

**Институт информационных и вычислительных технологий  
МОН РК**



**МАТЕРИАЛЫ**  
**научной конференции**  
**ИИВТ МОН РК**  
**«Современные проблемы информатики и**  
**вычислительных технологий»**  
**2 - 5 июля 2018 года**

**Алматы 2018**

УДК 004(063)  
ББК 32.973  
С56

Главный редактор:  
**Калимолдаев М.Н.** - генеральный директор ИИВТ, академик НАН РК, доктор физико-математических наук, профессор

Ответственные редакторы:  
**Юничева Н.Р.** - ученый секретарь ИИВТ МОН РК, кандидат технических наук, доцент  
**Мамырбаев О.Ж.** - заместитель генерального директора ИИВТ, доктор PhD  
**Магзом М.М.** - заместитель генерального директора ИИВТ, доктор PhD

**С56** **Современные проблемы информатики и вычислительных технологий:** Мат. науч. конф. (2 - 5 июля 2018 г). – Алматы: ИИВТ МОН РК, 2018. – 347 с.

ISBN 978-601-332-124-0

В сборнике представлены материалы конференции ИИВТ МОН РК «Современные проблемы информатики и вычислительных технологий».

В сборнике опубликованы научные материалы сотрудников института, докторантов PhD и магистрантов, а также научных работников, принимающих участие в проектах по грантовому и программно-целевому финансированию.

Рассмотрены актуальные вопросы в области информационных и вычислительных технологий, информационной безопасности, проблем распознавания образов, классификации и теории принятия решений, оптимизации и оптимального управления динамическими системами, разработки моделей методов и информационной технологии построения интеллектуальных экспертных систем прогнозирования и управления сложными объектами, математического моделирования и управления динамическими, техническими и экономическими системами, теории моделей и спецификаций информационных систем, робототехнических систем, проблем синтеза и распознавания казахской речи.

Материалы сборника предназначены для научных работников, докторантов и магистрантов, а также студентов старших курсов.

УДК 004(063)  
ББК 32.973

ISBN 978-601-332-124-0

© Институт информационных и вычислительных технологий  
МОН РК, 2018

## **Программный комитет**

### **Председатель:**

Калимолдаев М.Н., академик НАН РК, д.ф.-м.н., профессор,  
генеральный директор ИИВТ МОН РК

### **Сопредседатели:**

Мамырбаев О.Ж., доктор PhD, зам. ген. директора ИИВТ МОН РК  
Магзом М.М., доктор PhD, зам. ген. директора ИИВТ МОН РК

### **Ученый секретарь:**

Юничева Н.Р., к.т.н., доцент, ученый секретарь ИИВТ МОН РК

### **Члены Программного комитета:**

Амиргалиев Е.Н., д.т.н., профессор  
Арсланов М.З., д.ф.-м.н., профессор  
Ашимов А.А., академик НАН РК, д.т.н.  
Бияшев Р.Г., д.т.н., профессор  
Калижанова А.У., к.ф.-м.н., доцент  
Карабеков Б.С., к.ф.-м.н., доцент  
Кудайкулов А.К., д.ф.-м.н., профессор  
Мазаков Т.Ж., д.ф.-м.н., профессор  
Мусабаев Р.Р., к.т.н.  
Мустафин С.А., к.т.н., доцент  
Найзабаева Л.К., д.т.н., профессор  
Нуртазин А.Т., к.т.н., доцент  
Нысанбаева С.Е., д.т.н., профессор  
Пак И.Т., академик РАЕН, д.т.н., профессор  
Самигулина Г.А., д.т.н., профессор  
Тлеубергенов М.И., д.ф.-м.н., профессор  
Утепбергенов И.Т., д.т.н., профессор

## **Организационный комитет**

### **Председатель:**

Сахариев Б.Б., член-корр. МАИН, зав. лабораторией ИИВТ МОН РК

### **Члены Организационного комитета:**

Абдилдаева А.А., доктор PhD  
Анищенко Л.Н.  
Ахметжанов М.А., доктор PhD, зам. ген. директора ИИВТ МОН РК  
Калиева Г.С.  
Капалова Н.А., к.т.н.  
Кулемзин А.А.  
Мажитов Ш.С.  
Шахмаев Р.А.

*Место проведения:* комн. 105 ИИВТ

# **CREDIT RISK ASSESSMENT AND MODELING METHODS USING RECOGNITION ALGORITHMS BASED ON CALCULATION OF ESTIMATES**

**Arslanov M.Z., Mustafin S.A. Naizabayeva L.K.**

*Institute of information and computational technologies, Almaty, Kazakhstan*

*The work is devoted to the study of the problems of determining the credit risk in lending to individuals and the development of a rating model of assessing borrowers based on the precedent approach.*

Timely detection of possible bankruptcy signs allows management of credit institutions (hereinafter banks) to take prompt measures to remedy the financial condition and reduce the risk of bankruptcy.

