258–280.
16. Lobo J.L. et al. Spiking neural networks and online learning: An overview and perspectives // *Neural Networks*. 2020. vol. 121. pp. 88–100.
17. Kasabov N. et al. Dynamic evolving spiking neural networks for on-line spatio-and spectro-temporal pattern recognition // *Neural Networks*. 2013. vol. 41. pp. 188–201.
18. Kasabov N. *Time-Space, Spiking Neural Networks and Brain-Inspired Artificial Intelligence* // Berlin: Springer-Nature. 2018. 738 p.
19. Khan G.M., Miller J.F. In search of intelligence: evolving a develop-mental neuron capable of learning // *Connection Science*. 2014. vol. 26. no. 4. pp. 297–333.
20. Bakhshiev A.V., Gundelakh F.V. Mathematical Model of the Impulses Transformation Processes in Natural Neurons for Biologically Inspired Control Systems Development // *CEUR Workshop Proceedings*. 2015. vol. 1452. pp. 1–12.
21. Bakhshiev A., Demcheva A., Stankevich L. CSNM: The Compartmental Spiking Neuron Model for Developing Neuromorphic Information Processing Systems // *International Conference on Neuroinformatics*. 2021. pp. 327–333.
22. Бахшиев А.В., Корсаков А.М., Астапова Л.А., Станкевич Л.А. Структурная адаптация сегментной спайковой модели нейрона // *Нелинейная динамика в когнитивных исследованиях - 2021 : Труды VII Всероссийской конференции, Нижний Новгород, 20–24 сентября 2021 года*. 2021. С. 30–33.
23. Park S. et al. T2FSNN: deep spiking neural networks with time-to-first-spike coding // *2020 57th ACM/IEEE Design Automation Conference (DAC)*. 2020. pp. 1–6.
24. Lin P. et al. SpikeCD: a parameter-insensitive spiking neural network with clustering degeneracy strategy // *Neural Computing and Applications*. 2019. vol. 31. no. 8. pp. 3933–3945.
25. UCI Machine Learning Repository: Iris Data Set [Electronic resource]. URL: <https://archive.ics.uci.edu/ml/datasets/iris> (дата обращения: 22.04.2021).
26. Astopova L. A. et al. Compartmental spiking neuron model for pattern classification // *Journal of Physics: Conference Series*. 2021. vol. 2094. no. 3. pp. 032032.
27. MNIST handwritten digit database, Yann LeCun, Corinna Cortes and Chris Burges [Electronic resource]. URL: <http://yann.lecun.com/exdb/mnist/> (дата обращения: 12.10.2021).
28. Щур Н.А., Половко С.А., Деулин А.А. Применение методов вычислительной гидродинамики для получения характеристик переходных процессов АНПА // *Робототехника и техническая кибернетика*. 2020. Т. 8. №4. С. 287–295.
29. Горюнов В.В., Половко С.А., Щур Н.А. Разработка схемотехнического решения и конструкции емкостной матрицы датчиков давления для применения в робототехнике // *Робототехника и техническая кибернетика*. 2020. Т. 8. №4. С. 308–318.
30. Bakhshiev A.V. et al. The architecture of a software platform for growing spiking neural networks simulator developing // *Journal of Physics: Conference Series*. 2020. vol. 1679. no. 4. pp. 042001.

**Корсаков Антон Михайлович** — старший научный сотрудник, лаборатория «технологий искусственного интеллекта», Центральный научно-исследовательский и опытно-конструкторский институт робототехники и технической кибернетики (ЦНИИ

РТК). Область научных интересов: нейроморфные системы и спайковые нейронные сети, математические основы управления робототехническими системами, системы технического зрения. Число научных публикаций — 34. anton\_korsakov@mail.ru; Тихорецкий проспект, 21, 194064, Санкт-Петербург, Россия; р.т.: +7(911)784-2747.

**Астапова Любовь Алексеевна** — инженер-программист, Санкт-Петербургский политехнический университет Петра Великого (СПбПУ). Область научных интересов: нейроморфные системы и спайковые нейронные сети, математические основы управления робототехническими системами, системы технического зрения. Число научных публикаций — 8. astapova.la@yandex.ru; улица Политехническая, 29, 195251, Колпино, Россия; р.т.: +7(981)785-2931.

**Бахшиев Александр Валерьевич** — канд. техн. наук, доцент, Санкт-Петербургский политехнический университет Петра Великого (СПбПУ). Область научных интересов: применение методов искусственного интеллекта в решении задач технического зрения, развитие биологически-инспирированных подходов к созданию систем управления и обработки информации, разработка новых архитектур спайковых нейронных сетей. Число научных публикаций — 108. palexab@gmail.com; Политехническая улица, 29, 195251, Санкт-Петербург, Россия; р.т.: +7(921)748-0010.

**Поддержка исследований.** Работа проводилась в рамках выполнения государственного задания Минобрнауки России № 075-01623-22-00 «Исследование и разработка биоподобной системы управления поведением мобильных роботов на базе энергоэффективных программно-аппаратных нейроморфных средств».



A. KORSAKOV, L. ASTAPOVA, A. BAKHSHEV  
**APPLICATION OF A COMPARTMENTAL SPIKING NEURON  
MODEL WITH STRUCTURAL ADAPTATION FOR SOLVING  
CLASSIFICATION PROBLEMS**

*Korsakov A., Astapova L., Bakhshiev A. Application of a Compartmental Spiking Neuron Model with Structural Adaptation for Solving Classification Problems.*

**Abstract.** The problem of classification using a compartmental spiking neuron model is considered. The state of the art of spiking neural networks analysis is carried out. It is concluded that there are very few works on the study of compartmental neuron models. The choice of a compartmental spiking model is justified as a neuron model for this work. A brief description of such a model is given, and its main features are noted in terms of the possibility of its structural reconfiguration. The method of structural adaptation of the model to the input spike pattern is described. The general scheme of the compartmental spiking neurons' organization into a network for solving the classification problem is given. The time-to-first-spike method is chosen for encoding numerical information into spike patterns, and a formula is given for calculating the delays of individual signals in the spike pattern when encoding information. Brief results of experiments on solving the classification problem on publicly available data sets (Iris, MNIST) are presented. The conclusion is made about the comparability of the obtained results with the existing classical methods. In addition, a detailed step-by-step description of experiments to determine the state of an autonomous uninhabited underwater vehicle is provided. Estimates of computational costs for solving the classification problem using a compartmental spiking neuron model are given. The conclusion is made about the prospects of using spiking compartmental models of a neuron to increase the bioplausibility of the implementation of behavioral functions in neuromorphic control systems. Further promising directions for the development of neuromorphic systems based on the compartmental spiking neuron model are considered.

**Keywords:** neuromorphic systems, spiking neuron, spiking neural networks, classification task, autonomous underwater vehicle.

**Korsakov Anton** — Senior researcher, Laboratory of artificial intelligence technologies, Russian state scientific center for robotics and technical cybernetics (RTC). Research interests: neuromorphic systems and spike neural networks, mathematical foundations of robotic systems control, computer vision. The number of publications — 34. anton\_korsakov@mail.ru; 21, Tikhoretsky Av., 194064, St. Petersburg, Russia; office phone: +7(911)784-2747.

**Astapova Lyubov** — Software engineer, Peter the Great St.Petersburg Polytechnic University (SPbPU). Research interests: neuromorphic systems and spike neural networks, mathematical foundations of robotic systems control, computer vision. The number of publications — 8. astapova.la@yandex.ru; 29, Polytechnic St., 195251, Kolpino, Russia; office phone: +7(981)785-2931.

**Bakhshiev Aleksandr** — Ph.D., Associate professor, Peter the Great St.Petersburg Polytechnic University (SPbPU). Research interests: application of artificial intelligence methods in solving computer vision problems, development of biologically inspired approaches to the creation of information management and processing systems, development of new spike neural network architectures. The number of publications — 108.

palexab@gmail.com; 29, Polytechnic St., 195251, St. Petersburg, Russia; office phone: +7(921)748-0010.

**Acknowledgements.** This work was done as a part of the state task of the Ministry of Education and Science of Russia No. 075-01623-22-00 «Research and development of a biosimilar system for controlling the behavior of mobile robots based on energy-efficient software and hardware neuromorphic tools».

## References

1. Prieto A. et al. Neural networks: An overview of early research, current frame-works and new challenges. *Neurocomputing*. 2016. vol. 214. pp. 242–268.
2. Haykin S.S. *Neural networks and learning machines*. Pearson Education. 2009. 934 p.
3. Walter F., Röhrbein F., Knoll A. Computation by time. *Neural Processing Letters*. 2016. vol. 44. no. 1. pp. 103–124.
4. Gerstner W., Kistler W.M. *Spiking neuron models: Single neurons, populations, plasticity*. Cambridge University Press. 2002. 496 p.
5. Izhikevich E.M. Which model to use for cortical spiking neurons. *IEEE Transactions on Neural Networks*. 2003. vol. 14. pp. 1569–1572.
6. Lin X., Zhang T. Dynamical properties of piecewise linear spiking neuron model. *Tien Tzu Hsueh Pao/Acta Electronica Sinica*. 2009. vol. 37. no. 6. pp. 1270–1276.
7. Ghosh-Dastidar S., Adeli H. Spiking neural networks. *International Journal of Neural Systems*. 2009. vol. 19. no. 4. pp. 295–308.
8. Maass W. Lower bounds for the computational power of networks of spiking neurons. *Neural computation*. 1996. vol. 8. no. 1. pp. 1–40.
9. Kulkarni S.R., Rajendran B. Spiking neural networks for handwritten digit recognition — Supervised learning and network optimization. *Neural Networks*. 2018. vol. 103. pp. 118–127.
10. Almási A.D. et al. Review of advances in neural networks: Neural design technology stack. *Neurocomputing*. 2016. vol. 174. pp. 31–41.
11. Denève S., Alemi A., Bourdoukan R. The brain as an efficient and robust adaptive learner. *Neuron*. 2017. vol. 94. no. 5. pp. 969–977.
12. Glaser J.I. et al. The roles of supervised machine learning in systems neuroscience. *Progress in neurobiology*. 2019. vol. 175. pp. 126–137.
13. Rummelhart, D.E. Learning representations by back-propagating errors. *Nature*. 1986. vol. 323. no. 9. pp. 533–536.
14. Lin X. et al. Supervised learning algorithms for spiking neural networks: a review. *Acta electronica sinica*. 2015. vol. 43. no. 3. pp. 577.
15. Wang X., Lin X., Dang X. Supervised learning in spiking neural networks: A review of algorithms and evaluations. *Neural Networks*. 2020. vol. 125. pp. 258–280.
16. Lobo J.L. et al. Spiking neural networks and online learning: An overview and perspectives. *Neural Networks*. 2020. vol. 121. pp. 88–100.
17. Kasabov N. et al. Dynamic evolving spiking neural networks for on-line spatio-and spectro-temporal pattern recognition. *Neural Networks*. 2013. vol. 41. pp. 188–201.
18. Kasabov N. *Time-Space, Spiking Neural Networks and Brain-Inspired Artificial Intelligence*. Springer-Nature. 2018. 738 p.
19. Khan G.M., Miller J.F. In search of intelligence: evolving a developmental neuron capable of learning. *Connection Science*. 2014. vol. 26. no. 4. pp. 297–333.
20. Bakhshiev A.V., Gundelakh F.V. Mathematical Model of the Impulses Transformation Processes in Natural Neurons for Biologically Inspired Control Systems Development. *CEUR Workshop Proceedings*. 2015. vol. 1452. pp. 1–12.

21. Bakhshiev A., Demcheva A., Stankevich L. CSNM: The Compartmental Spiking Neuron Model for Developing Neuromorphic Information Processing Systems. International Conference on Neuroinformatics. 2021. pp. 327–333.
22. Bakhshiev A.V. et al. Strukturnaja adaptacija segmentnoj spajkovoju modeli nejrona. Nelinejnaja dinamika v kognitivnyh issledovanijah - 2021 : Trudy VII Vserossijskoj konferencii, Nizhnij Novgorod, 20–24 sentjabrja 2021 goda. 2021. pp. 30–33. (In Russ.).
23. Park S. et al. T2FSNN: deep spiking neural networks with time-to-first-spike coding. 2020 57th ACM/IEEE Design Automation Conference (DAC). 2020. pp. 1–6.
24. Lin P. et al. SpikeCD: a parameter-insensitive spiking neural network with clustering degeneracy strategy. Neural Computing and Applications. 2019. vol. 31. no. 8. pp. 3933–3945.
25. UCI Machine Learning Repository: Iris Data Set [Electronic resource]. URL: <https://archive.ics.uci.edu/ml/datasets/iris> (accessed: 22.04.2021).
26. Astapova L. A. et al. Compartmental spiking neuron model for pattern classification. Journal of Physics: Conference Series. 2021. vol. 2094. no. 3. pp. 032032.
27. MNIST handwritten digit database, Yann LeCun, Corinna Cortes and Chris Burges [Electronic resource]. URL: <http://yann.lecun.com/exdb/mnist/> (accessed: 12.10.2021).
28. Tschur N., Polovko S., Deulin A. Application of the computational fluid dynamics methods to obtain the characteristics of AUV transient responses. Robotics and Technical Cybernetics. 2020. vol. 8. no. 4. pp. 287-295. (In Russ.).
29. Goryunov V., Polovko S., Tschur N. echnology of cybernetic models' creating for the synthesis and refinement of autonomous underwater vehicle motion control system controllers. Robotics and Technical Cybernetics. 2020. vol. 8. no. 4. pp. 308-318. (In Russ.).
30. Bakhshiev A.V. et al. The architecture of a software platform for growing spiking neural networks simulator developing. Journal of Physics: Conference Series. 2020. vol. 1679. no. 4. pp. 042001.

Д.Т. ГАЛЕЕВ, В.С. ПАНИЩЕВ  
**ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ ЯЗЫКОВЫХ  
МОДЕЛЕЙ "ТРАНСФОРМЕР" В ЗАДАЧЕ НАХОЖДЕНИЯ  
ОТВЕТА НА ВОПРОС В РУССКОЯЗЫЧНОМ ТЕКСТЕ**

*Галеев Д.Т., Панищев В.С. Экспериментальное исследование языковых моделей "трансформер" в задаче нахождения ответа на вопрос в русскоязычном тексте.*

**Аннотация.** Целью исследования является получение более легковесной языковой модели, которая сравнима по показателям EM и F-меры с лучшими современными языковыми моделям в задаче нахождения ответа на вопрос в тексте на русском языке. Результаты работы могут найти применение в различных вопросно-ответных системах, для которых важно время отклика. Поскольку более легковесная модель имеет меньшее количество параметров чем оригинальная, она может быть использована на менее мощных вычислительных устройствах, в том числе и на мобильных устройствах. В настоящей работе используются методы обработки естественного языка, машинного обучения, теории искусственных нейронных сетей. Нейронная сеть настроена и обучена с использованием библиотек машинного обучения Torch и Hugging face. В работе было проведено обучение модели DistilBERT на наборе данных SberQUAD с применением дистилляции и без. Произведено сравнение работы полученных моделей. Обученная в ходе дистилляции модель DistilBERT (EM 58,57 и F-мера 78,42) смогла опередить результаты более крупной генеративной сети ruGPT-3-medium (EM 57,60 и F-мера 77,73) притом, что ruGPT-3-medium имеет в 6,5 раз больше параметров. Также модель продемонстрировала лучшие показатели EM и F-мера, чем та же модель, но к которой применялось только обычное дообучение без дистилляции (EM 55,65, F-мера 76,51). К сожалению, полученная модель сильнее отстаёт от более крупной дискриминационной модели ruBERT (EM 66,83, F-мера 84,95), которая имеет в 3,2 раза больше параметров. Предложены направления для дальнейшего исследования.

**Ключевые слова:** машинное обучение, глубокое обучение, нейронные сети, обработка естественного языка, трансформер.

**1. Введение.** С появлением архитектуры «трансформер» вся индустрия обработки естественного языка получила значительный скачок в результатах. В работах [1-3] показано каких успехов смогли добиться различные модели данной архитектуры. Результаты «трансформера» поспособствовали появлению таких моделей как: BART [4], T5 [5], Pegasus [6], ProphetNet [7]. Также исследования показали, что использование двух отдельных частей трансформера, кодировщика и декодера, отдельно друг от друга также позволяют получить выдающийся результат. Это способствовало появлению большого числа моделей, состоящих только из кодировщика: BERT [8], ALBERT [9], RoBERTa [10], а также способствовало появлению моделей, состоящих только из декодера: GPT-3 [11], CTRL [12], Transformer-XL [13].

Одной из основных проблем использования данных моделей является большое количество параметров. Чем больше количество параметров, тем больше ресурсов необходимо затратить на обучение сети и непосредственно на её применение. Данный нюанс может быть критичен для систем, в которых важна скорость отклика. Также из-за необходимости в большой вычислительной мощности применение данных моделей может быть ограничено только мощным оборудованием, что исключает применение моделей на мобильных устройствах. Поэтому одним из направлений изучения в области обработки естественного языка является нахождение способов по уменьшению размерности моделей при сохранении качества результатов. Одним из основных способов выполнить данную задачу является дистилляция [14].

Суть дистилляции заключается в том, что можно обучить более легковесную модель, которая будет имитировать поведение более сложной модели-учителя. В качестве учителя может выступать ансамбль моделей. Модель DistilBERT показала отличные результаты по сравнению с оригинальной версией [15].

Для языковых моделей существует огромное количество задач по обработке естественного языка, среди которых: текстовый поиск, машинный перевод, написание краткого содержания текста, распознавание именованных сущностей и т.д. Также одной из основных задач для обработки естественного языка при помощи нейронных сетей является создание модели, которая ищет ответ на вопрос в тексте. Под поиском ответа на вопрос в тексте будем подразумевать: наличие текста и вопроса к тексту, система должна выбрать в качестве ответа на вопрос непрерывный фрагмент из данного текста. В англоязычной литературе данный тип задачи называется Extractive Question Answering (извлечение ответа на вопрос). Для моделей типа «кодировщик» данная задача является задачей классификации, в которой они пытаются найти ответы на вопросы: «Является ли данное слово из текста началом ответа на заданный вопрос?», «Является ли данное слово из текста концом ответа на заданный вопрос?». Для данной задачи на английском языке основным набором данных является SQuAD [16]. Для русского языка компанией Сбер был создан набор данных SberQuAD [17].

Целью исследования является получение более легковесной языковой модели, которая несильно уступает лучшим современным языковым моделям в задаче нахождения ответа на вопрос в тексте на русском языке.

## 2. Материалы и методы

**2.1. Набор данных для обучения модели.** Одним из стандартных наборов данных для вопросно-ответных систем на английском языке является Stanford Question Answering Dataset (SQuAD) [16]. Набор данных основан на статьях с сайта Wikipedia. Статьи, которые были взяты в набор, охватывают широкий спектр тем – от музыкальных знаменитостей до абстрактных понятий. Вопросы основаны на содержании статьи, и человек способен ответить на них, прочитав текст статьи. Некоторые статьи могут содержать несколько вопросов. SQuAD часто используют как бенчмарк для новых языковых моделей.

На данный момент в интернете можно найти следующие версии этого набора: SQuAD 1.1 и SQuAD 2.0.

SQuAD 1.1 содержит 107785 вопросно-ответных пар, основанных на 536 статьях. В данной версии набора присутствуют вопросы, на которые есть ответы в представленных текстах. Данный набор разбит на 3 части: тренировочную (80%), а валидационную (10%) и проверочную (10%). Средняя длина текста составляет ~755 символа (120 токенов), средняя длина вопроса составляет ~60 символа (10 токенов), средняя длина ответа составляет ~19 символов (3 токена).

SQuAD 2.0 содержит все данные из версии 1.1, но также содержит дополнительные 50 тысяч вопросов, на которые нет ответов в представленных текстах. Это бы сделано для того, чтобы модель училась понимать ситуации, когда в тексте нет необходимой информации для ответа на вопрос.

Основным набором данных для вопросно-ответных систем на русском языке является SberQuAD. Он основан на русских статьях с сайта Wikipedia. Форматом ответов и вопросов он совпадает с SQuAD 1.1. Данный набор содержит 45328 тренировочных наборов из текста, вопроса и ответа, 5036 валидационных наборов и 23936 проверочных наборов. К сожалению, ответы на проверочные данные не представлены публично, поэтому результаты работы модели сравниваются на валидационном наборе.

Большинство вопросов в наборе данных SberQuAD начинаются либо с вопросительного слова, либо с предлога. Далее представлены десять наиболее распространенных начальных слов: «что», «в», «как», «кто», «какие», «когда», «какой», «где», «сколько», «на». Средняя длина текста составляет ~754 символа (102 токена), средняя длина вопроса составляет ~64 символа (9 токенов), средняя длина ответа составляет ~26 символов (4 токена).

Конкретный экземпляр набора данных был взят из библиотеки Hugging face. Каждый экземпляр данных содержит следующие поля: context, question, answers.

В поле context находится текст, к которому будет задан вопрос, question содержит вопрос к тексту из поля context, а answers включает в себя поле answer\_start, которое содержит индекс начала ответа на вопрос, и поле text, которое содержит полный текст ответа.

Пример тренировочных данных представлен на рисунке 1.

```
{
  'answers': {'answer_start': [166], 'text': ['9 млрд рублей в год']},
  'context': 'Город Байконур и космодром Байконур вместе образуют комплекс Байконур , арендованный Россией у Казахстана на период до 2050 года. Эксплуатация космодрома стоит около 9 млрд рублей в год (стоимость аренды комплекса Байконур составляет 115 млн долларов – около 7,4 млрд рублей в год; ещё около 1,5 млрд рублей в год Россия тратит на поддержание объектов космодрома), что составляет 4,2 % от общего бюджета Роскосмоса на 2012 год. Кроме того, из федерального бюджета России в бюджет города Байконура ежегодно осуществляется безвозмездное поступление в размере 1,16 млрд рублей (по состоянию на 2012 год). В общей сложности космодром и город обходятся бюджету России в 10,16 млрд рублей в год.',
  'id': 18340,
  'question': 'Сколько стоит эксплуатация космодрома?',
  'title': 'SberChallenge'
}
```

Рис. 1. Пример тренировочных данных

**2.2. Выбор модели.** Для опытов было решено использовать модели семейства BERT. Архитектура Bidirectional Encoder Representations from Transformers является одной из самых популярных моделей для обработки естественного языка [8], поскольку она показывает одни из лучших результатов среди аналогов. В данной модели присутствует кодировщик из архитектуры трансформер [1]. Изначально модель BERT обучалась одновременно на двух задачах:

- 1) предсказание пропущенных слов в тексте (masked language modeling);
- 2) определения является ли вторая часть текста логичным продолжением первой (next sentence prediction).

В оригинальной статье было описано две версии модели:

- 1) BERT BASE. Данная версия модели содержит 12 блоков трансформера. В каждом блоке трансформера содержится 12 «голов» внимания. Для каждого входного вектора на выход подаётся вектор с длиной 768. Модель содержит 110 миллионов параметров.

- 2) BERT LARGE. Данная версия модели содержит 24 блока трансформера. В каждом блоке трансформера содержится 16 «голов» внимания. Для каждого входного вектора на выход подаётся вектор с длиной 1024. Модель содержит 340 миллионов параметров.

Особенностью современных языковых моделей является то, что они обычно выкладываются в сеть после тренировки на различных задачах языкового моделирования или близкие к ним. Данное обучение происходит на огромных корпусах текстов. Это значит, что данные сети уже содержат векторные представления для слов из своего словаря и способны хорошо решать задачи языкового моделирования. И последующая работа с данными моделями заключается в добавлении поверх сети дополнительного полносвязного нейронного слоя с необходимой функцией активации. После этого происходит процесс дообучения модели под конкретную задачу.

Основные этапы работы модели BERT:

- 1) Токенизация входного текста. Добавление первым элементов специального токена [CLS].
- 2) Векторизация полученных токенов.
- 3) Применение позиционного кодирования для полученных векторов.
- 4) Передача векторов на вход стека блоков трансформера.
- 5) Для каждого входного вектора на выход подается результирующий вектор (размерность которого равна 768 в базовой модели BERT) и, в зависимости от задачи, различные результирующие вектора подаются на добавленный «сверху» слой нейронов.

Стоит отметить, что версия DistilBERT имеет ту же общую архитектуру, что и BERT, но в которой произведены некоторые упрощения, а количество слоев было уменьшено в 2 раза [15].

**2.3. Токенизация.** Одним из современных алгоритмов для токенизации текстов на естественном языке является Byte Pair Encoding (BPE) [18]. Основными шагами BPE являются:

1. создание словаря из всех символов языка;
2. представление слов из текста как списка символов;
3. подсчёт количества вхождений каждой пары символов;
4. объединение самых частотных пар в токен и добавление данного токена в словарь;
5. повторение пункта 4 до тех пор, пока не будет получен словарь заданного размера.

У данного алгоритма есть варианты с различными улучшениями, например BPE-Dropout [19].

Также одним из популярных алгоритмов токенизации является WordPiece [20]. WordPiece устроен похожим образом, так же как и BPE, только объединяются не самые частотные пары токенов, а



максимизирующие правдоподобие униграммной языковой модели. Именно данный алгоритм используется в модели BERT.

Применение алгоритмов таких как BPE и WordPieces позволяет бороться с проблемой отсутствия слова в словаре (out of vocabulary). Данная проблема случается, когда модель не знает входящего слова, поскольку его не было в её словаре на этапе обучения. Теперь минимальной частью слова является символ и все символы добавляются в словарь, а это значит, что любое слово сможет быть разбито на последовательность токенов, которые знает модель.

**2.4. Векторизация.** После того как текст был токенизирован, необходимо представить полученные токены в формат наиболее понятный для модели. Чаще всего эта задача решается сопоставлением каждого токена с вектором.

Одним из самых популярных способов для векторизации токенов является word2vec [21]. Word2vec — представляет собой малослойную искусственную нейронную сеть состоящую из двух слоев, которая обрабатывает текст, преобразуя его в «векторизованные» представления. Входными данными для данной сети являются большие корпуса текстов, из которых на выходе получается пространство векторов, размерность которых обычно не превышает несколько сотен (нет проблемы с большой размерностью векторов), где каждый токен в корпусе представлен вектором из сгенерированного пространства. Данные векторы учитывают семантическую близость слов (нет основной проблемы one-hot векторов). Данный способ может быть применён к множеству различных языков в различных задачах [22, 23].

Данная модель обычно обучается выполнять задачу языкового моделирования, т.е. модель пытается угадать одно или несколько слов, пропущенных в тексте. Для word2vec характерно 2 подхода обучения (рисунок 2):

1) Continuous Bag of Words (CBOW) — это метод, в котором модель пытается предсказать целевое слово по словам вокруг него. CBOW обычно хорошо работает на небольших наборах данных.

2) Skip-gram — это метод, в котором модель пытается предсказать слова вокруг данного целевого слова, что в точности противоположно CBOW. Skip-gram лучше работает на больших наборах данных.



Рис. 2. Методы CBOW и Skip-gram для обучения векторов токенов word2vec

Также для улучшения показателей векторов применяется негативный отбор (negative sampling). Данный подход подразумевает включать в процесс обучения пары слов, которые точно не являются соседями. Иначе модель в процессе обучения будет видеть только слова, которые являются соседями.

Также стоит отметить, что модель BERT обучает контекстно-зависимые представления. Это значит, что в зависимости от слов вокруг, слову будет сопоставляться нужный вектор. Данный подход позволяет сопоставлять омонимы с различными векторными представлениями.

**2.5. Блок трансформера.** Каждый блок трансформера состоит из следующих последовательных слоёв (рисунок 3):

- 1) слой многоголового самовнимания (multi-head self-attention);
- 2) слой нормализации;
- 3) слой прямого распространения;
- 4) слой нормализации.

Все блоки трансформера идентичны по структуре, но имеют разные веса.

Изначально на вход блока трансформера идут все вектора токенов предложения. Сам блок трансформера выдаёт точно такое же количество векторов, которое он получил на вход. Следовательно, все слои внутри блока возвращают то же самое количество векторов, которое они получили на входе.



Рис. 3. Структура блока трансформера

Слой прямого распространения необходим для выявления нелинейных зависимостей. Слой многоголового самовнимания использует только линейные функции. Нет смысла передавать результат от одной линейной функции к другой по цепочке, так как добавление новых линейных слоев не позволит выявить нелинейные зависимости в данных. Эта проблема решается добавлением слоя с нелинейными функциями активации между линейными слоями внимания. Обычно применяется функция активации ReLU (1).

$$ReLU(x) = \max(0, x). \quad (1)$$

Само преобразование из слоя прямого распространения применяется отдельно к каждому входному вектору.

На каждый слой нормализации [24] подаётся сумма векторов из результата предыдущего слоя (т.е. слоя многоголового самовнимания или прямого распространения) с вектором, который был до преобразований из предыдущего слоя. Применение слоя нормализации необходимо для уменьшения времени обучения модели, а суммирование векторов до и после слоя с преобразованиями (skip-connection) необходимо для того, чтобы бороться с проблемой затухающих градиентов.

**2.6. Слой самовнимания.** Слой самовнимания кодирует взаимосвязь каждого слова с каждым другим словом в том же самом предложении, сосредотачивая большее внимание на самых значимых словах. А поскольку уделяется внимание предложением самому себе то механизм называется самовниманием. На рисунке 4 показано, как

часть механизма внимания для слова «он» фокусируется на слове «дядя» в предложении «Мой дядя самых честных правил, когда не в шутку занемог, он уважать себя заставил и лучше выдумать не мог».

Вначале слой самовнимания создаёт три матрицы для входного предложения: матрицу запросов (*query*), матрицу ключей (*key*) и матрицу значений (*value*). Эти матрицы создаются с помощью перемножения векторов токенов на три матрицы, которые были обучены во время процесса обучения нейронной сети. Для того чтобы получить выходной результат для данной «головы» самовнимания необходимо провести следующие преобразования (2) над полученными матрицами (рисунок 5):

$$Attention(query, key, value) = softmax\left(\frac{query \cdot key^T}{\sqrt{d_{key}}}\right) \cdot value, \quad (2)$$

где  $d_{key}$  – размерность вектора *key*.



Рис. 4. Демонстрация работы слоя самовнимания (программа для визуализации взята в [25])



Рис. 5. Механизм самовнимания

Большое распространение получила техника использования нескольких параллельных блоков самовнимания называемая многоголовым самовниманием (рисунок 6). Каждая «голова» представляет собой отдельный экземпляр механизма самовнимания (рисунок 5). Применение множества «голов» улучшает производительность слоя самовнимания за счет того, что разное большее количество «голов» позволяет устанавливать больше связей между токенами в предложении.



Рис. 6. Техника многоголового самовнимания

Результаты работы нескольких «голов» внимания конкатенируются друг с другом и затем умножаются на матрицу весов, которая была обучена вместе с моделью. Полученный результат передаётся в слой нормализации.

**2.7. Позиционное кодирование.** У применения механизма самовнимания есть один серьёзный недостаток – данный механизм не учитывает порядок слов в предложении. Часто в естественных языках порядок слов в предложении важен и не может быть отброшен или изменён. Авторы архитектуры трансформер предложили использовать «позиционное кодирование» для решения этой проблемы. Для этого добавляют специальный вектор в каждый входящий вектор токена. Эти векторы имеют определенный шаблон, который запоминает модель и который помогает определить позицию каждого слова или расстояние между разными словами в предложении. Расчёт вектора происходит по следующим формулам (3, 4):

$$PE_{(pos,2i)} = \sin\left(\frac{pos}{10000^{\frac{2i}{d_{model}}}}\right), \quad (3)$$

$$PE_{(pos,2i+1)} = \cos\left(\frac{pos}{10000^{\frac{2i}{d_{model}}}}\right), \quad (4)$$

где,  $pos$  – это позиция токена,  $i$  – это измерение в векторе токена,  $d_{model}$  – это длина вектора токена.

**2.8. Описание подхода к поиску ответа на вопрос в тексте с использованием языковых моделей на базе BERT.** Под поиском ответа на вопрос в тексте будем подразумевать: наличие текста и вопроса к тексту, модель должна выбрать в качестве ответа на данный вопрос непрерывный фрагмент из данного текста. Так как в постановке задачи ответ является непрерывным отрывком из текста, то он может быть однозначно задан позициями начала и конца. Получается, что модель должна ответить на два вопроса (5) для каждого вектора токена: «Является ли данный токен из текста началом ответа на заданный вопрос?», «Является ли данный токен из текста концом ответа на заданный вопрос?» (рисунок 7):

$$Model(\theta, w_{start}, w_{end}) = -\sum_k \left( \log p_{start}(y_k^{start} | C, Q; \theta, w_{start}) + \log p_{end}(y_k^{end} | C, Q; \theta, w_{end}) \right), \quad (5)$$

где  $C$  - текст,  $Q$  - вопрос,  $\theta$  - параметры языковой модели,  $w_{start}$  и  $w_{end}$  - обучаемые параметры для предсказания позиции начала и конца ответа на вопрос,  $y_k^{start}$  и  $y_k^{end}$  - позиции начала и конца ответа на вопрос для примера с индексом  $k$ . Вероятности  $p_{start}$  и  $p_{end}$  при использовании модели BERT определяются следующим образом (6, 7):

$$p^{start}(y) = softmax(\left[ \left\{ w_{start}, BERT_0^N \right\}, \left\{ w_{start}, BERT_1^N \right\}, \dots, \left\{ w_{start}, BERT_L^N \right\} \right]), \quad (6)$$

$$p^{end}(y) = softmax(\left[ \left\{ w_{end}, BERT_0^N \right\}, \left\{ w_{end}, BERT_1^N \right\}, \dots, \left\{ w_{end}, BERT_L^N \right\} \right]), \quad (7)$$

где  $softmax$  – функция софтмакс,  $L$  — число токенов во входной последовательности,  $\{ \cdot, \cdot \}$  — скалярное произведение,  $w_{start}$  и  $w_{end}$  - обучаемые параметры для предсказания позиции начала и конца ответа на вопрос,  $BERT_i^N$  — выход с последнего слоя  $N$  для  $i$ -ого токена во входной последовательности.

Вопрос и текст перед передачей в BERT разделяются специальным токеном [SEP].

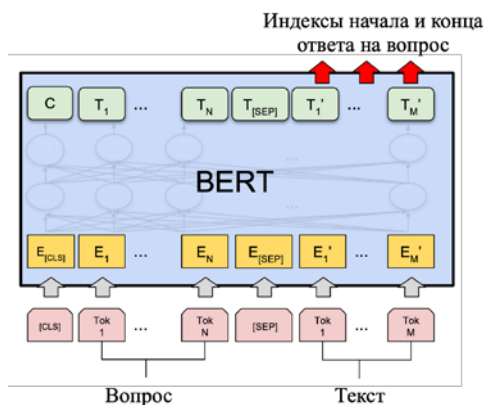


Рис. 7. Использование модели BERT для нахождения ответа на вопрос в тексте (рисунок взят из работы [8])

**2.9. Дистилляция.** Несмотря на то, что современные модели показывают внушительные результаты в основных задачах обработки естественного языка, данные модели часто имеют один недостаток – свой размер. Большое количество параметров модели может сильно ограничить пропускную способность конвейера, на котором используется данная модель, кроме того, большая размерность не позволяет эффективно использовать данные модели на мобильных устройствах. Поэтому одним из важных направлений исследований в области естественного языка является исследование подходов к уменьшению размера модели при сохранении той же точности работы. Одним из набирающих популярность способов уменьшения размерности моделей является дистилляция знаний.

Дистилляция знаний заключается в том, что вначале обучается большая модель, затем на основе предсказаний данной модели-учителя обучается более легковесная модель-ученик [14]. Данный подход показал хорошие результаты (полученные в ходе дистилляции модели могут совсем незначительно уступать в результатах), из-за чего появилось большое количество дистиллированных версий популярных сетей (таких как BERT, GPT). Данные дистиллированные версии содержат упрощения архитектур и меньшее количество параметров, затем данные версии моделей обучаются на основе своих больших собратьев. Также полученные модели можно дообучать под необходимую задачу.

Для задач классификации процесс дистилляции заключается в том, что модели ученика и учителя вычисляют свои распределения вероятностей по классам для конкретного экземпляра данных, а затем необходимо найти различие между этими вероятностями при помощи расхождения Кульбака-Лейблера. Полученное расхождение используется для расчета значения ошибки на текущей итерации работы сети.

### 3. Проведение опытов

**3.1. Показатели.** Основными показателями качества работы вопросно-ответных моделей являются EM и F-мера.

Exact match (EM) — точное совпадение, доля ответов системы, которые полностью совпадают с одним из правильных ответов с точностью до пунктуации и регистра.

F-мера — представляет собой совместную оценку полноты и точности. Данный показатель вычисляется по следующей формуле (8):

$$F = 2 \cdot \frac{\text{Точность} \cdot \text{Полнота}}{\text{Точность} + \text{Полнота}}, \quad (8)$$

Полнота (recall) вычисляется по следующей формуле (9):

$$\text{Полнота} = \frac{TP}{TP + FN}. \quad (9)$$

Точность (precision) вычисляется по следующей формуле (10):

$$\text{Точность} = \frac{TP}{TP + FP}, \quad (10)$$

где TP – Истинноположительные предсказания модели (модель правильно отнесла объект к классу), TN - Истинноотрицательные предсказания модели (модель правильно не отнесла объект к классу),



FN – Ложноотрицательные предсказания модели (модель неправильно не отнесла объект к классу), FP – Ложноположительные предсказания модели (модель неправильно отнесла объект к классу).

**3.2. Выбор модели нейронной сети.** Основной библиотекой для работы с моделями для обработки естественного языка является Hugging face. В данной библиотеке можно найти огромное количество разнообразных моделей с архитектурой «трансформер» и производными от неё. На момент написания статьи число доступных моделей в библиотеке составляет 22022. Также в данной библиотеке можно найти большое количество наборов данных. На момент написания статьи число наборов данных составляет 2059. Данная библиотека часто используется в крупных компаниях. У себя внутри библиотека

Hugging face полагается на такие библиотеки как Jax, PyTorch и TensorFlow, которые являются стандартом для всей индустрии машинного обучения. Также на сайте библиотеки в онлайн можно проверить результаты работы большинства моделей (без непосредственной установки библиотеки на персональный компьютер).

На момент написания статьи среди русскоязычных моделей для дистилляции было всего 6:

- 1) DeepPavlov/distilrubert-base-cased-conversational<sup>1</sup>;
- 2) DeepPavlov/distilrubert-tiny-cased-conversational<sup>2</sup>;
- 3) Geotrend/distilbert-base-ru-cased<sup>3</sup>;
- 4) Geotrend/distilbert-base-en-ru-cased<sup>4</sup>;
- 5) Geotrend/distilbert-base-en-el-ru-cased<sup>5</sup>;
- 6) Geotrend/distilbert-base-en-fr-nl-ru-ar-cased<sup>6</sup>.

Первые две не подошли к решаемой задаче, поскольку, по словам авторов, данные модели стоит использовать, только если данные решаемой задачи имеют разговорную структуру или содержат неформальный язык. Модели с 3 по 6 были основаны на мультиязычной модели distilbert-base-multilingual-cased (содержит  $135 \times 10^6$  параметров), которая в свою очередь представляет дистиллированную версию bert-base-multilingual-cased (содержит  $177 \times 10^6$  параметров). Данные модели были получены путём

<sup>1</sup> <https://huggingface.co/DeepPavlov/distilrubert-base-cased-conversational>

<sup>2</sup> <https://huggingface.co/DeepPavlov/distilrubert-tiny-cased-conversational>

<sup>3</sup> <https://huggingface.co/Geotrend/distilbert-base-ru-cased>

<sup>4</sup> <https://huggingface.co/Geotrend/distilbert-base-en-ru-cased>

<sup>5</sup> <https://huggingface.co/Geotrend/distilbert-base-en-el-ru-cased>

<sup>6</sup> <https://huggingface.co/Geotrend/distilbert-base-en-fr-nl-ru-ar-cased>

уменьшения поддержки 108 языков до поддержки фиксированного набора языков [26]. Было принято решение использовать модель под номером 3 поскольку она должна поддерживать только русский язык, т.е. иметь меньшее количество параметров по сравнению с остальными версиями ( $54 \times 10^6$  параметров в Geotrend/distilbert-base-ru-cased, против  $72 \times 10^6$  в Geotrend/distilbert-base-en-ru-cased).

В качестве модели-учителя была выбрана ruBert-base поскольку среди дискриминационных моделей она показала лучшие результаты (таблица 1).

**3.3. Описание стенда для обучения нейронной сети.** Обучение сети проводилось на сервисе Google Colab [27]. Данный сервис предоставляет в бесплатное пользование компьютеры с производительными видеокартами (уровня NVIDIA Tesla K80) на ограниченное время (12 часов). Также доступ может быть прекращен раньше из-за простоя компьютера. Взаимодействие с компьютерами осуществляется через программу Jupyter notebook, которая поддерживает язык Python.

**3.4. Обучение нейронной сети без применения процесса дистилляции.** Процесс обычного обучения сети представлял собой следующие этапы:

- 1) Загрузку модели Geotrend/distilbert-base-ru-cased.
- 2) Загрузку набора данных SberQuAD.
- 3) Проведение тренировки модели на наборе данных со следующими параметрами обучения:
  - Шаг обучения (learning rate):  $2 \times 10^{-5}$ ;
  - Размер блока: 16;
  - Длины входных последовательностей: 384;
  - Количество эпох обучения: 3.
- 4) Проверка полученной модели на валидационном наборе данных.
- 5) Расчёт показателей EM и F-меры.
- 6) Время обучения модели заняло 3 часа 27 минут.

**3.5. Обучение нейронной сети с применением процесса дистилляции.** Обучение дистиллированной модели включало в себя следующие шаги:

- 1) Загрузку обученной ранее модели ruBert-base.
- 2) Загрузку модели Geotrend/distilbert-base-ru-cased.
- 3) Загрузку набора данных SberQuAD.
- 4) Проведение дистилляции модели Geotrend/distilbert-base-ru-cased на основе результатов, полученных из обученной ранее модели ruBert-base (2 эпохи обучения).

5) Проверка полученной модели на валидационном наборе данных.

6) Расчёт показателей EM и F-меры.

Время обучения модели заняло 4 часа 15 минут.

**4. Результаты и их обсуждение.** В результате обучения модели без дистилляции были получены следующие показатели на валидационном наборе данных: EM = 55,65 и F-мера = 76,51.

В результате обучения модели с дистилляцией были получены следующие показатели на валидационном наборе данных: EM = 58,57 и F-мера = 78,42.

Для сравнения были взяты модели, которые показали лучшие результаты на наборе данных SberQuAD. Данные модели представляют не только кодировщики, но еще декодеры и полные трансформеры. Стоит отметить, что для сравнения не использовались другие дистиллированные версии моделей. Результаты сравнения моделей представлены в таблице 1.

Таблица 1. Результаты сравнения моделей

Название модели	EM	F-мера	Количество параметров в модели
ruBERT-base	66,83	84,95	178×10 <sup>6</sup>
ruBERT-large	67,25	85,25	427×10 <sup>6</sup>
ruRoBERTa-large	65,23	85,45	355×10 <sup>6</sup>
ruT5-base	66,26	84,56	222×10 <sup>6</sup>
ruT5-large	68,57	86,73	737×10 <sup>6</sup>
<b>ruGPT-3-medium</b>	<b>57,60</b>	<b>77,73</b>	<b>356×10<sup>6</sup></b>
<b>ruGPT-3-large</b>	<b>59,57</b>	<b>79,51</b>	<b>760×10<sup>6</sup></b>
mT5-base (Google)	64,03	83,40	390×10 <sup>6</sup>
mT5-large (Google)	69,63	87,06	973×10 <sup>6</sup>
<b>Модель, обученная без применения дистилляции</b>	<b>55,65</b>	<b>76,51</b>	<b>54×10<sup>6</sup></b>
<b>Модель, обученная с применением дистилляции</b>	<b>58,57</b>	<b>78,42</b>	<b>54×10<sup>6</sup></b>

Модель, обученная с применением дистилляции, смогла опередить модель, которая была обучена обычным способом (EM 58,57 и F-мера 78,42 против EM 55,65 и F-мера 76,51). Также она смогла опередить генеративную модель ruGPT-3-medium (EM 58,57 и F-мера 78,42 против EM 57,60 и F-мера 77,73). Модель ruGPT-3-medium имеет 356 миллионов параметров, а модель DistilBERT имеет всего 54 миллиона параметров. Но полученная модель имеет большее отставание от более крупных

дискриминационных моделей, например ruBERT (EM 66,83, F-мера 84,95), которая имеет 178 миллионов параметров.