A risk is the probability of a loss that could affect the performance of an economic entity or an economic transaction. Moreover, since the purpose of the bank's activities is to maximize profits, it must pay great attention to the implementation of its operations with the minimum possible risks. To avoid bankruptcy, banks need to seek and apply effective methods and tools to manage these risks in order to achieve and maintain a stable position in the banking services market.

Numerous models for calculating credit risk evaluation of the borrower indicate the existence of problems in determining credit risk. In modern banking practice, the problem of developing a system that would allow a credit manager, having set current parameters, to obtain an estimate the repayment of the loan and interest, is topical. A computational system based on the ideas of the pattern recognition theory is proposed. It allows determining the borrower's risk class. Such a system can also be used in making various decisions in the course of monitoring real contracts - whether to conclude a contract, what measures to take in case of violations of the repayment schedule or interest payments, whether to prolong the contract, how to react to depreciation of collateral, etc.

Management of banking operations is essentially the management of risks associated with the bank portfolio, with a set of assets that provide the bank with income from its activities. The main part of the bank portfolio is loans to businesses and individuals, which is connected with a risk of full or partial loss of the bank's resources.

When developing the system, the following risk management algorithm was used:

1. Qualitative analysis - identification of the full range of risks, description of risks, analysis of initial assumptions, classification and grouping.
2. Quantitative analysis - formalization of uncertainty; calculation, assessment and accounting of risks.
3. Minimizing risks by designing risk management strategies, selecting the optimal strategy and implementing it.
4. Risk control - monitoring risks, reassessing and adjusting risks, as well as making operational decisions on deviations.

Risk management of the loan product is carried out at all stages of the life cycle of the risk evaluation system through monitoring and control actions in cases of deviation from the specified project parameters. Thus, the special nature of risk management as a stage of system design is to monitor continuously in dynamics and with adjustments in order to avoid

deviations from the parameters specified by the bank because of exposure to risks in the process of participating in the system design cycle.

The main types of risk are liquidity risk, interest rate risk, risk of default on a loan.

The latter type of risk is especially important, since the non-return of borrowed funds brings large losses to banks and can serve as one of the reasons for bankruptcy.

The form, the amount of lending, the method of repaying the loan and the collateral requirements depend on the lending risk.

In this regard, the task of prompt and objective evaluation of the lending risk is urgent.

The degree of risk of lending transactions is expressed in the highest interest rate for transactions that are of a credit nature (loans, guarantees) in comparison with other assets. Loan rates should compensate the bank for the value of the funds provided for the term, the risk of changing the value of collateral and the risk of the borrower failing to fulfill its obligations.

The risk of non-fulfillment of obligations by the borrower is determined by a large number of factors integrated in the concept of the client's creditworthiness.

The creditworthiness of a bank customer is its ability to fully and timely pay off on its obligations. Evaluation of the borrower's creditworthiness, as the task of determining the financial stability of the borrower. It is important both at the stage of selection of potential borrowers, and at the stage of control over the course of the loan repayment.

When assessing creditworthiness, risk factors are taken into account:

- 1) The nature of the transaction being negotiated;
- 2) Customer credit history;
- 3) The state of the industry and the region;
- 4) Position of the client in the specified industry;
- 5) The financial condition of the client;
- 6) The ability of the client to provide the property as collateral;
- 7) The social status of the client.

Analysis of credit risks is a complex task, including: developing methodological approaches to the analysis of a potential borrower; monitoring and analysis of changes in financial stability of a potential borrower, identification of unfavorable trends and their causes at the earliest possible stage.

To determine and analyze the risk, there are indicators for assessing the risk of lending. Given the multiplicity of risk indicators, the difference in the level of their critical assessments there is the need to evaluate existing techniques and determine how effective they are.

Traditional models of risk evaluation. Analysis of traditional models and methods of evaluating banking risks (such as the financial condition of the borrower, including the model of the borrower's classification by the point system, a two-factor model for assessing the probability of the bankruptcy of the borrower, assessing the probability of the borrower's bankruptcy based on the Altman Z-account, the model of Roman Lis for determining the financial state, the assessment of the borrower's financial position by U.Biver's indicators, the R-model of forecasting the risk of bankruptcy, etc.) makes it possible to identify the main shortcomings of such evaluation approaches.