Текст	Сверхкороткие импульсы лазерного излучения используются в лазерной химии для запуска и анализа химических реакций. Здесь лазерное излучение позволяет обеспечить точную локализацию, дозированность, абсолютную стерильность и высокую скорость ввода энергии в систему. В настоящее время разрабатываются различные системы лазерного охлаждения, рассматриваются возможности осуществления с помощью лазеров управляемого термоядерного синтеза. Лазеры используются и в военных целях, например, в качестве средств наведения и прицеливания. Рассматриваются варианты создания на основе мощных лазеров боевых систем защиты воздушного, морского и наземного базирования.'
Вопрос	Что используются и в военных целях, например, в качестве средств наведения и прицеливания?
Правильный ответ	Лазеры
Ответ модели	Лазеры

Рис. 8. Демонстрация работы модели

На рисунке 8 представлен один вопросный набор из валидационной части набора данных SberQUAD и форматированный результат работы полученной модели. Оригинальный формат ответа модели представлен на рисунке 9.

```
{
  'score': 0.870181143283844,
  'start': 435,
  'end': 441,
  'answer': 'Лазеры'
}
```

Рис. 9. Формат ответа обученной модели

**5. Заключение.** Обученная в ходе дистилляции модель DistilBERT (EM 58,57 и F-мера 78,42) смогла опередить результаты более крупной генеративной сети ruGPT-3-medium (EM 57,60 и F-мера 77,73) притом, что ruGPT-3-medium имеет в 6,5 раз больше параметров. К сожалению, полученная модель сильнее отстает от более крупной дискриминационной модели ruBERT (EM 66,83, F-мера 84,95), которая имеет в 3,2 раза больше параметров.

Учитывая большое количество готовых полных моделей следует изучить, какие результаты может показать ансамбль данных моделей [28]. Дальнейшая дистилляция полученного ансамбля может позволить получить модель, которая будет иметь значительно меньшее количество параметров и превосходить результаты моделей-учителей,

работающих по отдельности. Также можно произвести дообучение/дистилляцию на более крупных и более мелких моделях.

### Литература

1. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., Polosukhin, I. Attention is all you need // *Advances in Neural Information Processing Systems* 30. 2017. pp. 5998-6008.
2. Yang Z., Keung J., Yu X., Gu X., Wei Z., Ma X., Zhang M. A Multi-Modal Transformer-based Code Summarization Approach for Smart Contracts // *The 2021 International Conference on Program Comprehension*. 2021. pp. 1-12.
3. Juraska J., Walker M. Attention Is Indeed All You Need: Semantically Attention-Guided Decoding for Data-to-Text NLG // *Proceedings of the 14th International Conference on Natural Language Generation*. 2021. pp. 416-431.
4. Lewis M., Liu Y., Goyal N., Ghazvininejad M., Mohamed A., Levy O., Stoyanov V., Zettlemoyer L. BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension // *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. 2020. pp. 7871-7880.
5. Raffel C., Shazeer N., Roberts A., Lee K., Narang S., Matena M., Zhou Y., Li W., Liu P.J. Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer // *Journal of Machine Learning Research*, Volume 21. 2020. pp. 1-67.
6. Zhang J., Zhao Y., Saleh M., Liu P. J. PEGASUS: Pre-training with Extracted Gap-sentences for Abstractive Summarization // *Proceedings of the 37th International Conference on Machine Learning*. 2020. pp. 11328-11339.
7. Qi W., Yan Y., Gong Y., Liu D., Duan N., Chen J., Zhang R., Zhou M. ProphetNet: Predicting Future N-gram for Sequence-to-Sequence Pre-training // *Findings of the Association for Computational Linguistics: EMNLP 2020*. 2020. pp. 2401-2410.
8. Devlin J., Chang M., Lee K., Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding // *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. 2019. pp. 4171-4186.
9. Lan Z., Chen M., Goodman S., Gimpel K., Sharma P., Soricut R. ALBERT: A Lite BERT for Self-supervised Learning of Language Representations. ArXiv. 2019. URL: <https://arxiv.org/abs/1909.11942> (дата обращения: 12.11.2021).
10. Liu Y., Ott M., Goyal N., Du J., Joshi M., Chen D., Levy O., Lewis M., Zettlemoyer L., Stoyanov V. RoBERTa: A Robustly Optimized BERT Pretraining Approach. ArXiv. 2019. URL: <https://arxiv.org/abs/1907.11692> (дата обращения: 12.11.2021).
11. Brow T. B., Mann B., Ryder N., Subbiah M., Kaplan J., Dhariwal P., Neelakantan A., Shyam P., Sastry G., Askell A., Agarwal S., Herbert-Voss A., Krueger G., Henighan T., Child R., Ramesh A., Ziegler D. M., Wu J., Winter C., Hesse C., Chen M., Sigler E., Litwin M., Gray S., Chess B., Clark J., Berner C., McCandlish S., Radford A., Sutskever I., Amodei D. Language Models are Few-Shot Learners // *Advances in Neural Information Processing Systems* 33 (NeurIPS 2020). 2020. pp. 1877-1901.
12. Keskar N.S., McCann B., Varshney L. R., Xiong C., Socher R. CTRL: A Conditional Transformer Language Model for Controllable Generation. ArXiv. 2019. URL: <https://arxiv.org/abs/1909.05858> (дата обращения: 12.11.2021).
13. Dai Z., Yang Z., Yang Y., Carbonell J., Le Q.V., Salakhutdinov R. Transformer-XL: Attentive Language Models Beyond a Fixed-Length Context // *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 2019. pp. 2978-2988.

14. Hahn S., Choi H. Self-Knowledge Distillation in Natural Language Processing // Proceedings of the International Conference on Recent Advances in Natural Language Processing, Varna, Bulgaria, September 2-4, 2019. 2019. pp. 423-430.
15. Sanh V., Debut L., Chaumond J., Wolf T. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. ArXiv. 2019. URL: <https://arxiv.org/abs/1910.01108> (дата обращения: 12.11.2021).
16. Rajpurkar P., Zhang J., Lopyrev K., Liang P. SQuAD: 100,000+ Questions for Machine Comprehension of Text // Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. 2016. pp. 2383–2392.
17. Efimov P., Chertok A., Boytsov L., Braslavski P. SberQuAD - Russian Reading Comprehension Dataset: Description and Analysis // Experimental IR Meets Multilinguality, Multimodality, and Interaction - 11th International Conference of the CLEF Association, CLEF 2020, Thessaloniki, Greece, September 22-25, 2020, Proceedings. 2020. pp. 3-15.
18. Sennrich R., Haddow B., Birch A. Neural Machine Translation of Rare Words with Subword Units // Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2016. pp. 1715–1725.
19. Provilkov I., Emelianenko D., Voita E. BPE-Dropout: Simple and Effective Subword Regularization. ArXiv. 2020 // Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020. pp. 1882–1892.
20. Schuster M., Nakajima K. Japanese and Korea voice search // 2012 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2012, Kyoto, Japan, March 25-30, 2012. 2012. pp. 5149-5152.
21. Mikolov T., Chen K., Corrado G., Dean J. Efficient Estimation of Word Representations in Vector Space. ArXiv. 2013. URL: <https://arxiv.org/pdf/1301.3781.pdf> (дата обращения: 12.11.2021).
22. Фат Х.Н., Ань Н.Т.М. Алгоритм классификации вьетнамского текста с использованием долгой краткосрочной памяти и Word2Vec // Информатика и автоматизация. 2020. № 6 (19). С. 1255-1279.
23. Алтаф С., Iqbal S., Soomro M.W. Эффективный алгоритм классификации естественного языка обнаружения повторяющихся контролируемых признаков // Информатика и автоматизация. 2021. № 3 (20). С. 623-653.
24. Lei Ba J., Kiros J.R., Hinton G.E. Layer Normalization. ArXiv. 2016. URL: <https://arxiv.org/pdf/1607.06450.pdf> (дата обращения: 12.11.2021).
25. URL: <https://github.com/jessevig/bertviz> (дата обращения: 12.11.2021).
26. Abdaoui A., Pradel C., Sigel G. Load What You Need: Smaller Versions of Multilingual BERT. ArXiv. 2020. URL: <https://arxiv.org/abs/2010.05609> (дата обращения: 12.11.2021).
27. URL: <https://colab.research.google.com/notebooks/welcome.ipynb?hl=ru> (дата обращения: 12.11.2021).
28. Li S., Li R., Peng V. Ensemble ALBERT on SQuAD 2.0. ArXiv. 2021. URL: <https://arxiv.org/abs/2110.09665> (дата обращения: 12.11.2021).

**Галеев Денис Талгатович** — аспирант, ФГБОУ ВО Юго-Западный государственный университет (ЮЗГУ). Область научных интересов: искусственный интеллект, машинное обучение, обработка естественного языка. Число научных публикаций — 15. [ga3wvw@mail.ru](mailto:ga3wvw@mail.ru); улица 50 лет Октября, 94, 305040, Курск, Россия; р.т.: +7(4712)222-665.

**Панищев Владимир Славиевич** — канд. техн. наук, доцент, кафедра вычислительной техники, ФГБОУ ВО Юго-Западный государственный университет (ЮЗГУ). Область научных интересов: нейронные сети, обработка изображений. Число научных публикаций — 150. [gskunk@yandex.ru](mailto:gskunk@yandex.ru); улица 50 лет Октября, 94, 305040, Курск, Россия; р.т.: +7(4712)222626.

D. GALEEV, V. PANISHCHEV  
**EXPERIMENTAL STUDY OF LANGUAGE MODELS OF  
"TRANSFORMER" IN THE PROBLEM OF FINDING THE  
ANSWER TO A QUESTION IN A RUSSIAN-LANGUAGE TEXT**

*Galeev D., Panishchev V. Experimental Study of Language Models of "Transformer" in the Problem of Finding the Answer to a Question in a Russian-Language Text.*

**Abstract.** The aim of the study is to obtain a more lightweight language model that is comparable in terms of EM and F1 with the best modern language models in the task of finding the answer to a question in a text in Russian. The results of the work can be used in various question-and-answer systems for which response time is important. Since the lighter model has fewer parameters than the original one, it can be used on less powerful computing devices, including mobile devices. In this paper, methods of natural language processing, machine learning, and the theory of artificial neural networks are used. The neural network is configured and trained using the Torch and Hugging face machine learning libraries. In the work, the DistilBERT model was trained on the SberQUAD dataset with and without distillation. The work of the received models is compared. The distilled DistilBERT model (EM 58,57 and F1 78,42) was able to outperform the results of the larger ruGPT-3-medium generative network (EM 57,60 and F1 77,73), despite the fact that ruGPT-3-medium had 6,5 times more parameters. The model also showed better EM and F1 metrics than the same model, but to which only conventional training without distillation was applied (EM 55,65, F1 76,51). Unfortunately, the resulting model lags further behind the larger robert discriminative model (EM 66,83, F1 84,95), which has 3,2 times more parameters. The application of the DistilBERT model in question-and-answer systems in Russian is substantiated. Directions for further research are proposed.

**Keywords:** machine learning, deep learning, neural networks, natural language processing, transformer.

**Galeev Denis** — Ph.D., Graduate student, Southwest State University (SWSU). Research interests: artificial intelligence, machine learning, natural language processing. The number of publications — 15. ra3www@mail.ru; 94, 50 let Oktyabrya St., 305040, Kursk, Russia; office phone: +7(4712)222-665.

**Panishchev Vladimir** — Ph.D., Associate professor, Department of computer engineering, Southwest State University (SWSU). Research interests: neural networks, image processing. The number of publications — 150. gskunk@yandex.ru; 94, 50 let Oktyabrya St., 305040, Kursk, Russia; office phone: +7(4712)222626.

## References

1. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., Polosukhin, I. Attention is all you need // *Advances in Neural Information Processing Systems* 30. 2017. pp. 5998-6008.
2. Yang Z., Keung J., Yu X., Gu X., Wei Z., Ma X., Zhang M. A Multi-Modal Transformer-based Code Summarization Approach for Smart Contracts // *The 2021 International Conference on Program Comprehension*. 2021. pp. 1-12.
3. Juraska J., Walker M. Attention Is Indeed All You Need: Semantically Attention-Guided Decoding for Data-to-Text NLG // *Proceedings of the 14th International Conference on Natural Language Generation*. 2021. pp. 416-431.

4. Lewis M., Liu Y., Goyal N., Ghazvininejad M., Mohamed A., Levy O., Stoyanov V., Zettlemoyer L. BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension // Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020. pp. 7871-7880.
5. Raffel C., Shazeer N., Roberts A., Lee K., Narang S., Matena M., Zhou Y., Li W., Liu P.J. Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer // Journal of Machine Learning Research, Volume 21. 2020. pp. 1-67.
6. Zhang J., Zhao Y., Saleh M., Liu P. J. PEGASUS: Pre-training with Extracted Gap-sentences for Abstractive Summarization // Proceedings of the 37th International Conference on Machine Learning. 2020. pp. 11328-11339.
7. Qi W., Yan Y., Gong Y., Liu D., Duan N., Chen J., Zhang R., Zhou M. ProphetNet: Predicting Future N-gram for Sequence-to-Sequence Pre-training // Findings of the Association for Computational Linguistics: EMNLP 2020. 2020. pp. 2401-2410.
8. Devlin J., Chang M., Lee K., Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding // Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers). 2019. pp. 4171-4186.
9. Lan Z., Chen M., Goodman S., Gimpel K., Sharma P., Soricut R. ALBERT: A Lite BERT for Self-supervised Learning of Language Representations. ArXiv. 2019. URL: <https://arxiv.org/abs/1909.11942> (accessed: 12.11.2021).
10. Liu Y., Ott M., Goyal N., Du J., Joshi M., Chen D., Levy O., Lewis M., Zettlemoyer L., Stoyanov V. RoBERTa: A Robustly Optimized BERT Pretraining Approach. ArXiv. 2019. URL: <https://arxiv.org/abs/1907.11692> (accessed: 12.11.2021).
11. Brow T. B., Mann B., Ryder N., Subbiah M., Kaplan J., Dhariwal P., Neelakantan A., Shyam P., Sastry G., Askell A., Agarwal S., Herbert-Voss A., Krueger G., Henighan T., Child R., Ramesh A., Ziegler D. M., Wu J., Winter C., Hesse C., Chen M., Sigler E., Litwin M., Gray S., Chess B., Clark J., Berner C., McCandlish S., Radford A., Sutskever I., Amodei D. Language Models are Few-Shot Learners // Advances in Neural Information Processing Systems 33 (NeurIPS 2020). 2020. pp. 1877-1901.
12. Keskar N.S., McCann B., Varshney L. R., Xiong C., Socher R. CTRL: A Conditional Transformer Language Model for Controllable Generation. ArXiv. 2019. URL: <https://arxiv.org/abs/1909.05858> (accessed: 12.11.2021).
13. Dai Z., Yang Z., Yang Y., Carbonell J., Le Q.V., Salakhutdinov R. Transformer-XL: Attentive Language Models Beyond a Fixed-Length Context // Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. 2019. pp. 2978-2988.
14. Hahn S., Choi H. Self-Knowledge Distillation in Natural Language Processing // Proceedings of the International Conference on Recent Advances in Natural Language Processing, Varna, Bulgaria, September 2-4, 2019. 2019. pp.423-430.
15. Sanh V., Debut L., Chaumond J., Wolf T. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. ArXiv. 2019. URL: <https://arxiv.org/abs/1910.01108> (accessed: 12.11.2021).
16. Rajpurkar P., Zhang J., Lopyrev K., Liang P. SQuAD: 100,000+ Questions for Machine Comprehension of Text // Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing. 2016. pp. 2383-2392.
17. Efimov P., Chertok A., Boytsov L., Braslavski P. SberQuAD - Russian Reading Comprehension Dataset: Description and Analysis // Experimental IR Meets Multilinguality, Multimodality, and Interaction - 11th International Conference of the CLEF Association, CLEF 2020, Thessaloniki, Greece, September 22-25, 2020, Proceedings. 2020. pp. 3-15.



18. Sennrich R., Haddow B., Birch A. Neural Machine Translation of Rare Words with Subword Units // Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2016. pp. 1715–1725.
19. Provilkov I., Emelianenko D., Voita E. BPE-Dropout: Simple and Effective Subword Regularization. ArXiv. 2020 // Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. 2020. pp. 1882–1892.
20. Schuster M., Nakajima K. Japanese and Korea voice search // 2012 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2012, Kyoto, Japan, March 25-30, 2012. 2012. pp. 5149-5152.
21. Mikolov T., Chen K., Corrado G., Dean J. Efficient Estimation of Word Representations in Vector Space. ArXiv. 2013 URL: <https://arxiv.org/pdf/1301.3781.pdf> (accessed: 12.11.2021).
22. Phat H.N., Anh N.T.M. Vietnamese Text Classification Algorithm using Long Short Term Memory and Word2Vec // Informatics and Automation. 2020. № 6 (19). pp. 1255-1279.
23. Altaf S., Iqbal S., Soomro M.W. Efficient natural language classification algorithm for detecting duplicate unsupervised features // Informatics and Automation. 2021. № 3 (20). pp. 623-653.
24. Lei Ba J., Kiros J.R., Hinton G.E. Layer Normalization. ArXiv. 2016. URL: <https://arxiv.org/pdf/1607.06450.pdf> (accessed: 12.11.2021).
25. URL: <https://github.com/jessevig/bertviz> (accessed: 12.11.2021).
26. Abdaoui A., Pradel C., Sigel G. Load What You Need: Smaller Versions of Multilingual BERT. ArXiv. 2020. URL: <https://arxiv.org/abs/2010.05609> (accessed: 12.11.2021).
27. URL: <https://colab.research.google.com/notebooks/welcome.ipynb?hl=ru> (accessed: 12.11.2021).
28. Li S., Li R., Peng V. Ensemble ALBERT on SQuAD 2.0. ArXiv. 2021. URL: <https://arxiv.org/abs/2110.09665> (accessed: 12.11.2021).

В.И. ПИМЕНОВ, И.В. ПИМЕНОВ  
**АНАЛИЗ И ВИЗУАЛИЗАЦИЯ ДАННЫХ В ЗАДАЧАХ  
МНОГОКРИТЕРИАЛЬНОЙ ОПТИМИЗАЦИИ ПРОЕКТНЫХ  
РЕШЕНИЙ**

*Пименов В.И., Пименов И.В. Анализ и визуализация данных в задачах многокритериальной оптимизации проектных решений.*

**Аннотация.** Накопление данных о процессах управления проектами и типовых решениях сделало актуальными исследования, связанные с применением методов инженерии знаний для многокритериального поиска вариантов, которые задают оптимальные настройки параметров проектной среды. Цель: разработка методики поиска и визуализации групп проектов, которые могут быть оценены на основе концепции доминирования и интерпретироваться в терминах проектных переменных и показателей эффективности. Методы: обогащение выборки с сохранением неязной связи между проектными переменными и показателями эффективности осуществляется с помощью прогнозирующей нейросетевой модели. Для обнаружения фронта Парето в многомерном критериальном пространстве используется набор генетических алгоритмов. Онтология проектов определяется после кластеризации вариантов в пространстве решений и преобразования кластерной структуры в критериальное пространство. Автоматизация поиска в многомерном пространстве зоны наибольшей кривизны фронта Парето, определяющей равновесные проектные решения, их визуализация и интерпретация осуществляются с помощью плоского дерева решений. Результаты: плоское дерево строится при любой размерности критериального пространства и имеет структуру, которая имеет топологическое соответствие с проекциями разделяемых образов кластеров из многомерного пространства на плоскость. Для различных видов преобразований и корреляций между показателями эффективности и проектными переменными показано, что участки наибольшей кривизны фронта Парето определяются либо содержанием целого кластера, либо частью вариантов, представляющих “лучший” кластер. Если на плоском дереве к правому верхнему углу примыкает неразделенный прямоугольный кластер, то его представители в критериальном пространстве хорошо отделены от остальных кластеров и при максимизации показателей эффективности наиболее приближены к идеальной точке. Все представители такого кластера являются эффективными решениями. Если кластер-победитель содержит внутри дерева решений доминируемые варианты, то “лучший” кластер представляют оставшиеся варианты, которые задают оптимальные настройки проектных переменных. Практическая значимость: предложенная методика поиска и визуализации групп проектов может найти применение при выборе условий ресурсного и организационно-экономического моделирования проектной среды, обеспечивающих оптимизацию рисков, стоимостных, функциональных и временных критериев.

**Ключевые слова:** управление проектом, многокритериальная классификация, нейронная сеть, генетический алгоритм, фронт Парето, кластерная структура, визуализация, плоское дерево решений.

**1. Введение.** Отличительное свойство, которое выделяет проект среди существующих экономических, промышленных, коммерческих, информационных процессов, – новизна целей и условий реализации. В основе управления проектами лежит традиционный треугольник

контроля качества, времени и затрат [1], что, однако, не исчерпывает всей многомерности проекта как объекта управления [2]. Агрегированное описание  $X$  проектной среды разделяется на две основные группы: группу ресурсного моделирования и группу организационно-экономического моделирования [3]. Кроме того, оценивать эффективность проекта требуется на разных стадиях жизненного цикла [4].

Технологические и технические ресурсы определяют условия реализации производственного процесса: условия труда, время производственного цикла, использование основных и оборотных средств, состояние научно-технического развития и удовлетворение потребностей потребительского рынка. Человеческие ресурсы характеризуют обеспечение производственного процесса требуемой численностью работников соответствующей специальности и квалификации, нормой рационального использования кадров, организацией системы оплаты труда, наличием системы стимулирования и мотивации работников. Информационные ресурсы связаны с необходимостью получения достоверных данных и их обработкой с использованием современных программных средств. Пространственно-временные ресурсы определяются географическими, геологическими, геометрическими и временными переменными.

Группа показателей себестоимости и ценообразования определяется с учетом производственных и рыночных факторов, а также в соответствии с требованиями налоговой системы. Организационные факторы включают в себя все аспекты управления структурой организации, производственным процессом, а также финансовыми и правовыми механизмами. Финансовые показатели раскрывают источники капитала для реализации проекта, а также возможности по предоставлению этого капитала. Показатели экономической эффективности определяют целесообразность инвестирования на каждом этапе реализации проекта на основе нормативных, статистических, методических и экспертных позиций.

Для оценки эффективности проекта обычно выбираются стоимостные, функциональные и временные показатели  $Y$ . Управление проектом сводится к организации и направлению ресурсов таким образом, чтобы проект был завершен в рамках бюджетных затрат, достижению заданного уровня качества продукта или услуги, функциональности программно-технической системы, а также к минимизации рисков и потерь. Социальная эффективность проявляется в улучшении среды проживания, доходов и качества жизни населения.

Процесс формирования проектного решения представлен на рисунке 1.

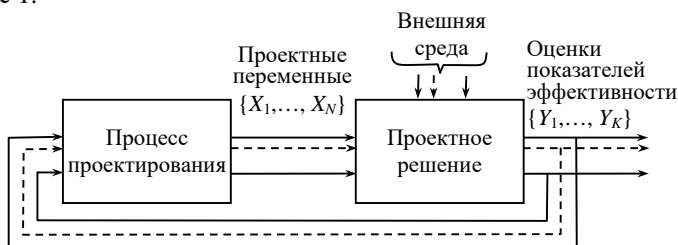


Рис. 1. Формирование проектного решения

На начальном этапе, после рассмотрения конкретного содержания проекта, его причинно-следственных связей, определяются существенные входные переменные  $\{X_1, \dots, X_N\}$ , оказывающие влияние на состояние проекта, а также набор показателей  $\{Y_1, \dots, Y_K\}$ , по которым оценивается его эффективность. Примеры содержательного анализа различных прикладных проектов и критериального аппарата, предоставляющего возможность объективного сравнения вариантов их реализации, приведены в [1, 3–10].

Процессы проектирования опираются на методики и инструкции, определяющие диапазоны варьируемых проектных переменных и показателей конфигурации, которые образуют допустимую область решений  $\mathbf{X} \in \mathbf{X}$ .

Воздействия внешней среды на проект ограничивают проектные процессы под влиянием социальных норм, культуры производства, военно-политической обстановки, рисков, инфляции и пр.

Показатели эффективности проектного решения являются теми выходными характеристиками, которые позволяют классифицировать и интерпретировать проектное решение в критериальном пространстве  $\mathbf{Y}$ .

Ключевым этапом при формировании проектного решения является построение модели преобразования  $\mathbf{X} \rightarrow \mathbf{Y}$ . Его выполнение является трудноформализуемой задачей по ряду признаков: принципиальной невозможности точного описания всех взаимосвязей (часто нелинейных) между проектными переменными и показателями эффективности, большой размерности пространства  $\mathbf{X} \times \mathbf{Y}$ , динамически изменяющимся информационным потокам на разных стадиях жизненного цикла, ограниченным финансовым ресурсам.

Более того, при наличии альтернативных проектных решений (вариантов реализации проекта), многомерная модель должна предоставлять возможность семантической интерпретации предлагаемых вариантов, формулировать те законы и правила, по которым происходит отнесение проектного решения к определенному уровню эффективности, визуализировать варианты реализации с сохранением топологического соответствия между их образами в многомерных пространствах.

Объектом исследования является процесс моделирования преобразования  $X \rightarrow Y$  при формировании проектного решения. Предметом исследования являются методы анализа вариантов проектных решений, их визуализации и интерпретации.

Предполагается, что поиск, оценка и визуализация групп проектов, как сложных многомерных объектов, могут быть осуществлены на основе концепции доминирования и методов машинного обучения, в рамках которых решаются задачи автоматической классификации вариантов, построения модели взаимосвязей  $X \rightarrow Y$  и визуализации кластерной структуры с упорядочением групп проектов по уровню эффективности.

## **2. Оценивание показателей эффективности проекта.**

Расчетные методики оценки эффективности научно-технического, организационного, экономического, социального или ИТ-проекта базируются на соотношении затрат (включающих связанные с резервированием и страхованием, а также санкций) и результатов. Расчет статических и динамических показателей  $Y$  включает в себя выполнение требуемой нормы доходности проекта, анализ структуры цены и капитала, оценивание значимости рисков, учет инфляции и пр. [11]. Оценки эффективности инвестиций в ИТ-проекты, связанные с электронной формой предоставляемых товаров и услуг, могут быть получены через положительный эффект, достигаемый производственной надсистемой, в интересах которой внедряются информационные технологии [5].

При оценивании инвестиционных проектов, занимающих значительную часть проектной деятельности современной компании, в силу неопределенности воздействия экзогенных факторов, нет возможности просчитать количественные оценки некоторых финансовых показателей и получить экономическую эффективность проекта.

Оценить потенциальную привлекательность проекта и инвестиционные риски можно на основе методов имитационного моделирования [12–13]. В ряде случаев оценивание вероятности того

или иного решения при различных условиях реализации проекта может произвести проект-менеджер [14]. Однако прогнозирование значений показателей эффективности может быть выполнено экспертом при воздействии на проектную ситуацию не более 3–5 входных переменных, причем слабо зависимых между собой [15]. Кроме того, качество проектного решения определяется набором показателей, что приводит к построению многокритериальных моделей оптимизации.

При создании проекта в технической сфере имеет место значительное количество компромиссных конфигураций выпускаемого промышленного продукта и параметров, оказывающих влияние на эффективность разработки [8], что делает объекты проектирования не только многомерными, но и структурно неопределенными. Особенность многокритериальной оптимизации – в том, что показатели эффективности проектов являются взаимозависимыми и при измерении степени достижения одной или нескольких целей демонстрируют противоречивое влияние [9].

Известный подход к выбору компромиссного варианта, использующий свертывание системы показателей эффективности и оптимизацию интегрального показателя, основан на предпочтениях экспертов [16].

Вместе с тем, энтропия проектной ситуации в условиях изменчивости проектной среды может быть существенно снижена на основе данных, накопленных в ранее реализованных проектах [17–19].

**3. Подход к анализу проектных решений на основе концепции доминирования и методов машинного обучения.** Непрерывный сбор данных о процессах управления проектами и использование инструментов бенчмаркинга при исследовании лучших практик, типовых решений позволяют выделить отдельные параметры проекта и оценить эффекты их изменения по группе эталонных проектов, например, на основе дисперсионного анализа [2].

Возможность исследовать механизм конкуренции между показателями эффективности и исключить субъективизм при выборе решения дает рассмотрение пространственного распределения репрезентативного набора вариантов из множества Парето [20].

Пусть  $\chi$  – допустимая область решений (пространство возможных реализаций проекта).

Требуется:

1) найти такое множество векторов  $P^*$ ,  $P^* \in \chi$ , для которого выполняется условие:

$$Y(X) < Y(X) \forall X \in P^*$$

при наличии массива накопленных данных:

$$(X^{<1>}, \dots, X^{<N>}, Y^{<1>}, \dots, Y^{<K>})_{n \times (N+K)},$$

в которых в неявном виде представлены ограничения на значения проектных переменных  $X$  (задающие пространство возможных реализаций проекта  $\chi$ ) и зависимости  $X \rightarrow Y$ ;

2) выполнить визуализацию множества  $P^*$  в критериальном пространстве:

$$FP^* = \{Y(X) = [Y_1(X), Y_2(X), \dots, Y_K(X)] : X \in P^*\},$$

которая дает возможность исследовать распределение элементов множества  $FP^*$  при размерности пространства  $K > 3$  и выбрать решение  $X^*$ , оптимизирующее по Парето вектор показателей эффективности:

$$Y(X) = [Y_1(X), Y_2(X), \dots, Y_K(X)].$$

Каждый вариант представляется в двух векторных пространствах – пространстве решений  $X$  и критериальном пространстве  $Y$ . Образ допустимой области решений  $Y(\chi)$  может быть исследован для выявления множества Парето только при получении знания об отображении пространства  $X$  в  $Y$ , которое зачастую является нелинейным [21].

При наличии данных требуемые знания фиксируются в обученной нейронной сети. Поскольку числовые, порядковые и номинальные показатели предварительно соответствующим образом нормализуются, интерпретация образов векторов осуществляется с использованием евклидовой метрики, как для количественных признаков, не имеющих разрывов значений. Таким образом, непрерывность отображения  $X \rightarrow Y$  между показателями проекта обеспечивает превращение близких точек –  $N$ -мерных векторов решения в близкие точки –  $m$ -мерные векторы показателей эффективности.

После применения автоматической классификации, которая основана на гипотезе компактности, распределение близких вариантов

по однородным группам должно сохранять в пространствах  $X$  и  $Y$  топологическое подобие (рисунок 2). При оптимизации показателей эффективности поиск осуществляется в пространстве решений таким образом, что каждый вариант может быть прослежен в критериальном пространстве и оценен на основе концепции доминирования.

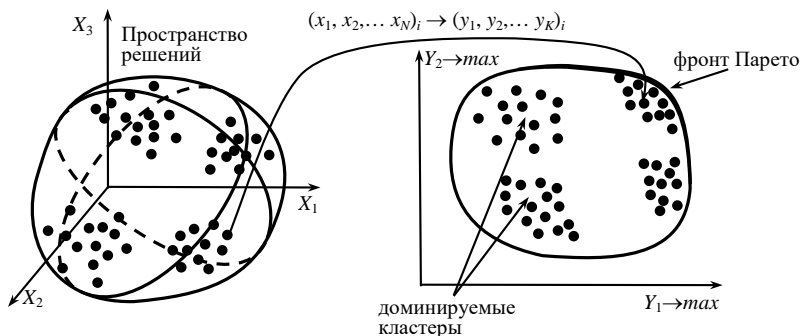


Рис. 2. Преобразование пространства решений в критериальное пространство  $X \rightarrow Y$  ( $N = 3$ ,  $K = 2$ )

В простейшем случае, при одномерной классификации, классы решений упорядочены по оценкам значения заданного показателя эффективности: ценности кредитного проекта, его степени риска, состояния спроса на производимую продукцию, уровня загрузки производственных мощностей, качества ремонта технической системы и т.д.

Многомерная же классификация может осуществляться на основе агрегирования оценок отдельных параметров проектной среды  $X$ , например, в случае оценивания кредитов по группам риска, полученных от банковских работников, или на основе экспертной оценки кредитным работником показателей  $Y$ .

Процедура классификации, как отнесение исследуемых вариантов к категориям проектов, призвана снизить неопределенность в процессе принятия решений.

**4. Методика анализа и визуализации проектных решений с использованием автоматической классификации данных.** Основные этапы анализа вариантов проектных решений на основе автоматической классификации данных приведены на рисунке 3.



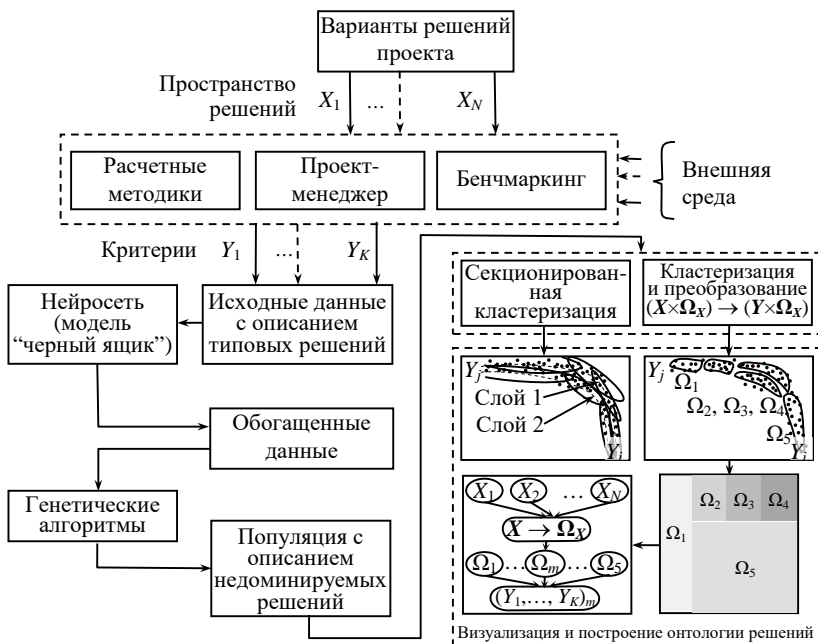


Рис. 3. Анализ и визуализация проектных решений

Автоматическая классификация в пространствах  $X$  или  $Y$  требует наличия данных значительного объема, отсутствующих, как правило, в проектной практике. Обогащение выборки с сохранением неявной связи между проектными переменными и показателями эффективности может выполняться с помощью прогнозирующей нейросетевой модели, обученной на наборе доступных прецедентов [22].

Чтобы прогнозирующая нейросетевая модель была повторяемой и согласованной во всем пространстве проектирования, обучение следует проводить с привлечением методов активного планирования эксперимента, когда обучающая выборка представляет объекты, распределенные по всем диапазонам значений переменных  $\{X_1, \dots, X_N\}$  для заполнения пространства проектирования. Сгенерированные таким образом варианты будут заполнять как области с низкими значениями показателей эффективности, так и области с высокими значениями показателей.

Исследование вариантов выполняется среди недоминируемых решений, которые образуют множество Парето  $P^*$ . В нем каждый вариант может быть улучшен альтернативным решением как минимум по одному показателю эффективности без ухудшения других показателей.

Образ множества Парето в критериальном пространстве определяет фронт Парето  $FP^*$ . При максимизации показателей  $Y_j, j = \overline{1, K}$ , фронт Парето находится в многомерном критериальном гиперпространстве над облаком вариантов со стороны идеальной точки, координаты которой определяются оптимальными значениями локальных показателей эффективности.

Для уплотнения ядра Парето-оптимальных проектов, представляющих в многомерном пространстве  $Y$  фронт Парето, используется набор генетических алгоритмов [23–25]. В качестве начальной популяции используется обогащенная выборка. В хромосоме особи зашифровывается комбинация значений проектных переменных. В каждом поколении с помощью обученной нейросетевой модели находят оценки показателей эффективности и уточняется фронт Парето.

Дальнейшее сужение пространства поиска связано со структуризацией недоминируемых решений.

В ситуациях, когда множество эффективных решений оказывается незначительным, обосновано применение нескольких слоев Парето. Например, при оценке акций по нескольким критериям [26].

В первый слой попадают Парето-оптимальные варианты. Варианты, полученные после исключения попавших в слой 1 объектов, принадлежат слою 2, затем слою 3 и так далее. Варианты, отнесенные ко второму и третьему слоям, хотя и доминируются элементами предшествующего слоя, как правило, имеют достаточно высокие значения показателей эффективности, и их целесообразно рассматривать наряду с Парето-оптимальными вариантами слоя 1, если число объектов во множестве Парето недостаточно.

При максимизации двух показателей первые слои располагаются в правой верхней части диаграммы размещения вариантов в пространстве критериев. Последующие слои располагаются ниже и левее. Включение каждого дополнительного слоя означает добавление вариантов с худшими значениями показателей, чем в предыдущих слоях. Поэтому, с одной стороны, увеличение количества слоев приводит к снижению среднего значения

каждого показателя по всем отобранным вариантам. Однако, с другой стороны, коэффициент вариации (отношение стандартного отклонения показателя к его среднему) также имеет тенденцию к падению с увеличением количества слоев. В экономических приложениях это означает снижение риска создаваемого портфеля инвестиционных решений. Поэтому количество используемых слоев определяется компромиссом между средней прибыльностью решений и риском.

Другой подход к структуризации фронта Парето заключается в кластеризации данных. В этом случае осуществляется выделение кластера, из которого выбирается образ желаемого решения [27, 28], либо представитель выбирается в каждом кластере [10], а образ желаемого решения выбирается среди установленного подмножества вариантов – представителей от каждого кластера.

Методология секционированной кластеризации множества Парето в критериальном пространстве [8] позволяет найти группы проектов, имеющих общий профиль по показателям эффективности  $\{Y_1, \dots, Y_K\}$ . Например, один кластер будет включать проекты, которые имеют высокое качество по одному показателю, и низкое качество по другим. Другой кластер может включать усредненные проекты, которые не являются лидерами ни по одному показателю. На практике проекты, реализующие экстремальные конфигурации промышленных изделий, не превосходят варианты, основанные на более сбалансированных конструкциях. Подобная неиерархическая кластеризация [29], основанная на заданном числе кластеров, игнорирует сходство вариантов в пространстве проектных переменных  $X$ .

Вместе с тем, на ранней стадии проектирования гораздо легче реализовать процедуры оценки схожести различных решений на основе обработки данных о проектных переменных. Дальнейшая интерпретация проектных решений возможна при наличии значимой связи между данными в пространствах  $X$  и  $Y$ . Степень влияния пространства решений  $X$  на критериальное пространство  $Y$  устанавливается с помощью канонического корреляционного анализа [30].

Для выбора окончательного решения требуется интерпретация фронта Парето. Предпочтения эксперта, зачастую используемые на практике, дают субъективно наилучшее решение. Переход в режим автоматического выбора решения возможен на основе минимизации расстояния до идеальной точки, координаты которой определяются оптимальными значениями показателей [31, 32]. При вычислении расстояний используют простейшую евклидову метрику. Образ

полученного таким методом решения находится в зоне наибольшей кривизны кривой, аппроксимирующей фронт Парето [33, 27].

Геометрия фронта Парето, сформированного при пополнении популяции на основе генетической мутации особей, определяет распределение компромиссных (равновесных) решений между так называемыми якорными точками. Такие точки соответствуют образам асимптотических решений задач однокритериальной оптимизации и задают диапазоны изменения каждого показателя на Парето-фронте [25]. В случае исследования проектов в технической сфере, точки, лежащие в окрестности якорных точек, характеризуют узкоспециальные варианты с большими значениями одного из показателей эффективности. Варианты со средними значениями показателей являются более универсальными.

Набор наиболее эффективных решений позволяет сформировать выделение зоны максимальной кривизны кривой, аппроксимирующей фронт Парето [34]. Поиск и исследование данной зоны возможны при использовании специальных средств визуализации данных.

Разработанный в Федеральном исследовательском центре «Информатика и управление» РАН визуализатор Pareto Front Viewer [35] позволяет выполнять ортографическое построение фронта Парето для ограниченного числа показателей эффективности. Например, фронт Парето строится на плоскости  $Y_1 - Y_2$  с использованием цветовой шкалы для третьего показателя  $Y_3$ .

Для выявления зоны наибольшей кривизны фронта Парето в многомерном критериальном пространстве неограниченной размерности предлагается использовать результаты кластеризации данных в области решений. Далее, группы проектов, имеющих близкий профиль  $\Omega_m, m = \overline{1, M}$ , восстанавливаются в критериальном пространстве с помощью обученной нейросети. На заключительном этапе дискретизированное ортографическое представление и интерпретация эффективных вариантов осуществляются с помощью плоского дерева решений.

Преобразование кластерной структуры  $\Omega_x = \{\Omega_1, \dots, \Omega_M\}$  в критериальное пространство  $Y$ :

$$(X \times \Omega_x) \rightarrow (Y \times \Omega_x),$$

позволяет построить плоское дерево решений  $Y \rightarrow \Omega_x$ , которое дает возможность выбрать кластер эффективных проектов  $\Omega_m^*$  и указать

для него интервалы значений проектных переменных и показателей эффективности.

**5. Проверка методики при визуализации и интерпретации проектных решений с помощью плоского дерева.** В рамках определенного проекта можно зафиксировать знания лишь о конкретном преобразовании  $X \rightarrow Y$ . Поэтому, чтобы продемонстрировать возможности ортографического представления и интерпретации эффективных вариантов при различных преобразованиях, сформируем набор данных (датасет) и выполним имитацию различных отображений  $X \rightarrow Y$  линейного и нелинейного вида, включая парные зависимости  $X_j - Y_j$  как с положительной, так и отрицательной корреляциями. Пусть допустимая область решений  $X$  определена в пространстве трех числовых входных переменных, с разделением данных на 4 кластера (рисунок 4). Расстояние до центра кластера кодируется размером символа – большая метка означает, что вариант находится ближе к центру кластера.

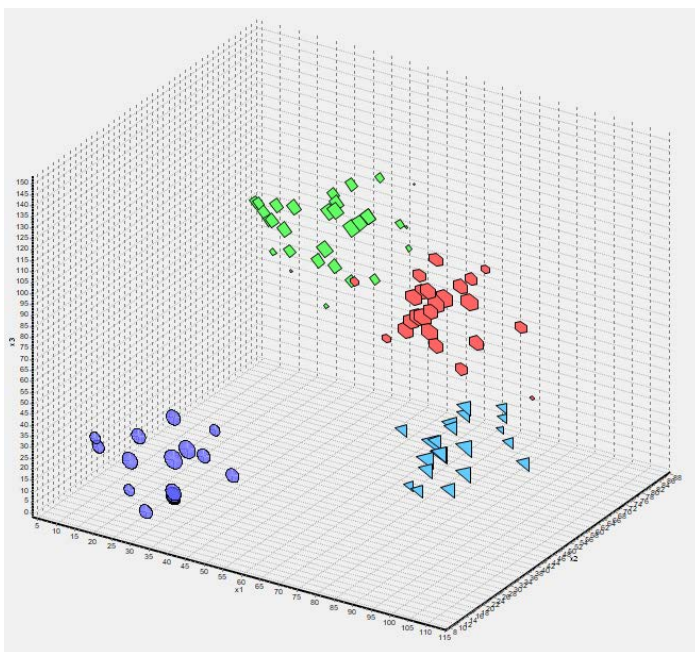


Рис. 4. Кластеризация набора данных в пространстве решений  $X$

Будем считать, что улучшению качества решения соответствует увеличение значений числовых показателей. Чтобы продемонстрировать соответствие найденных с помощью плоского дерева решений зоне наибольшей кривизны фронта Парето, ограничимся мощностью  $|Y| = 2$ .

Покажем, как с помощью канонического корреляционного анализа устанавливается влияние пространства решений на критериальное пространство. Выполним моделирование нелинейного преобразования трехмерного пространства  $X$  в двумерное пространство  $Y$  с помощью системы уравнений:

$$\begin{aligned} y_1 &= 92 - (x_1 + 10x_2 + 0,1x_3^2)/32, \\ y_2 &= 92 - 4(50/x_1 + 50/x_2 + 50/x_3). \end{aligned}$$

В статистическом приложении найдем линейные комбинации двух групп  $\{X_i\}$ ,  $i = \overline{1,3}$ , и  $\{Y_j\}$ ,  $j = \overline{1,2}$ , которые имеют наибольшую корреляцию. Первая каноническая переменная представлена линейной комбинацией трех стандартизированных входных переменных:

$$L_1 = -0,009x_1 - 0,174x_2 - 0,953x_3,$$

вторая – комбинацией двух стандартизированных показателей эффективности:

$$L_2 = 0,91y_1 - 0,15y_2.$$

Графическое отображение набора данных на плоскости  $L_1 - L_2$  подтверждает наличие влияния группы переменных  $\{X_1, X_2, X_3\}$  на группу показателей  $\{Y_1, Y_2\}$  (рисунок 5).