The reasons for this are in the following. First, these models provide an incomplete, one-sided evaluation of the financial condition of the borrower, and therefore too large deviations of the forecast from reality are possible. Secondly, the models were developed for certain economic and political conditions. Thirdly, existing models of classification by the point system are usually developed conditionally and appropriate refinements are needed to apply them. Therefore, in order to obtain a more objective evaluation of the financial and

economic state of the borrower, there is a need to create a more effective evaluation methodology.

None of the decision rules can be recognized as universal and free of shortcomings that limit the scope of their application. Moreover, the creation of universal decision rules, apparently, is impossible in principle. This is due to the fact that, depending on the purposes, the system of preferences and the possibilities of obtaining information on preferences, various decision rules can be created [1].

**The proposed model of risk evaluation.** We propose a formalized approach that makes it possible to implement the decision-making process on the appropriateness of issuing a loan using pattern recognition techniques. The task of evaluating the financial condition is reduced to the classification of borrowers on the basis of sufficiently weak requirements for the initial data. Borrowers classified as belonging to the same class of patterns will have common properties, i.e. will have approximately the same level of financial stability, creditworthiness, and be equal.

When solving the task of classifying a borrower, the expert determines the classes of risk evaluation, characterized as a class of borrowers with a very high risk probability, a class of borrowers with an average risk probability, a class of borrowers with a very low risk probability. Borrowers who wish to take out a loan are referred to one of the previously defined classes, based on the level of their financial condition. Depending on the chosen risk class, the system determines which credit policy the manager should select.

**Substantial statement of the problem.** State classes are defined, each of which contains borrowers with the same risk evaluation. Each borrower is described by a set of features that characterize its state. It is required to put the defined borrower to one of the risk classes.

In the language of the mathematical theory of pattern recognition, this problem is defined as follows.

Formal setting of the problem. Input information is a set of vectors

$$X = \{x_j | j = 1, M\} \subset R^n, \text{ где } x_j = (x_j^{(1)}, \dots, x_j^{(n)}).$$

Here  $M$  – is the number of vectors  $n$  – is the number of characteristics,  $x_j$  -  $j$ -th is the vector ( $j$ -th observation).

The task of training with the teacher. Suppose that unknown  $q$  classes of objects  $X_1, X_2, \dots, X_q$ , that in a given set  $X$  (the training set) are represented by finite subsets

$$X_1, X_2, \dots, X_q, X = \bigcup_{i=1}^q X_i, X_i \cap X_j = \emptyset \quad \forall i \neq j.$$

It is necessary to create a rule by means of which it is possible to classify with a sufficient degree of reliability objects belonging to classes  $X_1, X_2, \dots, X_q$ , but not contained in the training set  $X$ . The criterion for the quality of the sought-for decision rule is the percentage of correctly recognized objects with a priori known classification, but not participating in training. A set of such objects is called a check or validation sample. Forming a validation sample from the set of all the objects presented for training is also a very important task.

The recognition of an object is based on a previously conducted training - a stage that precedes the recognition phase. The aim of the training is to select such thresholds, which

result in a minimum number of errors on the training set of objects with the known membership of classes (training). The next stage after the training is the stage of training control. It consists in checking the training, calculating the error in recognizing objects of the examination set of objects that belong to known classes (the exam). In addition, the last stage is the recognition of an unknown object.

At present, a large number of algorithms for object recognition have been developed and extensive experience in solving applied problems in various fields of science and technology has been accumulated. The most developed and well-known model of recognition is the model of algorithms for calculating assessments. An algebraic theory of pattern recognition is created [2]. Multiparametric algorithms of this model were used as the basis for the decision support system for determining the size of the loan.

The class of algorithms based on the calculation of assessments is given by describing the six elements defining it:

1. A system of reference sets,
2. Proximity function,
3. Calculation of estimates for the rows of a fixed reference set,
4. Calculation of the estimate for the class from the reference set,
5. An estimate for a class in the system of reference sets,
6. The decision rule.

Any recognition algorithm, the task of which consists of these six elements, will be an algorithm for calculating assessments, which are a set of all possible recognition algorithms that can be specified by the six elements under consideration.

Thus, we can say that by choosing a particular system of reference sets, defining the proximity function, specifying the rules for calculating assessments for the rows of a fixed reference set, for a class over a reference set and for a system of reference sets, also assigning a decision rule, we get some specific algorithm for computing evaluations. The class of recognition algorithms based on the calculation of assessments includes all possible algorithms that can be created from the six elements considered.

Advantages of this model are as follows.