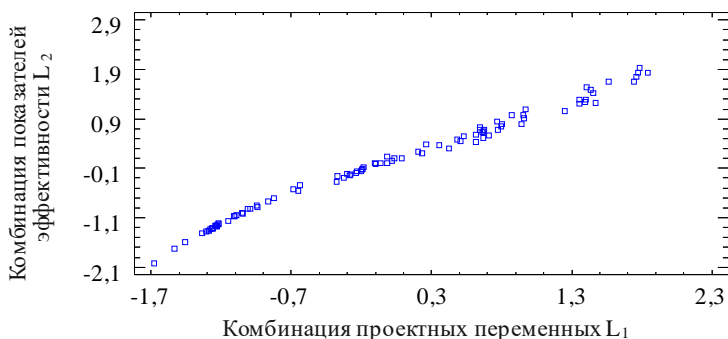


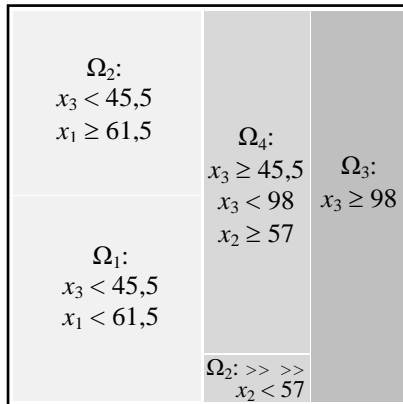
Рис. 5. Диаграмма рассеяния на плоскости канонических переменных (нелинейные преобразования  $X \rightarrow Y$ )

Установив наличие значимой связи между данными в пространствах  $X$  и  $Y$ , выполним дальнейшую визуализацию вариантов решений и их образов в критериальном пространстве, с помощью которых осуществляется выделение кластера наиболее эффективных проектов.

Выполним визуализацию набора данных в пространстве решений (рисунок 4) с помощью плоского дерева [36]. В общем случае такое ортографическое представление может быть построено при любой размерности пространства решений.

Чтобы плоское дерево имело дискретную структуру, которая имеет топологическое соответствие с проекциями разделяемых образов кластеров из многомерного пространства решений на плоскость, используется адаптивное квантование пространства  $X$  с помощью генетического алгоритма [37].

Для данных рассматриваемого тестового примера дискретная модель  $X \rightarrow \Omega_X$  представлена плоским деревом, осуществляющим преобразование комбинации значений трех проектных переменных в один из четырех кластеров (рисунок 6). Каждая ветвь традиционного решающего дерева соответствует прямоугольнику, внутри которого могут находиться дочерние ветви или листья. Количество объектов, попавших в узел, определяет размеры прямоугольников.

Рис. 6. Плоское дерево решений  $X \rightarrow \Omega_X$ 

Для поиска наиболее эффективных проектов используем плоское дерево решений, которое визуализирует кластерную структуру  $\Omega_X$  в критериальном пространстве  $Y$ .

Соответствие зоны наибольшей кривизны фронта Парето найденному на дереве кластеру проверим при различных преобразованиях пространств  $X \rightarrow Y$  линейного и нелинейного вида, моделируя парные зависимости “вход–выход” как с положительной, так и с отрицательной корреляциями. При использовании методики анализа и визуализации проектных решений преобразование данных  $X \rightarrow Y$  фиксируется в обученной нейросетевой модели.

Распределение кластеризованных вариантов в критериальном пространстве после линейного преобразования  $X \rightarrow Y$ , при наличии корреляций одного знака между показателями эффективности и проектными переменными, приведено на рисунке 7.

Компромиссные решения со средними значениями показателей лежат в зоне наложения кластеров  $\Omega_2$  и  $\Omega_4$ . При максимизации показателей  $Y_1, Y_2$  доминирующими будут решения, ассоциированные с кластером  $\Omega_2$ .



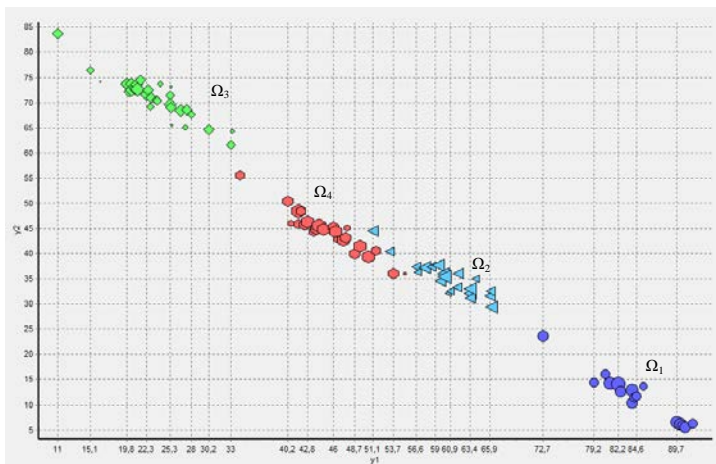


Рис. 7. Представление кластерной структуры  $\Omega_X$  в критериальном пространстве  $Y$  при линейном отображении  $(X_1, X_2, X_3) \rightarrow (Y_1, Y_2)$

Соответствующее плоское дерево решений  $Y \rightarrow \Omega_X$  представлено на рисунке 8. Образ идеальной точки расположен в правом верхнем углу. Эффективными будут проекты, относящиеся ко второму кластеру, прямоугольный образ которого примыкает к идеальной точке. Таким образом, решение на основе плоского дерева совпадает с результатами анализа фронта Парето.



Рис. 8. Плоское дерево решений  $Y \times \Omega_X$  при линейном преобразовании  $X \rightarrow Y$

Распределение кластеризованных вариантов в критериальном пространстве после нелинейного преобразования  $X \rightarrow Y$ , при наличии

корреляций одного знака между показателями эффективности и проектными переменными, приведено на рисунке 9.

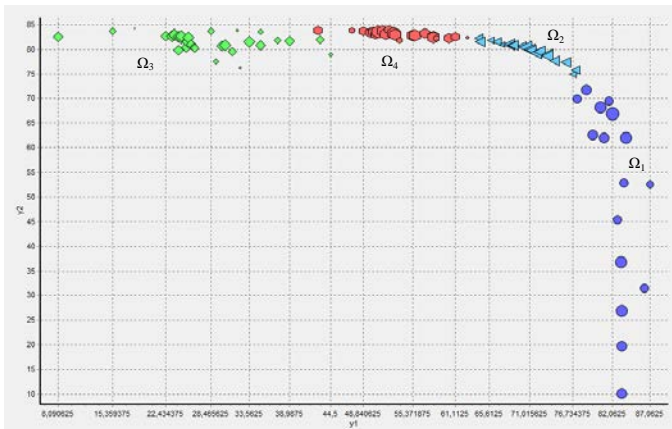


Рис. 9. Представление кластерной структуры  $\Omega_X$  в пространстве целей  $Y$  при нелинейном отображении  $(X_1, X_2, X_3) \rightarrow (Y_1, Y_2)$

Компромиссные решения распределены в зоне максимальной кривизны фронта Парето и полностью ассоциируются в данном случае с кластером  $\Omega_2$ . Решение на основе плоского дерева также совпадает с результатами анализа фронта Парето: возле идеальной точки размещается единственный прямоугольник второго кластера (рисунок 10).

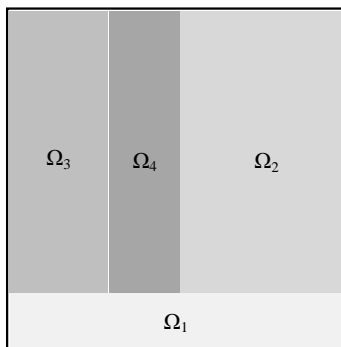


Рис. 10. Плоское дерево решений  $Y \times \Omega_X$  при нелинейном преобразовании  $X \rightarrow Y$

Распределение кластеризованных вариантов в критериальном пространстве после линейного преобразования  $X \rightarrow Y$ , при наличии корреляций разного знака между показателями эффективности и проектными переменными, приведено на рисунке 11.

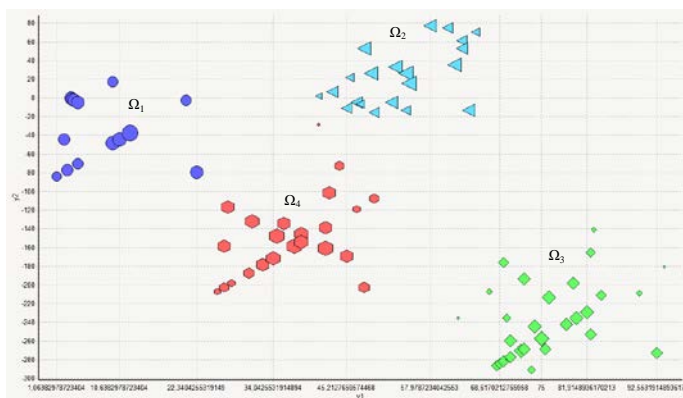


Рис. 11. Представление кластерной структуры  $\Omega_X$  в пространстве целей  $Y$  при линейном отображении  $(X_1, X_2, X_3) \rightarrow (Y_1, Y_2)$  и “смешанной” корреляции

Компромиссные решения в данном случае распределены в нескольких слоях Парето и полностью ассоциируются с кластером  $\Omega_2$ . Решение на основе плоского дерева также совпадает с результатами анализа фронта Парето: возле идеальной точки размещается единственный прямоугольник второго кластера (рисунок 12).

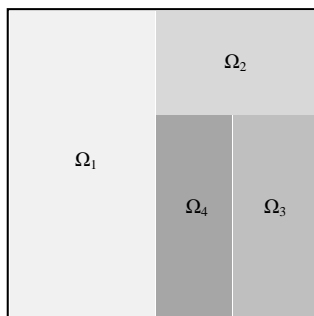


Рис. 12. Плоское дерево решений  $Y \times \Omega_X$  при линейном преобразовании  $X \rightarrow Y$  и “смешанной” корреляции

Распределение кластеризованных вариантов в критериальном пространстве после нелинейного преобразования  $X \rightarrow Y$ , при наличии корреляций разного знака между показателями эффективности и проектными переменными, приведено на рисунке 13.

Компромиссные решения в данном случае распределены в нескольких слоях Парето. Доминирующими являются решения, ассоциированные с кластером  $\Omega_3$ .

Существующая методология секционированной кластеризации в критериальном пространстве и выбор кластера, попавшего в зону перегиба [8], не позволяют дифференцировать решения по группам  $\Omega_1, \Omega_2$  и  $\Omega_3$ , имеющим сходство в пространстве  $X$ .

Предложенная, с использованием нейронной сети, методика переноса кластерной структуры из пространства проектных переменных в критериальное пространство и визуализация кластеров с помощью плоского дерева дают решение, которое совпадает с результатами анализа фронта Парето: возле идеальной точки размещаются прямоугольники кластера  $\Omega_3$  (рисунок 14).

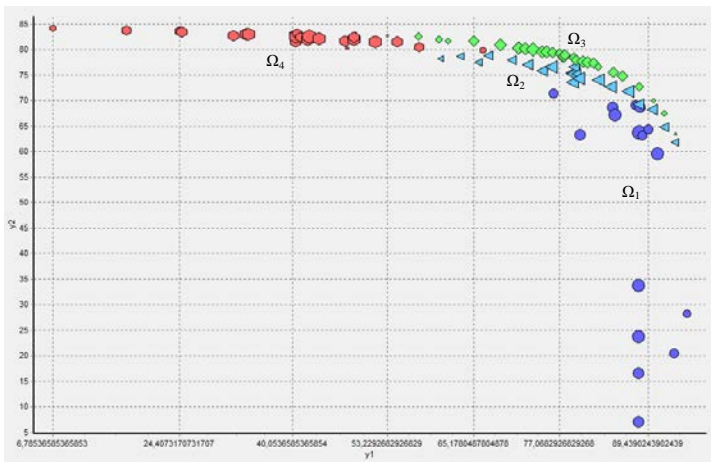


Рис. 13. Представление кластерной структуры  $\Omega_X$  в пространстве целей  $Y$  при нелинейном отображении  $(X_1, X_2, X_3) \rightarrow (Y_1, Y_2)$  и “смешанной” корреляции

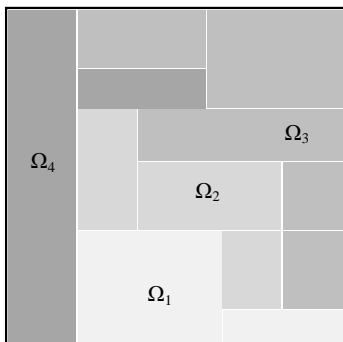


Рис. 14. Плоское дерево решений  $Y \times \Omega_x$  при нелинейном преобразовании  $X \rightarrow Y$  и “смешанной” корреляции

Как видно из диаграмм рассеяния, участки перегиба фронта Парето определяются либо содержимым целого кластера (рисунки 9, 13), либо частью вариантов, представляющих “лучший” кластер (рисунки 7, 11).

Решение об отборе всех или части представителей кластера принимается после анализа плоского дерева решений  $Y \rightarrow \Omega_x$ . Если на плоском дереве к правому верхнему углу примыкает неразделенный прямоугольник кластера (рисунок 12), то его представители в критериальном пространстве хорошо отделены от остальных кластеров и, при максимизации показателей  $Y_j, j = \overline{1, K}$ , наиболее приближены к идеальной точке. Все представители такого кластера являются эффективными решениями. Если на плоском дереве кластер-победитель разделяется на несколько прямоугольников, то недоминируемые образы решений с лучшими значениями проектных переменных определяются примыкающим к правому верхнему углу прямоугольником. Подгруппы доминируемых вариантов кластера определяются прямоугольниками, которые находятся внутри дерева решений – как прямоугольники кластера  $\Omega_2$  на рисунке 8.

Построенная в критериальном пространстве на обогащенной выборке модель регрессионной гиперповерхности показывает внутреннюю зависимость между показателями эффективности  $Y_j, j = \overline{1, K}$ .

**6. Заключение.** Многомерность проекта как объекта управления включает агрегированное описание ресурсных и организационно-экономических показателей проектной среды, а также

стоимостные, функциональные и временные показатели эффективности. Данные о процессах управления проектами и использование инструментов бенчмаркинга при исследовании типовых решений позволяют установить взаимосвязь показателей эффективности и исключить субъективизм при выборе решения. При многокритериальной оптимизации поиск осуществляется в пространстве решений таким образом, что каждый вариант может быть прослежен в критериальном пространстве и оценен на основе концепции доминирования.

При решении данной задачи получены следующие результаты.

Предложен подход к анализу проектных решений на основе методов машинного обучения.

Поскольку, согласно гипотезе компактности, близкие варианты проектов чаще располагаются в одном кластере, то множество наилучших решений в критериальном пространстве чаще всего соответствует определенному кластеру в пространстве решений.

Так как в проектной практике, как правило, отсутствуют выборки достаточного объема, обогащение данных с сохранением неявной связи между проектными переменными и показателями эффективности может выполняться с помощью прогнозирующей нейросетевой модели, обученной на наборе доступных прецедентов.

Разработана методика анализа, визуализации и интерпретации проектов с помощью плоского дерева решений.

Для обнаружения фронта Парето в многомерном критериальном пространстве используется набор генетических алгоритмов. Наилучший компромисс определяется теми решениями, образы которых распределены в зоне максимальной кривизны кривой, аппроксимирующей фронт Парето. Существующая методология секционированной кластеризации множества Парето в критериальном пространстве позволяет найти группы проектов, имеющих общий профиль по значениям показателей эффективности, однако игнорирует сходство вариантов в пространстве проектных переменных.

Для интерпретации проектных решений при наличии значимой связи между данными в проектном и критериальном пространствах и автоматизации поиска в многомерном пространстве зоны наибольшей кривизны фронта Парето, предложено использовать плоское дерево решений, визуализирующее в критериальном пространстве  $Y$  кластерную структуру  $\Omega_x$ , установленную в пространстве проектных переменных  $X$ .

Плоское дерево может быть построено после преобразования  $X \rightarrow Y$  при любой размерности критериального пространства и иметь

структуру, которая имеет топологическое соответствие с проекциями разделяемых образов кластеров из многомерного пространства на плоскость.

На тестовом наборе данных показано, что участки наибольшей кривизны фронта Парето определяются либо содержимым целого кластера, либо частью вариантов, представляющих “лучший” кластер. Решение об отборе представителей кластера принимается после анализа плоского дерева решений. Если на плоском дереве к правому верхнему углу примыкает неразделенный прямоугольник кластера, то его представители в критериальном пространстве хорошо отделены от остальных кластеров и, при максимизации показателей  $Y_j, j = \overline{1, K}$ , наиболее приближены к идеальной точке. Все представители такого кластера являются эффективными решениями. Если кластер-победитель содержит внутри дерева решений подгруппу доминируемых вариантов, то “лучший” кластер представляют оставшиеся варианты, которые задают оптимальные настройки проектных переменных  $X_j, j = \overline{1, N}$ .

Дальнейшее развитие предложенного подхода связано с оцениванием эффективности проекта на разных стадиях жизненного цикла, когда на первый план выходит задача прогнозирования и проактивного управления. Поскольку цели, определяющие развитие каждого  $i$ -го этапа жизненного цикла, корректируются, на разных этапах развития проекта используются различные пространства  $X^{<i>} \times Y^{<i>}$ . Прогнозная аналитика опирается как на текущие, так и ретроспективные большие данные. В этом случае многофакторное прогнозирование показателей  $Y^{<i>}$  может быть осуществлено привлечением методов снижения размерности подаваемых на вход динамически меняющихся данных и рекуррентных нейронных сетей глубокого обучения, не имеющих ограничений по характеру входных данных и способных решать задачи практически любой сложности.

### Литература

1. Perrier N., Benbrahim S.-E., Pellerin R. The core processes of project control: A network analysis // *Procedia Computer Science*. 2018. vol. 138. pp. 697–704.
2. Laursen Markus, Svejvig Per, Gerstrøm Rode Anna Le. Four Approaches to Project Evaluation // *The 24th Nordic Academy of Management Conference (NFF-2017)*. 2017. pp. 1–25.
3. Makeev Vladimir, Isaev Albert, Kulikov Sergey, Stratan Dmitry, Shevkunov Nikolay. Modeling and assessing the effectiveness of investment projects in the agricultural sector // *XII International Scientific Conference on Agricultural Machinery Industry*. 2019. 10–13 September. vol. 403: 012077.

4. Budeli Lalamani, Wichers J.H. Evaluating aspects of power plant performance using Project Success Life Cycle Model (PSLCM) // *PM World Journal*. 2018, vol. VII. iss. XI. pp. 1–25.
5. Юсупов Р.М., Мусаев А.А. Особенности оценивания эффективности информационных систем и технологий // *Труды СПИИРАН*. 2017. Вып. 2(51), С. 5–34.
6. Chereshtin D., Royzenon G., Britkov V. Multidimensional classifier of risk analysis methods // 11th World Conference «Intelligent Systems for Industrial Automation» (WCIS-2020). 2020. vol. 1323. pp. 529–536.
7. Koledina K.F., Koledin S.N., Karpenko A.P., Gubaydullin I.M., Vovdenko M.K. Multi-objective optimization of chemical reaction conditions based on a kinetic model // *Journal of Mathematical Chemistry*. 2019. vol. 57, p.p. 484–493.
8. Serafini Marco, Furini Francesco, Colombo Giorgio, Rizzi Caterina. Optimized development: defining design rules through product optimization techniques // *Computer-Aided Design & Applications*. 2016. vol. 13. no. 5. pp. 600–609.
9. Zhao Menglong, Huang Shengzhi, Huang Qiang, Wang Hao, Leng Guoyong, Liu Siyuan, Wang Lu. Copula-Based Research on the Multi-Objective Competition Mechanism in Cascade Reservoirs Optimal Operation // *Water*. 2019. vol. 11. iss. 995. pp. 1–19.
10. Wicaksono Albert, Jeong Gimoon, Kang Doosun. Water–Energy–Food Nexus Simulation: An Optimization Approach for Resource Security // *Water*. 2019. vol. 11. iss. 4: 667. pp. 1–19.
11. Akhanova M.A., Eropkina A.S., Ovchinnikova S.V., Skifskaya A.L. Methodology of estimating an IT project efficiency // *International Journal of Mechanical Engineering and Technology (IJMET)*. 2018. vol. 9. iss. 13. pp. 803–809.
12. Орехова Н.Ю. Построение математической модели инвестиционного проекта // *Труды СПИИРАН*. 2003. Вып. 1. Т. 3, С. 187–195.
13. Kharchenko Volodymyr, Kharchenko Hanna. Simulation Modeling in Assessing the Effectiveness and Risk of Investment Projects // *Modern Economics*. 2020. vol. 22(1). pp. 119–124.
14. Glukhikh I.N., Pisarev M.O., Nonieva K.Z. Developing an Automated System for Assessing an Innovative Project's Economic Efficiency for an Oil and Gas Industry Case // *International Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon)*, IEEE. 2020. October 6–9. access. no. 20256255.
15. Мусаев А.А., Нозик А.А., Русинов Л.А. Прогностический анализ безопасности промышленного предприятия // *Известия Санкт-Петербургского государственного технологического института (технического университета)*. 2016. № 34(60). С. 87–93.
16. Aliyev Elchin, Rzayev Ramin, Ali Adila. Multi-criteria Evaluation of Investment Projects Using the Fuzzy Method of Weighted Maximin Convolution // 11th International Conference on Theory and Application of Soft Computing, Computing with Words and Perceptions and Artificial Intelligence (ICSCCW–2021). 2022. pp. 671–679.
17. Vanhoucke Mario, Batselier Jordy. A Statistical Method for Estimating Activity Uncertainty Parameters to Improve Project Forecasting // *Entropy*. 2019. vol. 21. iss. 952. pp. 1–28.
18. Chernyakhovskaya Liliya, Nizamutdinov Marsel. Development of Knowledge Base for Intellectual Decision Support in the Process of Innovative Project Management // *IEEE XXI International Conference Complex Systems: Control and Modeling Problems (CSCMP)*. 2019. access. no. 19318924.
19. Piterska Varvara, Shakhov Anatoliy, Lohinov Oleh, Lohinova Liliia. The Method of Transfer of Research Project Results of Institution of Higher Education // *IEEE 14th*



- International Conference on Computer Sciences and Information Technologies (CSIT). 2019. access. no. 19250579.
20. Ayaz Ahsan, Rasheed Ashhad. Multi-Objective Design Optimization of Multicopter using Genetic Algorithm // IEEE International Bhurban Conference on Applied Sciences and Technologies (IBCAST). 2021. access. no. 20633290.
  21. Sanjeev Kumar Dash, Ajit Kumar Behera, Satchidananda Dehuri, Sung-Bae Cho. Radial basis function neural networks: a topical state-of-the-art survey // Open Computer Science. 2016. vol. 6. iss. 1. pp. 33–63.
  22. Пименов В.И., Кофнов О.В., Пименов И.В. Оценка эффективности проектов на основе применения моделей машинного обучения и методов многокритериальной оптимизации // Совершенствование математического образования – 2020: состояние и перспективы развития: Материалы XI междунар. науч.-методич. конф. Тирасполь. 2020. С. 115–119.
  23. Pimenov V.I., Pimenov I.V. Interpretation of a trained neural network based on genetic algorithms // Информационно-управляющие системы. 2020. № 6. С. 12–20.
  24. Qin Shufen, Sun Chaoli, Jin Yaochu, Tan Ying, Fieldsend Jonathan. Large-Scale Evolutionary Multiobjective Optimization Assisted by Directed Sampling // IEEE Transactions on Evolutionary Computation. 2021. vol. 25. iss. 4. pp. 724–738.
  25. Li Yuping, Petrov D.A., Sherlock Gavin. Single nucleotide mapping of trait space reveals Pareto fronts that constrain adaptation // Nature Ecology & Evolution. 2019. vol. 3. pp. 1539–1551.
  26. Израйлевич С.В., Цудикман В.Я. Опционы: системный подход к инвестициям. Критерии оценки и методы анализа торговых возможностей // М.: Альпина Паблишер. 2008. 280 с.
  27. Handl Julia, Knowles Joshua. Modes of Problem Solving with Multiple Objectives: Implications for Interpreting the Pareto Set and for Decision Making // Multiobjective Problem Solving from Nature. 2008. pp.131–151.
  28. Chikumbo Oliver, Granville Vincent. Optimal Clustering and Cluster Identity in Understanding High-Dimensional Data Spaces with Tightly Distributed Points // Machine learning & knowledge extraction. 2019. vol. 1. iss. 2. pp. 715–744.
  29. Mercioni Marina Adriana, Holban Ștefan. Evaluating hierarchical and non-hierarchical grouping for develop a smart system // IEEE International Symposium on Electronics and Telecommunications (ISETC). 2018. access. no. 18326471.
  30. Ida Masaaki. Consideration on the variation of financial data of institutions for canonical correlation analysis // IEEE 21st International Conference on Advanced Communication Technology (ICACT). 2019. access. no. 18636837.
  31. Georgioudakis Manolis, Fragiadakis Michalis. Selection and Scaling of Ground Motions Using Multicriteria Optimization // Journal of Structural Engineering (ASCE). 2020. vol. 146. iss. 11: 04020241.
  32. Zhu Yun, Wang Jun, Liang Shuang. Multi-Objective Optimization Based Multi-Bernoulli Sensor Selection for Multi-Target Tracking // Sensors. 2019. vol. 19. iss. 4: 980. pp. 1–18.
  33. Титов В.Г., Залазинский А.Г., Крючков Д.И., Нестеренко А.В. Многокритериальная оптимизация методом «идеальной точки» состава сырья для изготовления композиционной заготовки // Известия вузов. Порошковая металлургия и функциональные покрытия. 2019. №2. С. 49–56.
  34. Xiaoping Fang, Yaoming Cai, Zhihua Cai, Xinwei Jiang, Zhikun Chen. Sparse Feature Learning of Hyperspectral Imagery via Multiobjective-Based Extreme Learning Machine // Sensors. 2020. vol. 20. iss. 5: 1262. pp. 1–19.
  35. Березкин В.Е., Каменев Г.К., Лотов А.В. Программа для визуализации многомерной границы Парето в невыпуклых задачах многокритериальной

- оптимизации (PFV-II). Свидетельство о государственной регистрации программы для ЭВМ № RU 2019664809 от 13.11.2019.
36. Long Lim Kian, Hui Lim Chien, Fook Gim Yeong, Wan Zainon Wan Mohd Nazmee. A Study on the Effectiveness of Tree-Maps as Tree Visualization Techniques // *Procedia Computer Science*. 2017. iss. 124. pp. 108–115.
37. Пименов В.И., Пименов И.В. Применение генетического алгоритма для оптимизации дискретной структуры решающего дерева // *Вестник Санкт-Петербургского государственного университета технологии и дизайна. Серия 1. Естественные и технические науки*. 2020. № 3. С. 55–60.

**Пименов Виктор Игоревич** — д-р техн. наук, старший научный сотрудник, заведующий кафедрой, кафедра информационных технологий, Санкт-Петербургский государственный университет промышленных технологий и дизайна. Область научных интересов: распознавание образов, интеллектуальный анализ данных, когнитивные технологии. Число научных публикаций — 200. v\_rim@mail.ru; Вознесенский проспект, 46, 190068, Санкт-Петербург, Россия; р.т.: +7(812)312-2567.

**Пименов Илья Викторович** — канд. техн. наук, доцент, кафедра вычислительных систем и информатики, Государственный университет морского и речного флота имени адмирала С.О. Макарова. Область научных интересов: моделирование процессов принятия решений, интеллектуальные системы, распознающие базы знаний, многомерный анализ данных. Число научных публикаций — 56. i-rim@mail.ru; Двинская улица, 5/7, 198035, Санкт-Петербург, Россия; р.т.: +7(812)748-9635.

V. PIMENOV, I. PIMENOV

**DATA ANALYSIS AND VISUALIZATION IN THE TASKS OF THE PROJECT SOLUTIONS MULTICRITERIA OPTIMIZATION**

*Pimenov V., Pimenov I. Data Analysis and Visualization in the Tasks of the Project Solutions Multicriteria Optimization.*

**Abstract.** The accumulation of data on project management processes and standard solutions has made relevant research related to the use of knowledge engineering methods for a multi-criteria search for options that set optimal settings for project environment parameters. Purpose: Development of a method for searching and visualizing groups of projects that can be evaluated based on the concept of dominance and interpreted in terms of project variables and performance indicators. Methods: The enrichment of the sample while maintaining an implicit link between the project variables and performance indicators is carried out using a predictive neural network model. A set of genetic algorithms is used to detect the Pareto front in the multidimensional criterion space. The ontology of projects is determined after clustering options in the solution space and transforming the cluster structure into the criterion space. Automation of the search in the multidimensional space of the Pareto front greatest curvature zone, which determines the equilibrium design solutions, their visualization and interpretation are carried out using a tree map. Results: A tree map is constructed at any dimension of the criterion space and has a structure that has a topological correspondence with projections of shared cluster images from a multidimensional space onto a plane. For various types of transformations and correlations between performance indicators and project variables, it is shown that the areas of the Pareto front greatest curvature are determined either by the contents of the whole cluster or by part of the variants representing the "best" cluster. If an undivided rectangle of a cluster is adjacent to the upper right corner of a tree map, then its representatives in the criterion space are well separated from the rest of the clusters and, when maximizing performance indicators, are closest to the ideal point. All representatives of such a cluster are effective solutions. If the winning cluster contains dominant options inside the decision tree, then the "best" cluster is represented by the remaining options that set the optimal settings for the project variables. Practical relevance: The proposed methods of searching and visualizing groups of projects can be used when choosing the conditions of resource and organizational and economic modeling of the project environment, ensuring the optimization of risks, cost, functional, and time criteria.

**Keywords:** project management, multi-criteria classification, neural network, genetic algorithm, Pareto front, cluster structure, visualization, tree map.

**Pimenov Viktor** — Ph.D., Dr.Sci., Senior researcher, head of the department, Department of information technologies, Saint-Petersburg State University of Industrial Technologies and Design. Research interests: image recognition, data mining, cognitive technologies. The number of publications — 200. v\_pim@mail.ru; 46, Voznesensky Av., 190068, St. Petersburg, Russia; office phone: +7(812)312-2567.

**Pimenov Ilya** — Ph.D., Associate professor, Department of computer systems and computer science, Admiral Makarov State University of Maritime and Inland Shipping. Research interests: data mining, decision-making modelling, intelligent system, knowledge base, multivariate data analysis. The number of publications — 56. i-pim@mail.ru; 5/7, Dvinskaya St., 198035, St. Petersburg, Russia; office phone: +7(812)748-9635.

## References

1. Perrier Nathalie, Benbrahim Salah-Eddine, Pellerin Robert. The core processes of project control: A network analysis. *Procedia Computer Science*. 2018. vol. 138. pp. 697–704.
2. Laursen Markus, Svejvig Per, Gerstrøm Rode Anna Le. Four Approaches to Project Evaluation. The 24th Nordic Academy of Management Conference (NFF-2017). 2017. pp. 1–25.
3. Makeev Vladimir, Isaev Albert, Kulikov Sergey, Stratan Dmitry, Shevkunov Nikolay. Modeling and assessing the effectiveness of investment projects in the agricultural sector. XII International Scientific Conference on Agricultural Machinery Industry. 2019. vol. 403: 012077.
4. Budeli Lalamani, Wichers J.H. Evaluating aspects of power plant performance using Project Success Life Cycle Model (PSLCM). *PM World Journal*. 2018. vol. VII, iss. XI, pp. 1–25.
5. Yusupov R.M., Musaev A.A. [Efficiency of Information Systems and Technologies: Features of Estimation]. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2017. iss. 2(51). pp. 5–34. (In Russ.).
6. Chereskin D., Royzenon G., Britkov V. Multidimensional classifier of risk analysis methods. 11th World Conference «Intelligent Systems for Industrial Automation» (WCIS-2020). 2020. vol. 1323. pp. 529–536.
7. Koledina K.F., Koledin S.N., Karpenko A.P., Gubaydullin I.M., Vovdenko M.K. Multi-objective optimization of chemical reaction conditions based on a kinetic model. *Journal of Mathematical Chemistry*. 2019. vol. 57, pp. 484–493.
8. Serafini Marco, Furini Francesco, Colombo Giorgio, Rizzi Caterina. Optimized development: defining design rules through product optimization techniques. *Computer-Aided Design & Applications*. 2016. vol. 13. no. 5. pp. 600–609.
9. Zhao Menglong, Huang Shengzhi, Huang Qiang, Wang Hao, Leng Guoyong, Liu Siyuan, Wang Lu. Copula-Based Research on the Multi-Objective Competition Mechanism in Cascade Reservoirs Optimal Operation. *Water*. 2019. vol. 11. iss. 995. pp. 1–19.
10. Wicaksono Albert, Jeong Gimoon, Kang Doosun. Water–Energy–Food Nexus Simulation: An Optimization Approach for Resource Security. *Water*. 2019. vol. 11. iss. 4: 667. pp. 1–19.
11. Akhanova M.A., Eropkina A.S., Ovchinnikova S.V., Skifskaya A.L. Methodology of estimating an IT project efficiency. *International Journal of Mechanical Engineering and Technology (IJMET)*. 2018. vol. 9. iss. 13. pp. 803–809.
12. Orekhova N.U. [Mathematical simulation of investment project in base of algorithmic networks]. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2003. iss. 1. vol. 3. pp. 187–195. (In Russ.).
13. Kharchenko Volodymyr, Kharchenko Hanna. Simulation Modeling in Assessing the Effectiveness and Risk of Investment Projects. *Modern Economics*. 2020. vol. 22(1). pp. 119–124.
14. Glukhikh I.N., Pisarev M.O., Nonieva K.Z. Developing an Automated System for Assessing an Innovative Project's Economic Efficiency for an Oil and Gas Industry Case. *IEEE International Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon)*. 2020. access. no. 20256255.
15. Musaev A.A., Nozick A.A., Rusinov L.A. [Predictive analysis of safety of an industrial enterprise]. *Izvestiya SPbGTU(TU) – Bulletin of St. Petersburg State Technological Institute (Technical University)*. 2016, no. 34(60). pp. 87–93. (In Russ.).
16. Aliyev Elchin, Rzayev Ramin, Ali Adila. Multi-criteria Evaluation of Investment Projects Using the Fuzzy Method of Weighted Maximin Convolution. 11th

- International Conference on Theory and Application of Soft Computing, Computing with Words and Perceptions and Artificial Intelligence (ICSCCW–2021), 2022. pp. 671–679.
17. Vanhoucke Mario, Batselier Jordy. A Statistical Method for Estimating Activity Uncertainty Parameters to Improve Project Forecasting. *Entropy*. 2019. vol. 21. iss. 952. pp. 1–28.
  18. Chernyakhovskaya Liliya, Nizamutdinov Marsel. Development of Knowledge Base for Intellectual Decision Support in the Process of Innovative Project Management. *IEEE XXI International Conference Complex Systems: Control and Modeling Problems (CSCMP)*. 2019. access. no. 19318924.
  19. Piterska Varvara, Shakhov Anatoliy, Lohinov Oleh, Lohinova Liliia. The Method of Transfer of Research Project Results of Institution of Higher Education. *IEEE 14th International Conference on Computer Sciences and Information Technologies (CSIT)*. 2019. access. no. 19250579.
  20. Ayaz Ahsan, Rasheed Ashhad. Multi-Objective Design Optimization of Multi-copter using Genetic Algorithm. *IEEE International Bhurban Conference on Applied Sciences and Technologies (IBCAST)*. 2021. access. no. 20633290.
  21. Sanjeev Kumar Dash, Ajit Kumar Behera, Satchidananda Dehuri, Sung-Bae Cho. Radial basis function neural networks: a topical state-of-the-art survey. *Open Computer Science*. 2016. vol. 6. iss. 1. pp. 33–63.
  22. Pimenov V.I., Kofnov O.V., Pimenov I.V. [Evaluating the effectiveness of projects based on the use of machine learning models and multi-criteria optimization methods]. *Sovershenstvovanie matematicheskogo obrazovanija – 2020: sostojanie i perspektivy razvitiija: Materialy XI mezhdunar. nauch.-metodich. konf. [Improvement of mathematical education – 2020: state and prospects of development: Materials of the XI Int. Sc. and Meth. Conf.]*. Tiraspol. 2020. pp. 115– 119. (In Russ.).
  23. Pimenov V.I., Pimenov I.V. Interpretation of a trained neural network based on genetic algorithms. *Informatsionno-upravliaiushchie sistemy – Information and Control Systems*. 2020. no. 6. pp. 12–20.
  24. Qin Shufen, Sun Chaoli, Jin Yaochu, Tan Ying, Fieldsend Jonathan. Large-Scale Evolutionary Multiobjective Optimization Assisted by Directed Sampling. *IEEE Transactions on Evolutionary Computation*. 2021. vol. 25. iss. 4. pp. 724–738.
  25. Li Yuping, Petrov D.A., Sherlock Gavin. Single nucleotide mapping of trait space reveals Pareto fronts that constrain adaptation. *Nature Ecology & Evolution*. 2019. vol. 3. pp. 1539–1551.
  26. Izrajlevich S. V., Cudikman V. Ja. Opciony: sistemnyj podhod k investicijam. Kriterii ocenki i metody analiza torgovyh vozmozhnostej [Options: a systematic approach to investment. Evaluation criteria and methods for analyzing trading opportunities]. M.: Alpina Publ. 2008. 280 p. (In Russ.).
  27. Handl Julia, Knowles Joshua. Modes of Problem Solving with Multiple Objectives: Implications for Interpreting the Pareto Set and for Decision Making. *Multiobjective Problem Solving from Nature*. 2008. pp.131–151.
  28. Chikumbo Oliver, Granville Vincent. Optimal Clustering and Cluster Identity in Understanding High-Dimensional Data Spaces with Tightly Distributed Points. *Machine learning & knowledge extraction*. 2019. vol. 1. iss. 2. pp. 715–744.
  29. Mercioni Marina Adriana, Holban Ștefan. Evaluating hierarchical and non-hierarchical grouping for develop a smart system. *IEEE International Symposium on Electronics and Telecommunications (ISETC)*. 2018. access. no. 18326471.
  30. Ida Masaaki. Consideration on the variation of financial data of institutions for canonical correlation analysis. *IEEE 21st International Conference on Advanced Communication Technology (ICACT)*. 2019. access. no. 18636837.

31. Georgioudakis Manolis, Fragiadakis Michalis. Selection and Scaling of Ground Motions Using Multicriteria Optimization. *Journal of Structural Engineering (ASCE)*. 2020. vol. 146. iss. 11: 04020241.
32. Zhu Yun, Wang Jun, Liang Shuang. Multi-Objective Optimization Based Multi-Bernoulli Sensor Selection for Multi-Target Tracking. *Sensors*. 2019. vol. 19. iss. 4: 980. pp. 1–18.
33. Titov V.G., Zalazinsky A.G., Kryuchkov D.I., Nesterenko A.V. [Multi-criteria optimization by the «ideal point» method of raw material composition for composite blank manufacturing]. *Izvestiya vuzov. Poroshkovaya metallurgiya i funktsional'nye pokrytiya – News of universities. Powder metallurgy and functional coatings*. 2019. no. 2. pp. 49–56. (In Russ.).
34. Xiaoping Fang, Yaoming Cai, Zhihua Cai, Xinwei Jiang, Zhikun Chen. Sparse Feature Learning of Hyperspectral Imagery via Multiobjective-Based Extreme Learning Machine. *Sensors*. 2020. vol. 20. iss. 5: 1262. pp. 1–19.
35. Berezkin V.E., Kamenev G.K., Lotov A.V. [A program for visualizing the multidimensional Pareto boundary in nonconvex multicriteria optimization problems (PFV-II)]. *Programma dlja vizualizacii mnogomernoj granicy Pareto v nevyuklykh zadachah mnogokriterial'noj optimizacii. Svidetel'stvo o gosudarstvennoj registracii programmy dlja JeVM*. 2019. no. RU 2019664809. (In Russ.).
36. Long Lim Kian, Hui Lim Chien, Fook Gim Yeong, Wan Zainon Wan Mohd Nazmee. A Study on the Effectiveness of Tree-Maps as Tree Visualization Techniques. *Procedia Computer Science*. 2017. iss. 124. pp. 108–115.
37. Pimenov V.I., Pimenov I.V. [Applying a genetic algorithm to optimize the discrete structure of a decision tree]. *Vestnik Sankt-Peterburgskogo gosudarstvennogo universiteta tekhnologii i dizajna. Seriya 1. Estestvennye i tekhnicheskie nauki – Bulletin of the Saint Petersburg state University of technology and design. Series 1. Natural and technical Sciences*. 2020. no. 3. pp. 55– 60. (In Russ.).

А.Г. БОРОДИНОВ, В.В. МАНОЙЛОВ, И.В. ЗАРУЦКИЙ, А.И. ПЕТРОВ,  
В.Е. КУРОЧКИН, А.С. САРАЕВ  
**МАШИННОЕ ОБУЧЕНИЕ В ЗАДАЧАХ BASE-CALLING ДЛЯ  
МЕТОДОВ СЕКВЕНИРОВАНИЯ НОВОГО ПОКОЛЕНИЯ**

*Бородинов А.Г., Манойлов В.В., Заруцкий И.В., Петров А.И., Курочкин В.Е., Сараев А.С.*  
**Машинное обучение в задачах base-calling для методов секвенирования нового поколения.**

**Аннотация.** Развитие технологий секвенирования следующего поколения (NGS) внесло существенный вклад в тенденции снижения затрат и получения массивных данных секвенирования. В Институте аналитического приборостроения РАН разрабатывается аппаратно-программный комплекс (АПК) для расшифровки последовательности нуклеиновых кислот методом массового параллельного секвенирования (Нанофор СПС). Алгоритмы обработки изображений, входящие в состав АПК, играют существенную роль в решении задач расшифровки генома. Финальной частью такого предварительного анализа сырых данных является процесс base-calling. Base-calling — это процесс определения нуклеотидного основания, которое генерирует соответствующее значение интенсивности в каналах флуоресценции для различных длин волн на кадрах изображения проточной ячейки для различных циклов секвенирования методом синтеза. Приведен обширный анализ различных подходов к решению задач base-calling и сводка распространенных процедур, доступных для платформы Illumina. Рассмотрены различные химические процессы, включенные в технологию секвенирования методом синтеза, вызывающие смещения в значениях регистрируемых интенсивностей, включая эффекты фазирование / префазирование (phasing/prephasing), затухания сигнала (signal decay) и перекрестные помехи (cross-talk). Определена обобщенная модель, в рамках которой рассматриваются возможные реализации. Рассмотрены возможные подходы машинного обучения (machine learning) для создания и оценки моделей, реализующих этап обработки base-calling. Подходы ML принимают различные формы, включая обучение без учителя (unsupervised), обучение с частичным привлечением учителя (semi-supervised), обучение с учителем (supervised). В работе показана возможность применения различных алгоритмов машинного обучения на основе платформы Scikit-learn. Отдельной важной задачей является оптимальное выделение признаков, выделенных в обнаруженных кластерах на проточной ячейке для машинного обучения. Наконец, на ряде данных секвенирования для приборов MiSeq Illumina и Нанофор СПС показана перспективность метода машинного обучения для решения задачи base-calling.

**Ключевые слова:** секвенирование нового поколения, base-calling, биоинформатика, машинное обучение.

**1. Введение.** Последние достижения в технологии высокопроизводительного секвенирования позволяют одновременно секвенировать миллионы фрагментов ДНК, создавая огромную пропускную способность, что в свою очередь требует более эффективных и точных методов анализа [1]. Поскольку стоимость секвенирования продолжает быстро снижаться, перспективы геномики расширяются. В то же время необходимо решить ряд проблем,

связанных с получением данных и статистическим анализом. Например, получение определенной последовательности нуклеотидов из данных интенсивностей флуоресценции усложнено из-за химических процессов, а также из-за многочисленных оптических аппаратных искажений. Переход от зашумленных данных интенсивностей флуоресценции, как результата процесса обработки изображений к последовательностям нуклеотидных оснований получил название *base-calling*. Точность определения каждого нуклеотида (*base-call*) представлена в виде показателя качества (*quality score*) и соотносится с каждым определенным нуклеотидным основанием. Под *ридом* (*read*, прочтение) понимается отдельная последовательность нуклеотидов, полученная в результате секвенирования. Эти статистические показатели далее используются в процессе выравнивания (*alignment*) полученных ридов с известным эталонным геномом в процессе таких видов анализа, как анализ ChIP-seq [2] или RNA-seq [3]. Следовательно, точность и качество процедуры *base-calling* могут напрямую влиять на дальнейший анализ последовательностей ДНК. В работе представлен обзор решений задачи *base-calling* и апробация нового алгоритма *base-calling*, основанного на методах машинного обучения.