1. Controllable model parameters;
2. Availability of training;
3. The analysis models are parametric;
4. Possibility of finding significant objects of classes;
5. Presence of weights of objects, signs;
6. Joint accounting of heterogeneous information;
7. Ability to select significant signs;
8. Evaluation of expert breakdown of objects into classes;

Using this approach allows us to obtain a more accurate and objective result. This helps to reduce the risk of loss and make a profit, as well as to give reasons for the decisions made when working with lenders and partners.

The ultimate goal of the work is the construction of an experimental computer system for determining the risk of lending based on the borrower's condition. For this purpose, the information system "Risk evaluation in retail lending" (hereinafter referred to as the "RISK" system) was developed, on the input of which the borrower's descriptions submitted by the vector-descriptions and divided by the expert into classes on risks are submitted. For a recognizable borrower, the system determines the class of its credit risk.

**Integrated software environment** [3-7]. The "RISK" system includes the following modules: a characteristic normalization module, a training module, an examination module,

a recognition module. In addition, the system includes a module for analyzing the borrower's risk based on a dynamic set of characteristics, a module for selecting significant characteristics, a module for selecting financially stable borrowers, a module for selecting the standards for each risk class, and a module for integrating dissimilar remote systems.

**Conclusion.** In this work, we touched on the issues of credit risks in banking operations and showed the possibility of using recognition algorithms based on the calculation of evaluations in them. There are also other approaches to the definition of bank risks in the issuance of loans. We note that a lot of research has been devoted to the theoretical foundations of risk in multicriterion and gaming problems.

### References

1. Озерной В.М. Проблемы и методы принятия решений при векторном критерии // Вопросы кибернетики. Теория принятия решений. Вып. 8. – 1974. – С. 53-60.
2. Журавлев Ю.И. Алгебраический подход к решению задач распознавания и классификации // Проблемы кибернетики. – 1978. - Вып. 33. - С. 5-68.
3. Джуматова Г.К., Мустафин С.А. Комплекс алгоритмов анализа и распознавания объектов // Известия НАН РК, сер. физ.-мат., 2007, №3. - С. 64-69.
4. Мустафин С.А., Зейнуллина А., Масимханова Ж. Анализ многомерных данных в задачах оптимизации // Новости науки Казахстана, № 2. – Алматы: 2013, С. 63-68
5. Мустафин С.А., Джуматова Г.К. О методике оценки кредитного риска физических лиц // Экономика: стратегия и практика. Институт экономики КН МОН РК. Научное издание. 2010, №4, С. 45-50.
6. Мустафин С.А., Джуматова Г.К. Оценка кредитоспособности предприятия // Новости науки Казахстана, № 2. – Алматы: 2008, С. 63-68
7. Мустафин С.А., Джуматова Г.К. Банковская оценка кредитоспособности физического лица. Сборник материалов Международной научно-практической конференции: 2-3 декабря 2010 г., Омский институт (филиал) РГТЭУ, 2010. – С.477-480

## ENERGY METHOD RESEARCH THERMO-STRESS- DEFORMATION STATE IN A ROD OF VARIABLE CROSS SECTION

**Kudaykulov A.K.<sup>1</sup>, Tashev A.A.<sup>1</sup>, Arshidinova M.T.<sup>1,2</sup>, Begaliyeva K.B.<sup>1,2</sup>**

<sup>1</sup>*Institute of Information and Computing Technologies CS MES RK*

<sup>2</sup>*Al-Farabi Kazakh National University*

e-mail: *mukaddas\_arshidi@mail.ru, kalamkas\_b@mail.ru,*

*kudaykulov2006@mail.ru, azattash@mail.ru*

**Abstract.** *A horizontal rod of limited length is considered. Radius of the rod varies linearly along its length. The cross-sectional area of the left end is greater than the right one. The lateral surface of the test rod is completely insulated. The heat flow is fed to the cross-sectional area of the left end. Through the cross-sectional area of the right end of the rod,*



heat exchange takes place with the surrounding medium. The field of distribution of temperature, displacement, three components of deformation and stresses are determined in the work, provided that both ends of the rod are rigidly fixed. And also, the magnitude of the elongation of the rod is determined when one end of the rod is fixed and when the other is free. In the case of fixing the two ends of the rod, the magnitude of the resulting axial compressive force is also calculated. When studying the rod, the fundamental laws of conservation of energy were used.