## **2. Основы метода**

**2.1. Биологические основы.** Хромосома эукариот образуется из единственной и чрезвычайно длинной молекулы ДНК, которая содержит линейную последовательность множества генов. Хромосомы состоят из более мелких субъединиц, называемых нуклеотидами, каждая из которых содержит пентозный сахар, фосфатную группу и одно из четырех азотистых оснований (А, С, G, Т). Общепринято использование букв А, С, G и Т для обозначения нуклеотида, который содержит соответствующее азотистое основание. Важно отметить, что А связывается с Т и С связывается с G, поэтому без потери общности одна нить ДНК определяет другую (комплементарна ей).

**2.2. Высокопроизводительное секвенирование (NGS).** В-первых, следует получить несколько образцов генома из организма (стадия сбора или *acquisition*). Поскольку геномы многих организмов, как правило, очень длинные, проанализировать последовательность от начала до конца не представляется возможным. Поэтому отобранная ДНК разрезается на более мелкие фрагменты и делается несколько копий для усиления полезного сигнала. Эти шаги называются фрагментация (*fragmentation*) и амплификация (*amplification*), соответственно. Предыдущие шаги образуют стадию подготовки



геномного образца (genomic sample preparation) или подготовки библиотеки (library preparation) [4].

Основания нуклеотидов флуоресцентно помечены, и отдельные изображения в различных каналах флуоресценции получаются для каждого из четырех оснований А, С, G и Т. Каналы флуоресценции отличаются друг от друга длинами волн возбуждающего сигнала. Изображения далее обрабатываются для получения значений интенсивности для каждого нуклеотида [5]. Каждый из нуклеотидов генерирует сигнал флуоресценции на определенной длине волны. Base-calling — это процесс определения нуклеотидного основания, которое генерирует соответствующее значение интенсивности в каналах А, С, G и Т. Существует неопределенность при проведении процедуры base-calling, поэтому соответствующая оценка качества присваивается каждому выбранному основанию. Этот показатель качества (quality score) является функцией вероятности неправильного выбора нуклеотидного основания. В таблице 1 приведен список распространенных методов base-calling. Поскольку произошла фрагментация цепочки ДНК, информация о позициях расположения ридов теряется. Так как нас интересует весь геном или очень большие части генома, мы должны попытаться собрать фрагменты вместе, чтобы получить исходные позиции нуклеотидов в полной цепочке ДНК. Если существует эталонный геном, он может помочь нам определить местоположение секвенированных участков ДНК [6].

Несмотря на всю кажущуюся вариабельность методов base-calling наиболее широко используемым базовым алгоритмом является Bustard, несколько алгоритмов были построены с использованием Bustard в качестве отправной точки. Алгоритм Bustard основан на параметрической модели и применяет алгоритм Маркова для определения вероятности моделирования переходной матрицы фазинга/префазинга (phasing/prephasing), и матрицы cross-talk. Алгоритм Bustard предполагает, что матрица перекрестных помех постоянна для данного цикла секвенирования, и что эффекты фазинга одинаковы относительно всех нуклеотидных оснований.

В методах BayesCall and naiveBayesCall для оценки всех неизвестных параметров используется алгоритм максимума правдоподобия (EM), а полученная максимальная апостериорная вероятность используется в base-calling. При этом naiveBayesCall существенно выигрывает в оптимизации скорости вычислений. Основной мотивацией для модификации OnlineCall является создание вычислительно эффективного алгоритма с базовой моделью BayesCall. Оценка параметров выполняется с помощью unsupervised онлайн-

алгоритма EM, а полученные апостериорные вероятности используются для base-calling. В алгоритме Softy используется также обобщенная модель base-calling, но апостериорные вероятности, используемые для процедуры base-calling, получаются либо с помощью алгоритма прямого-обратного прохода (FB), либо алгоритма soft-output Витерби (SOVA).

Таблица 1. Сводка распространенных процедур base-calling

Наименование	Год	Формат ввода	Мера качества	Тип модели
BlindCall [7]	2014	RTA	None	Blind deconvolution
3Dec [8]	2017	RTA	Phred	Parametric
freelbis [9]	2013	IPAR, Firecrest, Bustard reads	Phred	Nonparametric SVM
Softy [10]	2013	Firecrest	Probability	Parametric
AYB [11]	2012	RTA	Phred	Nonparametric
OnlineCall [12]	2012	Firecrest	Probability	Parametric
BM-BC [13]	2012	Firecrest	Unknown	Parametric
ParticleCall [14]	2012	Firecrest	Probability	Parametric
TotalReCaller [15]	2011	Firecrest, RTA	None	Parametric
naiveBayesCall [16]	2010	Firecrest	Probability	Parametric
Srfim [17]	2009	IPAR, Firecrest	Phred	Parametric
BayesCall [18]	2009	Firecrest	Probability	Parametric
Ibis [19]	2009	IPAR, Firecrest, Bustard reads	Phred	Nonparametric SVM
Rolexa [20]	2008	IPAR, Firecrest	Phred	Parametric
Alta-Cyclic [21]	2008	Firecrest c Bustard reads	Unknown	Nonparametric SVM

Метод Rolexa предлагает новый алгоритм base calling, использующий кластеризацию на основе параметрических моделей и статистические оценки для выявления неоднозначных нуклеобаз и кодирования их символами IUPAC (International Union of Pure and Applied Chemistry). Используются также оптимальные вложенные теги (sub-tags), используя оценку, основанную на информационном содержании, чтобы удалить неопределенные основания ближе к концу прочтения.

All Your Base (AYB) base-caller — это еще один метод, который полностью отличается от Bustard и семейства модификаций Bustard. Он основан на модели, но не предполагает наличия определенных

распределений наблюдаемых интенсивностей, что делает его непараметрическим методом. Истинная матрица нуклеобаз получается с помощью алгоритма Витерби, который находит наиболее вероятную последовательность нуклеотидов, а алгоритм прямого/обратного распространения (Forward/Backward) используется для получения апостериорных вероятностей, что приводит к оценкам показателей качества.

Метод улучшенной базовой идентификации (Improved Base Identification System, Ibis) для проведения base-calling использует схему статистического обучения. Этот подход использует метод опорных векторов (SVM) для поиска закономерностей в данных. SVM используют полиномиальные ядра с входными данными текущего, предыдущего и последующего циклов. Этап обучения находит оптимальную гиперплоскость, которая может разделить паттерны между интенсивностями четырех каналов. Для обучения SVM требуются достаточно объемные результаты выравнивания прочтений относительно референтного генома.

Новейший метод base-calling 3Dec учитывает эффект cross-talk между ближайшими кластерами.

**3. Base-calling в секвенировании методом синтеза (Sequencing-by-synthesis).** Чтобы прояснить процедуру base-calling обсудим принципиально новые методы химической подготовки анализируемых образцов к проведению генетического анализа., используемую в платформе Illumina и называемую секвенированием методом синтеза (Sequencing-by-synthesis) [22].

После получения образцов генома двухцепочечная ДНК случайным образом фрагментируется посредством обработки ультразвуком. Короткие известные последовательности, называемые адаптерами, лигированы к концам двухцепочечных фрагментов. Адаптеры используются для прикрепления фрагментов к поверхности, на которой располагаются короткие последовательности ДНК. Чтобы генерировать достаточное количество ДНК-материала для последовательности используют стадии полимеразной цепной реакции для создания нескольких дубликатов молекул ДНК / адаптера. Эти молекулы денатурируются, превращаясь в одноцепочечные, а затем связываются с поверхностью стеклянной проточной ячейки, содержащей плотное множество (газон) прикрепленных олигопраймеров. Проточные кюветы (Flowcells), как правило, содержат 8 дорожек (lanes), а меньшие субъединицы, называемые плитками (tiles), составляют одну дорожку. Количество плиток варьируется между версиями Illumina, но обычно содержит до 100

плиток. Олигопраймеры комплементарны одному концу одноцепочечной цепи матрицы. Полимеразный фермент завершит построение комплементарной цепи, и исходный шаблон будет смыт. Затем происходит "стыковочная" амплификация (bridge amplification) с целью несколько раз скопировать шаблоны, чтобы получить почти 1000 копий идентичных одиночных нитей и достаточно плотный кластер. Последовательности синтезируются по одному основанию за раз параллельно во всей проточной ячейке. Процесс присоединения одного азотистого основания (нуклеотида) будем называть циклом. Для каждого цикла добавляется ДНК-полимераза и молекулы, состоящие из флуоресцентно меченных оснований с присоединенным обратимым терминатором, которые позволяют последовательно проводить процесс, предотвращая присоединение более чем одного основания. После присоединения одного основания к цепочке лазер возбуждает кластер, генерируя излучение флуоресценции на определенной длине волны.

В системе параллельного секвенирования прибора «Нанофор СПС» используются четыре видеокамеры по числу типов нуклеотидов. Каждая из видеокамер настроена на регистрацию одного из типов нуклеотидов: «А», «С», «G» или «Т». Сигнал флуоресценции возбуждается двумя лазерами в определенном диапазоне излучения видимого света. Регистрируемое излучение пропускается через различные светофильтры, соответствующие длинам волн флуоресценции каждого из четырех красителей, которыми специфично помечены нуклеотиды. Таким образом, каждая из видеокамер регистрирует изображения кластеров молекул ДНК, на конце которых расположены нуклеотиды определенной «буквы».

Четыре набора изображений фиксируются через четыре различных фильтра, один для каждого из красителей, используемых соответственно для каждого из оснований. Терминаторы удаляются, чтобы обеспечить включение основания в следующий цикл секвенирования. Это продолжается до тех пор, пока полное прочтение (read) не будет полностью выполнено.

Обработка изображения в заданном цикле дает множество из четырех интенсивностей, где каждое из четырех значений представляет интенсивность, считанную через определенный оптический фильтр. Выходные данные этапа обработки изображений содержат значения интенсивности для всех прочтений (ридов) и циклов, (x,y) -координаты позиций кластера на проточной ячейке и, наконец, нуклеотидные основания, получаемые как результат процедуры base-calling. На рисунке 1 показан пример значений

интенсивности для одного прочтения. Самый элементарный метод base-calling выдал бы последовательность на основе значения максимальной интенсивности, и в этом случае это будет последовательность GGAAAATGAG. В некоторых циклах максимум очевиден, но в других случаях самые большие и вторые по величине значения интенсивностей достаточно близки, см. циклы 7 и 3 соответственно. На соотношениях максимальной и остальных интенсивностей в различных каналах строится построение фильтров, отсекающих кластеры недостаточного качества (chastity, purity). Только статистические методы помогут моделировать эти сложные ситуации и разрешить коллизии. Несмотря на то, что технология секвенирования бурно развивается в последние годы, особенности в химии протекания процессов в проточной кювете продолжают вызывать сложности в процессе определения нуклеотидов из данных зафиксированных интенсивностей.

```
> intensities
  1  2  3  4  5  6  7  8  9 10
A -17.7 16.5 847.7 1077.6 1044.7 1039.9 17.4 55.6 1015.9 63.5
C 9.2 34.8 651.8 835.4 754.6 708.4 38.1 50.8 736.1 36.5
G 1121.5 955.8 -6.4 15.4 9.9 3.9 37.2 1146.9 37.4 1234.4
T 588.9 494.9 14.8 3.6 5.4 25.6 639.2 647.4 30.6 670.7
```

Рис. 1. Пример значений интенсивности для первых 10 циклов чтения. Строки представляют интенсивности в определенных каналах, а столбцы представляют циклы. Элементарный base-caller выдал бы последовательность для прочтения GGAAAATGAG

Химические процессы, включенные в технологию секвенирования методом синтеза, вызывают некоторые смещения в значениях регистрируемых интенсивностей, включая эффекты фазирование / префазирование (phasing/prephasing), затухания сигнала (signal decay) и перекрестные помехи (cross-talk) [23]. В любом данном кластере возможно, что небольшая часть синтезированных цепочек будет отставать в синхронизации присоединения нуклеотидов по сравнению с остальными цепями. Например, ферменты могут не сработать, в результате чего основание не включится в цепочку ДНК. Нить ДНК может просто отставать, перегонять планируемый процесс присоединения по одному основанию или полностью становиться неактивной, что приводит к неточным показаниям интенсивностей в цикле. Это явление называется фазированием (phasing или lagging). Префазирование (prephasing или leading) происходит, когда небольшая часть цепей забегает вперед и включает сразу два нуклеотида. Prephasing имеет те же последствия, что и phasing. Как phasing, так и prephasing могут привести к неправильному определению

нуклеотидного основания. На рисунке 2 показано, как phasing и prephasing могут выглядеть с точки зрения анализа интенсивностей. Интенсивности канала A для одного кластера отображаются в зависимости от цикла, а в нижней части каждой строки показано основание. Следовательно, более высокие значения интенсивности соответствуют нуклеотидному основанию, полученному из стандартной процедуры base-calling от Illumina. Фиолетовая стрелка указывает на то, что эффект phasing произошел, поскольку после нуклеотида A наблюдается более высокий сигнал, а черная стрелка указывает на то, что эффект prephasing произошел, поскольку существует более высокий сигнал от нуклеотидных оснований, не являющихся A. Во время процесса секвенирования проточная ячейка промывается несколько раз, и возможно, что сам секвенируемый материал также будет смыт. Кроме того, неспособность ферментов приводит к неактивности в некоторых нитях секвенируемого ДНК. Такие потери вызывают снижение интенсивности сигнала и увеличение шума в процессе секвенирования. Это явление известно как затухание сигнала (signal decay), и, как очевидно, существует корреляция между длиной циклов секвенирования и количеством потерянного материала. Затухание сигнала можно увидеть на рисунке 3. Этот график показывает максимальную и минимальную из четырех интенсивностей, которые мы можем рассматривать как сигнал и шум соответственно, за несколько прочтений и циклов. Наклон сигнала и шума показывают тенденцию к уменьшению сигнала и увеличению шума по всему циклу. Флуоресцентные красители, используемые в платформе Illumina, имеют частично перекрывающуюся частоту излучения, что приводит к некоторой корреляции показаний интенсивностей. На рисунке 4 показана диаграмма интенсивностей по каналам A и C. Красное, синее и серое облако точек показывают корреляцию интенсивностей A/C для оснований A, C, G и T соответственно. По мере увеличения интенсивности A интенсивность C также увеличивается для тех оснований, которые были названы A, и, когда интенсивность C увеличивается, интенсивность A также увеличивается для тех оснований, которые названы C. Корреляция между каналами интенсивности A/C для оснований A и C не обязательно совпадают.

Частота излучения используемых красителей частично перекрывается, что приводит к корреляции показаний интенсивностей. Это приводит к тому, что с ростом интенсивности A растет и интенсивность C, и наоборот. Подобное явление происходит и с каналами G и T.

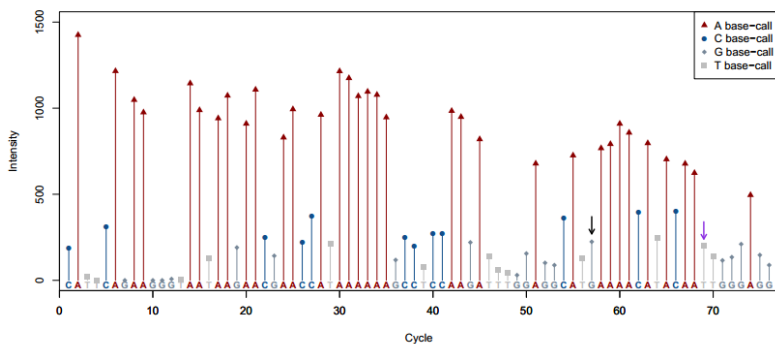


Рис. 2. На этом рисунке показано, где могут происходить phasing и prephasing и как могут выглядеть при этом значения интенсивностей. Phasing может произойти в цикле 69 (фиолетовая стрелка), потому что интенсивность в канале A составляет около 200, даже если соответствующий этому циклу нуклеотид был T. Следует обратить внимание, что предыдущий нуклеотид был A, и аналогично в цикле 57 интенсивность в канале также составляет около 200, даже если соответствующее каналу основание было G. Это показывает, что процесс prephasing мог произойти, поскольку нуклеотидом в следующем цикле является A [23]

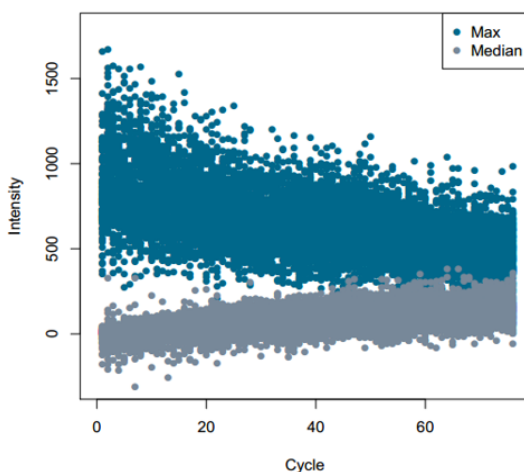


Рис. 3. Максимум из четырех интенсивностей и медиана оставшихся трех значений нанесены на график зависимости от цикла секвенирования. Тенденция к снижению сигнала и тенденция к увеличению шума показывают, как затухание сигнала влияет на значения интенсивности в разных циклах [23]

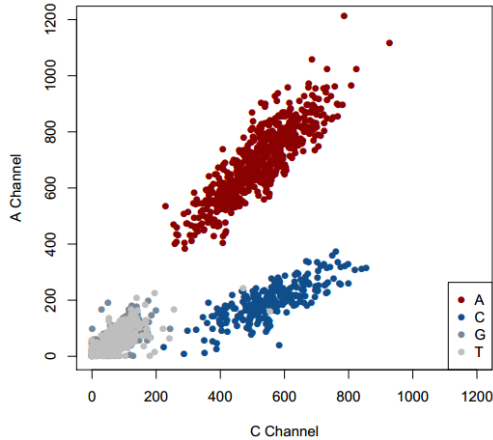


Рис. 4. Интенсивности каналов А и С нанесены для множества прочтений определенном цикле 30. Цвет показывает нуклеотид, поэтому мы можем видеть три облака точек; красный - для А, синий - для С, а серые - для оснований G и T. Поскольку интенсивность А увеличивается для оснований А, интенсивность С также увеличивается. Эта корреляция значений интенсивности является результатом процесса cross-talk [23]

## 4. Модель Base-calling

**4.1. Нотация.** Сначала определим математические обозначения для модели base-calling. В таблице 2 даны обозначения, которые будут использоваться для описания модели base-calling. В таблице 3 приведен пример того, как четверки интенсивностей представлены в математических обозначениях. Затем определим обобщенную модель base-calling.

Таблица 2. Обозначения, которые будут использоваться для описания единой статистической модели base-calling

<b>i</b>		Индекс прочтений (read index), $i = 1, 2, \dots, N$
<b>j</b>		Индекс циклов (cycle index), $j = 1, 2, \dots, J$
<b>k</b>		Индекс каналов (signal channel index), $k = A, C, G, T$
<b>Z<sub>i</sub></b>	<b>4 x J</b>	Массив интенсивностей после коррекции Illumina
<b>B</b>	<b>4 x J</b>	Массив коррекции фона (background correction)
<b>Y<sub>i</sub></b>	<b>4 x J</b>	Массив на Наблюдаемые интенсивности
<b>M</b>	<b>4 x 4</b>	Cross-talk матрица
<b>X<sub>i</sub></b>	<b>4 x J</b>	Массив истинных интенсивностей
<b>P</b>	<b>J x J</b>	Phasing/prephasing матрица
<b>D</b>	<b>J x J</b>	Signal decay матрица
<b>E<sub>i</sub></b>	<b>4 x J</b>	Error term матрица



Таблица 3. Пример представления интенсивностей согласно введенной нотации

Наблюдаемые интенсивности  $Y_i$

	1	2	3	4	5	6	7	8	9	10
A	$Y_{i1A}$	$Y_{i2A}$	$Y_{i3A}$	$Y_{i4A}$	$Y_{i5A}$	$Y_{i6A}$	$Y_{i7A}$	$Y_{i8A}$	$Y_{i9A}$	$Y_{i10A}$
C	$Y_{i1C}$	$Y_{i2C}$	$Y_{i3C}$	$Y_{i4C}$	$Y_{i5C}$	$Y_{i6C}$	$Y_{i7C}$	$Y_{i8C}$	$Y_{i9C}$	$Y_{i10C}$
G	$Y_{i1G}$	$Y_{i2G}$	$Y_{i3G}$	$Y_{i4G}$	$Y_{i5G}$	$Y_{i6G}$	$Y_{i7G}$	$Y_{i8G}$	$Y_{i9G}$	$Y_{i10G}$
T	$Y_{i1T}$	$Y_{i2T}$	$Y_{i3T}$	$Y_{i4T}$	$Y_{i5T}$	$Y_{i6T}$	$Y_{i7T}$	$Y_{i8T}$	$Y_{i9T}$	$Y_{i10T}$

↓

	1	2	3	4	5	6	7	8	9	10
A	-17.7	16.5	847.7	1077.6	1044.7	1039.9	17.4	55.6	1015.9	63.5
C	9.2	34.5	651.8	835.4	754.6	708.4	38.1	50.8	736.1	36.5
G	1121.5	955.8	-6.4	15.4	9.9	3.9	37.2	1146.9	37.4	1234.4
T	588.9	494.9	14.8	3.6	5.4	25.6	639.2	647.4	30.6	670.7

**4.2. Обобщённая модель base-calling.** Все методы base-calling работают со значениями интенсивности, полученными во время секвенирования. После рассмотрения различных методов base-calling становится ясно, что в методах, используемых для моделирования интенсивностей, есть немало общего. Эти методы варьируются от параметрических до непараметрических и статистических моделей, основанных на полностью эмпирических методах машинного обучения. Попытаемся определить общую модель, которая объединяет подавляющее большинство методов base-calling. Классификация методов base-calling приведена на рисунке 5. Сначала определим следующую обобщенную модель для base-calling как:

$$Z_i - B = Y_i = MX_iPD + E_i \tag{1}$$

В формуле (1) использованы обозначения согласно таблице 2, содержащей краткие описания параметров, а также предполагаемые размеры матриц. Используются следующие индексы; прочтение (read) / кластер  $i$  для  $i = 1, 2, \dots, N$ , цикл  $j$  для  $j = 1, 2, \dots, J$  и сигнальный канал  $k$  для  $k$  из  $\{A, C, G, T\}$ . Обычно индексы 1, 2, 3 и 4 используются взаимозаменяемо с A, C, G и T соответственно.  $Z_i$  - это интенсивности перед коррекцией фона Illumina B.  $Y_i$  - это наблюдаемые интенсивности. Наш интерес заключается в получении значений  $X_i$ , интенсивности для данного нуклеотидного основания. Например, если истинный нуклеотид представляет собой G, то соответствующий столбец  $X_i$  будет  $(0,0,1,0)^T$ . Строки и столбцы как  $Y_i$ , так и  $X_i$  представляют каналы и циклы соответственно. Cross-talk, phasing/prephasing и signal decay моделируются матрицами M, P и D

соответственно. Матрица  $E_i$  предназначена для представления статистической ошибки.

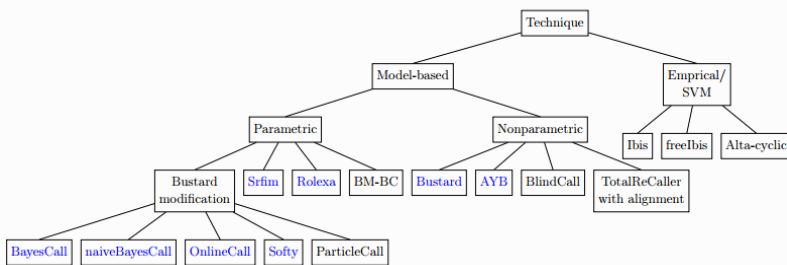


Рис. 5. Различные методы моделирования base-calling, используемые в настоящее время

Рассмотрим уравнение (1) с матрицей  $M$ :

$$M = \begin{pmatrix} 1 & m_{12} & m_{13} & m_{14} \\ m_{21} & 1 & m_{23} & m_{24} \\ m_{31} & m_{32} & 1 & m_{34} \\ m_{41} & m_{42} & m_{43} & 1 \end{pmatrix}, \quad (2)$$

где каждый элемент  $m_{rs}$  матрицы перекрестных помех  $M$  указывает величину наблюдаемой интенсивности в канале  $s$ , генерируемую сигналом от нуклеотида  $r$ ; для каждого  $r, s = 1, 2, 3, 4$ . Элементы  $M$  оцениваются посредством итеративного процесса для оценки элементов матрицы, предложенного в работе [24]. Необработанные интенсивности флуоресценции рассматриваются как линейные комбинации вклада флуоресцентных красителей и каналов. Подход состоит в том, чтобы оценивать каждый элемент  $m_{rs}$  и  $m_{sr}$ , сначала рассматривая только эти каналы,  $r$  и  $s$ . Производится разбиение множества точек на бины по квантилям. Для тех значений, чей  $r$ -й компонент попадает в данный интервал квантилей, возьмем пару интенсивностей в каналах, имеющую минимальное значение в  $s$ -компоненте. По этим выбранным парам точек строится регрессия по норме  $L1$  и получаем оценку наклона кривой. Наклон является оценкой  $m_{rs}$ . Поменяем местами компоненты интенсивностей и повторим процедуру, чтобы получить оценку  $m_{sr}$ . Сделаем это для всех пар  $r$  и  $s$  с  $r \neq s$ , чтобы получить оценки элементов матрицы cross-talk (2).

Phasing and prephasing моделируются с помощью матрицы вероятности перехода  $Q$  ( $J \times J$ ), отслеживающей положение

обратимых терминаторов. Напомним, что Phasing происходит, когда позиция терминатора отстает, а rephasing происходит, когда позиция терминатора опережает небольшую долю в ДНК составе шаблона. Таким образом, элементы  $Q$  в терминах  $u$  (текущей) и  $v$  (следующей) позиции терминатора, могут быть смоделированы как:

$$Q_{uv} = \begin{cases} p & v = u \\ 1 - p - q & v = u + 1 \\ q & v = u + 2 \\ 0 & \text{в ином случае} \end{cases}, \quad (3)$$

где  $p$  – вероятность phasing,  $q$  – вероятность rephasing, а  $1 - p - q$  – вероятность нормального включения нуклеотида. Поскольку считается, что phasing и rephasing происходит не только в границах одного цикла от текущего цикла, и что эти эффекты могут сохраняться в течение нескольких циклов до и после, рассматривается матрица вероятностей  $t$ -шагового перехода,  $Q^t$ . При этом,  $(u, v)$ -й элемент  $Q^t$  представляет вероятность того, что в данной цепочке ДНК шаблона в цикле  $u$  переместится в цикл  $v$  после  $t$  циклов. Таким образом,  $(v, t)$ -й элемент матрицы phasing / rephasing уравнения (1) принимает вид:

$$p * P_{v-1,t} + (1 - p - q) * P_{v-1,t-1} + q * P_{v-1,t-2}, \quad (4)$$

для  $v = 2, \dots, J$  и  $t = 1, 2, \dots, J$  с  $P_n = 1 - p - q$ ,  $P_{12} = q$ , а остальные столбцы первой строки равны 0. Параметры  $p$  и  $q$  оцениваются по возрастающей корреляции интенсивностей в течение первых нескольких циклов секвенирования. Потери сигнала оцениваются диагональной матрицей:

$$D = \left[ \text{diag} \left( \frac{\overline{W}_1}{W_1}, \frac{\overline{W}_2}{W_2}, \dots, \frac{\overline{W}_j}{W_j} \right) \right]^{-1}, \quad (5)$$

$$\overline{W}_j = \sum_{i=1}^N (Y'_{ijA} + Y'_{ijC} + Y'_{ijG} + Y'_{ijT}),$$

$$Y'_{ij} = M^{-1}Y_{ij},$$

где диагональные элементы представляют собой результат перенормировки концентраций путем взятия среднего значения интенсивностей с поправкой на перекрестные помехи и использования его в качестве нормализующего фактора.

**5. Машинное обучение в задаче base-calling.** Введение в задачи ДНК секвенирования прикладного машинного обучения (machine learning, ML) включает в себя создание и оценку моделей, использующих алгоритмы, способные распознавать, классифицировать и прогнозировать определенные результаты на основе данных. Подходы ML принимают различные формы, включая обучение без учителя (unsupervised), обучение с частичным привлечением учителя (semi-supervised), обучение с учителем (supervised) (рисунок 6) [25]. Например, часто целью supervised ML, применяемого к данным секвенирования, является построение правила принятия решения (т.е. модели) из набора собранных наблюдений для прогнозирования метки ответа немеченого образца с использованием набора измерений. Входные переменные часто при этом называют признаками (features), а соответствующие выборки – наблюдениями (observations).

Принципиальное различие между unsupervised (USML) и supervised машинным обучением (SML) заключается в том, что в USML образцы разделяются с использованием функций без какой-либо ссылки на метки ответов и строится прогноз к какому кластеру может принадлежать ответ, тогда как SML строит гиперповерхности, которая разделяет базовое векторное пространство на наборы, по одному для каждого класса [26] (рисунок 7).

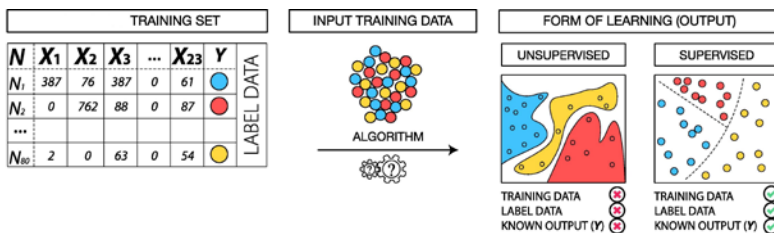


Рис. 6. Схематическое представление unsupervised и supervised форм ML и нескольких конкретных методов, предсказывающих три метки условного ответа (синий/красный/желтый) [25]. Рисунок изображает общую матрицу, содержащую наблюдения или образцы ( $N$ ), признаки ( $X_1, \dots, X_{23}$ ) и метки нескольких классов ( $Y$ ). Входные данные обрабатываются, чтобы либо предсказать к какому множеству принадлежит то или иное наблюдение (unsupervised), либо найти наилучшую границу, разделяющую множества (supervised). При этом следует понимать отличие множества от класса, хотя оба содержат объекты, близкие по своим свойствам. Классы и их свойства задаются априорно, в то время как множества формируются исключительно на основе близости значений признаков объектов, а свойства выясняются в процессе их содержательной интерпретации

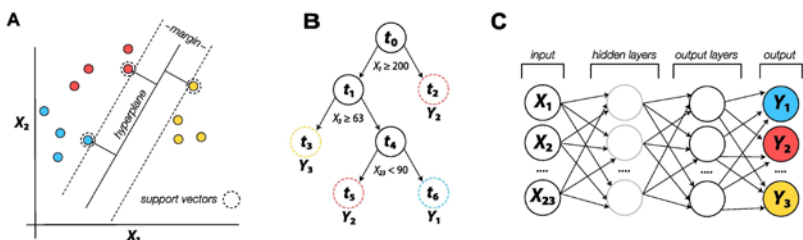


Рис. 7. Различные методы машинного обучения [25]. (А) Линейный классификатор метода опорных векторов SVM, демонстрирующий разделение между метками классов, где гиперплоскость максимизирует расстояние между ближайшими точками на множестве данных для обучения (training set). Опорные векторы относятся к трем векторам положения, проведенным из начала точек выборки (пунктирная окружность) с целью максимизации расстояния между оптимальной гиперплоскостью и опорными векторами с целью проведения границы решения. (В) Дерево решений, построенное для классификации выборок в  $Y$  на основе значений входных признаков. Деревья начинаются с корневого узла ( $t_0$ ) и растут до различных листовых узлов (заштрихованные круги), чтобы закончиться конечным узлом, так что агрегированные прогнозы по конечным узлам усредняются по  $k$ -деревьям для наилучших прогнозов  $\hat{Y}$ . (С) Нейронная сеть, отображающая структуру последовательных слоев. Входные значения  $X$  передаются на следующий скрытый уровень, который передает взвешенные соединения на выходной уровень для предсказания  $\hat{Y}$

**5.1. Обучение без учителя (USML).** Методы обучения без учителя часто используются для первоначального исследовательского анализа многомерных данных секвенирования и для выработки гипотез с целью последующего анализа, поскольку они помогают в визуализации и прояснении структуры данных, которые не имеют заранее определенных меток ответов, присвоенных наблюдениям. Эти методы работают с целью выявления однородных подгрупп путем кластеризации данных или для обнаружения аномалий путем поиска закономерностей с помощью методов уменьшения размерности (DR).

**5.1.1. Метод k-средних (K-means clustering).** Цель метода k-средних [27] состоит в том, чтобы сгруппировать выборки в определенное количество ( $k$ ) непересекающихся подгрупп (кластеров) с использованием расстояний, рассчитанных между объектами, чтобы каждая точка данных принадлежала только к одной группе. Этот метод назначает точки данных кластеру таким образом, чтобы сумма квадратов расстояний между точками данных и центроидом (среднее значение всех точек данных, представленными геометрическим центром кластера) была минимизирована. За счет уменьшения

внутрикластерных вариаций точки данных располагаются так, чтобы построить кластер, который принимает сферическую форму, окружающую центр тяжести, и это позволяет различным подгруппам данных оставаться как можно дальше друг от друга. Недостатком K-средних является то, что они не могут хорошо строить кластеры на точках данных, которые образуют данные более сложной формы нежели круговые. Дополнительным ограничением является то, что требуется предварительно определить определенное количество кластеров, что иногда является серьезным ограничением.

**5.1.2. Метод главных координат (Principal Coordinates Analysis, PCoA).** В анализе PCoA [28] данные разбиваются на компоненты, чтобы максимизировать линейную корреляцию между точками данных в матрице различий, задаваемых входными признаками. Посредством «преобразования координат» количество точек данных  $x$  заменяется новыми полученными координатами  $y$ , что снижает размерность набора данных за счет отбрасывания координат, которые могут не удовлетворять заданному порогу дисперсии данных секвенирования. Этот метод сохраняет глобальную структуру данных, проецируя их на пространство с меньшим количеством измерений. Точки размещаются так, чтобы попарные расстояния между ними в новом пространстве как можно меньше отличались от эмпирически измеренных расстояний в пространстве признаков изучаемых объектов. Метод главных координат PCoA, или многомерное шкалирование (MDS, multidimensional scaling), во многом похож на метод главных компонент PCA [29] -, но вместо корреляционной матрицы выполняет вычисление собственных значений и собственных векторов симметричной матрицы расстояний. Так можно компенсировать некоторые отклонения от предпосылок в отношении статистического распределения данных, принятых для корреляционного анализа, но одновременно возникает проблема выбора подходящей метрики дистанции.

**5.1.3. Стохастическое вложение соседей с  $t$ -распределением (t-distributed Stochastic Neighbor Embedding, t-SNE).** Стохастическое вложение соседей с  $t$ -распределением — это метод машинного обучения визуализации данных, разработанный Лоренсом ван дер Маатеном и Джеффри Гинтоном [30]. Это удобный метод нелинейного снижения размерности путем вложения многомерных данных в двух- или трехмерное пространство для дальнейшей визуализации. В частности, он отображает каждую точку многомерного пространства в двух или трехмерную точку евклидова пространства так, что подобные объекты располагаются рядом, а непохожие объекты

соответствуют удаленным точкам с высокой вероятностью. Алгоритм t-SNE состоит из двух основных этапов. Первоначально, t-SNE создает распределение вероятностей по парам многомерных объектов таким образом, что подобные объекты имеют высокую вероятность быть выбранными, в то время как непохожие точки имеют очень малую вероятность быть выбранными вместе. Далее, t-SNE определяет подобное распределение вероятностей для точек в карте низкомерного пространства и минимизирует разногласия по расстоянию Кульбака-Лейблера между двумя распределениями по месту расположения точек на карте.

**5.2. Машинное обучение с учителем (SML).** Машинное обучение с учителем (SML) — это более сложная форма изучения наборов данных секвенирования, поскольку, в отличие от неконтролируемых методов, метки ответов (**Y**) назначаются каждому образцу в наборе данных, группируя их в значимые категории. При этом более целенаправленное исследование данных может быть достигнуто, поскольку модель обучается на специальном наборе признаков (**X**) (training set) для создания правил, в которых они могут служить предикторами явлений или результатов. Другими словами, с помощью некоторой функции происходит установление соответствия между набором признаков (**X**) и метками ответов (**Y**). После обучения эта модель может принимать новые немаркированные образцы с аналогичными признаками (testing set) и прогнозировать их результат (**Y**) на основе того, что она узнала из обучающего набора. SML можно использовать с непрерывными числовыми выходными данными или категориальными выходными данными. В следующем разделе представлен обзор некоторых из наиболее распространенных алгоритмов SML для задач прогнозирования на основе данных секвенирования.

**5.2.1. Случайные леса (Random Forests, RF).** Случайные леса [31] широко используются для решения различных задач биоинформатики. Этот метод строит несколько лесов, состоящих из деревьев решений, используя информацию, содержащуюся во входных функциях, для последовательного разделения выборок на основе присвоенных им значений (**Y**). Леса управляются начальной загрузкой и критерием разделения узлов, который использует информацию, содержащуюся в случайном подмножестве признаков. Тот факт, что в каждом лесу строятся сотни или тысячи деревьев решений с использованием подмножества как выборок, так и признаков, позволяет получить совокупное среднее значение прогнозов, сделанных в каждом конечном узле. Таким образом, RF является

идеальной основой для последовательного выявления «истинных эффектов» в сложных и разнородных данных. Дополнительными факторами, которые делают RF привлекательным на практике, является то, что они являются готовыми, поддающимися вычислительной обработке и высокопроизводительными классификаторами, устойчивыми к выбросам, зашумленным и нелинейным данным и ошибкам в метках ответов. Метод RF менее подвержен эффекту переобучения (overfitting), чем другие методы SML, что способствует его привлекательности.

### **5.2.2. Градиентный бустинг (Gradient Boosting, GB).**

Градиентный бустинг [26] или метод повышения градиента, когда он используется для деревьев решений, представляет собой ансамблевый метод, в котором используется процесс повышения (бустинг) для последовательного объединения отдельных алгоритмов обучения (деревьев решений) с целью получения более удачного решения (learner). Деревья с усилением градиента отличаются от деревьев RF тем, что каждое дерево решений строится последовательно в попытке уменьшить ошибки предыдущего дерева, а не параллельно. Кроме того, каждое дерево, построенное в GB, имеет фиксированный размер и соответствует исходным данным, а не выборкам начальной загрузки, как это делается в RF. Подобно RF, можно использовать как числовые, так и категориальные признаки, но на практике может быть сложнее найти оптимальные параметры настройки для хорошей подгонки модели, такие как количество оценщиков дерева.

**5.2.3. Метод опорных векторов (Support Vector Machines, SVM).** Целью метода SVM [32] является нахождение наилучшего обобщенного линейного разделения меток ответа (Y) гиперплоскостью, которая максимизирует разницу между различными значениями Y (или каждым классом) в помеченных метками данных. Граница решения находится таким образом, что каждый класс отделяется при максимально возможном расстоянии от ближайших выборок (называемых опорными векторами и определяющими эту границу решения). SVM относятся к категории линейных дискриминантных методов SML. Алгоритм основан на допущении, что чем больше расстояние между параллельными разделяющими гиперплоскостями, тем меньше будет средняя ошибка классификатора. Эти модели могут работать с различными типами объектов, но по своей природе их трудно интерпретировать, поскольку они не дают прямых оценок вероятности в полученной оценке.

**5.2.4. Логистическая регрессия.** В отличие от обычной регрессии метод логистической регрессии [33] не предсказывает



значение числовой переменной на основе выборки начальных значений. Вместо этого значение функции представляет собой вероятность того, что данное исходное значение принадлежит определенному классу. Основная идея логистической регрессии заключается в том, что пространство начальных значений может быть разделено линейной границей (т.е. прямой) на две области, соответствующие классам. При этом оптимально использовать регуляризацию, как метод, используемый для уменьшения переобучения. Гребневая регрессия или ридж-регрессия [34] удовлетворяет модели, которая уменьшает дисперсию без увеличения систематической ошибки и это достигается путем наложения ограничений на сложность параметров. Этот метод добавляет штрафной член к функции потерь, позволяя ограничить сложность параметра. Гребневая регрессия может использоваться как для классификации, так и для регрессии, но может быть весьма требовательной к вычислительным ресурсам в случае большого входного пространства признаков.

**5.2.5. Нейронные сети (Neural Networks).** Нейронные сети [35] используют архитектуру построения иерархической модели, в которой несколько структурированных сетей взаимосвязанных узлов (нейронов) строятся с весами, соответствующими каждому ребру сети, чтобы обеспечить сопоставление входных данных  $X$  с ответами  $Y$ . Сети связаны между собой через механизм распространения с прямой связью (feed-forward propagation), где каждый нейрон получает входные данные от предыдущих нейронов. Сеть начинается с входных слоев, которые связаны с каждым нейроном в одном или нескольких скрытых слоях, которые используют алгоритм обратного распространения (backpropagation algorithm) для максимизации весов, размещенных на каждом нейроне, для улучшения прогнозирования. Этот процесс является итеративным, когда последний скрытый слой встречается с выходным слоем для получения прогнозируемого вывода ответа ( $Y$ ). Нейронные сети очень перспективны в своей способности идентифицировать сложную структуру в многомерных и сложных наборах данных. Нейронные сети часто называют методами «черного ящика», поскольку бывает сложно интерпретировать способы принятия решения.

**6. Применение методов машинного обучения для решения задачи base-calling.** Для применения различных алгоритмов машинного обучения используется платформа Scikit-learn [36], поддерживая простой в использовании интерфейс, тесно интегрированный с языком Python. Основным объектом при этом

является *estimator*, который реализует метод *fit*, отвечающий за подгонку модели на предоставленной обучающей выборке. Классы модели, принадлежащие классу методов обучения с учителем, например, *SVM*, также реализуют метод *predict*, обеспечивающий составление прогноза по обучающей выборке. Кроме того, некоторые наследники класса *estimator*, т.н. *transformer*-классы (как, например, *PCA*) имеют метод *transform*, который позволяет менять входные данные модели. Класс *estimator* в общем случае реализует метод *score*, который позволяет посчитать показатель выбранной метрики качества. Другим важным классом *Scikit-learn* является *cross-validation iterator*, предоставляющий разнообразные методы скользящего контроля.

Последним этапом процесса секвенирования NGS является анализ изображения, суть которого состоит в идентификации кластеров с последующим проведением операции *base-calling*. В работе [5] подробно описывались подходы для выполнения задачи обнаружения кластеров. Положения и оценённые радиусы такого кластера затем используются для извлечения ряда характеристик из каждого кластера и его непосредственного фона. Эти функции затем используются в качестве атрибутов, то есть служат входными данными для алгоритмов машинного обучения. Извлечение признаков выполняется путем изучения интенсивности света каждого пикселя определенной прямоугольной области кластера и некоторого окружающего фона. Затем вычисляются восемь статистических показателей: четыре для фона (*background*) и четыре для сердцевины кластера (*foreground*). Этот процесс показан на рисунке 8.

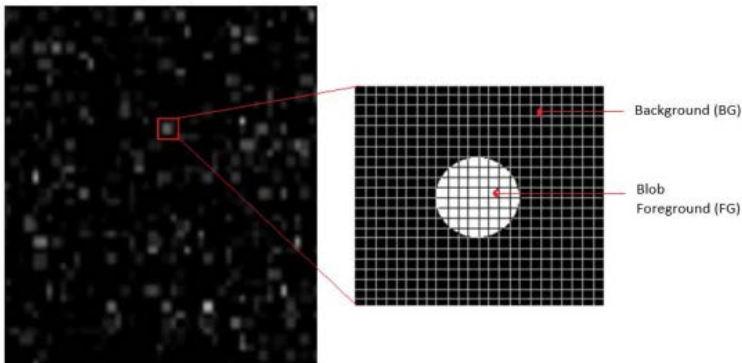


Рис. 8. Отбор признаков из “сырых” данных секвенирования NGS [37]

Извлекаются следующие статистические показатели: для фона (background) - max, mean, median и mode, для центральной зоны кластера (foreground) - max, mean, pct90 и pct99, где max — максимальное значение интенсивности, mean — среднее арифметическое значение, mode — наиболее часто встречающееся значение и pct90, и pct99 — 90-й и 99-й проценти соответственно. Каждый набор изображений, используемых для процесса base-calling одного цикла, состоит из четырех изображений, по одному для каждого канала флуоресценции. Поскольку из каждого такого изображения извлекается восемь статистических показателей, для каждого кластера имеется в общей сложности тридцать два статистических показателя, которые используются в качестве атрибутов в алгоритмах машинного обучения.

Одним из наиболее фундаментальных рабочих принципов секвенирования NGS является распараллеливание, что означает, что ДНК фрагментируется на короткие риды, которые секвенируются одновременно. При этом в каждом кластере будут разные фрагменты ДНК. Однако для целей данного проекта машинного обучения мы используем заранее секвенированные последовательности, предварительно отображенные (мэппированные) на известные референтные последовательности. Иногда это достаточно короткие индексные последовательности, короткие последовательности ампликонов (искусственно синтезированные известные последовательности нуклеотидов), иногда – короткие риды из ДНК последовательности бактериофага Phix174. Это позволяет сформировать обучающую выборку с известным нуклеотидным составом коротких прочтений.

После того, как данные надлежащим образом исследованы и преобразованы (в частности, нормализованы), набор данных необходимо случайным образом разделить на две части: обучающий и тестовый набор. Обычно около восьмидесяти процентов, выбирается в качестве обучающих данных, а оставшиеся данные составляют тестовый набор.

### **7. Результаты машинного обучения в задаче base-calling.**

Различные методы машинного обучения были апробированы на разных наборах данных секвенирования, таких как, набор индексов, небольшие ампликоны, данные секвенирования эталонного референтного генома Phix174. Обработываемые данные были получены на приборе MiSeq фирмы Illumina (США) и на опытном образце отечественного прибора «Нанофор СПС».

Для машинного обучения на данных секвенирования использовались следующие методы:

- логистическая регрессия (logit model) с различными видами регуляризации;
- метод опорных векторов (support vector machine);
- деревья принятия решений (decision tree);
- случайные леса (random forest);
- ансамблевые методы, в частности, бэггинг классификаторы с голосованием.

Типичный отчет со сравнением точности предсказания правильной буквы в нуклеотидной последовательности выглядит следующим образом (где box-plot диаграмма в удобной форме показывает медиану, среднее, нижний и верхний квартили, минимальное и максимальное значение выборки и выбросы) (рисунок 9) [38].

В нашем случае каждый из четырех каналов для определения нуклеотидов имеет 8 признаков. Четыре на полезный сигнал и четыре на фон. Соответствующая box-plot диаграмма для 32 признаков (признаки 0 – 7 для канала А, 8 – 15 для канала С, признаки 16 – 23 для канала G, признаки 24 – 31 для канала Т) представлена на рисунке 10. Последовательность признаков для каждого канала: foreground - max, mean, pct90 и pct99; background - max, mean, median и mode.

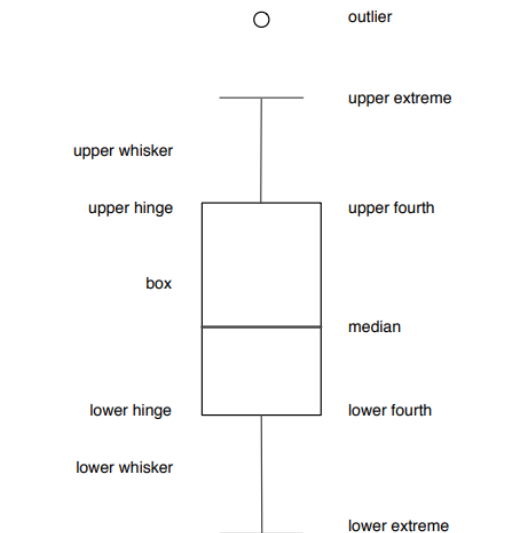


Рис. 9. Структурная схема box-plot диаграммы

Вероятности правильного предсказания для данных, полученных на приборе MiSeq фирмы Illumina (США) и на опытном образце отечественного прибора «Нанофор СПС оказались примерно одинаковыми. Наибольшую вероятность правильного предсказания показал метод логистической регрессии. Для логистической регрессии для повышения точности предсказания на тестовом наборе была использована L1-регуляризация (lasso regularization).

Распределение распределение всех нуклеотидов по каналам для образца Phix 174 примерно одинаковое и приблизительно равно 25%, что очевидно из рисунка 10. Диаграммы типа box-plot, подобные представленному на рисунке 10, дают наглядное представление о совокупности распределения интенсивностей кластеров и шумов в ближайшем окружении кластеров по всем каналам нуклеотидов и позволяют выбрать набор признаков, отвечающих за разделение интенсивностей в кластерах по каналам.

```
LogisticRegression 0.9924318030605456
SVC 0.9735528942115769
DecisionTreeClassifier 0.9762974051896207
RandomForestClassifier 0.97895874916833
ExtraTreesClassifier 0.9693945442448436
BaggingClassifier 0.9757984031936128
VotingClassifier 0.9807884231536926
```

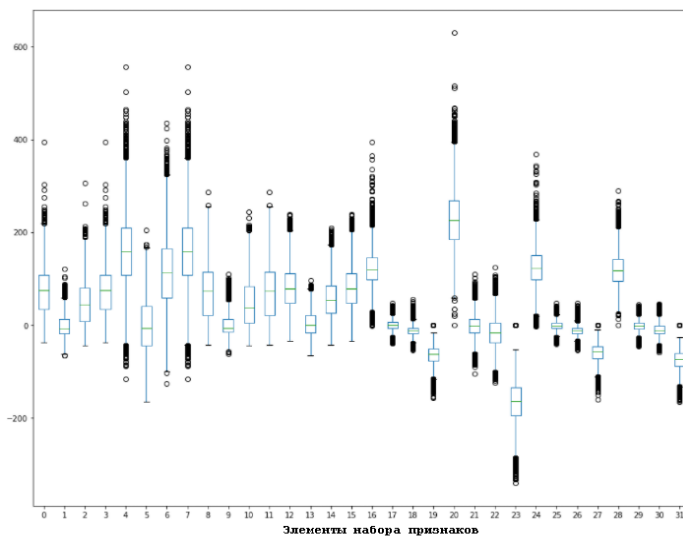


Рис. 10. Типичный отчет со сравнением точности предсказания методами машинного обучения с соответствующей box-plot диаграммой. В левом верхнем углу представлены вероятности правильного предсказания для различных методов машинного обучения. По оси x представлен набор признаков

Прогноз, полученный методом Decision Tree, позволяет путем анализа узлов дерева решений выделить наиболее значащие для прогнозирования признаки и снизить количество обрабатываемой информации.

В частности, такой анализ позволил для большинства методов машинного обучения снизить количество признаков для прогнозирования с 32 до 8, ограничившись только двумя первыми характеристиками для foreground и background. При этом точность прогноза даже повысилась.

Анализ box-plot диаграмм отдельно для каждого канала изображений и для пар каналов, вносящих основной вклад в эффект cross-talk позволяет судить о тех комбинациях признаков, которые определяют работу методов обучения.

Наконец, ожидаемый выигрыш от применения ансамблевых методов пока не был получен либо из-за неоптимального выбора соответствующих гиперпараметров методов, либо просто из-за недостаточности объема выборки для обучения.

Таким образом, использование методов машинного обучения в задаче построения последовательностей нуклеотидов на основе измеренных интенсивностей сигналов флуоресценции показало возможность получения достоверных результатов без использования традиционных процедур первичного анализа данных, таких как коррекция влияния перекрестных помех (cross-talk), фильтрация амплитуд кластеров на основе сравнения интенсивностей сигналов в различных каналах (purity и chastity) и других.

## **8. Выводы.**

1. Приведенный обзор методов машинного обучения для обработки информации при построении последовательностей нуклеотидов на основе статистической модели base-calling позволяет раскрыть вычислительные алгоритмы этой стадии обработки данных в приборах для генетического анализа. В виду особенностей задачи base-calling (типичная задача классификации по четырем классам нуклеотидов) особо перспективным кажется применение таких методов машинного обучения, как логистическая регрессия и нейронные сети. Для выделения особо значимых признаков перспективны также методы на основе деревьев решений (Decision Tree) и метод главных компонент.

2. Данный обзор является одним из первых в русскоязычной литературе, в котором обсуждаются вычислительные алгоритм обработки информации на этапе base-calling и особенности их реализации. В тоже же время многократно возросшие вычислительные

мощности раскрывают новые горизонты применения искусственного интеллекта.

3. Приведенные результаты применения разнообразных моделей машинного обучения к решению задачи base-calling показали, что все они на тестовой выборке дают результаты построения последовательностей нуклеотидов, совпадающие на 97...99.5% с последовательностями нуклеотидов в референтных геномах. При этом обучающая выборка была достаточно ограниченной по объему.

4. Ожидается, что дальнейшее применение более объемной обучающей выборки даст еще более лучшие результаты в решении задачи base-calling. На следующей стадии обработки важное значение имеют алгоритмы для оценки показателей качества определения как отдельных нуклеотидов, так и ридов.

5. Методы машинного обучения типа деревьев поиска и уменьшение размерности пространства признаков, по которым проводится обучение позволяет выделить параметры, определяющие вероятность статистической ошибки при высокопроизводительном секвенировании.

6. Безусловно, полезным при накоплении результатов секвенирования на отечественном полногеномном секвенаторе является сравнение результатов машинного обучения на этапе base-calling с данными, полученными на секвенаторе Illumina.

7. Ожидается, что дальнейшее применение более объемной обучающей выборки даст еще более лучшие результаты в решении задачи base-calling. На следующей стадии обработки важное значение имеют алгоритмы для оценки показателей качества определения как отдельных нуклеотидов, так и ридов [38]. Методы машинного обучения типа деревьев поиска и уменьшение размерности пространства признаков, по которым проводится обучение, позволяют выделить параметры, определяющие вероятность статистической ошибки при высокопроизводительном секвенировании.

Для успешного внедрения рассмотренной практики машинного обучения предполагается стадия обучения при известной циклограмме эксперимента на некотором эталонном образце с известным геномным составом (например, Phix174). Далее полученная обученная модель может быть использована для решения задачи base-calling уже на любых образцах при сходных циклограммах эксперимента. Поскольку для обучения в качестве обучающих признаков взяты параметры изображений в одном цикле, предполагается, что такая модель сможет учесть эффекты decay и cross-talk. При включении в модель признаков

предыдущих и последующих циклов возможно учесть эффекты phasing/prephasing.

### Литература

1. Бородинов А. Г., Манойлов В. В., Заруцкий И. В., Петров А. И., Курочкин В. Е. Поколения методов секвенирования ДНК (ОБЗОР) // Научное приборостроение. 2020. т. 30. № 4. С. 3—20
2. Wenxiu Ma, Wing Hung Wong The analysis of ChIP-Seq data // *Methods Enzymol.* 2011. vol. 497. pp. 51-73.
3. Zhong Wang, Mark Gerstein, Michael Snyder RNA-Seq: a revolutionary tool for transcriptomics // *Nat Rev Genet.* 2009. vol.10. no. 1. pp. 57-63.
4. Syed, F., Grunenwald, H. & Caruccio, N. Next-generation sequencing library preparation: simultaneous fragmentation and tagging using in vitro transposition // *Nat Methods.* 2009. vol. 6. pp. i-ii.
5. Манойлов В. В., Бородинов А. Г., Заруцкий И. В., Петров А. И., Курочкин В. Е. Алгоритмы обработки сигналов флуоресценции массового параллельного секвенирования нуклеиновых кислот // *Труды СПИИРАН.* 2019. т. 18. № 4. С. 1010–1036.
6. Schilbert H.M., Rempel A., Pucker B. Comparison of Read Mapping and Variant Calling Tools for the Analysis of Plant NGS Data // *Plants.* 2020. vol. 9. p. 439.
7. Ye C., Hsiao C., Corrada-Bravo H. BlindCall: ultra-fast base-calling of high-throughput sequencing data by blind deconvolution // *Bioinform.* 2014. vol. 30. no. 9. pp. 1214–1219.
8. Wang B, Wan L, Wang A, Li L.M. An adaptive decorrelation method removes Illumina DNA base-calling errors caused by crosstalk between adjacent clusters // *Sci Rep.* 2017. vol. 7.
9. Renaud G., Kircher M., Stenzel U., Kelso J. FreeIbis: an efficient basecaller with calibrated quality scores for Illumina sequencers // *Bioinformatics.* 2013. vol. 29. pp. 1208–1209.
10. Das S., Vikalo H. Base calling for high-throughput short-read sequencing: dynamic programming solutions // *BMC Bioinformatics.* 2013. vol. 14. p. 129.
11. Massingham T., Goldman T. All your base: a fast and accurate probabilistic approach to base calling // *Genome Biol.* 2012. vol. 13. p. R13.
12. Das S., Vikalo H. OnlineCall: fast online parameter estimation and base calling for illumina's next-generation sequencing // *Bioinformatics.* 2012. vol. 28. no. 13. pp. 1677–1683.
13. Ji Y., Mitra R., Quintana F., Jara A., Mueller P., Liu P., Lu Y., Liang S. BM-BC: a Bayesian method of base calling for Solexa sequence data // *BMC Bioinformatics.* 2012. vol. 13. p. S6.
14. Shen X., Vikalo H. ParticleCall: A particle filter for base calling in next-generation sequencing systems // *BMC Bioinformatics.* 2012. vol. 13. p. 160.
15. Menges F., Narzisi G., Mishra B. TotalReCaller: improved accuracy and performance via integrated alignment and base-calling // *Bioinformatics.* 2011. vol. 27. no. 17. pp. 2330-2337.
16. Kao W.C., Song Y.S. naiveBayesCall: an efficient model-based base-calling algorithm for high-throughput sequencing // *J Comput Biol.* 2011. vol.18. no. 3. pp. 365-377.
17. Corrada-Bravo H., Irizarry R.A. Model-based quality assessment and base-calling for second-generation sequencing data // *Biometrics.* 2009. vol. 3. pp. 665–674.
18. Kao W.C., Stevens K., Song Y.S. BayesCall: a model-based basecalling algorithm for high-throughput short-read sequencing // *Genome Res.* 2009. vol. 19. pp. 1884–1895.



19. Kircher M., Stenzel U., Kelso J. Improved base calling for the Illumina Genome analyzer using machine learning strategies // *Genome Biol.* 2009. vol. 10. pp. R83.1–9.
20. Rougemont J., Amzallag A., Iseli C. Probabilistic base calling of Solexa sequencing data // *BMC Bioinformatics.* 2008. vol. 9. p. 431.
21. Erlich Y., Mitra P.P., Delabastide M., et al. Alta-cyclic: a self-optimizing base caller for next-generation sequencing // *Nat Methods.* 2008. vol. 5. pp. 679–682.
22. Зубов В. В., Чемерис Д. А., Василев Р. Г., Курочкин В. Е., Алексеев Я. И. Краткая история методов высокопроизводительного секвенирования нуклеиновых кислот // *Биомика.* 2021. т. 13. № 1. С. 27–46.
23. Cacho A. Base-Calling of High-Throughput Sequencing Data Using a Random Effects Mixture Model // *UC Riverside.* 2016. 91 p.
24. Li L., Speed T. An estimate of the crosstalk matrix in four-dye fluorescence-based DNA sequencing // *Electrophoresis.* 1999. vol. 20. pp. 1433–1442.
25. Ghannam R., Techtmann S. Machine learning applications in microbial ecology, human microbiome studies, and environmental monitoring // *Computational and Structural Biotechnology Journal.* 2021. vol. 19. pp. 1092–1107.
26. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: data mining, inference, and prediction // *Springer Science & Business Media.* 2009. 745 p.
27. Forgy E.W. Cluster analysis of multivariate data: efficiency versus interpretability of classifications // *Biometrics.* 1965. vol. 21. pp. 768–769.
28. Mohammadi S.A., Prasanna B.M. Review and Interpretation Analysis of Genetic Diversity in Crop Plants — Salient Statistical Tools // *Crop Science.* 2003. vol. 43. pp. 1235–1248.
29. Jackson J.E. A User's Guide to Principal Components // *John Wiley & Sons.* 1991.
30. Van der Maaten L., Hinton G. Visualizing Data using t-SNE // *Journal of Machine Learning Research.* 2008. vol. 9. pp. 2579–2605.
31. Breiman L. Random forests // *Machine Learn.* 2001. vol. 45. no. 1. pp. 5–32.
32. Suykens J.A., Vandewalle J. Least squares support vector machine classifiers // *Neural Process Letters.* 2004. vol. 9. no. 3. pp. 293–300.
33. Tolles J, Meurer W.J. Logistic Regression: Relating Patient Characteristics to Outcomes // *JAMA.* 2016. vol. 316. no. 5. pp. 533–534.
34. Hoerl A.E., Kennard R.W. Ridge regression: biased estimation for nonorthogonal problems // *Technometrics.* 1970. vol. 12. no. 1. pp. 55–67.
35. LeCun Y., Bengio Y., Hinton G. Deep learning // *Nature.* 2015. vol. 521. pp. 436–444.
36. About us — scikit-learn 0.20.1 documentation. URL: <https://scikit-learn.org>. (дата обращения 18.03.2022).
37. Tegfalk E. Application of machine learning techniques to perform base-calling in next-generation DNA sequencing // *KTH, SCI.* 2020.
38. Wickham H., Stryjewski L. 40 years of boxplots. URL: <https://vita.had.co.nz/papers/boxplots.pdf>. (дата обращения 23.03.2022).

**Бородин Андрей Геннадьевич** — канд. физ.-мат. наук, начальник, сектор информационных проектов, АО "Научные приборы". Область научных интересов: математическая статистика, проблемы анализа, обработки и представления данных, искусственный интеллект. Число научных публикаций — 15. borodinov@gmail.com; улица Ивана Черных, 31-33, 198095, Санкт-Петербург, Россия; р.т.: +7(911)212-0895.

**Манойлов Владимир Владимирович** — д-р техн. наук, доцент, заведующий лабораторией, лаборатории автоматизации измерений и цифровой обработки сигналов, Институт аналитического приборостроения Российской академии наук (ИАП РАН).

Область научных интересов: представление и обработка сигналов и изображений в аналитических приборах. Число научных публикаций — 76. mapoi1ov-vv@mail.ru; улица Ивана Черных, 31-33, 198095, Санкт-Петербург, Россия; р.т.: +7(812)363-0750.

**Заруцкий Игорь Вячеславович** — канд. техн. наук, старший научный сотрудник, лаборатория автоматизации измерений и цифровой обработки сигналов, Институт аналитического приборостроения Российской академии наук (ИАП РАН). Область научных интересов: представление и обработка сигналов и изображений в аналитических приборах. Число научных публикаций — 45. igorzv@yandex.ru; улица Ивана Черных, 31-33, 198095, Санкт-Петербург, Россия; р.т.: +7(812)363-0720.

**Петров Александр Иванович** — канд. техн. наук, заведующий сектором электроники и программного обеспечения, лаборатория методов и приборов иммунного и генетического анализа, Институт аналитического приборостроения Российской академии наук (ИАП РАН). Область научных интересов: представление и обработка сигналов и изображений в аналитических приборах. Число научных публикаций — 25. fataip@mail.ru; улица Ивана Черных, 31-33, 198095, Санкт-Петербург, Россия; р.т.: +7(812)363-0720.

**Курочкин Владимир Ефимович** — д-р техн. наук, профессор, заведующий лабораторией, лаборатория методов и приборов иммунного и генетического анализа, Институт аналитического приборостроения Российской академии наук (ИАП РАН). Область научных интересов: исследования и оптимизация электромиграционных методов анализа, развитие аналитических методик для капиллярного электрофореза, исследование оп-тических методов детектирования, разработка методов и приборов для ДНК анализа, разработка методик подготовки проб и специализированных реактивов. Число научных публикаций — 210. lavrovav@yandex.ru; улица Ивана Черных, 31-33, 198095, Санкт-Петербург, Россия; р.т.: +7(812)363-0719.

**Сараев Алексей Сергеевич** — инженер 2 категории, лаборатория методов и приборов иммунного и генетического анализа, Институт аналитического приборостроения Российской академии наук (ИАП РАН). Область научных интересов: моделирование процессов в аналитических приборах. Число научных публикаций — 1. alex.niisrb@yandex.ru; улица Ивана Черных, 31-33, 198095, Санкт-Петербург, Россия; р.т.: +7(812)363-0720.

**Поддержка исследований.** Работа выполнена в рамках государственного задания Министерства науки и высшего образования номер гос. регистрации 122032300337-4 от 23.03.22.

A. BORODINOV, V. MANOILOV, I. ZARUTSKY, A. PETROV, V. KUROCHKIN,  
A. SARAEV

## MACHINE LEARNING IN BASE-CALLING FOR NEXT- GENERATION SEQUENCING METHODS

*Borodinov A., Manoilov V., Zarutsky I., Petrov A., Kurochkin V., Saraev A.* **Machine Learning in Base-Calling for Next-Generation Sequencing Methods.**

**Abstract.** The development of next-generation sequencing (NGS) technologies has made a significant contribution to the trend of reducing costs and obtaining massive sequencing data. The Institute for Analytical Instrumentation of the Russian Academy of Sciences is developing a hardware-software complex for deciphering nucleic acid sequences by the method of mass parallel sequencing (Nanofor SPS). Image processing algorithms play an essential role in solving the problems of genome deciphering. The final part of this preliminary analysis of raw data is the base-calling process. Base-calling is the process of determining a nucleotide base that generates the corresponding intensity value in the fluorescence channels for different wavelengths in the flow cell image frames for different synthesis sequencing runs. An extensive analysis of various base-calling approaches and a summary of the common procedures available for the Illumina platform are provided. Various chemical processes included in the synthesis sequencing technology, which cause shifts in the values of recorded intensities, are considered, including the effects of phasing / prephasing, signal decay, and crosstalk. A generalized model is defined, within which possible implementations are considered. Possible machine learning (ML) approaches for creating and evaluating models that implement the base-calling processing stage are considered. ML approaches take many forms, including unsupervised learning, semi-supervised learning, and supervised learning. The paper shows the possibility of using various machine learning algorithms based on the Scikit-learn platform. A separate important task is the optimal selection of features identified in the detected clusters on a flow cell for machine learning. Finally, a number of sequencing data for the MiSeq Illumina and Nanofor SPS devices show the promise of the machine learning method for solving the base-calling problem.

**Keywords:** next-generation sequencing, base-calling, bioinformatics, machine learning.

**Borodinov Andrew** — Ph.D., Head of sector, Sector of information projects, Scientific Instruments Joint Stock Company. Research interests: mathematical statistics, problems of analysis, processing and presentation of data, artificial intelligence. The number of publications — 15. borodinov@gmail.com; 31-33, Ivana Chernykh St., 198095, St. Petersburg, Russia; office phone: +7(911)212-0895.

**Manoilov Vladimir** — Ph.D., Dr.Sci., Associate Professor, Head of laboratory, Laboratory of automation of measurements and digital signal processing, Institute for Analytical Instrumentation Russian Academy of Sciences (IAI RAS). Research interests: representation and processing of signals and images in analytical devices. The number of publications — 76. manoilov-vv@mail.ru; 31-33, Ivana Chernykh St., 198095, St. Petersburg, Russia; office phone: +7(812)363-0750.

**Zarutsky Igor** — Ph.D., Senior researcher, Laboratory of automation of measurements and digital signal processing, Institute for Analytical Instrumentation Russian Academy of Sciences (IAI RAS). Research interests: representation and processing of signals and images in

analytical devices. The number of publications — 45. igorzv@yandex.ru; 31-33, Ivana Chernykh St., 198095, St. Petersburg, Russia; office phone: +7(812)363-0720.

**Petrov Alexander** — Ph.D., Head of the sector of electronics and software, Laboratory of methods and instruments for immune and genetic analysis, Institute for Analytical Instrumentation Russian Academy of Sciences (IAI RAS). Research interests: representation and processing of signals and images in analytical devices. The number of publications — 25. fataip@mail.ru; 31-33, Ivana Chernykh St., 198095, St. Petersburg, Russia; office phone: +7(812)363-0720.

**Kurochkin Vladimir** — Ph.D., Dr.Sci., Professor, Head of the laboratory, Laboratory of methods and instruments for immune and genetic analysis, Institute for Analytical Instrumentation of the Russian Academy of Sciences (IAI RAS). Research interests: research and optimization of electromigration analysis methods, the development of analytical methods for capillary electrophoresis, the study of optical methods of detection, the development of methods and instruments for DNA analysis, the development of methods for preparing samples and specialized reagents. The number of publications — 210. lavrovas@yandex.ru; 31-33, Ivana Chernykh St., 198095, St. Petersburg, Russia; office phone: +7(812)363-0719.

**Saraev Aleksey** — 2nd category engineer, Laboratory of methods and devices for immune and genetic analysis, Institute for Analytical Instrumentation of the Russian Academy of Sciences (IAI RAS). Research interests: modeling of processes in analytical instruments. The number of publications — 1. alex.niispb@yandex.ru; 31-33, Ivana Chernykh St., 198095, St. Petersburg, Russia; office phone: +7(812)363-0720.

**Acknowledgements.** This research was performed within the framework of the state number registration 122032300337-4 dated 03/23/22, Ministry of Science and Higher Education of the Russian Federation.

## References

1. Borodinov A.G., Manoilov V.V., Zarutsky I. V., Petrov A. I., Kurochkin V. E. [Generations of DNA sequencing methods (REVIEW)]. *Nauchnoe priborostroenie - Nauchnoe priborostroenie*. 2020. vol. 30. no. 4. pp. 3-20. (In Russ.).
2. Wenxiu Ma, Wing Hung Wong. The analysis of ChIP-Seq data. *Methods Enzymol*. 2011. vol. 497. pp. 51-73.
3. Zhong Wang, Mark Gerstein, Michael Snyder. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet*. 2009. vol.10. no. 1. pp. 57-63.
4. Syed, F., Grunenwald, H. & Caruccio, N. Next-generation sequencing library preparation: simultaneous fragmentation and tagging using in vitro transposition. *Nat Methods*. 2009. vol. 6. pp. i-ii.
5. Manoilov V.V., Borodinov A.G., Petrov A.I., Zarutsky I.V., Kurochkin V.E. [Algorithms of processing fluorescence signals for mass parallel sequencing of nucleic acids]. *Trudy SPIIRAN - SPIIRAS Proceedings*. 2019. vol. 18. no. 4. pp. 1010–1036.
6. Schilbert H.M., Rempel A., Pucker B. Comparison of Read Mapping and Variant Calling Tools for the Analysis of Plant NGS Data. *Plants*. 2020. vol. 9. p. 439.
7. Ye C., Hsiao C., Corrada-Bravo H. BlindCall: ultra-fast basecalling of high-throughput sequencing data by blind deconvolution. *Bioinform*. 2014. vol. 30. no. 9. pp. 1214–1219.
8. Wang B, Wan L, Wang A, Li L.M. An adaptive decorrelation method removes Illumina DNA base-calling errors caused by crosstalk between adjacent clusters. *Sci Rep*. 2017. vol. 7.

9. Renaud G., Kircher M., Stenzel U., Kelso J. FreeIbis: an efficient basecaller with calibrated quality scores for Illumina sequencers. *Bioinformatics*. 2013. vol. 29. pp. 1208–1209.
10. Das S., Vikalo H. Base calling for high-throughput short-read sequencing: dynamic programming solutions. *BMC Bioinformatics*. 2013. vol. 14. p. 129.
11. Massingham T., Goldman N. All your base: a fast and accurate probabilistic approach to base calling. *Genome Biol*. 2012. vol. 13. p. R13.
12. Das S., Vikalo H. OnlineCall: fast online parameter estimation and base calling for illumina's next-generation sequencing. *Bioinformatics*. 2012. vol. 28. no. 13. pp. 1677–1683.
13. Ji Y., Mitra R., Quintana F., Jara A., Mueller P., Liu P., Lu Y., Liang S. BM-BC: a Bayesian method of base calling for Solexa sequence data. *BMC Bioinformatics*. 2012. vol. 13. p. S6.
14. Shen X., Vikalo H. ParticleCall: A particle filter for base calling in next-generation sequencing systems. *BMC Bioinformatics*. 2012. vol. 13. p. 160.
15. Menges F., Narzisi G., Mishra B. TotalReCaller: improved accuracy and performance via integrated alignment and base-calling. *Bioinformatics*. 2011. vol. 27. no. 17. pp. 2330–2337.
16. Kao W.C., Song Y.S. naiveBayesCall: an efficient model-based base-calling algorithm for high-throughput sequencing. *J Comput Biol*. 2011. vol.18. no. 3. pp. 365–377.
17. Corrada-Bravo H., Irizarry R.A. Model-based quality assessment and base-calling for second-generation sequencing data. *Biometrics*. 2009. vol. 3. pp. 665–674.
18. Kao W.C., Stevens K., Song Y.S. BayesCall: a model-based basecalling algorithm for high-throughput short-read sequencing. *Genome Res*. 2009. vol. 19. pp. 1884–1895.
19. Kircher M., Stenzel U., Kelso J. Improved base calling for the Illumina Genome analyzer using machine learning strategies. *Genome Biol*. 2009. vol. 10. pp. R83.1–9.
20. Rougemont J., Amzallag A., Iseli C. Probabilistic base calling of Solexa sequencing data. *BMC Bioinformatics*. 2008. vol. 9. p. 431.
21. Erlich Y., Mitra P.P., Delabastide M., et al. Altacyclic: a selfoptimizing base caller for next-generation sequencing. *Nat Methods*. 2008. vol. 5. pp. 679–682.
22. Zubov V.V., Chemeris D.A., Vasilov R.G., Kurochkin V.E., Alekseev Ya.I. [Brief history of high-throughput nucleic acid sequencing methods]. *Biomika - Biomics*. 2021. vol. 13. no. 1. pp. 27–46. (In Russ.)
23. Cacho A. Base-Calling of High-Throughput Sequencing Data Using a Random Effects Mixture Model. UC Riverside, 2016. 91 p.
24. Li L., Speed T. An estimate of the crosstalk matrix in four-dye fluorescence-based DNA sequencing. *Electrophoresis*. 1999. vol. 20. pp. 1433–1442.
25. Ghannam R., Techtmann S. Machine learning applications in microbial ecology, human microbiome studies, and environmental monitoring. *Computational and Structural Biotechnology Journal*. 2021. vol. 19. pp. 1092–1107.
26. Hastie T., Tibshirani R., Friedman J. The elements of statistical learning: data mining, inference, and prediction. Springer Science & Business Media, 2009. 745 p.
27. Forgy E.W. Cluster analysis of multivariate data: efficiency versus interpretability of classifications. *Biometrics*. 1965. vol. 21. pp. 768–769.
28. Mohammadi S.A., Prasanna B.M. Review and Interpretation Analysis of Genetic Diversity in Crop Plants —Salient Statistical Tools. *Crop Science*. 2003. vol. 43. pp. 1235–1248.
29. Jackson J.E. A User's Guide to Principal Components. John Wiley & Sons, 1991.
30. Van der Maaten L., Hinton G. Visualizing Data using t-SNE. *Journal of Machine Learning Research*. 2008. vol. 9. pp. 2579–2605.
31. Breiman L. Random forests. *Machine Learn*. 2001. vol. 45. no. 1. pp. 5–32.

32. Suykens J.A., Vandewalle J. Least squares support vector machine classifiers. *Neural Process Letters*. 2004. vol. 9. no. 3. pp. 293–300.
33. Tolles J, Meurer W.J. Logistic Regression: Relating Patient Characteristics to Outcomes. *JAMA*. 2016. vol. 316. no. 5. pp. 533-534.
34. Hoerl A.E., Kennard R.W. Ridge regression: biased estimation for nonorthogonal problems. *Technometrics*. 1970. vol. 12. no. 1. pp. 55–67.
35. LeCun Y., Bengio Y., Hinton G. Deep learning. *Nature*. 2015. vol. 521. pp. 436–444.
36. About us — scikit-learn 0.20.1 documentation. Available at: <https://scikit-learn.org>. (accessed 18.03.2022).
37. Tegfalk E. Application of machine learning techniques to perform base-calling in next-generation DNA sequencing. KTH, SCI, 2020.
38. Wickham H., Stryjewski L. 40 years of boxplots. Available at: <https://vita.had.co.nz/papers/boxplots.pdf>. (accessed 23.03.2022).

А.Ю. ПЕРЕВАРЮХА  
**ДИНАМИЧЕСКАЯ МОДЕЛЬ ПОПУЛЯЦИОННОЙ ИНВАЗИИ С  
ЭФФЕКТОМ ДЕПРЕССИИ**

*Переварюха А.Ю. Динамическая модель популяционной инвазии с эффектом депрессии.*

**Аннотация.** Статья посвящена исследованию актуального сценария развития популяционных процессов в современных нестабильных биосистемах методами компьютерного моделирования. Биологические инвазии стали чрезвычайно распространенным явлением из-за изменений климата, целенаправленной деятельности с задачей улучшения продуктивности экосистем и случайного стечения обстоятельств. Динамика ситуаций после вселения чужеродного вида чрезвычайно разнообразна. Далеко не всегда вселенец гладко занимает экологическую нишу, как в логистических моделях. В отдельных случаях реализуется явление вспышки численности вплоть до начала разрушения видом своей новой среды. Развитие ситуации после инвазии зависит от суперпозиции биотических и абиотических факторов. На динамику численности вселенца влияют благоприятность сложившихся условий, возможность реализации репродуктивного потенциала и сопротивление биотического окружения. Противодействие развивается с запаздыванием и проявляется при достижении вселенцем значительной численности. Обоснована и разработана непрерывная модель инвазионного процесса с резким переходом в состояние депрессии численности. Стадия популяционного кризиса завершается с переходом к равновесию, так как оказываемое биотической средой сопротивление в модельном сценарии адаптивно и пороговым образом зависит от численности вида-вселенца. Применение вычислительного феноменологического описания сценария с активным, но запаздывающим противодействием среды практически целесообразно для оценки ситуаций при выработке мер искусственного противодействия нежелательному вселенцу. В модели существует режим сохранения колебаний после выхода из стадии депрессии, если эффективность подавления вселенца оказывается недостаточной.

**Ключевые слова:** динамические модели, вычислительные сценарии, активные инвазии, кризис и депрессия, имитационные модели в экологии, управление в биосистемах.

**1. Введение.** Рассматривается моделирование особых ситуаций, когда регулируемое противодействие агрессивно размножающемуся виду в биологическом сообществе вырабатывается с запаздыванием и приводит к резкому переходу в фазу депрессии численности вселенца. Для остановки распространения вредоносного инвазивного вида сейчас часто проводится специальная интродукция вида-антагониста, но эффективность такого метода подавления на практике непостоянная. Необходимо обоснование модели для специфического сценария популяционной динамики, описывающей управляемую активность вида-антагониста против чужеродной популяции, но с механизмом их коадаптации.

Цель работы – разработка модели популяционного процесса для вида с большим репродуктивным потенциалом в ситуации с адаптивной формой противодействия его распространению в среде. После первичной

вспышки численность вселенца всегда снижается, но иногда обычное снижение резко переходит в фазу депрессии и даже с риском исчезновения – вариант «бутылочного горлышка». Давление антагонистов усиливается при достижении вселенцем пороговых значений численности, но не остается постоянным.

Основной результат работы – в вычислительном эксперименте получен сценарий преодоления глубокой депрессии для атакуемого вида с высоким репродуктивным коэффициентом и с дальнейшим установлением равновесного состояния. Примером описанной ситуации служит динамика растущей колонии бактерий, задействовавших адаптационный механизм защиты после вселения нового вируса-бактериофага. Преодоление кризиса в данном взаимодействии начинается, если работает эндонуклеаза системы *CRISPR/CAS 9* (сейчас этот механизм стал средством геномного ДНК-редактирования). Через некоторое время у выживших бактерий *E. coli* эндонуклеаза *CAS 9* обучается распознавать выбранные особые короткие фрагменты ДНК у данного штамма вируса и разрывать в этом месте цепочку нуклеотидов. Потому терапия инфекций с применением бактериофагов теряет эффективность.

В классической модели вспышки динамику определяют переходы между стационарными состояниями [1]. Для инвазионных процессов стационарных состояний еще не существует. Развиваются резкие переходные режимы, а устойчивые равновесия появляются позднее при эволюционном становлении обновленной биосистемы [2]. Сценарии с сериями вспышек и промежуточных кризисов остаются за рамками традиционных популяционных моделей равновесной или регулярной циклической динамики. В предыдущих работах мы разрабатывали дискретные итерационные модели популяционных процессов на основе гибридных структур [3]. Нами были предложены критерии адекватности поведения при возникновении нелинейных эффектов и хаотического поведения траектории моделей с функциональными итерациями для регулирования эксплуатации запасов осетровых рыб Каспия [4]. Для модели инвазии с эффектом кризиса требуется учет времени запаздывания и в регуляции и в реакции окружения на нового конкурента. Выберем для исследуемой ситуации непрерывное моделирование – дифференциальные уравнения с отклоняющимся аргументом, где запаздывающая регуляция диверсифицирована по своему генезису.

Новизной работы является применение уравнения со сосредоточенным запаздыванием в трех функциональных компонентах и описание краткой формы популяционной депрессии, отличной от системного экологического кризиса «бутылочного горлышка».



## 2. Моделирование цикличности и запаздывающая регуляция.

Цикличность активности многих популяций остается не полностью изученным природным феноменом. Цикличность численности видов наблюдается на островах Арктики, в Индийском океане и в экспериментальных лабораторных аквариумах. Наличие противостояния жертвы и хищника не является обязательным условием для колебаний [5]. Внутривидовая конкуренция между стадиями развития достаточна. Для математического описания колебаний, которые не следуют из взаимодействий видов-антагонистов, Г. Хатчинсоном в теоретическом обзоре [6] предложено при моделировании рассмотреть влияние запаздывания – предшествующего состояния биосистемы  $N(t - \psi(t))$  на скорость современного воспроизводства.

Модели с запаздыванием развивали одновременно несколько специалистов. Математическая модель в форме уравнения с  $y(t - 1)$  предложена У. Райтом:  $y'(t) = -\alpha y(t - 1)[1 + y(t)]$  (оригинальные обозначения Райта) [7]. Позднее похожее уравнение с двумя параметрами обосновано Р. Мэем в таком виде [8]:

$$\frac{dN(t)}{dt} = rN(t) \left( 1 - \frac{N(t - \tau)}{K} \right), \quad (1)$$

где параметр  $r > 0$  — традиционно репродуктивный потенциал популяции в феноменологических моделях.  $N(t)$  — текущая численность, для  $N(t - \tau)$  используется константная функция-предыстория. Параметр  $K$  перенесен из модели ограниченного роста  $N(t) \rightarrow K$ , где отражает уровень насыщения стабилизированной экологической ниши. Уравнение (1) это обобщение  $\dot{N} = rf(N(t - \tau))$  с запаздыванием  $\tau$  для уравнения Ферхюльста [9]:

$$\frac{dN(t)}{dt} = rf(N(t)) = rN(t) \left( 1 - \frac{N(t)}{K} \right). \quad (2)$$

Заемствованный потом в другие модели хрестоматийный параметр  $K$  в (2) устанавливал доступный уровень неразрушающего заполнения среды: при  $0 < N(0) < K$  выполняется  $\max N(t) = K$ , но это теоретическое условие.

Теоретически важную величину  $K$  следует трактовать с различными аспектами для разных ситуаций инвазий, но в зависимости от параметров и свойств выбранной  $f(N)$ . За вариантами трактовки  $K$  стоит теория об экологической регуляции [10]. Ёмкость среды может воздейство-

вать на смертность опосредованно [11]. Теоретически фундаментальную  $K$  и реализованную  $\mathcal{K}$  экологические ниши вида надо различать [12].

Сущность параметра  $K$  в моделях для анализа циклических и хаотических режимов неоднозначно интерпретировалась. При развитии экстремальных процессов в биосистемах численность агрессивного вида-вселенца долго превышает гипотетическую балансовую ёмкость, возможную для биосистемы:  $N(t) \gg \hat{K}$ . Для летального патогена или смертоносного паразита балансовой ёмкости  $\lim_{t \rightarrow \infty} N(t) \rightarrow K$  не существует и в ряде других ситуаций  $K \neq const$ .

Циклическую траекторию обозначим  $N_*(t; \mathbf{R})$ , где  $\mathbf{R}$  – множество параметров, влияющих на амплитуду колебаний. Модель (1) использовалась для объяснения цикличности  $\forall N(0) > 0$ . При  $r\tau > \pi/2$  решение демонстрирует бифуркацию появления цикла  $N_*(t)$  [13]. Предельная ёмкость заполнения видом своей экологической ниши  $K$  при  $r\tau \gg \pi/2$  будет точкой центра для возникшего цикла  $N_*(t; r\tau)$  со свойством  $\forall N(0), |\max N_*(t) - \min N_*(t)| \approx K$ . Уравнение (1) анализировалось в [14, 15] в однопараметрической форме:  $\dot{x} = \lambda x(t)(1 - x(t - \tau))$ .

Мейнард Смит предложил вместо (1) для флуктуаций экспериментальных популяций насекомых модель с разными величинами отставания в регуляции  $\tau_1, \tau_2$  и убылью от внешних факторов среды  $\delta$  [16]:

$$\frac{dN(t)}{dt} = rN(t - \tau_1)\left(1 - \frac{N(t - \tau_2)}{K}\right) - \delta N(t). \quad (3)$$

Существуют осложнения для описания ограниченной динамики роста [17], с коэффициентами варьирования ёмкости ниши  $\vartheta, \gamma$  и двумя показателями жесткости внутривидовой регуляции  $\Theta, \kappa$ :

$$\frac{dN(t)}{dt} = rN(t - \tau) \frac{(1 - N(t)/(K + \vartheta N(t)))^\Theta}{(1 - N(t))^\kappa / K(1 - \gamma)}. \quad (4)$$

Многие «логистические» модели не имеет смысла дополнять рассмотрением  $t - \tau$ . Запаздывание  $\tau_1$  избыточно по экологическим представлениям вносить в репродуктивный компонент логистических моделей:  $rN(t - \tau_1)$ . Отличие у разных кривых ограниченного и лимитированного средой увеличения численности – положение точки перегиба кривой  $N_p \neq 0$  на графике решения  $N(t)$ . Для модели (2) ордината у точки перегиба  $N_p = K/2$ , а абсцисса этой точки  $N_p$  будет вычисляться нами следующим образом:  $t_p = r^{-1} \ln(K - N(0))/N(0)$ . Свойство нахождения на кривой точки перегиба  $N_p, f''(t_p) = 0$  требуется установить для проведения оптимальной эксплуатации биоресурсов и анализа сценариев с изъятием

$\dot{N} = rf(N(t)) - Q$ . Прирост  $dN/dt$  в  $N_p$  максимален, – излишек запаса  $N(t) - N_p$  можно использовать для промысла, но только при условии отсутствия порогового эффекта. Для (1) предложены модификации и обобщения уравнения [18]. Следующее уравнение [19] подходит для конкуренции внутривидовых объединений, например, сезонных нерестовых групп осетровых рыб Каспийского моря:

$$\frac{dN(t)}{dt} = rN(t - \tau_3) \left( 1 - \frac{c_1N(t - \tau_2) + c_2N(t - \tau_1)}{K} \right). \quad (5)$$

Можно применить  $K$  вместе с относительной величиной насыщения ниши, как было показано нами ранее:

$$\frac{dN(t)}{dt} = rN(t) \left( \frac{K - N(t - \tau)}{K + cN(t - \tau)} \right). \quad (6)$$

Модель  $\dot{N} = rf(N(t - \tau)) - F(N)$  с запаздывающей регуляцией, но без параметра ниши  $K$  с  $f(x) = rxe^{-bx}$  предложена на основе изучения поведения лабораторных популяций, где  $b$  отражает уровень воздействия негативных факторов конкуренции, а  $\delta$  дополнительный фактор убыли:

$$\frac{dN(t)}{dt} = rN(t - \tau) \exp(-bN(t - \tau)) - \delta N(t). \quad (7)$$

В опытах создана конкуренция за ресурсы между стадиями развития насекомых [20], что вызвало колебания большой амплитуды.  $\delta$  в (7) убыль от факторов каннибализма. При увеличении  $r\tau$  траектория (7) демонстрирует релаксационные колебания, но со значениями  $\min N_*(t; r\tau) \rightarrow \epsilon < \sigma$ .