**Keywords.** Elongation, axial force, cross-section, temperature, displacement, deformation, stress.

## I. INTRODUCTION

Many load-bearing elements of gas-generator, nuclear and thermal power stations, jet engines and the processing industry are rods of variable cross-section. To ensure reliable operation of these equipments, it is necessary to provide the thermal strength of load-bearing elements in the form of variable-section rods that operate with the simultaneous action of dissimilar kinds of heat sources. Because of the variability of the cross section, nonlinear thermomechanical processes appear in such rods. To study the nature of such processes, consider a horizontal rod of limited length, of variable cross-section. In this case, the radius of the section varies linearly along the length of the investigated rod, i.e.  $r = ax + b$ , ( $0 \leq x \leq l$ ), where is the  $l$  - length of the rod,  $a, b - const$ . The cross-sectional area of the rod varies nonlinearly along the length of the rod in the following manner  $F(x) = \pi(ax + b)^2$  [ $\text{cm}^2$ ]. The lateral surface of the test rod along the entire length is heat-insulated. On the cross-sectional area of the left end of the rod  $F(x=0) = \pi b^2$ , a heat flux with a constant intensity  $q$  [ $\frac{\text{watt}}{\text{cm}^2}$ ]. Through the cross-sectional area of the right end of the rod  $F(x=l) = \pi(al + b)^2$ , heat exchange takes place with the surrounding medium. At the same time, the heat transfer coefficient  $h$  [ $\frac{\text{watt}}{\text{cm}^2 \cdot ^\circ\text{C}}$ ], ambient temperature  $T_{oc}$  [ $^\circ\text{K}$ ], the physical and mechanical properties of the core material is characterized by the coefficient of thermal expansion  $\alpha$  [ $\frac{1}{^\circ\text{K}}$ ], thermal conductivity  $K_{xx}$  [ $\frac{\text{watt}}{\text{cm} \cdot ^\circ\text{K}}$ ] and modulus of elasticity  $E$  [ $\frac{\text{kgf}}{\text{cm}^2}$ ]. The scheme of the investigated rod is shown in Fig. 1.

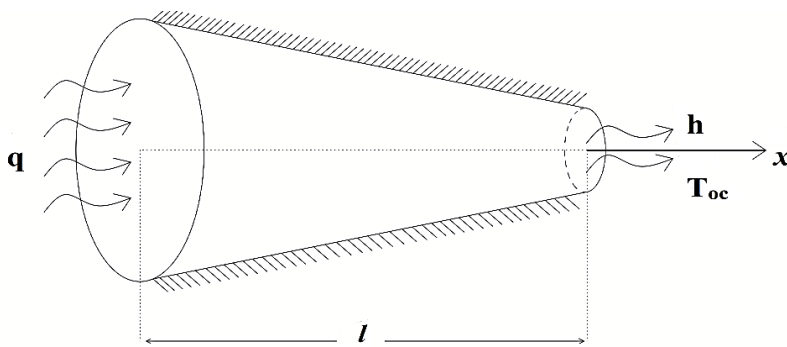


Figure 1. Scheme of the investigated rod

## II. OVERVIEW

In the presence of heat flow, heat insulation and heat transfer, the functional of the total thermal energy for the investigated rod has the form [1]:

$$J = \int_{F(x=0)} qTds + \int_V \frac{K_{xx}}{2} \left( \frac{\partial T}{\partial x} \right)^2 dv + \int_{F(x=l)} \frac{h}{2} (T - T_{oc})^2 ds, \quad (1)$$

where  $T=T(x)$  the field of distribution of temperatures along the length of the rod, which is approximated by a complete polynomial of the fourth order

$$T(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 = \phi_i(x)T_i + \phi_j(x)T_j + \phi_k(x)T_k + \phi_m(x)T_m + \phi_n(x)T_n \quad (2)$$

where  $\phi(x)$  are spline functions:

$$\begin{aligned} \phi_i(x) &= \frac{(3l^4 - 25l^3x + 70l^2x^2 - 80lx^3 + 32x^4)}{3l^4}, \\ \phi_j(x) &= \frac{(48l^3x - 208l^2x^2 + 288lx^3 - 128x^4)}{3l^4}, \\ \phi_k(x) &= \frac{(-36l^3x + 228l^2x^2 - 384lx^3 + 192x^4)}{3l^4}, \\ \phi_m(x) &= \frac{(16l^3x - 112l^2x^2 + 224lx^3 - 128x^4)}{3l^4}, \\ \phi_n(x) &= \frac{-3l^3x + 22l^2x^2 - 48lx^3 + 32x^4}{3l^4}, \end{aligned} \quad (3)$$

$0 \leq x \leq l$  where the nodal temperature values

$$T_j = T(x=0); \quad T_i = T(x=\frac{l}{4}); \quad T_k = T(x=\frac{l}{2}); \quad T_m = T(x=\frac{3l}{4}); \quad T_n = T(x=l); \quad (4)$$

Taking into account (2–4), minimizing (1) with  $T_i$ ,  $T_j$ ,  $T_k$ ,  $T_m$  and  $T_n$  we obtain a resolving system of algebraic equations taking into account existing natural boundary conditions. Solving the system we determine the nodal values of temperature (4), and by (2) we construct the field of temperature distribution along the length of the rod. If one end of the rod is fixed and the other end is free, then the length of the rod  $\Delta L_T$  [cm] is determined according to the general law of thermophysics [1]

$$\Delta L_T = \int_0^l \alpha T(x) dx. \quad (5)$$

If both ends of the rod are rigidly fixed, then an axial compressive force  $R$  [kG] arises in the rod, which is determined from the compatibility condition of the deformation [1]

$$R = \frac{\Delta l_T \cdot E \int_0^l F(x) dx}{l^2}. \quad (6)$$

In this case, a distribution field of the thermo-elastic component of the voltage  $\sigma(x) [\frac{\kappa \Gamma}{\text{CM}^2}]$  arises in the rod:

$$\sigma(x) = \frac{R}{F(x)}, \quad 0 \leq x \leq l. \quad (7)$$

Then, according to Hooke's law, we can determine the distribution field of the thermo-elastic deformation component  $\varepsilon(x)$  [dimensionless]:

$$\varepsilon(x) = \frac{\sigma(x)}{E}. \quad (8)$$

The temperature component of deformations  $\varepsilon_T(x)$  [dimensionless] is determined according to the general law of thermophysics [1]:

$$\varepsilon_T(x) = -\alpha T(x). \quad (9)$$

Then, according to Hooke's law, the field of distribution of the temperature component of the stress

$$\sigma_T(x) \left[ \frac{\kappa \Gamma}{\text{CM}^2} \right]:$$

$$\sigma_T(x) = E \cdot \varepsilon_T(x) = -\alpha E \cdot T(x). \quad (10)$$

According to the theory of thermoelasticity, the laws of distribution of elastic components of deformations  $\varepsilon_x(x)$  [dimensionless] and stresses  $\sigma_T(x) [\frac{\kappa \Gamma}{\text{CM}^2}]$ :

$$\varepsilon_x(x) = \varepsilon(x) - \varepsilon_T(x), \quad (11)$$

$$\sigma_x(x) = E \cdot \varepsilon_x(x) = \sigma(x) - \sigma_T(x). \quad (12)$$

The potential energy of elastic deformations is used to determine the displacement field [4]:

$$\Pi = \int_V \frac{\sigma_x(x)}{2} \varepsilon_x(x) dv - \int_V \alpha E \cdot T(x) \cdot \varepsilon_x(x) dv. \quad (13)$$

According to the Cauchy relation [4], we have:

$$\varepsilon_x(x) = \frac{\partial U}{\partial x}, \quad (14)$$

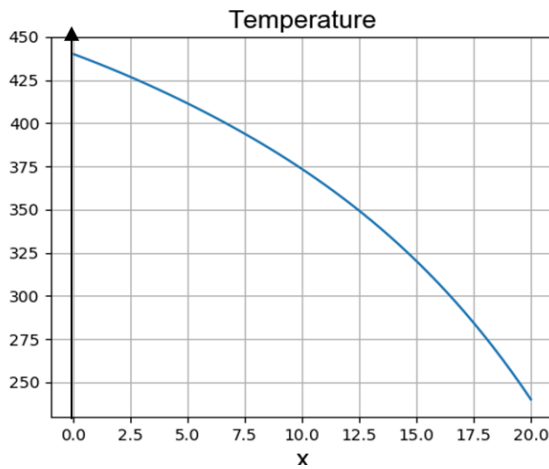
$$U = U(x) = \phi_1(x)U_i + \phi_2(x)U_j + \phi_3(x)U_k + \phi_m(x)U_m + \phi_n(x)U_n, \quad (15)$$

where  $U$  is the displacement field.