Важно сущностное толкование причин возникновения  $\tau$  запаздывания:  $t - \tau$  или обобщенно  $t - \psi(t)$ . Величина интервала  $\tau$  изначально по [6] относилась к регуляции эффективности восполнения запасов через задержку онтогенетического развития. Регуляция запаздывания в форме функции  $\tau = \psi(t)$  возникает при анализе динамики поколений с разной длиной стадий онтогенеза. Длина жизненного цикла вида  $T$  и интервалы  $\Delta T$  между пиками численности у разных его популяций часто несопоставимые величины. Автор предлагает запаздывание при интерпретации моделей  $\dot{N} = rf(N(t - \tau)) - F(N)$  разграничить как репродуктивное (онтогенетическое время  $T$ ), регуляционное из-за времени восстановления ресурсов и адаптивное – требуется для выработки ответной реакции иммунитета организма или среды в широком понимании.

Развитие и усложнение (1) привело к появлению направления в моделировании – разработки уравнений с отклоняющимся аргументом в задачах моделирования биологических процессов. Уравнения с  $N(t - \tau)$  стали актуальны не только для динамики популяций и эпидемий. Применяются в изучении механизмов транскрипции ДНК, синтеза и распада сложных органических молекул. Исследования моделей не закрыли многие проблемы в адекватном описании разнообразных популяционных колебаний. Развитие направления связано с проблемами, которые не имеют особого интереса с точки зрения теоретической математики, но значимы для вычислительного моделирования в биологической области. Развиваются модели с использованием дробного дифференцирования. Далее остановимся на важном проблемном аспекте адекватности метода.

**3. Исследование минимумов цикла  $N_*(t; r\tau)$  популяционных моделей.** Метод моделирования требует подтверждения адекватности, соответствия траектории наблюдениям и теории. Для корректного включения  $t - \tau$  нужно обсудить проблемы. Часто нам нужно увеличить значение  $r\tau > \pi/2$  в (2) или (4). Тогда цикл  $N_*(t; r\tau)$  приобретет форму негармонических колебаний. Релаксационный цикл с увеличивающейся амплитудой по мере возрастания  $r\tau$  получит форму далёких друг от друга  $\Lambda$ -образных максимумов  $\max N_*(t) \gg K$ . Соответственно, решение  $N_*(t)$  в минимумах сколь угодно близко будет приближаться к  $0 + \epsilon$  и для  $N_*(t)$  не соблюдается равномерная отдаленность траектории цикла  $N_*(t, r\tau)$  от нуля. Известно, что в однопараметрической форме  $\dot{N} = \lambda N(t)(1 - N(t - \tau))$  выполняется:

$$\min N_*(t, \lambda) = \exp \left( -e^\lambda + 2\lambda + \frac{1 + (1 + \lambda) \ln \lambda}{\lambda} + \mathcal{O} \left[ \frac{\ln^2 \lambda}{\lambda^2} \right] - 1 \right).$$

Интервалы вокруг минимумов  $N(t_{min}) \approx \min N_*(t; \lambda\tau)$  колебаний, где  $N_*(t; \lambda\tau)$  длительно демонстрирует околонулевые значения  $\min N_*(t; \lambda\tau) \rightarrow 0 + \epsilon$ , становятся нереально длинными и «затянутыми». Траектория  $N_*(t, r\tau)$  слишком прижата к оси абсцисс.

Интересный вариант уравнения изучен в [21]:

$$\frac{dN(t)}{dt} = \lambda N(t) f(N(t - 1)), \quad (8)$$

где безразмерный параметр  $\lambda \gg 1$ , а  $f(x)$  — дифференцируемая функция, всегда разложимая в асимптотический ряд. Для  $f(x) \forall x$  выполняются

ограничения:

$$f(0) = 1, f(x) = -a_0 + \sum_{k=1}^{\infty} \frac{a_k}{x_k}, x \rightarrow \infty, a_0 > 0. \quad (9)$$

Ограничениям (9) удовлетворяет в [21] функция  $f(x)$  при относительной регуляции воспроизводства с коэффициентом  $\varsigma$ :

$$f(x) = \frac{(1-x)}{(1+\varsigma x)}. \quad (10)$$

При анализе характеристик релаксационного цикла  $N_*(t)$  проведено преобразование: от уравнения (8) к сингулярно возмущенному  $\dot{x} = F(x(t-1), \epsilon)$  и к предельному релейному уравнению  $\dot{x} = R(x(t-1))$ . Так в (8) с (9) в [21] подтверждено присутствие орбитально устойчивого цикла  $N_*(t; \lambda\tau\varsigma)$ . Соответствие экстремальных точек цикла  $\min N_*(t)$  экосистемному процессу в (8) с (10) из вводимых ограничений (9) трудно обосновать, так как из утверждения о характеристиках колебаний в [21] получена оценка:

$$\min N_*(t, \lambda) \sim C_1 \exp(-\lambda a_0), C_1, a_0 = \text{const} > 0.$$

Значение  $\lambda$  декларируется в [21] большим. Проблема экосистемной интерпретируемости цикла  $N_*(t; \lambda\tau\varsigma)$  сохраняется в других модификациях моделей с  $N(t-\tau)$ . Длительный пульсирующие вспышки численности у вредителей нельзя прогнозировать при свойстве:  $\min_{0 < t < T_*} N_*(t, r) \rightarrow \epsilon$ , и  $\epsilon \ll 1$ . Это будет верно и для активных вселенцев при больших значениях  $\lambda$ . Цикл  $N_*(t; \lambda\tau)$  достигает малых величин  $0 + \epsilon$  для быстрого повторного установления высоких значений.

Все популяционные колебания с большой амплитудой («лемминговые циклы») – в действительности нестабильны и разрушаются агрессивным внешним воздействием факторов среды.

**4. Моделирование с явной минимальной численностью**  $N < L, N(t) \rightarrow 0$ . В некоторых специфических ситуациях значима минимальная численность для той группы [22], которая практически необходима для поддержания локальной популяции [23].

А. Д. Базыкин [24] предложил уравнение с противодействием  $-\delta N^2$  для описания сценария гибели популяции (из-за сильного эффекта Олли) с коэффициентами варьирования внутривидовой конкуренции  $\vartheta, \gamma$  и линейной естественной убылью  $\varsigma$ :

$$\frac{dN(t)}{dt} = r \frac{\gamma N(t)^2}{\gamma + \vartheta N(t)} - \varsigma N(t) - \delta N(t)^2. \quad (11)$$

Принцип «агрегированной группы» в экологии состоит в том, что для популяции есть оптимальный интервал численности взрослой части сообщества  $\Delta \bar{N}$ . Сущностно  $L$ -порог допустимого существования группы  $L < \inf \Delta \bar{N}$  из эффекта Олли напрямую не следует. Для общественных животных (насекомых, стайных рыб)  $L$ -порог несовместим со строгой функцией регуляции  $r f(N^k), k \geq 2$  в (3). Многомиллионные колонии социальных насекомых не страдают от высокой плотности. Минимально необходимое количество  $L$  рабочих насекомых действительно установлено для успешного существования семей медоносных пчел.

Обосновано предположив эффект ослабления конкуренции при  $N \approx L$  мы предлагаем включить сомножитель  $\times \sqrt[k]{(N - L)}$  с  $r f(N^\theta), 1 < \Theta < 2, k \in 2^i - 1$  в модификации модели Базыкина:

$$\frac{dN(t)}{dt} = rN(t) \left(1 - \frac{N(t)}{K}\right)^\Theta \times \sqrt[k]{(N(t) - L)}. \quad (12)$$

В нашей модели (12) диапазон воздействия негативного эффекта для малой группы начинается строго у окрестности порога:  $N(t) \rightarrow L + \epsilon$ .

В предложенной нами ранее модели в [25] кризисного процесса получен вариант разрушения неустановившихся колебаний при увеличенном  $r$ , где  $\mathcal{H} = 1/3K, \tau_1 = 2/3\tau, r\tau > \pi$ :

$$\frac{dN}{dt} = rN(t) \left(1 - \frac{N(t - \tau)}{K}\right) (\mathcal{H} - N(t - \tau_1)), \tau_1 < \tau. \quad (13)$$

Модель (13) описала вычислительный сценарий с «выбросом» траектории из цикла, что соответствует демографической катастрофе. После образования колебаний при превышении значения максимумом цикла предельного для экосистемы уровня траектория  $N(t) \rightarrow \infty$ . Сценарий разрушения цикла с (13) не требовал увеличения  $r \rightarrow R$  в ходе вычислительного эксперимента.

**5. Модель перехода инвазионной вспышки в депрессию.** Разработаем уравнение в форме  $\dot{N} = rf(N(t - \tau)) - F(N(t - \tau_1))$  для интересной ситуации в ходе извазионного процесса. Имитируем сценарий, где при большом значении  $r < \pi/2$  не исчерпан потенциал увеличения численности образующейся инвазивной популяции с  $N(t) \rightarrow K$ , но возрастает активация адаптирующегося все это время  $t_\tau$  биотического сопротивления.

Цикл  $N_*(t)$  для ситуации инвазии со вспышкой неактуален. Включим в новую модель фактор внешней убыли  $\delta N(t - \tau_1)$  с отложенным действием и иначе трактуемый регулирующий параметр  $\mathfrak{K}$ :

$$\frac{dN}{dt} = N(t)r \ln \left( \frac{\mathfrak{K}}{N(t - \tau)} \right) - \delta N(t - \tau_1). \quad (14)$$

Для  $\dot{N} = rN \ln(K/N)$  ордината точки перегиба  $N_p = K/e$  на кривой решения лежит ниже, чем  $K/2$ . Параметр ограниченной ниши у новой среды в нашей модели не аналогичен роли устойчивого равновесия в (2). В данной модификации  $\dot{N} = rf(N; \tau\mathfrak{K}) - F(N; \tau_1)$  мы используем в обозначениях  $\mathfrak{K}$ , так как достижение уровня  $N(t) = \mathfrak{K}$  возможно, но при больших  $r\tau$  только кратковременно. В вычислительном сценарии с (14) наблюдается гибель популяции агрессивного вселенца после двух максимумов осцилляций, как это было в опытах Гаузе с интродукцией хищника в колонии их жертвы [26]. Длительных колебаний  $\lim_{t \rightarrow \infty} |\max N_*(t) - \min N_*(t)| = const$  в серии опытов с двумя противоборствующими организмами инфузориями получено не было. Опыт заканчивался  $N_1(t_1) = 0, N(t_2) = 0, t_2 > t_1$ . Для ситуации, когда активность вселенного хищника не сдерживает иной фактор, то трофическая система из двух видов неустойчивая. При уменьшении  $r$ -параметра траектория (14) демонстрирует обычные гармонические колебания  $N_*(t; \tau r)$  как и (1). В (14) нарушается свойство  $\forall t, N(0) \geq 0, N(t) \geq 0$ , но это незначительная и решаемая нами сложность. Мы преодолевали схожую математическую проблему ранее в [27].

Для инвазионных процессов в конкурентной среде требуется более сложная форма противодействия, чем в (14) с  $-\delta N(t - \tau_1)$ , так как у инфузории жертвы не было шансов на адаптацию. Многовидовую регуляцию мы описать не сможем, так у паразитов и хищников в естественной среде есть и свои естественные враги, патогены и факторы-регуляторы. В результате усложнения схемы регуляции добиться полного уничтожения популяции вселенца в реальных экосистемах сложно. Включение нелинейности давления в форме  $\mathcal{F}(N) = -\delta N^k(t - \tau_1), k > 1, \tau \geq \tau_1$

обосновано тем, что текущее воздействие патогенных факторов будет определяться предшествующим состоянием популяции вселенцев. Учтем, что такое воздействие в свою очередь связано зависимостью  $\omega(t, \tau)$  с текущей концентрацией доступных им для атаки организмов.

**5.1. Описание фазы депрессии**  $\min N_*(t; r\tau) \rightarrow \epsilon, r\tau < \infty$  **после активной инвазии.** Популяции вселенцев вынуждены преодолевать развитие адаптационного кризиса  $\min N_*(t; r\tau) \rightarrow \epsilon$ . Будем моделировать ситуацию активного противодействия, которое формируется спустя определенное время против быстро размножающейся локальной группы. Опишем противоборство абстрактной биотической среды в форме уравнения  $\dot{N} = rN(t)f(N(t-\tau)) - F(N^k(t-\tau_1); J)$  для случая регулируемого размножения вселенца, который способен воздействовать на свою среду. Представим, что ситуация внезапного обострения конкурентного противоборства развивается для популяции уже при приближении к порогу численности, который обозначим как  $J$ . В модели  $J$  полагаем значительно меньше теоретически допустимого  $\mathcal{K}$  в случае инвазии предельного уровня  $\mathfrak{K}$  или предкритического  $\mathcal{H}$ . Отразим в новой модели два явления: пороговый эффект и адаптивность противодействия активному росту численности вселенца в уравнении при логарифмической саморегуляции:

$$\frac{dN(t)}{dt} = rN(t) \ln \left( \frac{\mathfrak{K}}{N(t-\tau)} \right) - \delta \frac{N^m(t-\tau_1)}{(J-N(t))^2} - qN(t), \quad (15)$$

$$\delta > q, m \geq 2, N(0) < J < \mathfrak{K}. \quad (16)$$

При приближении  $N(t)$  к порогу  $\max N(t) \rightarrow J - \epsilon$  конкуренция при внешнем воздействии обостряется, однако уровень давления  $\Upsilon$  зависит от состояния популяции нелинейно:

$$\Upsilon = \delta \frac{N^m(t-\tau_1)}{(J-N(t))^2} \rightarrow \min,$$

когда с возникновением глубокого кризиса давление антагонистов ослабляется, но  $\Upsilon$  не уменьшается до нулевой отметки:  $\Upsilon > \epsilon$ .

В вычислительном сценарии для (15) на первом этапе происходит «логистическое» увеличение численности небольшой группы  $N(0) < J$ , но рост останавливается после  $N(t) > N_p$ . Вместо обычной стабилизации  $\lim_{t \rightarrow \infty} |\max N(t) - \min N(t)| \rightarrow 0$  после краткого превышения равновесия далее начинается стадия резкой убыли численности с возрастанием  $F(N^2; J^{-1})$  при  $N \rightarrow J$ . Кризис наступает до момента  $t_J$ , когда потенциал роста будет исчерпан внутривидовой  $\ln_{\mathcal{K}}$ -регуляцией



из-за исчерпания ресурсов среды. Воздействие специфических биотических внешних, но зависимых от численности  $N(t - \tau_1)$  факторов создает порог для начала депрессии. При приближении к пороговому уровню численности популяция вселенца переходит в демографический кризис. Исследуемое развитие ситуации особой формы противоборства показано в вычислительном сценарии на рисунке 1<sup>1</sup>.

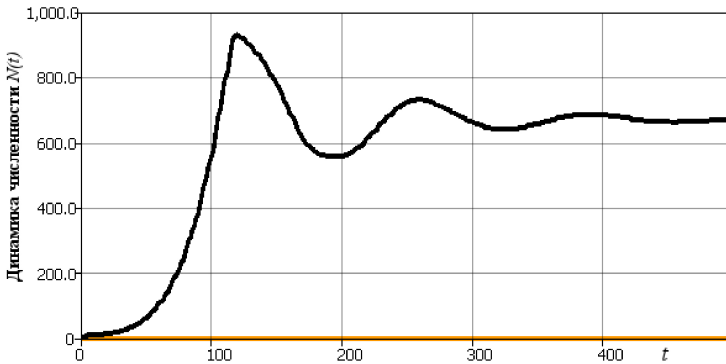


Рис. 1. Сценарий краткой индуцированной активностью антагонистов депрессии в новой модели (15) при  $m = 2, r = 7.19 \times 10^{-3}, J = 10^3, \mathfrak{K} = 15 \times 10^3, \tau = \tau_1 = 48, \delta = 10, q = 4.1 \times 10^{-3}, N(0) = 10$

Популяция проходит этап депрессии численности, так как реакция её противников при кризисе непостоянная и в это время ослабляется. В результате популяционный кризис после переходных колебаний стабилизируется на уровне  $\lim_{t \rightarrow \infty} N(t) = P$  и так, что  $P < J$  при малых  $q < \delta$ . Параметр  $\mathfrak{K}$  тут теоретическая предельная ёмкость истощаемых ресурсов среды, которая изначально малой локальной группой не достигается. В иммунологической интерпретации только очень высокая доза заражения  $N(0) > J$  тут способна стать летальной. Для патогенного вируса априори не существует равновесной ёмкости ниши, так как вирус разрушает организм – свою среду. Дополнительное внешнее воздействие  $q$  в модели (15) интерпретируется как влияние терапии или искусственных антропогенных мер воздействия, например интродукция паразитов против вредителей.

Модель (15) экологически применима для исходно малой группы  $N(0) < J/4$ . При  $N(0) = 2/3J, m > 2$  развитие покажет резкий эффект глубокого кризиса, но преодолимого  $N(t_m) \approx 0 + \epsilon, \lim_{t \rightarrow \infty} N(t) = P$ .

<sup>1</sup>Вычислительная среда «Rand Model Designer» алгоритм «Ovren—Zennaro».

Если уменьшать коэффициент воздействия окружения  $\delta$ , от которого зависит амплитуда сокращения численности при кризисе, то циклический режим сохраняется (вокруг уровня  $P$ ). Сценарии таких быстрых изменений на графиках имеют в англоязычной литературе образные названия: «хоккейная клюшка» (hockey stick graph), «бутылочное горлышко» (bottleneck), а полученный в модели (15) относится к образу «саксофон» (быстрое снижение с неполным восстановлением).

Рассмотренный сценарий отличается от ситуации прохождения у вновь образующейся популяцией стадии длительного минимума при стабильной малочисленной группе особей с относительно малым  $r$ . Длительное состояние минимальной реликтовой группы принципиально отлично по эволюционным аспектам от перехода к резкому кризису с восстановлением. Увеличение численности  $N(t) \rightarrow K$  в сценариях с длительным минимумом  $N(t) \approx L$  связано с нарастанием репродуктивного потенциала, где  $r \neq cost$ . Восстановление происходит не за счет увеличения индивидуальной плодовитости, но из-за возрастания выживаемости в ювенальном онтогенезе при появлении способов ухода от обнаружения.

**6. Экологическое обоснование свойств динамики вычислительного сценария.** Построению модели и вычислительному сценарию необходимо дать обоснование в реальных примерах. Известно из многих опытов еще с 1930-х гг., что вирусы бактериофаги не могут полностью подавить колонию бактерий, вызывая только временный кризис. В эксперименте с вселением бактериофага в колонию бактерий возникший эффект депрессии успешно преодолевался бактериями при адаптации их антивирусного механизма эндонуклеазы *CRISPR/CAS 9*. Специальный белок в клетке бактерии находил и разрезал молекулярными «ножницами» ДНК вируса, которую он опознавал по её специальному короткому фрагменту. На выделение подходящего образца чужеродной ДНК необходимо время. В результате методы терапии бактериальных инфекций с вирусами бактериофагами быстро теряли эффективность и не оправдали ожиданий [28], хотя полного восстановления колонии бактерий не наблюдалось. Сейчас этот механизм ножниц адаптивной защиты активно используется для точечного геномного редактирования.

Полученный в (15) сценарий описывает динамику локальных популяций оленей, интродуцированных на островах в Канаде. Колебания могут сохраняться в модели. Популяция при отсутствии хищников размножается, но олени при большой скученности уничтожают растительность Арктики и разрушают свою среду. Питание становится недостаточным, возникают эпизоотии, и смертность молодых особей увеличивается. Популяция входит в депрессию. Восстановление после  $\epsilon$ -кризиса зависит

от способности к регенерации их пищевых ресурсов, что происходит с запаздыванием. На острове Врангеля популяция оленей погибла полностью. Полной регенерации не происходит и так создается порог, который значительно меньше *изначальной* допустимой емкости экологической ниши  $K$ . В системе присутствуют и конкурируют два запаздывания: репродуктивный цикл оленей и темпы восстановления среды. Это создаст колебания со снижением среднего значения. Северная растительность восстанавливается медленно. Интересно, что при искусственном изъятии популяция оленей стабилизируется, но без изъятия  $q$  флуктуации продолжают, что и подтверждает разработанная нами модель. Регуляция искусственных популяций необходима.

В уравнении (15) феноменологически описано сложно регулируемое активное противодействие  $\Upsilon$ . Ослабление давления определено зависимостью от порога обнаружения, что соответствует возрастающей активности патогенов или паразитов в плотных скоплениях [30].

Предложенную функцию воздействия  $F(N^m(t - \nu); J)$  можно включать в модель затухающих колебаний активности вредителей леса:

$$\frac{dN(t)}{dt} = rN(t) \left( \frac{K - N^k(t - \tau)}{(K + cN^m(t - \tau))} \right), m = k + 1. \quad (17)$$

Таким способом мы смогли получить описание демпфирования серии всплеск в случае существования ограниченного лесного ресурса и противодействия естественных врагов-паразитов:

$$\frac{dN(t)}{dt} = rN(t) \left( \frac{K - N^2(t - \tau)}{(K + cN^3(t - \tau_1))} \right) - \delta \frac{N^m(t - \nu)}{(J - N(t))^2} - qN(t), \quad (18)$$

$$\delta > q, m \geq 2, N(0) < J < K. \quad (19)$$

Уравнение (18) можно использовать в составе «вольтерровских систем» для описания трофического взаимодействия с двумя пороговыми эффектами: пороговым сопротивлением со стороны среды и эффектом выживаемости агрегированной группы особей.

Существуют интересные примеры кризисной динамики и помимо популяционных наблюдений, например при терапии с применением иммунных клеток [31].

**7. Заключение.** Разработаны модели специфических популяционных процессов на основе уравнений с запаздыванием для ситуаций активного сопротивления, которое оказывает биотическое окружение (возможно искусственно индуцированное) инвазионному виду с высоким репродуктивным потенциалом. Модель (15) для сценария прохождения кризиса инвазивной популяцией разработана на основе двух явлений запаздывания: в регуляции из-за восстановления необходимых ресурсов и порогового запаздывающего противодействия. Моделируемый кризис резко начинается в фазе быстрого роста при активном сопротивлении биотического окружения при приближении к пороговой численности. Увеличение  $\tau_1$  делает кризис более выраженным. Явления резкой смены быстрого роста глубокой депрессией наблюдались в экодинамике разных чужеродных популяций [32].

Рассмотренная модель (15) может получить дальнейшее развитие в контексте запаздывающей выработки иммунного ответа от  $T$ -лимфоцитов  $CD8+$  на острую вирусную инфекцию. Полученный сценарий согласуется с вариантом хронизации очага инфекции [33]. Новые модели в дальнейшем могут применяться для прямого описания противодействия в системах уравнений, где в функции  $F(N(t), \Xi(t - \xi))$  указывается непосредственно численность вида-антагониста. Модель (18) подходит для ситуации взаимодействия мутирующего вируса с иммунитетом организма, реагирующего на неактуальные уже антигены, измененные участки белков вируса, что вызывает периодические рецидивы заболевания. Для случая вируса гриппа  $A1N1$  реакция иммунной системы часто зависит от начальной дозы заражения  $\Xi(0)$ .

Моделирование кризисных явлений традиционными численными алгоритмами актуально для современных задач биологического контроля. Нежелательные для экономики природопользования вселенцы на первых стадиях агрессивной инвазии часто не испытывают сопротивления со стороны биотического окружения. Быстрорастущая популяция в таких условиях оказывается вне области балансового равновесия  $\mathfrak{K} \pm \epsilon$  с ресурсами среды обитания, как для чужеродного хищного гребневика *Beroe ovata* в Азовском море согласно результатам Никитиной А.В. [34].

Некомпенсируемая репродуктивная активность вызывает резкие перепады численности. Активизация приспособившихся врагов происходит позднее, что резко меняет характер процесса стремительного роста численности вселенца, но редко останавливает инвазию. До момента появления приспособившегося эффективного вида-антагониста иногда требовались десятилетия. Так случилось с появлением в 1868 г. в Северной Америке нескольких особей бабочки непарного шелкопряда *Lymantria*

*dispar* [35], спорадически уничтожавшей потом леса. Дополнительно известно из наблюдений, если существуют пороговые состояния, то величина исходной группы проникших особей влияет на дальнейшее развитие процесса. Интерес представляет точечное включение стохастической компоненты в непрерывные модели. Влияние фактора стохастичности в биологических процессах достаточно ограниченное. возмущающее влияние, которое вносит долю неопределенности в варианты развивающегося процесса, но не делает его случайным. Перспективным методом описания не полностью предопределенных ситуаций при активации ответа среды является возмущение равномерно распределенной случайной величиной  $\hat{\gamma}$  значения запаздывания  $F(N(t - \gamma\tau_1))$  и получение возмущенной циклической траектории  $\widehat{N}_*(t)$  с непостоянной амплитудой колебаний.

### Литература

1. Ludwig D., Jones D.D., Holling C.S. Qualitative Analysis of Insect Outbreak Systems: The Spruce Budworm and Forest // *The Journal of Animal Ecology*. 1978. vol. 47. no. 1. pp. 315–332.
2. Myers J.H., Cory J.S. Biological Control Agents: Invasive Species or Valuable Solutions? // *Impact of Biological Invasions on Ecosystem Services*. 2017. vol. 12. pp. 20–26.
3. Переварюха А.Ю. Нелинейные эффекты и переходные режимы в динамике новых моделей управления биоресурсами // *Труды СПИИРАН*. 2011. № 16. С. 243–255.
4. Переварюха А.Ю. Об определении фрактальных объектов в динамике моделей управления биоресурсами // *Труды СПИИРАН*. 2012. № 24. С. 211–221.
5. Desharnais A., Laifu L. Stable Demographic Limit Cycles in Laboratory Populations of *Tribolium castaneum* // *Journal of Animal Ecology*. 1987. vol. 56. pp. 885–906.
6. Hutchinson G.E. Circular causal systems in ecology // *Ann. New York Acad. Sci.* 1948. vol. 50. pp. 221–248.
7. Wright E.M. A non-linear difference-differential equation // *J.Reine Angew. Math.* 1955. vol. 194. pp. 66–87.
8. May R.M., Conway G.R., Hassell M.P. Time delays, density-dependence and singlespecies oscillations // *J. Anim. Ecol.* 1974. vol. 43. pp. 747–770.
9. Nedorezov L.V. Approximation of time series of paramecia caudatum dynamics by the Verhulst and Gompertz models: a non-traditional approach // *Biophysics*. 2015. vol. 60. pp. 457–465.
10. Salesa L.P., Haywar M.W., Loyola R. What do you mean by «niche»? Modern ecological theories are not coherent on rhetoric about the niche concept // *Acta Oecologica*. 2021. vol. 110. pp. 103–1070.
11. Дубровская В.А., Трофимова И.В. Модель динамики структурированных субпопуляций осетровых рыб Каспия с учетом отклонений в темпах развития молоди // *Журнал Белорусского государственного университета. Биология*. 2017. № 3. С. 76–86.
12. Северцов А.С. Соотношение фундаментальной и реализованной экологических ниш // *Журнал общей биологии*. 2012. Т. 73. С. 323–333.
13. Kolesov A.Yu., Rozov N.Kh. The theory of relaxation oscillations for Hutchinson's equation // *Sb. Math.* 2011. vol. 202. pp. 829–858.

14. Сабатулина Т.Л. Об устойчивости обобщенного уравнения Хатчинсона с распределенным переменным запаздыванием // Вестник Пермского государственного технического университета. Механика. 2009. № 1. С. 46-56.
15. Глызин С.Д. Учет возрастных групп в уравнении Хатчинсона // Моделирование и анализ информационных систем. 2007. Т. 14. № 3. С. 29–42.
16. Smith M. *Mathematical ideas in biology*. Cambridge University Press, London, 1968, 168 p.
17. Liz E. Delayed logistic population models revisited // *Publ. Mat.* 2014. vol. 58. pp. 309–331.
18. Глызин Д.С., Кащенко С.А., Полстянов А.С. Пространственно-неоднородные периодические решения в распределенном уравнении Хатчинсона // Моделирование и анализ информационных систем. 2009. Т. 16. № 4. С. 77–85.
19. Кащенко И.С., Кащенко С.А. Динамика уравнения с двумя запаздываниями, моделирующего численность популяции // Известия вузов. Прикладная нелинейная динамика. 2019. Т. 27. С. 21–38.
20. Liu Y., Wei J. Bifurcation analysis in delayed Nicholson blowflies equation with delayed harvest // *Nonlinear Dynamics*. 2021. Vol. 105. P. 1805–1819.
21. Kolesov A.Y., Mishchenko E.F., Rozov N.K. A modification of Hutchinson's equation // *Computational Mathematics and Mathematical Physics*. 2010. vol. 50. pp. 1990–2002.
22. Никольский М.С. Об одной задаче оптимального управления, связанной с общей моделью А.Д. Базыкина «хищник–жертва» // Дифференциальные уравнения. 2011. Т. 47. С. 1638–1645.
23. Metzler W. *Mathematical Modelling Bifurcations of equilibria in Bazykin's predator-prey model* // *Mathematical Modelling*. 1985. vol. 6. pp. 111–123.
24. Bazykin A.D., Khibnik A.I. Bilocal model of dissipative structure // *Biophysics*. 1982. vol. 27. no. 1. pp. 133–138.
25. Perevaryukha A.Y. An iterative continuous-event model of the population outbreak of a phytophagous Hemipteran // *Biophysics*. 2016. vol. 61. pp. 334-341.
26. Gause G.F. *The struggle for existence*. Baltimore: Williams Wilkins. 1934, 163 p.
27. Perevaryukha A.Y. A continuous model for oscillating outbreaks of the population of a phytophagous moth, the tent caterpillar, *Malacosoma disstria* (Lepidoptera, Lasiocampidae) // *Biophysics*. 2020. vol. 65. № 1. pp. 118-130.
28. Lenski R.E. Dynamics of Interactions between Bacteria and Virulent Bacteriophage // *Advances in Microbial Ecology*. 1988. vol 10. pp. 1–44.
29. Forest and Wildlife Research Center report: <https://www.msudeer.msstate.edu/deer-habitat-carrying-capacity.php>
30. Morozov A.S., Rytova S.V., Thompson L.C. Introducing entomophagous insects to control pests: prediction of target species density // *Russian Entomological Journal*. 2003. vol. 12. pp. 441–445.
31. Kuznetsov V.A., Knott G.D. Modeling tumor regrowth and immunotherapy // *Mathematical and Computer Modelling*. 2001. vol. 33. no. 12–13. pp. 1275–1287.
32. Mikhailov V.V., Reshetnikov Yu.S. Model of Fish Population Dynamics with Calculation of Individual Growth Rate and Hydrological Situation Scenarios // *Information and Control Systems*. 2018. no. 4, pp. 31–38.
33. Luo C., Liu M. Dynamic changes and prevalence of SARS-CoV-2 IgG/IgM antibodies: Analysis of multiple factors // *International Journal of Infectious Diseases*. 2021. vol. 108. pp. 57–62.
34. Nikitina A.V., Sukhinov A.I. Optimal control of sustainable development in the biological rehabilitation of the Azov Sea // *Mathematical Models and Computer Simulations*. 2017. vol. 9. no. 1. pp. 101–107.

35. Andreadis T., Weseloh R. Discovery of *Entomophaga maimaiga* in North American gypsy moth, *Lymantria dispar* // Proc. Nat. Acad. Sci. U S A. 1990. vol. 87. pp. 2461—2465.

**Переварюха Андрей Юрьевич** — канд. техн. наук, старший научный сотрудник, лаборатория прикладной информатики и проблем информатизации общества, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук (СПб ФИЦ РАН). Область научных интересов: моделирование нелинейных явлений в биологических процессах: коллапс запасов промысловых рыб, анализ всплеск численности и депрессии численности насекомых вредителей, изменение темпов роста организмов, экстремальные инвазионные процессы, сценарии кризиса и депрессии, клеточные автоматы, экспертное управление. Число научных публикаций РИНЦ — 161. madelf@rambler.ru; 14-я линия В.О., 39, 199178, Санкт-Петербург, Россия; р.т.: +7(812)328-3337.

A. PEREVARYUKHA  
**DYNAMIC MODEL OF POPULATION INVASION WITH  
DEPRESSION EFFECT**

*Perevaryukha A. Yu.* **Dynamic Model of Population Invasion with Depression Effect.**

**Abstract.** The article is devoted to the study of one of the current scenarios for the development of population processes in contemporary ecological systems. Biological invasions have become extremely common due to climate change, economic activities to improve ecosystem productivity, and random events. The invader does not always smoothly occupy an ecological niche, as in logistic models. The dynamics of the situations we have chosen after the introduction of an alien species is extremely diverse. In some cases, the phenomenon of an outbreak of abundance is quickly realized up to the beginning of the destruction by the species of its new range. The development of the situation in the process of invasion depends on the superposition of biotic and abiotic factors. The dynamics of the abundance of the invader is affected by the favorable conditions and, to a greater extent, by the possibility of realizing the reproductive potential and the resistance of the biotic environment. Counteraction develops with a delay and manifests itself when the invader reaches a significant number. In the work, a continuous model of the invasive process with a sharp transition to a state of population depression has been developed. The stage of the population crisis ends with the transition to equilibrium, since the resistance in the model scenario depends adaptively and in a threshold way on the number. The problem of computational description of a scenario with active but delayed environmental resistance is practically relevant for situations of developing measures of artificial resistance to an undesirable invader. In the solution of our model, there is a mode of prolonged stable fluctuations after exiting the depression stage.

**Keywords:** dynamic models, computational scenarios, active invasions, crisis and depression, simulation models of ecology, control in biosystems.

**Perevaryukha Andrey** — Ph.D., Senior researcher, Laboratory of applied informatics and problems of information society, St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS). Research interests: modeling of nonlinear phenomena in biological processes: the collapse of commercial fish stocks, analysis of outbreaks and depressions in the number of insect pests, changes in the growth rates of organisms, extreme invasive processes, crisis and depression scenarios, cellular automata, expert management. The number of publications — 161. [madelf@rambler.ru](mailto:madelf@rambler.ru); 39, 14-th Line V.O., 199178, St. Petersburg, Russia; office phone: +7(812)328-1919.

## References

1. Ludwig D., Jones D.D., Holling C.S. Qualitative Analysis of Insect Outbreak Systems: The Spruce Budworm and Forest. *The Journal of Animal Ecology*. 1978. vol. 47. no. 1. pp. 315–332.
2. Myers J.H., Cory J.S. Biological Control Agents: Invasive Species or Valuable Solutions? *Impact of Biological Invasions on Ecosystem Services*. 2017. vol. 12. pp. 20–26.
3. Perevaryukha A. Yu. [Nonlinear effects and transitional regimes in the dynamics of new models of bioresource management]. *SPIIRAS Proceedings — Trudy SPIIRAN*. 2011. no. 16. pp. 243—255. (In Russ.).



4. Perevaryukha A.Yu. On the definition of fractal objects in the dynamics of bioresource management models. SPIIRAS Proceedings — Trudy SPIIRAN. 2012. no. 24. pp. 211—221.
5. Desharnais A., Laifu L. Stable Demographic Limit Cycles in Laboratory Populations of *Tribolium castaneum*. Journal of Animal Ecology. 1987. vol. 56. pp. 885–906.
6. Hutchinson G.E. Circular causal systems in ecology. Ann. New York Acad. Sci. 1948. vol. 50. pp. 221—248.
7. Wright E.M. A non-linear difference-differential equation. J. Reine Angew. Math. 1955. vol. 194. pp. 66—87.
8. May R.M., Conway G.R., Hassell M.P. Time delays, density-dependence and singlespecies oscillations. J. Anim. Ecol. 1974. vol. 43. pp. 747—770.
9. Nedorezov L.V. Approximation of time series of paramecia caudatum dynamics by the Verhulst and Gompertz models: a non-traditional approach. Biophysics. 2015. vol. 60. pp. 457–465.
10. Salesa L.P., Haywar M.W., Loyola R. What do you mean by «niche»? Modern ecological theories are not coherent on rhetoric about the niche concept Acta Oecologica. 2021. vol. 110. pp. 103–1070.
11. Dubrovskaya V.A., Trofimova I.V. [Model of the dynamics of structured subpopulations of Caspian sturgeons taking into account deviations in the rate of development of juveniles]. Journal of the Belarusian State University. Biology — Zhurnal Belorusskogo gosudarstvennogo universiteta. Biologiya . 2017. no. 3. pp. 76–86. (In Russ.).
12. Severtsov A.S. [Correlation between fundamental and realized ecological niches]. Journal of General Biology — Zhurnal obshchey biologii. 2012. T. 73. C. 323–333. (In Russ.).
13. Kolesov A.Yu., Rozov N.Kh. The theory of relaxation oscillations for Hutchinson's equation. Sb. Math. 2011. vol. 202. pp. 829—858.
14. Sabatulina T.L. [On the stability of the generalized Hutchinson equation with distributed variable delay]. Bulletin of the Perm State Technical University. Mechanics. — Vestnik Permskogo gosudarstvennogo tekhnicheskogo universiteta. Mekhanika. 2009. no. 1. pp. 46–56.
15. Glyzin S.D. [Accounting for age groups in the Hutchinson equation]. Modeling and analysis of information systems — Modelirovaniye i analiz informatsionnykh sistem. 2007. vol. 14. no. 3. pp. 29–42. (In Russ.).
16. Smith M. Mathematical ideas in biology. Cambridge University Press, London, 1968, 168 p.
17. Liz E. Delayed logistic population models revisited. Publ. Mat. 2014. vol. 58. pp. 309–331.
18. Glyzin D.S., Kashchenko S.A., Polstyanov A.S. [Spatially inhomogeneous periodic solutions in the distributed Hutchinson equation]. Modeling and analysis of information systems — Modelirovaniye i analiz informatsionnykh sistem. 2009. vol. 16. no. 4. pp. 77–85. (In Russ.).
19. Kashchenko I.S., Kashchenko S.A. [Dynamics of an equation with two delays simulating the size of a population]. Izvestiya vuzov. Applied nonlinear dynamics. – Izvestiya vuzov. Prikladnaya nelineynaya dinamika. 2019. vol. 27. pp. 21–38. (In Russ.).
20. Liu Y., Wei J. Bifurcation analysis in delayed Nicholson blowflies equation with delayed harvest. Nonlinear Dynamics. 2021. vol. 105. pp. 1805–1819.
21. Kolesov A.Y., Mishchenko E.F., Rozov N.K. A modification of Hutchinson's equation. Computational Mathematics and Mathematical Physics. 2010. vol. 50. pp. 1990–2002.
22. Nikolsky M.S. [On one optimal control problem related to the general model of A.D. Bazykin «predator–prey»]. Differential Equations. 2011. vol. 47. pp. 1638–1645. (In Russ.).

23. Metzler W. Mathematical Modelling Bifurcations of equilibria in Bazykin's predator-prey model. *Mathematical Modelling*. 1985. vol. 6. pp. 111–123.
24. Bazykin A.D., Khibnik A.I. Bilocal model of dissipative structure. *Biophysics*. 1982. vol. 27. no. 1. pp. 133–138.
25. Perevaryukha A.Y. An iterative continuous-event model of the population outbreak of a phytophagous Hemipteran. *Biophysics*. 2016. vol. 61. no. 2. pp. 334–341.
26. Gause G.F. *The struggle for existence*. Baltimore: Williams Wilkins. 1934, 163 p.
27. Perevaryukha A.Y. A continuous model for oscillating outbreaks of the population of a phytophagous moth, the tent caterpillar, *Malacosoma disstria* (Lepidoptera, Lasiocampidae). *Biophysics*. 2020. vol. 65. no. 1. pp. 118–130.
28. Lenski R.E. Dynamics of Interactions between Bacteria and Virulent Bacteriophage. *Advances in Microbial Ecology*. 1988. vol 10. pp. 1–44.
29. Forest and Wildlife Research Center report: <https://www.msudeer.msstate.edu/deer-habitat-carrying-capacity.php>
30. Morozov A.S., Rytova S.V., Thompson L.C. Introducing entomophagous insects to control pests: prediction of target species density. *Russian Entomological Journal*. 2003. vol. 12. pp. 441–445.
31. Kuznetsov V.A., Knott G.D. Modeling tumor regrowth and immunotherapy. *Mathematical and Computer Modelling*. 2001. vol. 33. no. 12–13. pp. 1275–1287.
32. Mikhailov V.V., Reshetnikov Yu.S. Model of Fish Population Dynamics with Calculation of Individual Growth Rate and Hydrological Situation Scenarios. *Information and Control Systems*. 2018. no. 4, pp. 31–38.
33. Luo C., Liu M. Dynamic changes and prevalence of SARS-CoV-2 IgG/IgM antibodies: Analysis of multiple factors. *International Journal of Infectious Diseases*. 2021. vol. 108. pp. 57–62.
34. Nikitina A.V., Sukhinov A.I. Optimal control of sustainable development in the biological rehabilitation of the Azov Sea. *Mathematical Models and Computer Simulations*. 2017. vol. 9. no. 1. pp. 101–107.
35. Andreadis T., Weseloh R. Discovery of *Entomophaga maimaiga* in North American gypsy moth, *Lymantria dispar*. *Proc Natl Acad Sci U S A*. 1990. vol. 87. pp. 2461–2465.

К.О. Логинов

**ЧИСЛЕННОЕ РЕШЕНИЕ ЗАДАЧИ ФИЛЬТРАЦИИ ОЦЕНОК  
ИНФОРМАЦИОННОГО ВОЗДЕЙСТВИЯ НА ЭЛЕКТОРАТ**

*Логинов К.О. Численное решение задачи фильтрации оценок информационного воздействия на электорат.*

**Аннотация.** В статье предложены постановка и численная схема решения задачи фильтрации оценок информационного воздействия средств масс-медиа на электорат, позволяющие с высокой степенью точности на заданном интервале наблюдения определить число индивидов в обществе, отдающих предпочтение определенному политическому субъекту (мнению). Основу постановки задачи составляет математическая модель оценки информационного воздействия на электорат при проведении выборных кампаний, которая сводится к решению стохастического дифференциального уравнения – уравнения состояния. Его исследование при составлении модели фильтрации оценок информационного воздействия предложено свести к численному решению уравнения Дункана–Мортенсена–Закаи при введении дополнительного уравнения наблюдения, которое получается из уравнения состояния при оценке его стохастических компонент (наблюдаемые интенсивности агитации от разнородных средств масс-медиа) методами полиспектрального анализа. Решение уравнения Дункана–Мортенсена–Закаи выполнено в проекционной постановке метода Галёркина при сведении к системе линейных дифференциальных уравнений и получении ее решения при дискретизации интервала анализа на подинтервалы и использования метода матричной экспоненты. Для уточнения особенностей алгоритмической реализации составленной численной схемы сформирована рекурсивная процедура численной фильтрации оценок информационного воздействия, представленная в виде псевдокода. Для наглядного сравнения результативности сформированного численного решения задачи нелинейной фильтрации оценок информационного воздействия средств масс-медиа на электорат проведены расчеты на тестовых выборках большого объема для различных значений количества политических субъектов (мнений) и числа подгрупп, на которые разделяется разнородный социум (электорат). Под результативностью решения понимается апостериорно вычисленная зависимость вероятности принятия ошибочного решения о победе политического субъекта-лидера от дисперсионных параметров нестационарных негауссовских шумов наблюдаемых интенсивностей агитации. Зависимости результативности предложенного решения сравнены с результатом непосредственного численного решения систем уравнений состояния и наблюдения.

**Ключевые слова:** численное решение, фильтрация оценок информационного воздействия, уравнение Дункана–Мортенсена–Закаи, оператор Фоккера–Планка–Колмогорова, полиспектральный анализ, гауссовы базисные функции.

**1. Введение.** В последнее время отмечается возрастающий интерес к исследованию технологий и ресурсов влияния социальных медиа на политические процессы, разработке математических методов и алгоритмических решений, позволяющих на основе наблюдений формировать обоснованные выводы о предпочтительности индивидов [1–7]. Указанный интерес обуславливается с одной стороны активным повсеместным внедрением интернет технологий, с другой стороны,

нетипичными для стандартных избирательных моделей агитационных особенностей социальных медиа. Описанная оригинальность связана с: распространением недостоверной информации; трудностями возложения ответственности, определения правового статуса, фиксации нарушений из-за пробелов в информационном праве; опасности разрешения имиджа при неумелом использовании социальных медиа; возможности быстрого получения обратной связи. Актуальность и своевременность указанных задач также подтверждается резко увеличившимся количеством сфальсифицированных новостей в интернете (по данным ТАСС в период пандемии в 2020 году число фейков в интернете возросло в 10 раз, за 2021 год их объем увеличен еще в 2 раза). Для обоснованного формирования способов противодействия указанному факту первостепенно необходимо системное понимание соответствующих механизмов информационных процессов, достижимое при изучении, формализации и моделировании особенностей информационного воздействия в социальных медиа.

Основные подходы математических представлений по получению оценок влияния информации, распространяемой социальными медиа в обществе, основываются на социальных сетях. Предполагается, что информационное воздействие формирует предпочтения индивидов в отношении политических субъектов (мнений). Социальная сеть представляется объединением социальных позиций (социальных акторов и их связей [8]) при представлении структуры социума за счет введения соответствующей топологии – социального графа. Основные задачи при формировании подобных моделей заключаются в оценке адекватности составляемых представлений, верификации корректности топологической структуры, определении ее устойчивости [9]. При этом топология реальных социальных сетей, в прямой постановке моделирующих процессы информационного противоборства в социальных медиа, обуславливается значительной количественной сложностью в сравнении с их качественными свойствами. Указанное определяет противоречие, которое устанавливает отсутствие баланса между сложностью и точностью к измерениям существующих математических моделей оценки информационного воздействия на электорат.

В направлении разрешения сформулированного противоречия при представлении социума численностью  $N_0$  совокупностью взаимодействующих  $M$  подгрупп ( $N_m < N_0$ ,  $N_m \geq 1$ ,  $m = \overline{1, M}$ ,

$\sum_{m=1}^M N_m = N_0$ , где  $N_m$  – число индивидов в  $m$ -й подгруппе социума) в работах [10–12], в дополнение к моделям социальных сетей и в развитие известных (например, из статьи [13]) математических моделей информационного противоборства в структурированном социуме, предложена математическая модель и численная схема оценки информационного воздействия на электорат при проведении выборных кампаний. Алгоритмические решения [10–12] сводятся к тому, что, зная параметры, начальные данные и периодичность информационного воздействия (характеризуется интенсивностями распространения положительной или отрицательной информации в социальных медиа), оценивается число сторонников того или иного политического субъекта (мнения). Основная отличительная особенность математической модели [10–12] заключается в учете стохастического характера интенсивностей распространения информации от внешних источников о политических субъектах (мнениях) при представлении модели в виде стохастической системы диффузионно-скачкообразного типа.

Вместе с тем в решениях [10–12] не уточняются вычислительные особенности, определяющие правила получения оценок наблюдаемых интенсивностей распространения информации при составлении вектора сноса и матрицы диффузии. Корректное определение последних существенным образом влияет на точность результирующего решения. Цель настоящей статьи состоит в уточнении указанного вопроса, требующего в том числе модификацию решений [10–12].

**2. Постановка задачи фильтрации оценок информационного воздействия на электорат.** Учитывая работы [10–12] и схематическое представление задачи математического моделирования процессов формирования информационного воздействия на электорат (рисунок 1), формализуем постановку задачи фильтрации.

В отношении  $K$  политических субъектов  $L$  внешних источников информации с интенсивностями  $\alpha_{kl}(t)$  и  $\gamma_{kl}(t)$  ( $l \in \{1, \overline{L}\}$ ;  $k \in \{1, \overline{K}\}$ ) распространяют положительные и отрицательные сведения соответственно. Значения интенсивностей  $\alpha_{kl}(t)$ ,  $\gamma_{kl}(t)$  может быть определено через отношение числа сообщений соответствующей тональности (положительная и отрицательная соответственно) к периоду времени наблюдения (в рассматриваемой задаче длительность интервала измеряется единицей времени – 1 час или 1 сутки) при учете их объема и

вероятностных характеристик оценки тональности. Для определения интенсивностей  $\alpha_{kl}(t)$ ,  $\gamma_{kl}(t)$  предлагается использовать программу для ЭВМ «Система анализа тональности публикаций СМИ «Тонус» [14]. В рамках современного информационного общества к  $L$  внешним источникам информации можно отнести:

1. Средства массовой информации и средства агитации: печать (газеты, журналы, альманахи, книги, брошюры, листовки, плакаты, стенды и билборды); радио (региональное, центральное); телевидение (центральное, региональное, местное); таргетированная реклама в сети Интернет; зарубежные средства массовой информации.
2. Социальные медиа: социальные сети (vk, facebook, TikTok, Одноклассники и пр.); мессенджеры (Telegram, WhatsApp, WeChat, QQ и пр.); видеохостинги (YouTube, Rutube, Dailymotion, Vzaar и пр.); блог-платформы (WordPress, LiveJournal, Blogger, Tumblr и пр.); тематические Интернет сайты и форумы; и прочее.
3. Информационные и аналитические агентства.

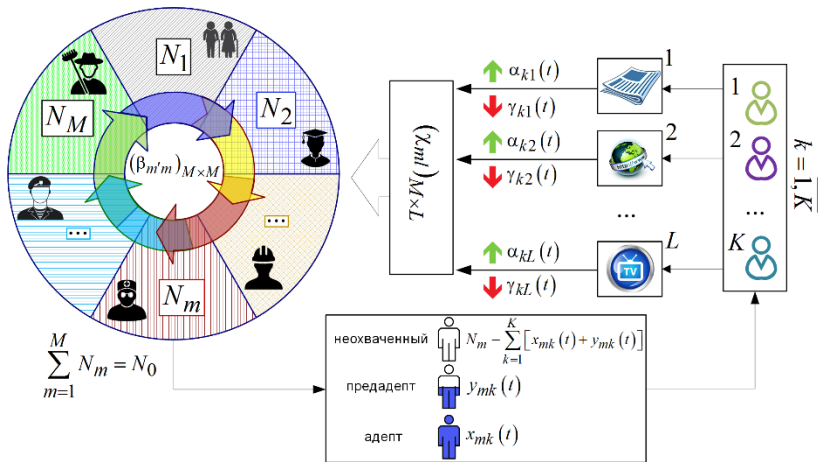


Рис. 1. Схематическое представление содержательной постановки задачи математического моделирования процессов формирования информационного воздействия на электорат

Под воздействием распространяемой информации на индивидов  $m$ -й подгруппы ( $m \in \{1, M\}$ ) социума (примеры признаков разделения социума на подгруппы: пол, возраст, образование, социальный статус, параметры региона проживания и пр.) общей

численностью  $N_0$  ( $N_m < N_0$ ,  $N_m \geq 1$ ,  $\sum_{m=1}^M N_m = N_0$ ) с учетом коэффициентов восприятия  $\chi_{ml} \in [0,1]$  (коэффициент восприятия  $\chi_{ml}$  определяет степень доверия индивидов  $m$ -й подгруппы к  $l$ -му источнику информации)  $l$ -х источников осуществляется информационное воздействие. При этом у индивидов формируется приверженность к  $k$ -му политическому субъекту при разделении каждой  $m$ -й подгруппы индивидов на три класса (рисунок 2): 1) неохваченные; 2) предадепты  $mk$ ; 3) адепты  $mk$ .

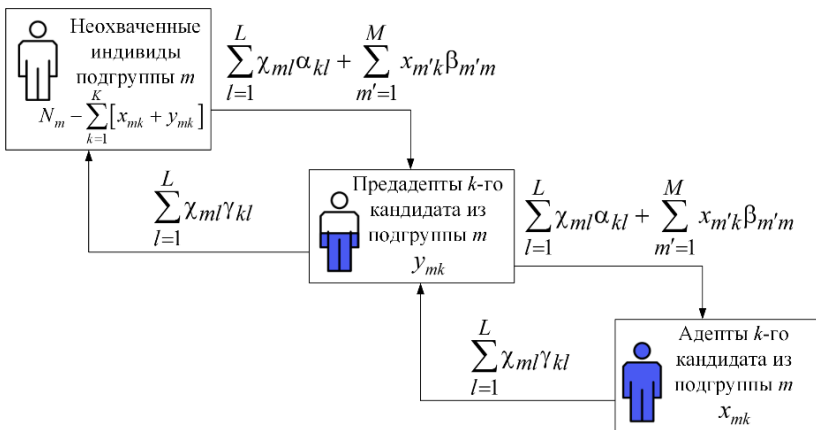


Рис. 2. Диаграмма состояний индивидов в рассматриваемой модели формирования информационного воздействия на электорат

У неохваченных индивидов отсутствуют предпочтения в отношении какого-либо кандидата.

Адептом  $mk$  является индивид из  $m$ -й подгруппы, который отдает предпочтение  $k$ -му кандидату и по средствам межличностной коммуникации с интенсивностью  $\beta_{mm'} \geq 0$  формирует предпочтение в отношении  $k$ -го кандидата среди индивидов  $m'$ -х ( $m' \in \{1, \overline{M}\}$ ) подгрупп либо за счет распространения положительной информации о  $k$ -м кандидате, либо путем распространения отрицательной информации о кандидатах  $k'$  ( $k' = \overline{1, K}, k' \neq k$ ). Уточним, что особенности формирования интенсивностей  $\alpha_{kl}(t)$  и  $\gamma_{kl}(t)$  адептами  $mk$  в социальных медиа не учитываются. Интенсивность  $\beta_{mm'}$

определяет степень восприятия информации индивидом  $m'$ -й подгруппы, поступающей при межличностной коммуникации от индивидов  $m$ -й подгруппы.

Преадептом  $mk$  является индивид из  $m$ -й подгруппы, который отдает предпочтение  $k$ -му кандидату, но никакой информации не распространяет. Первоначально неохваченный индивид  $m$ -й подгруппы под воздействием положительной информации из внешних источников и за счет межличностной коммуникации становится преадептом  $mk$ , а затем – адептом  $mk$ . Под воздействием отрицательной информации из внешних источников в отношении  $k$ -го кандидата происходит обратный переход. Адептом кандидата может стать только преадепт соответствующего кандидата, а неохваченным индивидом – преадепт. Число адептов  $mk$  в момент времени  $t \in [t_0, t_1]$  ( $[t_0, t_1]$  – интервал анализа) обозначим  $x_{mk}(t) \in [0, N_m]$ , преадептов –  $y_{mk}(t) \in [0, N_m]$ .

Уточним, что в отношении разрабатываемого алгоритмического решения особенности вычисления  $\chi_{ml}$  и  $\beta_{m'm}$  не рассматриваются, а их значения полагаются исходно заданным. В практических системах  $\chi_{ml}$  и  $\beta_{m'm}$  могут определяться экспертным оцениванием (задаваться по результатам социологических опросов или данным Всероссийского центра изучения общественного мнения) или в результате анализа взаимосвязности сетевых моделей социальных сетей [15].

Для введенных представлений решение исследуемой задачи сводится к определению в момент времени  $t \in [t_0, t_1]$  числа адептов  $\hat{x}_{mk}(t)$  и преадептов  $\hat{y}_{mk}(t)$  для заданных начальных условий  $x_{mk}(t_0)$ ,  $y_{mk}(t_0)$  и следующих допущений [10, 12]:

1. Значения коэффициентов восприятия  $\chi_{ml}$  и интенсивностей межличностной коммуникации  $\beta_{m'm}$  не зависят от времени  $t$ .
2. Переменные  $x_{mk}(t)$ ,  $y_{mk}(t)$  составляют непрерывный векторный Марковский процесс.
3. Интенсивности  $\alpha_{kl}(t)$ ,  $\gamma_{kl}(t)$  складываются из соответствующих истинных значений  $0 \leq \alpha_{kl}^0(t), \gamma_{kl}^0(t) < \infty$  и ошибок наблюдения  $\tilde{\alpha}_{kl}(t)$ ,  $\tilde{\gamma}_{kl}(t)$ , являющихся белым шумом с



соответствующими характеристиками:  $\mathbb{E}[\tilde{\alpha}_{kl}] = \mathbb{E}[\tilde{\gamma}_{kl}] = 0$ ;  $\text{cov}[\tilde{\alpha}_{kl}] = \text{cov}[\tilde{\gamma}_{kl}] = \delta(t - \tau)$ ;  $\text{cov}[d\tilde{\alpha}_{kl}] = [\varepsilon_{kl}^\alpha]^{-2}$ ;  $\text{cov}[d\tilde{\gamma}_{kl}] = [\varepsilon_{kl}^\gamma]^{-2}$ .

4. Скорость «вербовки» адептов  $mk$  пропорциональна числу предадептов  $mk$ , а скорость «вербовки» предадептов  $mk$  пропорциональна числу неохваченных индивидов  $m$ -й подгруппы с учетом перехода предадептов  $mk$  в адепты  $mk$  [13].

5. Величины начальных значений  $x_{mk}(t_0)$ ,  $y_{mk}(t_0)$  известны и принимаются равными 0.

Эволюция числа адептов  $x_{mk}(t)$  и предадептов  $y_{mk}(t)$  определяется следующей системой стохастических дифференциальных уравнений [10]:

$$d\vec{Z} = \vec{A}(\vec{Z}, t)dt + \Sigma(\vec{Z})d\vec{W}; \vec{Z}(t_0) = 0, \quad (1)$$

понимаемой в смысле Ито.

В выражении (1) приняты следующие обозначения [10–12]: вектор переменных  $\vec{Z} \in \Omega$  при блочном представлении  $\vec{Z} = (Z_i)_d = (\bar{z}^{(1)}, \bar{z}^{(2)}, \dots, \bar{z}^{(M)})$  определяется численностью адептов и предадептов относительно  $m$ -й подгруппы  $\bar{z}^{(m)}$ ;  $\Omega = \varepsilon^{(1)} \times \varepsilon^{(2)} \times \dots \times \varepsilon^{(M)} \subset \mathbb{R}^d$  –  $d$ -мерный ( $d = 2MK$ ) выпуклый многогранник (алгоритмические особенности параметризации  $\Omega$  уточнены в [11, 12]);  $\varepsilon^{(m)} \subset \mathbb{R}^{2K}$  – симплекс с  $2K + 1$  вершинами;  $\vec{A} = (A_i)_d = (\bar{a}^{(1)}, \bar{a}^{(2)}, \dots, \bar{a}^{(M)})$  – вектор сноса;  $\Sigma = (\Sigma_{ll'})_{d \times d} = \text{diag}\{\sigma_1, \sigma_2, \dots, \sigma_M\}$  – матрица диффузии;  $\vec{W} = (W_i)_d = (\bar{w}^{(1)}, \bar{w}^{(2)}, \dots, \bar{w}^{(M)})$  –  $d$ -мерное броуновское движение; (правила определения советуемых элементов  $\bar{z}^{(m)}$ ,  $\bar{a}^{(m)}$ ,  $\sigma_m$  представлены в приложении).

Оценивание числа  $\hat{z}^{(m)}$  адептов  $\hat{x}_{mk}(t)$  и предадептов  $\hat{y}_{mk}(t)$  выполняется при усреднении [10, 12]:  $\hat{Z}(t) = \int_{\Omega} \vec{Z} p(\vec{Z}, t) d\vec{Z}$ , где  $p(\vec{Z}, t) = \phi(\vec{Z}, t) / \int_{\Omega} \phi(\vec{Z}, t) d\vec{Z}$  – функция плотности распределения вероятности. В отличие от [10, 12] для повышения точности оценки

$\vec{A}(\vec{Z}, t)$ ,  $\Sigma(\vec{Z})$  эволюцию ненормированной функции плотности распределения вероятности  $\phi(\vec{Z}, t)$  будем определять из решения уравнения Дункана–Мортенсена–Закаи [16]:

$$d\phi(\vec{Z}, t)/dt = \mathbf{L}[\phi(\vec{Z}, t)] + \vec{v}(\vec{Z}, t)\mathbf{R}'^{-1}(t)\vec{Z}(t)\phi(\vec{Z}, t), \quad (2)$$

при дополнении уравнения состояния (1) уравнением наблюдения:

$$d\vec{Z} = \vec{v}(\vec{Z}, t)dt + d\vec{W}; \quad \vec{Z}(t_0) = 0. \quad (3)$$

В выражениях (2), (3) приняты следующие обозначения:  $\mathbf{L}[\phi] = -\sum_{i=1}^d \frac{\partial}{\partial Z_i} (A_i \phi) + \frac{1}{2} \sum_{i=1}^d \sum_{l=1}^d \frac{\partial^2 (D_{il} \phi)}{\partial Z_i \partial Z_l}$  – диффузионный оператор Фоккера–Планка–Колмогорова [17];  $\mathbf{D} = \Sigma \Sigma^T = (D_{il})_{d \times d}$  и  $\vec{A} = (A_i)_d$  – тензор диффузии и вектор сноса соответственно;  $\mathbf{R}' = \mathbf{R} \mathbf{R}^T$  – матрица условной интенсивности;  $\text{cov}[d\vec{W}] = \mathbf{R}(t)dt$ ;  $\vec{W}$  – измеримые при всех  $t \in [t_0, t_1]$  некоррелированные винеровские процессы;  $\vec{v}(\vec{Z}, t)$  – вектор наблюдения, задаваемый по правилам аналогичным (1) при составлении вектора сноса  $\vec{A}(\vec{Z}, t)$  через функции  $f_{mk}^{(1)}$ ,  $f_{mk}^{(2)}$ , за исключением того, что вместо значений  $A_{mk}^0$ ,  $\Gamma_{mk}^0$  подставляются наблюдаемые интенсивности  $A_{mk} = \sum_{l=1}^L \chi_{ml} \alpha_{kl}$ ,  $\Gamma_{mk} = \sum_{l=1}^L \chi_{ml} \gamma_{kl}$  (правила определения элементов  $\vec{A}(\vec{Z}, t)$  через функции  $f_{mk}^{(1)}$ ,  $f_{mk}^{(2)}$  уточнены в приложении).

Процедуру оценки  $A_{mk}^0$ ,  $\Gamma_{mk}^0$  по наблюдаемым на интервале  $[t_0, t_1]$  интенсивностям  $\alpha_{kl}$ ,  $\gamma_{kl}$  предлагается составить на основе методов полиспектрального анализа сигналов [18]. Рассмотрим кратко ее суть при обозначение интенсивностей  $\alpha_{kl}(t)$ ,  $\gamma_{kl}(t)$  некоторым сигналом  $s(t)$ .

Оценка полиспектра  $I(\xi)$  для  $s(t)$  ( $\xi \in \mathbb{W}$  – мультииндекс, где  $\mathbb{W} = \{ \xi = (\xi_1, \dots, \xi_\rho) : \xi_\rho \in \mathbb{Z}_+, \sum_{\rho \in [1; \rho]} \xi_\rho < N', \xi_{\rho+1} \geq \xi_\rho \}$  при  $\rho \geq 2$  – порядок полиспектра) выполняется прямым методом [19] при разбиении временного интервала наблюдения  $[t_0, t_1]$  на  $V$  одинаковых по длительности  $\Delta t = (t_1 - t_0)/V$  подынтервалов  $[t'_v, t'_{v+1}]$  ( $v \in \{0, \overline{V-1}\}$ ) при  $t'_0 = t_0$ ,  $t'_V = t_1$ ) и дискретизации наблюдений сигнала  $s(t)$  для  $[t'_v, t'_{v+1}]$  на  $N'$  отсчетов  $s_n^{(v)}$  ( $n' \in \{0, \overline{N'-1}\}$ ). Затем, с применением алгоритма быстрого прямого дискретного преобразования Фурье, вычисляется спектральное представление  $\tilde{s}_n^{(v)}$  каждого  $s_n^{(v)}$  относительно  $v$ -го подынтервала и, с учетом правила [19]  $I^{(v)}(\xi) = \{ \tilde{s}_{\|\xi\|}^{(v)} \}^* \prod_{\rho=1}^{\rho} \tilde{s}_{\xi_\rho}^{(v)}$  для  $\|\xi\| = \sum_{\rho=1}^{\rho} \xi_\rho$ , при средневывборочном усреднении определяется полиспектр  $I(\xi) = V^{-1} \sum_{v=0}^{V-1} I^{(v)}(\xi)$ .

Окончательная оценка  $\hat{s}(t)$  сигнала  $s(t)$  осуществляется в отношении его дискретных значений  $s_n = s(n(t_1 - t_0)/(N - 1))$  ( $n \in \{0, \overline{N-1}\}$ ,  $N = VN'$ ) по правилу  $\hat{s}_n = \sum_{n'=0}^{N'-1} s_{n+n'} \mathcal{G}_{n'}$ , где коэффициенты  $\mathcal{G}_{n'}$  фильтра определяются из быстрого обратного дискретного преобразования Фурье восстановленных оценок Фурье-спектра  $\tilde{\mathcal{G}}_{n'} = \zeta_{n'} e^{i\varphi_{n'}}$  по оценкам фазового (формируется по  $\varphi_{n'}$ ) и амплитудного (составляется по  $\kappa_{n'} = \ln \zeta_{n'}$ ) полиспектральных уравнений [19], задаваемых из  $\varphi'_\xi = \arg(I(\xi))$  и  $\kappa'_\xi = \ln |I(\xi)|$  соответственно:

$$\bar{\varphi} = (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \bar{\varphi}'; \quad \bar{\kappa} = (\tilde{\mathbf{G}}^T \tilde{\mathbf{G}})^{-1} \tilde{\mathbf{G}}^T \bar{\kappa}',$$

где  $\bar{\varphi} = (\varphi_{n'})_{N'}$ ;  $\bar{\kappa} = (\kappa_{n'})_{N'}$ ;  $\bar{\varphi}' = (\varphi'_\xi)_{|\mathbb{W}|}$ ;  $\bar{\kappa}' = (\kappa'_\xi)_{|\mathbb{W}|}$ ;  $\mathbf{G} = (G_{\xi n'})_{|\mathbb{W}| \times N'}$ ;  $\tilde{\mathbf{G}} = (\tilde{G}_{\xi n'})_{|\mathbb{W}| \times N'}$ . Элементы  $G_{\xi n'}$ ,  $\tilde{G}_{\xi n'}$  советующих матриц  $\mathbf{G}$ ,  $\tilde{\mathbf{G}}$  при исходном равенстве нулю формируются из последовательно

реализуемых условий  $G_{\xi_{n'}} = \sum_{\rho \in [1; \varphi]; \xi_{\rho} = n'} 1$ ;  $\tilde{G}_{\xi_{n'}} = \sum_{\rho \in [1; \varphi]; \xi_{\rho} = n'} 1$  при

последующем уточнении значений в соответствии с правилами:

$$G_{\xi_{n'}} = (G_{\xi_{n'}} - 1, \text{ if } \sum_{\rho \in [1; \varphi]} \xi_{\rho} = n'); \tilde{G}_{\xi_{n'}} = (\tilde{G}_{\xi_{n'}} + 1, \text{ if } \sum_{\rho \in [1; \varphi]} \xi_{\rho} = n').$$

По указанной процедуре производится оценка интенсивностей

$$\alpha_{kl}^0, \gamma_{kl}^0 \text{ и вычисляются искомые параметры } A_{mk}^0 = \sum_{l=1}^L \chi_{ml} \alpha_{kl}^0,$$

$$\Gamma_{mk}^0 = \sum_{l=1}^L \chi_{ml} \gamma_{kl}^0 \text{ для составления вектора сноса } \bar{A}(\bar{Z}, t) \text{ в (1). Затем по}$$

оцененным интенсивностям  $\alpha_{kl}^0, \gamma_{kl}^0$  и их наблюдению  $\alpha_{kl}, \gamma_{kl}$  выполняется определение советующих параметров  $\varepsilon_{kl}^{\alpha}, \varepsilon_{kl}^{\gamma}$  шумов  $\tilde{\alpha}_{kl}, \tilde{\gamma}_{kl}$  средневыборочными правилами [20] для составления матрицы диффузии  $\Sigma(\bar{Z})$  в (1).

Уточненное решение получения полиспектральных оценок  $A_{mk}^0, \Gamma_{mk}^0$  в (1) составляет начальное приближение в общей задачи фильтрации (2) при дополнительно вводимом уравнении наблюдения (3) с параметрами  $A_{mk}, \Gamma_{mk}$ . Ошибки оценки  $A_{mk}^0, \Gamma_{mk}^0, \varepsilon_{kl}^{\alpha}, \varepsilon_{kl}^{\gamma}$ , точность определения которых в конечном счете влияет на погрешность итогового решения  $\hat{Z}(t) = \int_{\Omega} \bar{Z}p(\bar{Z}, t) d\bar{Z}$ ,

минимизируются численным исследованием задачи (2) при составлении следующей рекурсивной схемы фильтрации оценок информационного воздействия на электродат.

**3. Схема численной фильтрации оценок информационного воздействия.** Следуя результатам [11], для проекционного решения уравнения Дункана–Мортенсена–Закаи (2) зададим аппроксимацию ненормированной функции плотности распределения вероятности:

$$\tilde{\phi}(\bar{Z}, t) = \sum_{u=1}^U \sum_{j \in M_{\rho}^{(u)}} c_j^{(u)}(t) \psi_j^{(u)}(\bar{Z}), \quad (4)$$

где  $\psi_j^{(u)}$  – базисные функции, определяемые на  $\omega^{(u)} \in \Omega$  при

построении разбиения  $\Omega = \bigcup_{u=1}^U \omega^{(u)}$  набором из  $U$  симплексов:

$$\omega^{(u)} = \left\{ \sum_{l=1}^{d+1} \zeta_l^{(u)} P_l^{(u)} : \left( \sum_{l=1}^{d+1} \zeta_l^{(u)} = 1 \right) \wedge \left( \forall l = \overline{1, d+1}, \zeta_l^{(u)} \geq 0 \right) \right\} \subset \mathbb{R}^d,$$

( $u = \overline{1, U}$ ) с  $d+1$  вершинами  $P_1^{(u)}, P_2^{(u)}, \dots, P_{d+1}^{(u)}$  и барицентрическими координатами  $\zeta_1^{(u)}, \zeta_2^{(u)}, \dots, \zeta_{d+1}^{(u)}$  при  $\omega^{(u)} \cap \omega^{(u')} = \emptyset$  для  $u \neq u'$ ;  $u, u' \in \{ \overline{1, U} \}$  (алгоритмические решения по разбиению  $\Omega$  на  $\omega^{(u)}$  уточнены в [11, 12] при построении барицентрической триангуляции).

Для составления устойчивого численного решения при реализации условий нормировки  $p(\bar{Z}, t) \geq 0, \int_{\Omega} p(\bar{Z}, t) d\bar{Z} = 1$  в отношении аппроксимации (4) для  $p(\bar{Z}, t) = \phi(\bar{Z}, t) / \int_{\Omega} \phi(\bar{Z}, t) d\bar{Z}$ , и ввиду наличия ненулевых первых и вторых частных производных в уравнении Дункана–Мортенсена–Закаи (2), базисные функции  $\psi_j^{(u)}$  зададим произведением (рисунок 3, а):

$$\psi_j^{(u)} = \sqrt{\frac{2}{\pi}} (r+1) \prod_{l=1}^{d+1} \varphi_{j_l}. \tag{5}$$

Гауссовых базисных функций [21] (рисунок 3, б):

$$\varphi_{j_l} = e^{-\frac{[2j_l+1-2\zeta_l(r+1)]^2}{2(d+1)}}. \tag{6}$$

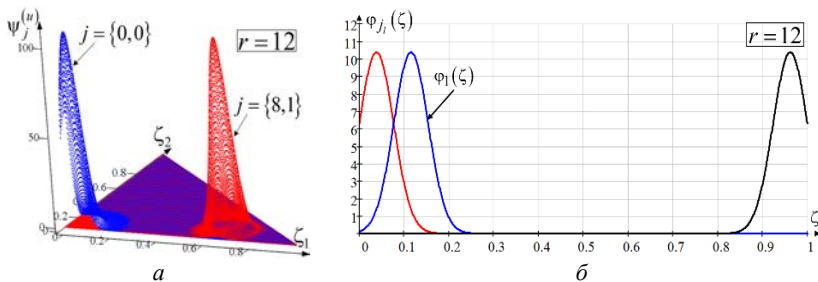


Рис. 3. Примеры представления базисных функций  $\psi_j^{(u)}$  (а) на симплексе для  $K = 2, M = 1$  и их образующих Гауссовых базисных функций  $\varphi_{j_l}$  (б)

При обозначении  $\langle \eta, \phi \rangle = \int_{\Omega} \eta(\bar{Z}) \phi(\bar{Z}) d\bar{Z}$  скалярного

произведения некоторых интегрируемых на  $\Omega$  функций  $\eta$  и  $\phi$  для заданной аппроксимации (4) сведем уравнение Дункана–Мортенсена–Закаи (2) к системе обыкновенных дифференциальных уравнений вида:

$$d\bar{C}(t)/dt = \mathbf{S}^{-1} [\mathbf{Q}^{(1)}(t) + \mathbf{Q}^{(2)}(t)] \bar{C}(t); \quad \bar{C}(t_0) = \mathbf{S}^{-1} \bar{W}, \quad (7)$$

где  $\bar{W} = \left( \left\langle p_0, \psi_j^{(u)} \right\rangle_{\bar{U}} \right)$ ;  $p_0 \equiv p(\bar{Z}, t_0)$ ;  $\bar{C} = \left( c_j^{(u)} \right)_{\bar{U}}$  – вектор искомых коэффициентов разложения ( $c_j^{(u)} \geq 0$ ) в аппроксимации (4), зависящих от  $t \in [t_0, t_1]$ ;  $\mathbf{S} = \left( \left\langle \psi_j^{(u)}, \psi_{j'}^{(u')} \right\rangle_{\bar{U} \times \bar{U}} \right)$ ;  $\mathbf{Q}^{(1)} = \left( \left\langle \psi_j^{(u)}, \mathbf{L}[\psi_{j'}^{(u')}] \right\rangle_{\bar{U} \times \bar{U}} \right)$ ;  $\mathbf{Q}^{(2)} = \left( \left\langle \psi_j^{(u)}, \bar{U} \mathbf{R}^{-1} \tilde{\mathbf{Z}} \psi_{j'}^{(u')} \right\rangle_{\bar{U} \times \bar{U}} \right)$ ;  $\bar{U} = U | \mathbb{M}_r^d$ ;  $\mathbb{M}_r^d$  – множество мультииндексов  $j, j' \in \mathbb{M}_r^d$  (вычислительно эффективный алгоритм формирования  $\mathbb{M}_r^d$  приведен в [22]):

$$\mathbb{M}_r^d = \left\{ j = (j_1, \dots, j_l, \dots, j_{d+1}) : j_l \in \mathbb{Z}_+, \sum_{l \in [1; d+1]} j_l = r \right\}, \quad (8)$$

где  $r \in \mathbb{N}$  – порядок аппроксимации на симплексе  $\omega^{(u)}$ ;  $\mathbb{Z}_+ = \mathbb{N} \cup \{0\}$ .

Решение системы обыкновенных дифференциальных уравнений (7) определим в виде:

$$\bar{C}(t) = \exp \left[ \mathbf{S}^{-1} \int_{t_0}^t \{ \mathbf{Q}^{(1)}(\tau) + \mathbf{Q}^{(2)}(\tau) \} d\tau \right] \mathbf{S}^{-1} \bar{W}, \quad (9)$$

где  $\exp[\cdot]$  обозначает матричную экспоненту.

Принимая во внимание доказательство [10] существования и единственности в смысле стохастической эквивалентности непрерывного с вероятностью 1 решения задачи (1), оценку сходимости численного решения уравнения Фоккера–Планка–Колмогорова в [12] при формировании аппроксимации искомой

функции плотности вероятностей гауссовыми базисными функциями, определим оценку сходимости метода Галёркина для аппроксимации (4) в решении уравнения Дункана–Мортенсена–Закаи (2).

**Теорема 1.** Пусть  $\tilde{\phi}(\bar{Z}, t) = \sum_{u=1}^U \sum_{j \in \mathbb{M}_r^d} c_j^{(u)}(t) \psi_j^{(u)}(\bar{Z})$ , тогда метод

Галёркина для уравнения (2) сходится, и справедлива оценка:

$$\|\phi - \tilde{\phi}\|_{\Omega} \leq O_3 \|\phi\|_{\Omega} \sqrt{h/(r+1)}, \quad (10)$$

где  $O_3$  – независимая от  $r$  положительная постоянная,  $h$  – максимальный линейный размер симплексов  $\omega^{(u)} \subset \Omega$ .

**Доказательство.** В [11] показано, что при представлении оператора Фоккера–Планка–Колмогорова в виде  $\mathbf{L} = \mathbf{G} + \mathbf{T}$  при

обозначении  $\mathbf{G}[\phi] = -\sum_{l=1}^d \frac{\partial}{\partial Z_l} (A_l \phi)$ ,  $\mathbf{T}[\phi] = \frac{1}{2} \sum_{l=1}^d \sum_{l'=1}^d \frac{\partial^2 (D_{ll'} \phi)}{\partial Z_l \partial Z_{l'}}$  операторы

$\mathbf{L}$  и  $\mathbf{G}$  имеют соответствующие ограниченные обратные операторы  $\mathbf{L}^{-1}$  и  $\mathbf{G}^{-1}$ , а оператор  $\mathbf{T}$  компактен. Тогда, следуя [11] и введя в рассмотрение оператор Дункана–Мортенсена–Закаи  $\mathbf{F} = \mathbf{L} + \mathbf{Z}$ , где  $\mathbf{Z}[\phi] = \bar{v}(\bar{Z}, t) \mathbf{R}^{-1}(t) \tilde{Z}(t) \phi(\bar{Z}, t)$ , с учетом теоремы 1 из [10] и определения уравнения наблюдения (3), получим, что операторы  $\mathbf{F}$  и  $\mathbf{L}$  имеют соответствующие ограниченные обратные операторы  $\mathbf{F}^{-1}$  и  $\mathbf{L}^{-1}$ , а оператор  $\mathbf{Z}$  компактен. Отсюда, следуя теореме Л. В. Канторовича [23], определим справедливость оценки:

$$\|\phi - \tilde{\phi}\|_{\Omega} \leq O_1 E_{\tilde{\Phi}}(\phi), \quad (11)$$

где  $O_1$  – независимая от  $r$  положительная постоянная;

$E_{\tilde{\Phi}}(\phi) = \inf_{\tilde{\phi} \in \tilde{\Phi}} \|\phi - \tilde{\phi}\|_{\Omega}$  – наилучшее приближение функции  $\phi$ ;

$\tilde{\Phi}$  – подпространство пространства  $\phi \in \Phi$  при  $\Phi: \Phi[\phi] = \tilde{\phi}$ ;

$\Phi$  – проекционный оператор.

Из неравенства (11) следует, что для определения сходимости численного решения уравнения Дункана–Мортенсена–

Закаи (2) методом Галёркина необходимо установить оценку  $\|f - \tilde{f}\|_{\Omega}$  наилучшего среднеквадратичного приближения  $\tilde{f}(\vec{Z}) = \sum_{u=1}^U \sum_{j \in \mathbb{M}_r^d} c_j^{(u)} \psi_j^{(u)}(\vec{Z})$  с коэффициентами разложения  $c_j^{(u)}$  непрерывно дифференцируемой на  $\Omega$  функции  $f(\vec{Z})$ . Тогда, принимая во внимание определение  $\Omega$ , метод замены переменных при интегрировании по симплексу [22] и лемму 2 из [11], получим:

$$\|f - \tilde{f}\|_{\Omega} \leq O_2 \|f\|_{\Omega} \sqrt{h/(r+1)}, \quad (12)$$

где  $O_2$  – независящая от  $r$  положительная постоянная.

Из (11), (12) при  $O_3 = O_1 O_2$  следует сходимость метода Галёркина для численного решения уравнения Дункана–Мортенсена–Закаи (2) и справедливость оценки (10), что и требовалось доказать.

В отношении заданного правила (9) вычисления эволюции вектора  $\vec{C}$  искомых коэффициентов разложения в аппроксимации (4) для  $\hat{Z}(t) = \int_{\Omega} \vec{Z} p(\vec{Z}, t) d\vec{Z} \approx \sum_{u=1}^U \sum_{j \in \mathbb{M}_r^d} c_j^{(u)}(t) \Psi_j^{(u)}$  ( $\Psi_j^{(u)} = \int_{\omega^{(u)}} \vec{Z} \psi_j^{(u)}(\vec{Z}) d\vec{Z}$ )

сформируем общую последовательность действий численной фильтрации оценок информационного воздействия при дискретизации интервала анализа  $[t_0, t_1]$  на  $N$  подынтервалов  $[\tau_n, \tau_{n+1}]$  ( $n \in \{0, \overline{N-1}\}$ ) при учете: 1) алгоритмических особенностей по построению  $\Omega$  и реализации процедур численного интегрирования по  $\Omega$ , утоненных в работах [11, 12]; 2) процедуры определения  $A_{mk}^0$ ,  $\Gamma_{mk}^0$ ,  $\varepsilon_{kl}^{\alpha}$ ,  $\varepsilon_{kl}^{\gamma}$  методом полиспектрального анализа; 3) алгоритма фильтрации, разработанного в [24] при формировании рекуррентно-итерационной модели адаптации.

*Шаг 1.* С применением методов численного интегрирования [25] на подынтервале  $[\tau_n, \tau_{n+1}]$  вычислить:

$$\vec{C}(\tau_{n+1}) = \exp \left[ \mathbf{S}^{-1} \int_{\tau_n}^{\tau_{n+1}} \{ \mathbf{Q}^{(1)}(\tau) + \mathbf{Q}^{(2)}(\tau) \} d\tau \right] \vec{C}(\tau_n), \quad (13)$$



где матрица  $\mathbf{Q}^{(2)}(\tau)$  определяется для исходно заданной ковариационной матрицы  $\mathbf{R}$  в момент времени  $\tau_n$  и рассчитываемого с применением численных методов [26] решения обыкновенных дифференциальных уравнений значения  $\tilde{\mathbf{Z}}(\tau)$  из уравнения наблюдения (3) при известном  $\hat{\mathbf{Z}}(\tau_n)$ .

*Шаг 2.* Выполнить нормировку  $\bar{\mathbf{C}}(\tau_{n+1})$  для  $j \in \mathbb{M}_r^d$ ,  $u \in \{\overline{1, U}\}$ :

$$c_j^{(u)} = \begin{cases} c_j^{(u)}, & c_j^{(u)} \geq 0; \\ 0, & c_j^{(u)} < 0; \end{cases} \quad \bar{\mathbf{C}} = \bar{\mathbf{C}} / \sum_{u=1}^U \sum_{j \in \mathbb{M}} c_j^{(u)} \tilde{\Psi}_j^{(u)}, \quad (14)$$

где  $\tilde{\Psi}_j^{(u)} = \int_{\omega^{(u)}} \psi_j^{(u)}(\bar{\mathbf{Z}}) d\bar{\mathbf{Z}}$ .

*Шаг 3.* Вычислить  $\tilde{\mathbf{Z}}(\tau_{n+1})$  из уравнения наблюдения (3) с применением численных методов [26] решения обыкновенных дифференциальных уравнений при  $\hat{\mathbf{Z}}(\tau_{n+1}) = \sum_{u=1}^U \sum_{j \in \mathbb{M}_r^d} c_j^{(u)}(\tau_{n+1}) \Psi_j^{(u)}$ .

*Шаг 4.* Для определенных  $\tilde{\mathbf{Z}}(\tau_{n+1})$  и  $\hat{\mathbf{Z}}(\tau_{n+1})$  средневыборочными правилами [27] скорректировать  $\mathbf{R}$  для момента времени  $\tau_{n+1}$ .

*Шаг 5.* Проверить условие  $n+1 < N-1$  и в случае его выполнения положить  $n = n+1$  и перейти к шагу 1. В противном случае завершить работу алгоритма и вывести результат фильтрации –  $\hat{\mathbf{Z}}_n \equiv \hat{\mathbf{Z}}(\tau_n)$  для  $n \in \{0, N-1\}$ .

Псевдокод алгоритма FILTERVOTERS фильтрации оценок информационного воздействия с учетом указанной общей последовательности действий приведен в виде листинга 1.

В сформированном алгоритме (листинг 1) предлагается производить вычисление интеграла (13) с применением кубатурных формул [25]. При этом для снижения вычислительных затрат (число вычислений матриц  $\mathbf{Q}^{(1)}$  и  $\mathbf{Q}^{(2)}$  при реализации процедур численного интегрирования на симплексах [26]) дискретизация интервала анализа

$[t_0, t_1]$  на  $N$  подынтервалов  $[\tau_n, \tau_{n+1}]$  выполняется таким образом, чтобы последний узел точки численного интегрирования для подынтервала  $n$  соответствовал первому узлу точки численного интегрирования для подынтервала  $n+1$  (в сформированном алгоритме используются две узловые точки при обеспечении третьего порядка точности численного интегрирования на подынтервале). Вспомогательные функции NORMVECTOR, ODESOLVE и SOLVECORREL реализуют условия нормировки (14), численное решение [26] системы обыкновенных дифференциальных уравнений наблюдения (3) и средневыворочное оценивание ковариационной матрицы  $\mathbf{R}$  [27] соответственно.

```

function FILTERVOTERS( $t_0, t_1, N, \bar{W}, \mathbf{S}, \mathbf{R}, \mathbf{Q}^{(1)}, \mathbf{Q}^{(2)}$ ){
     $k = 3^{-1/2}$ ;  $h = (t_1 - t_0) / [k(N - 1)]$ ;  $\mathbf{S}' = \mathbf{S}^{-1}$ ;
     $\bar{C} = \text{NORMVECTOR}(\mathbf{S}'\bar{W})$ ;  $\hat{Z}_0 = \sum_{u=1}^U \sum_{j \in \mathbb{M}_j^d} c_j^{(u)} \Psi_j^{(u)}$ ;
     $t = 0$ ;  $\tau = (0,5 - 0,5k)h$ ;  $\tilde{Z} = \text{ODESOLVE}(t, \tau, \hat{Z}_0)$ ;
     $\mathbf{Q}_1^{(1)} = \mathbf{Q}^{(1)}(\tau)$ ;  $\mathbf{Q}_1^{(2)} = \mathbf{Q}^{(2)}(\tau, \tilde{Z}, \mathbf{R})$ ;
     $\tau = (0,5 + 0,5k)h$ ;  $\tilde{Z} = \text{ODESOLVE}(t, \tau, \hat{Z}_0)$ ;
    for  $n = 1$  to  $N - 1$  do {
         $\mathbf{Q}_2^{(1)} = \mathbf{Q}^{(1)}(\tau)$ ;  $\mathbf{Q}_2^{(2)} = \mathbf{Q}^{(2)}(\tau, \tilde{Z}, \mathbf{R})$ ;
         $\mathbf{Q} = 0,5h(\mathbf{Q}_1^{(1)} + \mathbf{Q}_2^{(1)} + \mathbf{Q}_1^{(2)} + \mathbf{Q}_2^{(2)})$ ;
         $\mathbf{G} = \exp(\mathbf{S}'\mathbf{Q})$ ;  $\bar{C} = \text{NORMVECTOR}(\mathbf{G}\bar{C})$ ;
         $\hat{Z}_n = \sum_{u=1}^U \sum_{j \in \mathbb{M}_j^d} c_j^{(u)} \Psi_j^{(u)}$ ;  $\mathbf{R} = \text{SOLVECORREL}(\tilde{Z}, \hat{Z}_n)$ ;
         $t = \tau$ ;  $\tau = \tau + kh$ ;  $\tilde{Z} = \text{ODESOLVE}(t, \tau, \hat{Z}_0)$ ;
    }
    return  $\hat{Z}$ ;
}

```

Листинг 1. Псевдокод алгоритма FILTERVOTERS фильтрации оценок информационного воздействия на электротат

**4. Результаты вычислительного эксперимента.** Для наглядной демонстрации предпочтительности сформированной схемы численной фильтрации оценок информационного воздействия на

электорат при программной реализации полученного решения (программа для ЭВМ «Программа реализации алгоритма фильтрации оценок информационного воздействия на электорат при проведении выборных кампаний» [29]) проведена серия вычислительных экспериментов для различных значений чисел кандидатов  $K$  и подгрупп  $M$  разделения социума. По результатам чего получена сравнительная зависимость вероятности  $p_n$  принятия ошибочного решения о победе кандидата-лидера в момент времени  $t_1$  от величины среднеквадратической ошибки  $\varepsilon_{kl}^\alpha = 1,5 \cdot 10^{-3} \varepsilon$ ,  $\varepsilon_{kl}^\gamma = 5 \cdot 10^{-4} \varepsilon$  шумов  $\tilde{\alpha}_{kl}(t)$ ,  $\tilde{\gamma}_{kl}(t)$  соответственно. Основу алгоритма генерации  $\tilde{\alpha}_{kl}(t)$ ,  $\tilde{\gamma}_{kl}(t)$  составляет следующая последовательность действий, которую удобно представить при обозначении истинных значений интенсивностей  $\alpha_{kl}^0(t)$ ,  $\gamma_{kl}^0(t)$  и шумов  $\tilde{\alpha}_{kl}(t)$ ,  $\tilde{\gamma}_{kl}(t)$  некоторыми функциями  $\eta^0(t)$  и  $\tilde{\eta}(t)$  соответственно (среднеквадратическую ошибку, определяющую уровень шума  $\tilde{\eta}(t)$  обозначим  $\tilde{\varepsilon}$ ). Изначально на временном интервале  $[t_0, t_1]$ , разбиваемом на  $N$  одинаковых по длительности  $\Delta t = (t_1 - t_0)/N$  подынтервалов  $[t_n, t_{n+1}]$  ( $n \in \overline{0, N-1}$ ), выполняется генерация винеровского процесса  $w_n = w(t_n)$  по правилу  $w_n = \begin{cases} 0, & \text{if } n = 0; \\ w_{n-1} + \aleph(0, \tilde{\varepsilon}), & n \neq 0, \end{cases}$  где  $\aleph(0, \text{rnd}[\tilde{\varepsilon}])$  – функция генерации случайных чисел, подчиняющихся нормальному закону распределения со средним 0 и среднеквадратическим отклонением  $\text{rnd}[\tilde{\varepsilon}]$ ;  $\text{rnd}[\tilde{\varepsilon}]$  – функция генерации случайных чисел по равномерному закону распределения на интервале  $[0, \tilde{\varepsilon}]$ . Затем, применяя правила аппроксимации кубическими сплайнами [30], из дискретных значений  $w_n$  формируется гладкая функция  $w(t)$ . Далее, используя метод интерполяции Лагранжа для численного дифференцирования, вычисляются исходные дискретные значения шума  $\hat{\eta}_n = \hat{\eta}(t_n)$  по  $w(t)$  выражением [31]:

$$\hat{\eta}_n = \sum_{h=0}^{2H} w(\Delta\tau_h) \sum_{h_1=0}^{2H} \left\{ \left[ \frac{1}{\Delta\tau_h - \Delta\tau_{h_1}} \prod_{h_2=0}^{2H} \left\{ \frac{(t_n - \Delta\tau_{h_2})}{(\Delta\tau_h - \Delta\tau_{h_2})} \right\} \right], h_1 \neq h; \right. \\ \left. 0, h_1 = h, \right.$$

где  $\Delta\tau_h = t_n + \Delta t \rho_h$ ;  $2H$  – порядок интерполяции;  $\rho_h = \cos[\pi(1+2h)/(4H+2)]$  – корни многочлена Чебышёва первого рода порядка  $2H$  (для проводимых серий вычислительных экспериментов  $H$  выбрано равным 4) [32].