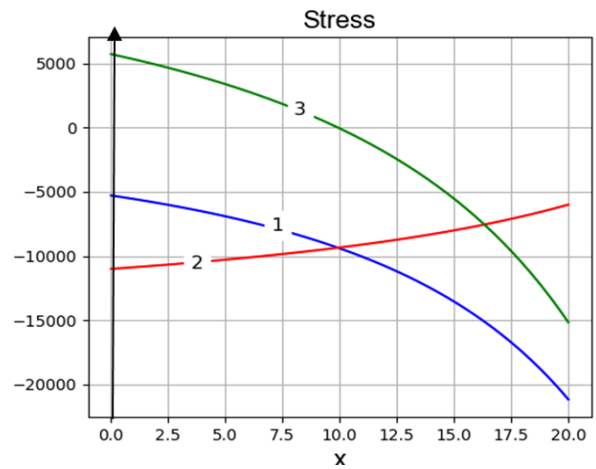
Minimizing  $\Pi$  from the nodal values of the displacement, a system of linear algebraic equations is constructed. To solve this system, it is necessary to specify the conditions for securing the two ends of the rod, i.e.  $U_i = U(x=0) = 0$  and  $U_n = U(x=l) = 0$ .

Further, defining  $U_i, U_j, U_k, U_m, U_n$  and substituting them in 15), a displacement field is constructed. For practical application of the above method and algorithm, we take the following initial data  $l=20\text{cm}$ ;  $a=\frac{1}{10}$ ;  $b=4\text{cm}$ ,  $\alpha=0,0000125\frac{1}{\text{K}}$ ;  $E=2\cdot 10^6\frac{\text{KG}}{\text{cm}^2}$ ;  
 $K_{xx}=100\frac{\text{watt}}{\text{cm}\cdot\text{K}}$ ;  $h=10\frac{\text{watt}}{\text{cm}^2\cdot\text{K}}$ ;  $T_{\infty}=40^{\circ}\text{K}$ ;  $q=-500\frac{\text{watt}}{\text{cm}^2}$ ;

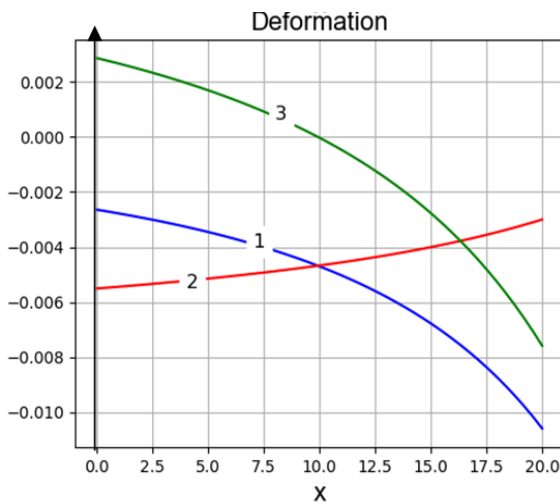
With these initial data, the solutions obtained are shown in Figure 2-5.



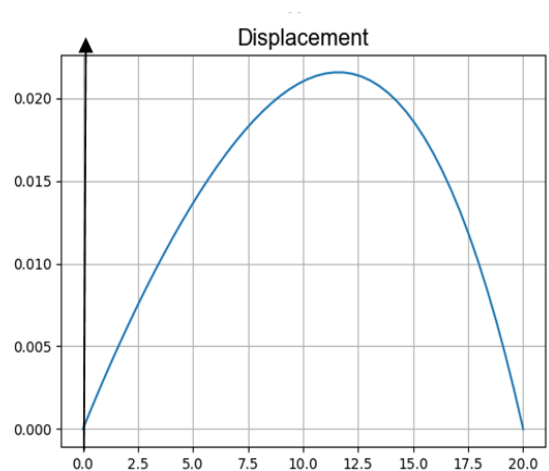
**Fig.2. Dependences of the temperature T along the length of the rod**



**Fig.3. Stress Dependencies along the length of the rod**



**Fig. 4. Dependence of the deformation along the length of the rod**



**Fig.5. Dependences of displacement along the length of the rod**

### III. DECISION

Figure 2 shows that the temperature is higher near the left end of the rod, where the heat flow is supplied. Due to the thermal insulation of the lateral surface, heat is lost minimally, so that the temperature at the right end of the rod is maintained at  $240^{\circ}\text{K}$ .

The stresses along the length of the rod are shown in Figure-3 (1– $\sigma(x)$ –thermoelastic, 2– $\sigma_T(x)$ – temperature, 3– $\sigma_x(x)$ – elastic component of the stress). It can be seen from the figure that the thermoelastic  $-\sigma(x)$  and temperature  $-\sigma_T(x)$  components of the stress along the entire length of the rod are of a compressive nature. While the elastic  $-\sigma_x(x)$  component of the stress in the region  $0 \leq x \leq \frac{l}{2}$  has a tensile character, and in the region  $\frac{l}{2} \leq x \leq l$  it is compressive.