Затем, применяя правила аппроксимации кубическими сплайнами [30], из дискретных значений  $\hat{\eta}_n$  формируется гладкая функция  $\hat{\eta}(t)$ . Для составленной функции  $\hat{\eta}(t)$  шум определяется по правилу  $\tilde{\eta}(t) = \theta(\eta^0(t) + \hat{\eta}(t)) - \eta^0(t)$ , где  $\theta(x) = \begin{cases} x, & x \geq 0; \\ 0, & x < 0. \end{cases}$  Пример гистограммы распределения вероятностей  $p_n$  и спектральная плотность мощности  $\Gamma_n$  шума  $\tilde{\eta}(t)$ , сгенерированного по указанной процедуре для  $\tilde{\varepsilon} = 2$  и  $N = 10^4$ , приведены на рисунке 4.

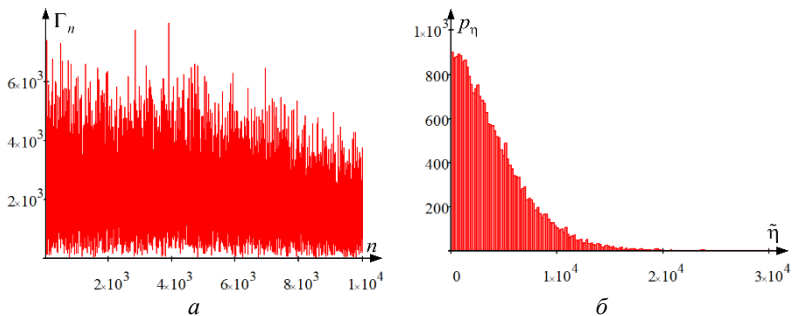


Рис. 4. Примеры гистограммы распределения вероятностей  $p_n$  (а) и спектральной плотности мощности  $\Gamma_0(f)$  (б) шума  $\tilde{\eta}(t)$

Число  $L$  внешних источников информации выбрано равным 3. Коэффициенты восприятия  $\chi_{ml}$  заданы следующими:  $\chi_{11} = 0,263$ ;  $\chi_{12} = 1,8$ ;  $\chi_{13} = 1,5$ ;  $\chi_{21} = 0,395$ ;  $\chi_{22} = 1,7$ ;  $\chi_{23} = 1,4$ ;  $\chi_{31} = 0,342$ ;  $\chi_{32} = 0,7$ ;  $\chi_{33} = 0,9$ . Интенсивности межличностной коммуникации  $\beta_{nm}$  выбраны:  $\beta_{11} = 6 \cdot 10^{-4}$ ;  $\beta_{12} = 3,8 \cdot 10^{-4}$ ;  $\beta_{13} = 4,4 \cdot 10^{-4}$ ;

$\beta_{21} = 2,6 \cdot 10^{-4}$ ;  $\beta_{22} = 6 \cdot 10^{-4}$ ;  $\beta_{23} = 3,6 \cdot 10^{-4}$ ;  $\beta_{31} = 2,2 \cdot 10^{-4}$ ;  
 $\beta_{32} = 3 \cdot 10^{-4}$ ;  $\beta_{33} = 4 \cdot 10^{-5}$  (значения для расчетных экспериментов в отношении  $\chi_{ml}$  и  $\beta_{nm}$  выбирались для соответствующего  $M \in \{1;3\}$ )

С целью обеспечения общности выводов по результатам оценки при относительно большом числе вычислительных экспериментов истинные значения интенсивностей  $\alpha_{kl}^0(t)$ ,  $\gamma_{kl}^0(t)$  для отдельного вычисления выбирались различными и формировались по следующей процедуре. Для заданных дисперсионных  $\sigma_{kl}^\alpha = 0,02$ ,  $\sigma_{kl}^\gamma = 5,3 \cdot 10^{-3}$  и начальных  $\mu_{kl}^\alpha$  ( $\mu_{11}^\alpha = 0,011$ ;  $\mu_{12}^\alpha = 0,029$ ;  $\mu_{13}^\alpha = 0,02$ ;  $\mu_{14}^\alpha = 0,021$ ;  $\mu_{21}^\alpha = 0,027$ ;  $\mu_{22}^\alpha = 0,017$ ;  $\mu_{23}^\alpha = 0,012$ ;  $\mu_{24}^\alpha = 0,02$ ;  $\mu_{31}^\alpha = 0,017$ ;  $\mu_{32}^\alpha = 0,018$ ;  $\mu_{33}^\alpha = 4,5 \cdot 10^{-3}$ ;  $\mu_{34}^\alpha = 7,5 \cdot 10^{-3}$ ),  $\mu_{kl}^\gamma$  ( $\mu_{11}^\gamma = 0,013$ ;  $\mu_{12}^\gamma = 9 \cdot 10^{-3}$ ;  $\mu_{13}^\gamma = 3,5 \cdot 10^{-3}$ ;  $\mu_{14}^\gamma = 5 \cdot 10^{-3}$ ;  $\mu_{21}^\gamma = 0,011$ ;  $\mu_{22}^\gamma = 4 \cdot 10^{-3}$ ;  $\mu_{23}^\gamma = 6 \cdot 10^{-3}$ ;  $\mu_{24}^\gamma = 7 \cdot 10^{-3}$ ;  $\mu_{31}^\gamma = 5 \cdot 10^{-3}$ ;  $\mu_{32}^\gamma = 10^{-3}$ ;  $\mu_{33}^\gamma = 4 \cdot 10^{-3}$ ;  $\mu_{34}^\gamma = 3,3 \cdot 10^{-3}$ ) значениях для подынтервалов  $[t_n, t_{n+1}]$  ( $n \in \overline{1, N-1}$ ) произведена генерация  $w_{kl}^\alpha(t_n)$ ,  $w_{kl}^\gamma(t_n)$  по выражению

$$w_{kl}^{\alpha,\gamma}(t_n) = \begin{cases} w_{kl}^{\alpha,\gamma}(t_{n-1}) + \sigma_{kl}^{\alpha,\gamma}, & \text{rnd}[1] > 0,5; \\ w_{kl}^{\alpha,\gamma}(t_{n-1}) - \sigma_{kl}^{\alpha,\gamma}, & \text{rnd}[1] \leq 0,5 \end{cases} \quad \text{при } w_{kl}^{\alpha,\gamma}(t_0) = \mu_{kl}^{\alpha,\gamma}. \text{ Затем,}$$

применяя правила аппроксимации кубическими сплайнами [30], из модулей советующих дискретных значений  $|w_{kl}^\alpha(t_n)|$  и  $|w_{kl}^\gamma(t_n)|$  формируются функции интенсивностей  $\alpha_{kl}^0(t)$  и  $\gamma_{kl}^0(t)$  соответственно.

Сравнение результативности решения задачи оценки информационного воздействия на электорат проводилось в отношении: 1 – непосредственного численного решения системы обыкновенных дифференциальных уравнений состояния (1) – прямое решение; 2 – решения системы обыкновенных дифференциальных уравнений  $d\vec{Z} = \vec{A}(\vec{Z}, t)dt$  для полученных оценок интенсивностей  $A_{mk}^0$ ,  $\Gamma_{mk}^0$  методами полиспектрального анализа; 3 – предложенной численной схеме нелинейной фильтрации.

Значение вероятности  $p_{п}$  оценивалось апостериорным исследованием по правилу  $p_{п} = N_{\text{ош}}/N_{\text{мл}}$ , где  $N_{\text{ош}}$  – число ошибочно принятых решений о победе кандидата-лидера;  $N_{\text{мл}}$  – общее число

проводимых экспериментов при фиксированном  $\varepsilon$ . Моделирование выполнено для следующих исходных данных: общее число проводимых экспериментов  $N_{\text{мд}} = 10^7$ ; порядок полиспектра  $\varphi = 3$  для оценки  $\alpha_{kl}^0, \gamma_{kl}^0$  наблюдаемых интенсивностей агитации  $\alpha_{kl}, \gamma_{kl}$ ; порядок аппроксимации  $r = 19$  в (4) при формировании множества мультииндексов (8); численность социума  $N_0 = 300$  при равномерном определении численности подгрупп  $N_m = N_0/M$  ( $m \in \{1, \overline{M}\}$ ) индивидов.

В таблицах 1–3 приведены советующие значения сравнительной зависимости вероятности ошибки  $p_n$  (выражена в %) от параметра  $\varepsilon$  среднеквадратической ошибки для трех указанных вариантов оценки.

Таблица 1. Сравнительная зависимость  $p_n$  в % от  $\varepsilon$  при  $K = 2$  и различных  $M$

$\varepsilon$	$M = 1$			$M = 2$			$M = 3$		
	1	2	3	1	2	3	1	2	3
1	0,11	1,12	2,44	0,84	1,33	2,76	0,95	2,40	3,02
5	1,86	3,19	3,52	9,62	8,75	6,40	9,13	9,11	7,12
10	4,26	5,26	4,88	10,94	10,69	8,40	14,20	13,62	9,28
15	6,93	6,88	5,76	12,25	11,61	8,85	14,66	14,26	9,92
20	8,87	8,68	6,42	13,18	12,53	9,22	14,91	14,44	10,18
25	9,88	9,32	6,81	13,98	13,44	9,67	15,21	14,50	10,40
30	11,05	10,45	7,40	14,72	14,00	9,95	15,39	14,79	10,52
35	11,72	11,26	7,72	15,09	14,35	10,19	15,67	14,90	10,60
40	12,34	11,78	8,12	15,41	14,60	10,46	16,05	15,11	10,68
45	12,81	12,33	8,52	15,63	14,80	10,59	16,39	15,25	10,76
50	13,50	12,81	8,89	15,83	14,92	10,73	16,66	15,49	10,87
55	13,81	13,15	9,13	16,09	15,13	10,80	16,97	15,76	10,97
60	14,27	13,58	9,45	16,37	15,28	10,84	17,23	16,12	11,05
65	14,68	13,98	9,61	16,55	15,49	10,93	17,33	16,28	11,20
70	15,11	14,23	9,77	16,79	15,78	10,98	17,56	16,35	11,32
75	15,21	14,47	9,93	16,98	15,92	11,04	17,98	16,52	11,38
80	15,50	14,74	10,09	17,15	16,15	11,12	18,09	16,67	11,50

Таблица 2. Сравнительная зависимость  $p_{\Pi}$  в % от  $\varepsilon$  при  $K = 3$  и различных  $M$ 

$\varepsilon$	$M = 1$			$M = 2$			$M = 3$		
	1	2	3	1	2	3	1	2	3
0,5	0,50	0,82	1,44	0,89	1,27	2,37	0,96	2,34	3,32
1	2,33	2,31	2,42	2,94	3,13	3,41	3,23	3,50	4,53
1,5	4,78	4,84	4,72	5,04	5,29	5,54	6,68	7,01	6,17
2,5	10,83	9,92	9,02	11,21	10,79	9,10	12,27	12,10	9,12
3,5	13,49	12,81	10,41	15,45	14,47	10,49	16,04	14,98	10,79
5	15,48	14,92	11,25	18,08	17,22	12,03	19,21	18,05	12,10
7,5	16,69	16,12	12,07	20,33	19,07	13,55	21,30	19,95	14,11
10	17,32	16,51	12,48	22,14	20,74	14,79	22,89	21,56	15,20
15	18,51	17,56	12,97	23,43	22,09	15,89	24,65	23,10	15,91
20	19,59	18,54	13,54	24,00	22,58	16,27	25,23	23,70	16,46
25	20,51	19,08	13,95	24,56	23,14	16,55	25,68	24,24	16,87
30	20,92	19,68	14,61	24,94	23,54	16,83	25,90	24,70	17,17
35	21,54	20,39	15,10	25,31	23,76	17,07	26,51	25,09	17,48
40	22,15	20,91	15,51	25,50	23,97	17,30	26,80	25,27	18,01

Таблица 3. Сравнительная зависимость  $p_{\Pi}$  в % от  $\varepsilon$  при  $K = 4$  и различных  $M$ 

$\varepsilon$	$M = 1$			$M = 2$			$M = 3$		
	1	2	3	1	2	3	1	2	3
0,5	0,76	1,18	2,14	1,10	1,49	3,11	1,18	3,06	3,77
1	4,37	4,38	4,56	4,56	5,05	5,69	5,31	6,00	6,79
1,5	8,37	8,18	7,95	8,94	8,83	9,07	11,49	11,60	10,20
2,5	14,87	13,51	12,00	16,38	15,64	13,85	17,39	17,01	13,91
3,5	19,32	18,27	14,47	22,65	21,11	16,95	23,24	21,61	17,01
5	23,64	22,75	16,91	26,74	25,44	19,36	28,88	27,11	20,14
7,5	27,68	26,65	19,08	31,80	29,73	21,82	34,32	32,03	23,23
10	29,53	28,12	20,43	33,47	31,01	22,77	36,82	34,07	24,88
15	30,31	28,25	20,83	33,49	31,36	22,78	37,53	34,60	25,17
20	31,20	29,11	21,20	34,14	31,77	23,22	37,68	35,14	25,27
25	31,69	29,11	21,30	34,32	31,94	23,24	37,79	35,27	25,21
30	31,84	29,79	21,52	34,42	32,22	23,34	37,93	35,40	25,36
35	32,32	30,22	22,06	34,76	32,22	23,67	38,59	35,60	25,83
40	32,51	30,61	22,26	34,76	32,60	23,77	39,02	35,78	26,11

Из полученных результатов серий вычислительных экспериментов (таблицы 1–3) следует, что: 1) предложенная схема фильтрации позволяет в среднем повысить точность оценки на 5,61 % в сравнении с решением систем обыкновенных дифференциальных

уравнений по сглаженным данным наблюдения методами полиспектрального анализа и на 6,94% в сравнении с прямым решением; 2) если значения интенсивностей  $\alpha_{kl}(t)$  и  $\gamma_{kl}(t)$  наблюдаются с низким уровнем шума предпочтительным по вычислительным затратам и точности (для малых  $\varepsilon$  ошибку оценки по предложенному алгоритму составляет вычислительная погрешность решения (9)) является непосредственное численное решение системы обыкновенных дифференциальных уравнений состояния (1); 3) если значения интенсивностей  $\alpha_{kl}(t)$  и  $\gamma_{kl}(t)$  наблюдаются с высоким уровнем нестационарного негауссовского шума [33] предпочтительность применения предложенного алгоритма (листинг 1) возрастает на 7,18 % в сравнении с решением системы обыкновенных дифференциальных уравнений состояния по сглаженным данным наблюдения методами полиспектрального анализа и на 8,92 % в сравнении с прямым решением; 4) с увеличением  $M$  при неизменной численности социума относительный выигрыш использования сформированной численной схемы в отношении минимизации вероятности ошибки  $p_n$  растет; 5) с увеличением  $K$  при неизменной численности социума относительный выигрыш использования сформированной численной схемы в отношении минимизации вероятности ошибки  $p_n$  при неизменном порядке аппроксимации  $r$  снижается.

**5. Заключение.** В целом, составленная численная схема фильтрации оценок информационного воздействия средств масс-медиа на электорат в условиях наблюдения интенсивностей агитации  $\alpha_{kl}(t)$ ,  $\gamma_{kl}(t)$  при наличии нестационарных негауссовских шумов  $\tilde{\alpha}_{kl}(t)$ ,  $\tilde{\gamma}_{kl}(t)$  в сравнении с типовыми решениями (таблицы 1–3) позволяет на заданном интервале наблюдения  $[t_0, t_1]$  в 1,28 раз точнее оценивать число индивидов в обществе, отдающих предпочтение определенному политическому субъекту (мнению). Основную отличительную новизну предложенной численной схемы фильтрации информационного воздействия составляет сведение исходного стохастического дифференциального уравнения (1) к численному решению уравнения Дункана–Мортенсена–Закаи (2) при введении дополнительного уравнения наблюдения (3), которое получается из уравнения состояния (1) при оценке его стохастических компонент (наблюдаемые интенсивности агитации  $\alpha_{kl}(t)$ ,  $\gamma_{kl}(t)$ ) методами полиспектрального анализа. Следуя результатам [11, 12] по исследованию сходимости



аппроксимации (4) функции плотности вероятностей базисными функциями (5) относительно оператора Фоккера–Планка–Колмогорова, решение уравнения Дункана–Мортенсена–Закаи (2) выполнено в проекционной постановке метода Галёркина при сведении к системе линейных дифференциальных уравнений (7) и получении ее решения в схеме рекурсивного оценивания (13) при дискретизации интервала анализа  $[t_0, t_1]$  на  $N$  подынтервалов  $[\tau_n, \tau_{n+1}]$  и использования метода матричной экспоненты (9). Конкретизация способа алгоритмической реализации (листинг 1) составленной численной схемы фильтрации оценок информационного воздействия средств масс-медиа на электорат составляет практическую значимость решения при уточнении вычислительных особенностей, определяющих правила получения оценок наблюдаемых интенсивностей агитации  $\alpha_{kl}(t)$ ,  $\gamma_{kl}(t)$  при составлении вектора сноса  $\bar{A}(\bar{Z}, t)$  и матрицы диффузии  $\Sigma(\bar{Z})$ , в развитие решений [10–12].

### 6. Приложение.

Вектор  $\bar{z}^{(m)} = (z_k^{(m)})_{2K} = (x_{m1}, y_{m1}, x_{m2}, y_{m2}, \dots, x_{mK}, y_{mK})$  характеризует численность adeptов  $x_{mk}$  и предадептов  $y_{mk}$  относительно  $m$ -й подгруппы; элементы вектора сноса  $\bar{A}$  в (1) формируются по правилу  $\bar{a}^{(m)} = (a_i^{(m)})_{2K} = (f_{m1}^{(1)}, f_{m1}^{(2)}, f_{m2}^{(1)}, f_{m2}^{(2)}, \dots, f_{mK}^{(1)}, f_{mK}^{(2)})$  при  $f_{mk}^{(1)} = y_{mk} \left[ A_{mk}^0 + \sum_{m'=1}^M x_{m'k} \beta_{m'm} \right] - x_{mk} \Gamma_{mk}^0$  ;  
 $f_{mk}^{(2)} = \left[ N_m - \sum_{k'=1}^K (x_{mk'} + y_{mk'}) - y_{mk} \right] \left[ A_{mk}^0 + \sum_{m'=1}^M x_{m'k} \beta_{m'm} \right] + \Gamma_{mk}^0 (x_{mk} - y_{mk})$  ;  
 $\Gamma_{mk}^0 = \sum_{l=1}^L \chi_{ml} \gamma_{kl}^0$  ;  $A_{mk}^0 = \sum_{l=1}^L \chi_{ml} \alpha_{kl}^0$  ; элементы вектора  $\bar{W}$  в (1) определяются  $\bar{w}^{(m)} = (w_i^{(m)})_{2K} = (B_{m1}^\gamma, B_{m1}^\alpha, B_{m2}^\gamma, B_{m2}^\alpha, \dots, B_{mK}^\gamma, B_{mK}^\alpha)$  компонентами  $B_{mk}^\gamma$  и  $B_{mk}^\alpha$  – измеримые при всех  $t \in [t_0, t_1]$  независимые винеровские процессы при  $dB_{mk}^\gamma(t) = \Gamma'_{mk} dt$  ;  $dB_{mk}^\alpha(t) = A'_{mk} dt$  ;  $\Gamma'_{mk} = \tilde{\Gamma}_{mk} / \sqrt{\sum_{l=1}^L (\chi_{ml} \varepsilon_{kl}^\gamma)^2}$  ;

$$A'_{mk} = \tilde{A}_{mk} / \sqrt{\sum_{l=1}^L (\chi_{ml} \varepsilon_{kl}^\alpha)^2}; \quad \tilde{\Gamma}_{mk} = \sum_{l=1}^L \chi_{ml} \tilde{\gamma}_{kl}; \quad \tilde{A}_{mk} = \sum_{l=1}^L \chi_{ml} \tilde{\alpha}_{kl}; \quad \text{элементы}$$

матрицы диффузии  $\Sigma$  задаются  $\sigma_m = \text{diag}\{\sigma_1^{(m)}, \sigma_2^{(m)}, \dots, \sigma_K^{(m)}\}$  при

$$\sigma_k^{(m)} = \begin{pmatrix} \bar{g}_{mk}^\gamma & g_{mk}^\alpha \\ \bar{g}_{mk}^\alpha & \bar{g}_{mk}^\gamma \end{pmatrix}; \quad g_{mk}^\alpha = y_{mk} \sqrt{\sum_{l=1}^L (\chi_{ml} \varepsilon_{kl}^\alpha)^2}; \quad g_{mk}^\gamma = -x_{mk} \sqrt{\sum_{l=1}^L (\chi_{ml} \varepsilon_{kl}^\gamma)^2};$$

$$\bar{g}_{mk}^\gamma = (x_{mk} - y_{mk}) \sqrt{\sum_{l=1}^L (\chi_{ml} \varepsilon_{kl}^\gamma)^2}.$$

### Литература

1. Mikhailov A.P., Petrov A.P., Proncheva O.G. A model of information warfare in a society with a piecewise constant function of the destabilizing Impact // *Mathematical Models and Computer Simulations*. 2019. Vol. 11. P.190-197. DOI: 10.1134/S2070048219020108
2. Chkhartishvili A.G. The Problem of Finding the Median Preference of Individuals in a Stochastic Model // *Automation and Remote Control*. 2021. Vol. 82. P.853-862. DOI: 10.1134/S000511792105009X
3. Gubanov D.A., Petrov I.V., Chkhartishvili A.G. Multidimensional model of opinion dynamics in social networks: polarization indices // *Automation and Remote Control*. 2021. Vol. 82. P.1802-1811. DOI: 10.1134/S0005117921100167
4. Byzov L.G., Gubanov D.A., Kozitsin I.V., Chkhartishvili A.G. A perfect politician for social networks: an approach to analyzing ideological preferences of users // *Automation and Remote Control*. 2021. Vol. 82. P.1614-1631. DOI: 10.1134/S0005117921090095
5. Gubanov D.A., Petrov I.V. Information Communities in Social Networks. Part I: From Concept to Mathematical Models // *Control Sciences*. 2021. No 1. P.13-20. DOI: 10.25728/cs.2021.1.2
6. Gubanov D.A., Petrov I.V. Information Communities in Social Networks. Part II: Networked Models of Formation // *Control Sciences*. 2021. No 2. P.16-28. DOI: 10.25728/cs.2021.2.2
7. Mikhailov A.P., Yukhno L.F. Dynamics of the dissemination of information in society under the conditions of hype // *Mathematical Models and Computer Simulations*. 2021. Vol. 13. P.716-722. DOI: 10.1134/S2070048221040165
8. Chkhartishvili A.G., Gubanov D.A., Novikov D.A. *Social Networks: Models of information influence, control and confrontation*. Springer, 2018
9. Полянский И.С., Беседин И.И., Панин Б.Л. Математическая модель комплекса инженерно-технических средств системы физической защиты объекта охраны // *Фундаментальные исследования*. 2013. № 6 (часть 6). С.1359-1365.
10. Полянский И.С., Логинов К.О., Ильин Н.И., Великих А.С. Математическая модель оценки информационного воздействия на электорат в социальных медиа при проведении выборных кампаний // *Математическое моделирование*. 2021. Т.33. №12. С.67-81. DOI: 10.20948/mm-2021-12-05
11. Ильинский А.С., Полянский И.С., Логинов К.О., Архипов Н.С. К вопросу численной оценки информационного воздействия на электорат при проведении выборных кампаний // *Прикладная Математика и информатика: Труды факультета ВМК МГУ им. М.В. Ломоносова. – Москва: МАКС Пресс. 2021. №68. С.15-28.*

12. Полянский И.С., Полянская И.В., Логинов К.О. Алгоритмические решения в задаче оценки информационного воздействия на электорат при проведении выборных кампаний // Физика волновых процессов и радиотехнические системы. 2021. Т.24. №4. С.72-80. DOI: 10.18469/1810-3189.2021.24.4.72-80
13. Михайлов А.П., Петров А.П., Маревцева Н. А., Третьякова И. В. Развитие модели распространения информации в социуме // Математическое моделирование. 2014. Т.26. №3. С.65-74.
14. Свид. 2008613616 Российская Федерация. Свидетельство об официальной регистрации программы для ЭВМ. Система анализа тональности публикаций СМИ «Тонус» / А.А. Кукушкин, И.С. Полянский; заявитель и правообладатель А.А. Кукушкин, И.С. Полянский. – Заявл. 01.10.2008, Реестр программ для ЭВМ. – 1 с.
15. Орлов Ю.Н. Панкратов А.С. К разработке модели эволюции структуры сетевого графа // Препринты ИПМ им. М.В. Келдыша. 2021. № 24. 16 с. DOI: 10.20948/prepr-2021-24
16. Chugai K.N., Kosachev I.M., Rybakov K.A. Approximate filtering methods in continuous-time stochastic systems // Advances in Theory and Practice of Computational Mechanics. Smart Innovation, Systems and Technologies, Spring. 2020. Vol. 173. P.351-371. DOI: 10.1007/978-981-15-2600-8\_24.
17. Yumiharu Nakano. Kernel-based collocation methods for Zakai equations // Stochastics and Partial Differential Equations: Analysis and Computations. 2019. Vol. 9. P. 476-494. DOI: 10.1007/s40072-019-00132-y.
18. Yonghua Jiang, Chao Tang, Xiaodi Zhang, Weidong Jiao, Gang Li, Taotao Huang. A Novel Rolling Bearing Defect Detection Method Based on Bispectrum Analysis and Cloud Model-Improved EEMD // IEEE Access. 2020. Vol. 8. P. 24323-24333. DOI: 10.1109/ACCESS.2020.2970813
19. Цифровая обработка сигналов и изображений в радиофизических приложениях / М.А. Басараб и [др.], под ред. В.Ф. Кравченко. – Москва : ФИЗМАТЛИТ. 2007. 544 с.
20. Borovkov A.A., Moullagaliev A. Mathematical Statistics. . – London: Routledge. 1998. 592 p. DOI: 10.1201/9780203749326
21. Karimi N., Kazem S., Ahmadian D., Adibi H. Ballestra L.V. On a generalized Gaussian radial basis function: Analysis and applications // Engineering Analysis with Boundary Elements. 2020. Vol. 112, P. 46-57.
22. Ильинский А.С., Полянский И.С., Степанов Д.Е. О сходимости баричесентрического метода в решении внутренних задач Дирихле и Неймана в  $R^2$  для уравнения Гельмгольца // Вестник Удмуртского университета. Математика. Механика. Компьютерные науки. 2021. Т. 31. № 1. С.3-18. DOI: 10.35634/vm210101
23. Даугавет И.К. Теория приближенных методов. Линейные уравнения. – 2-е изд., перераб. и доп. – Санкт-Петербург : БХВ-Петербург, 2006, 288 с.
24. Polyanskii I.S., Arkhipov N.S., Misyurin S.Yu. On solving the optimal control problem // Automation and Remote Control. 2019. Vol. 80. P.66-80. DOI: 10.1134/S0005117919010065
25. Epperson J.F. An Introduction to Numerical Methods and Analysis: Solutions Manual to Accompany, Third Edition. John Wiley & Sons, Inc. 2021. DOI:10.1002/9781119604570
26. Naranjo-Noda F.S., Jimenez J.C. Locally Linearized Runge-Kutta method of Dormand and Prince for large systems of initial value problems // Journal of Computational Physics. 2021. Vol. 426, P. 1-26. DOI: 10.1016/j.jcp.2020.109946.

27. Полянский, И.С., Патронов Д.Ю. Максимально правдоподобная оценка дисперсионно-ковариационной матрицы // Современные проблемы науки и образования. 2013. № 1.; URL: <https://science-education.ru/article/view?id=8516>
28. Lasserre J.B. Simple formula for integration of polynomials on a simplex // BIT Numerical Mathematics. 2021. Vol. 61, P. 523-533. DOI: 10.1007/s10543-020-00828-x
29. Свид. 2021682089 Российская Федерация. Свидетельство об официальной регистрации программы для ЭВМ. Программа реализации алгоритма фильтрации оценок информационного воздействия на электорат при проведении выборных кампаний / И.С. Полянский, К.О. Логинов, Ю.В. Санин; заявитель и правообладатель .С. Полянский, К.О. Логинов, Ю.В. Санин. – Заявл. 21.12.2021, Реестр программ для ЭВМ. – 1 с.
30. Pitolli F., Sorgentone C., Pellegrino E. Approximation of the Riesz–Caputo derivative by cubic splines // Algorithms. 2022. Vol. 15, No 2: 69. DOI: [doi.org/10.3390/a15020069](https://doi.org/10.3390/a15020069).
31. Полянский И.С., Полянская И.В., Фам Т.З. Математическая модель фильтрации канонических параметров спутника-ретранслятора при орбитальном движении // Физика волновых процессов и радиотехнические системы. 2019. Т.22. №4. С.50-57. DOI: 10.18469/1810-3189.2019.22.4.50-57
32. Encyclopedia of special functions: the Askey-Bateman project Volume 1 Univariate orthogonal polynomials / edited by Mourad H. Ismail (University of Central Florida) with assistance by Walter van Assche (KU Leuven, Belgium). Cambridge New York, NY Cambridge University Press, 2020. 388 p. DOI: 10.1017/9780511979156
33. Jiang Y., Tao J., Chen X. Simulation of Non-Gaussian/Non-stationary Random Vibration. In: Non-Gaussian Random Vibration Fatigue Analysis and Accelerated Test. Springer, Singapore. 2022. DOI: 10.1007/978-981-16-3694-3\_3

**Логинов Кирилл Олегович** — сотрудник, Управление информационно-телекоммуникационного обеспечения Службы специальной связи и информации Федеральной службы охраны Российской Федерации (УИТО Спецсвязи ФСО России). Область научных интересов: математическое моделирование, методы статистического анализа данных, динамические системы, оптимальное управление, стохастические дифференциальные уравнения. Число научных публикаций — 7. [kvirs@mail.ru](mailto:kvirs@mail.ru); переулоч Большой Кисельный, 4, 107031, Москва, Россия; р.т.: +7(920)298-03-78.

K. LOGINOV

**NUMERICAL SOLUTION OF THE PROBLEM OF FILTERING ESTIMATES INFORMATION IMPACT ON THE ELECTORATE**

*Loginov K. Numerical Solution of the Problem of Filtering Estimates Information Impact on the Electorate.*

**Abstract.** The formulation and numerical scheme for solving the problem of filtering estimates of the informational impact of mass media on the electorate, allowing with a high degree of accuracy at a given observation interval to estimate the number of individuals in society who prefer a certain political subject (opinion), are proposed in the article. A mathematical model for assessing the information impact on the electorate during election campaigns, which boils down to solving a stochastic differential equation – the equation of state, forms the basis of the formulation of the problem. When compiling a model for filtering information impact estimates, it is proposed to reduce the study of the equation of state to a numerical solution of the Duncan–Mortensen–Zakai equation by introducing an additional observation equation, which is obtained from the equation of state when evaluating its stochastic components (observed agitation intensities) by methods of polyspectral analysis. In the projection formulation of the Galerkin method, when reducing to a system of linear differential equations and obtaining its solution in a recursive estimation scheme when sampling the analysis interval into subintervals and using the matrix exponential method, the Duncan–Mortensen–Zakai equation is solved. For a visual comparison of the effectiveness of the generated numerical solution to the problem of filtering information impact assessments, calculations were carried out on test examples.

**Keywords:** numerical solution, filtering of information impact estimates, election campaign, Duncan–Mortensen–Zakai equation, polyspectral analysis, Gaussian basis functions.

**Loginov Kirill** — Employee, Special Communications Department of the FSO of Russia. Research interests: mathematical modeling, methods of statistical data analysis, dynamical systems, optimal control, stochastic differential equations. The number of publications — 7. kvirs@mail.ru; 4, Bolshoy Kiselny lane, 107031, Moscow, Russia; office phone: +7(920)298-03-78.

**References**

1. Mikhailov A.P., Petrov A.P., Proncheva O.G. A model of information warfare in a society with a piecewise constant function of the destabilizing Impact // *Mathematical Models and Computer Simulations*. 2019. Vol. 11. P.190-197. DOI: 10.1134/S2070048219020108
2. Chkhartishvili A.G. The Problem of Finding the Median Preference of Individuals in a Stochastic Model. *Autom Remote Control*. 2021 vol. 82. pp. 853–862. DOI: 10.1134/S000511792105009X
3. Gubanov D.A., Petrov I.V., Chkhartishvili A.G. Multidimensional model of opinion dynamics in social networks: polarization indices // *Automation and Remote Control*. 2021. Vol. 82. P.1802-1811. DOI: 10.1134/S0005117921100167
4. Byzov L.G., Gubanov D.A., Kozitsin I.V., Chkhartishvili A.G. A perfect politician for social networks: an approach to analyzing ideological preferences of users // *Automation and Remote Control*. 2021. Vol. 82. P.1614-1631. DOI: 10.1134/S0005117921090095

5. Gubanov D.A., Petrov I.V. Information Communities in Social Networks. Part I: From Concept to Mathematical Models // Control Sciences. 2021. No 1. P.13-20. DOI: 10.25728/cs.2021.1.2
6. Gubanov D.A., Petrov I.V. Information Communities in Social Networks. Part II: Networked Models of Formation // Control Sciences. 2021. No 2. P.16-28. DOI: 10.25728/cs.2021.2.2
7. Mikhailov A.P., Yukhno L.F. Dynamics of the dissemination of information in society under the conditions of hype // Mathematical Models and Computer Simulations. 2021. Vol. 13. P.716-722. DOI: 10.1134/S2070048221040165
8. Chkhartishvili A.G., Gubanov D.A., Novikov D.A. Social Networks: Models of information influence, control and confrontation. Springer. 2018.
9. Polyanskiy I.S., Besedin I.I., Panin B.L. [Mathematical model complex engineering and technical systems of physical protection of property]. Fundamental'nye issledovaniya. 2013. no. 6-6. pp. 1359–1365. (In Russ.)
10. Poljanskij I.S., Loginov K.O., Ilyin N.I., Velikh A.S. [Mathematical model assessing the information impact on the electorate in social media during election campaigns]. Matematicheskoe modelirovanie. 2021. vol. 33. no. 12. pp . 67–81. (In Russ.)
11. Il'inskii A.S., Polyanskii I.S., Loginov K.O., Arkhipov N.S. On the issue of numerical assessment of informational impact on the electorate during electoral campaigns. Prikladnaya Matematika i informatika: Trudy` fakul'teta VMK MGU im. M.V. Lomonosova, Moskva: MAKS Press. 2021. no. 68. pp. 15–28. (In Russ.)
12. Polyanskii I.S., Polyanskaya I.V., Loginov K.O. [Algorithmic solutions to the problem of assessing the information impact on the electorate during election campaigns]. Fizika volnovykh processov i radiotekhnicheskie sistemy. 2021. vol. 24. no. 4. pp. 72–80. (In Russ.)
13. Mikhailov A.P., Petrov A.P., Marevtseva N.A., Tretiakova I.V. [Development of a model of information dissemination in society]. Matematicheskoe modelirovanie. 2014. vol. 26. no. 3. pp. 65–74. (In Russ.)
14. Kukushkin A.A., Polyanskii I.S., Sistema analiza tonal'nosti publikacij SMI «Tonus» [Tonus Media Publication Tonality Analysis System]. Certificate RF. no. 2008613616. 2008.
15. Orlov Y.N., Pankratov A.S. To the evolution model of network graph structure construction. Preprinty` IPM im. M.V.Keldy'sha. 2021. no. 24. 16 p. DOI: 10.20948/prepr-2021-24 (In Russ.)
16. Chugai K.N., Kosachev I.M., Rybakov K.A. Approximate filtering methods in continuous-time stochastic systems. Advances in Theory and Practice of Computational Mechanics. Smart Innovation, Systems and Technologies. Spring. 2020. vol. 173. pp. 351–371.
17. Yumiharu Nakano. Kernel-based collocation methods for Zakai equations. Stochastics and Partial Differential Equations: Analysis and Computations. 2019. vol. 9. pp. 476–494.
18. Yonghua Jiang, Chao Tang, Xiaodi Zhang, Weidong Jiao, Gang Li, Taotao Huang. A Novel Rolling Bearing Defect Detection Method Based on Bispectrum Analysis and Cloud Model-Improved EEMD. IEEE Access. 2020. vol. 8. pp. 24323–24333.
19. Basarab M.A. [Digital processing of signals and images in radiophysical applications] Cifrovaya obrabotka signalov i izobrazhenij v radiofizicheskikh prilozheniyah. Moscow. Fizmatlit Publ. 2007. 544 p. (In Russ.)
20. Borovkov A.A., Moullagaliev A. Mathematical Statistics. – London: Routledge. 1998. 592 p. DOI: 10.1201/9780203749326

21. Karimi N., Kazem S., Ahmadian D., Adibi H. Ballestra L.V. On a generalized Gaussian radial basis function: Analysis and applications // *Engineering Analysis with Boundary Elements*. 2020. vol. 112. pp. 46–57.
22. Il'inskiia A.S., Polyanskii I.S., Stepanov D.E. [On the convergence of the barycentric method in solving internal Dirichlet and Neumann problems in  $R^2$  for the Helmholtz equation] *Vestn. Udmurtsk. Univ. Mat. Mekh. Komp. Nauki*. 2021. vol. 31. no. 1. pp. 3–18. (In Russ.)
23. Daugavet I.K. [Theory of approximate methods. Linear equations] *Teoriya priblizhennykh metodov. Linejny'e uravneniya. – 2-e izd., pererab. i dop.* Saint-Petersburg, BHV-Peterburg, 2006. 288 p. (In Russ.)
24. Polyanskii I.S., Arkhipov N.S., Misyurin S.Y. On solving the optimal control problem. *Automation and Remote Control*. 2019. vol. 80. no. 1. pp. 66–80.
25. Epperson J.F. *An Introduction to Numerical Methods and Analysis: Solutions Manual to Accompany, Third Edition*. John Wiley & Sons, Inc. 2021.
26. Naranjo-Noda F.S., Jimenez J.C. Locally Linearized Runge-Kutta method of Dormand and Prince for large systems of initial value problems // *Journal of Computational Physics*. 2021. vol. 426. pp. 1–26.
27. Polyanskii I.S., Patronov D.Yu. Maximum reasonable estimate variance-covariance matrices. *Sovremennyye problemy nauki i obrazovaniya*. 2013. no. 1. URL: <https://science-education.ru/ru/article/view?id=8516>
28. Lasserre J.B. Simple formula for integration of polynomials on a simplex // *BIT Numerical Mathematics*. 2021. vol. 61. pp. 523–533. DOI: 10.1007/s10543-020-00828-x (In Russ.)
29. Polyanskii I.S., Loginov K.O., Sanin Yu.V., Svidetel'stvo ob oficial'noj registracii programmy dlya E'VM. Programma realizacii algoritma fil'tracii ocenok informacionnogo vozdejstviya na e'lektorat pri provedenii vy'bornykh kampanij [The program for the implementation of the algorithm for filtering estimates of information impact on the electorate during election campaigns]. Certificate RF. no. 2021682089. 2021.
30. Pitolli F., Sorgentone C., Pellegrino E. Approximation of the Riesz–Caputo derivative by cubic splines. *Algorithms*. 2022. vol. 15. no. 2: 69. DOI: [doi.org/10.3390/a15020069](https://doi.org/10.3390/a15020069).
31. Polyanskii I.S., Polyanskaya I.V., Pham T.Z. [Mathematical model of filtering canonical parameters of a repeater satellite during orbital motion]. *Fizika volnovykh processov i radiotekhnicheskie sistemy*. 2019. vol. 22. no. 4. pp. 50–57. (In Russ.)
32. *Encyclopedia of special functions: the Askey-Bateman project Volume 1 Univariate orthogonal polynomials* / edited by Mourad H. Ismail (University of Central Florida) with assistance by Walter van Assche (KU Leuven, Belgium). Cambridge New York, NY Cambridge University Press, 2020. 388 p. DOI: 10.1017/9780511979156
33. Jiang Y., Tao J., Chen X. Simulation of Non-Gaussian/Non-stationary Random Vibration. In: *Non-Gaussian Random Vibration Fatigue Analysis and Accelerated Test*. Springer, Singapore. 2022. DOI: 10.1007/978-981-16-3694-3\_3

## Руководство для авторов

Взаимодействие автора с редакцией осуществляется через личный кабинет на сайте журнала «Информатика и автоматизация» <http://ia.spcras.ru/>. При регистрации авторам рекомендуется заполнить все предложенные поля данных. Подготовка статьи ведется с помощью текстовых редакторов MS Word 2007 и выше или LaTeX. Объем основного текста (до раздела Литература) - от 20 до 30 страниц включительно. Переносы разрешены. Номера страниц не проставляются. Основная часть текста статьи разбивается на разделы, среди которых являются обязательными: введение, хотя бы один «содержательный» раздел и заключение. Допускается также мотивированное содержанием и структурой материал а выделение подразделов. В основную часть опускается помещать рисунки, таблицы, листинги и формулы. Правила их оформления подробно рассмотрены на нашем сайте в разделе «Руководство для авторов».

## Author guidelines

Interaction between each potential author and the Editorial board is realized through the pesoal account on the website of the journal "Informatics and Automation" <http://ia.spcras.ru/>. At the registration the authors are requested to fill out all data fields in the proposed form. The submissions should be prepared using MS Word 2007, LaTeX. The text of the paper in the main part should not exceed 30 pages. Pages are not numbered; hyphenations are allowed. Certain figures, tables, listings and formulas are allowed in the main section, and their typography is considered in more detail at the journal web.

---

Signed to print 01.06.2022

Printed in Publishing center GUAP, 67 litera A, B. Morskaya, St. Petersburg, Russia, 190000

---

The journal is registered in the Russian Federal Agency for Communications and Mass-Media Supervision, certificate ПИ № ФС77-79228 dated September 25, 2020  
Subscription Index П5513, Russian Post Catalog

---

Подписано к печати 01.06.2022. Формат 60×90 1/16. Усл. печ. л. 11,6. Заказ № 249.

Тираж 300 экз., цена свободная.

Отпечатано в Редакционно-издательском центре ГУАП,  
190000, Санкт-Петербург, Б. Морская, д. 67, лит. А

---

Журнал зарегистрирован Федеральной службой по надзору в сфере связи и массовых коммуникаций, свидетельство ПИ № ФС77-79228 от 25 сентября 2020 г.

Подписной индекс П5513 по каталогу «Почта России»