Dependences of deformations along the length of the rod are shown in Figure 4 (1– $\sigma(x)$ –thermoelastic, 2– $\sigma_T(x)$ – temperature, 3– $\sigma_x(x)$ – elastic component deformation). The distribution field of the deformation components is proportional to the corresponding stresses.

Figure 5 shows the field of distribution of displacements of a rod fixed at two ends. From this it can be seen that all sections (except for exceptions) move in the direction of the  $x$  axis. The greatest amplitude of displacement corresponds to the coordinate of  $x \approx \frac{3l}{5}$ .

### IV. CONCLUSION

A numerical model of nonlinear thermomechanical processes in a rod of variable cross-section is developed, based on the fundamental law of conservation of energy. This allows to obtain reliable numerical results taking into account all natural boundary conditions. The results obtained are consistent with the corresponding laws of physics. This method can be used for the numerical solution of a class of problems determined by the steady-state thermomechanical state of load-bearing structural elements operating under the influence of dissimilar kinds of heat sources.

*This work is performed within grant financing of AP05131093 of Committee of Science of the Ministry of Education and Science of the Republic of Kazakhstan.*

### REFERENCES

1. A. Kudaykulov, A. Zhumadillayeva Numerical simulation of temperature distribution field in beam bulk in the simultaneous presens of heat insulation, heat flux and heat exchange // Acta physica polonica - A, 2016. – P. 335-336.
2. Tashenova, Z., Nurlybaeva, E. Kudaykulov, A. Method preparation and solution algorithm for resolving stationary problem of a rod under thermo - stressed condition restrained at both ends affected by heat exchange and heat flows // 2012 International Conference on Advanced Material and Manufacturing Science, ICAMMS 2012. Beijing. China.
3. Muheyat Niyazbek, Kuenssaule Talp, Kudaykulov A.K. Computer and Mathematical Modeling of Thermomechanical Processes in Element of Constructions. Acta Scientiarum Naturalium Universitatis Nankaiensis, 2017.-Vol 50, № 3. Tianjin. China.

4. Timoshenko S., Goodyear J. N. Theory of Elasticity / – McGRAW-Hill. Book. Company. Inc., 1987. – 567 p.

## ЖАЗЫҚ КРИСТАЛДЫҢ ЕКІӨЛШЕМДІ МОДЕЛІ

Амирханова Г.А., Дуйсенбаева А.Ж.

*Институт информационных и вычислительных технологий КН МОН РК,  
Казахстан*

e-mail: *gulshat.aa@gmail.com, aigerim.95.05@mail.ru*

*Аңдатпа.* Бұл мақалада көп қабатты бөлікті-біртектес материалдың жазық моделі және оңтайландыру есебі қарастырылады. Материал периодты бөлікті-біртектес көп қабатты құрылым түрінде ұсынылады, соның аясында әртүрлі қабаттардағы атом түрлері ерекшеленуі мүмкін. Жазық кристалды құрылымы бар материалдар қазіргі заманғы өнеркәсіпте кеңінен қолданыс табады. Мұндай материалдың жарқын мысалы ретінде графен ұсынылады - көміртегінің бір өлшемді кристалы. Соңғы уақытта фосфорен, силицен және т.с.с. басқа да жазық кристалдар синтезделді.

*Түйін сөздер:* материал, жазық кристалл, атом, бөлікті-біртектес материал, екі өлшемді модель.

**Кіріспе.** Заманауи ғылымды дамытудың басым бағыттарының бірі жаңа материалдарды құру болып табылады. Материалдар туралы ғылым және, атап айтқанда, нанотехнологияның дамуы арқасында, көптеген салаларда маңызды технологиялық серпінділікке қол жеткізілді. Заманауи наноматериалдар медицинада, ғарыштық және әскери өнеркәсіпте, әртүрлі өнім түрлерін өндіруде, құрылыста және т.б көптеген салаларда қолданылады [1]. Материалдар туралы ғылымда пайдаланылатын математикалық моделдеу әдістері, әртүрлі материалдардың қасиетін, нақты физикалық экспериментсіз сандық зерттеуге мүмкіндік береді. Мұндай тәсіл шығындардың айтарлықтай төмендеуіне қол жеткізуге және жаңа материалдарды жасау процесін жеделдетуге мүмкіндік береді.

### **Жазық кристалдың екіөлшемді моделіне сипаттама**

Бұл мақалада қарастырылатын көп қабатты бөлікті-біртектес материалдың жазық моделі [2] ұсынылған модельді анықтайды. Материал периодты түрде бөлікті-біртектес көп қабатты құрылым түрінде беріледі. Модель қабаттар құрылымына бірқатар шектеулер енгізеді:

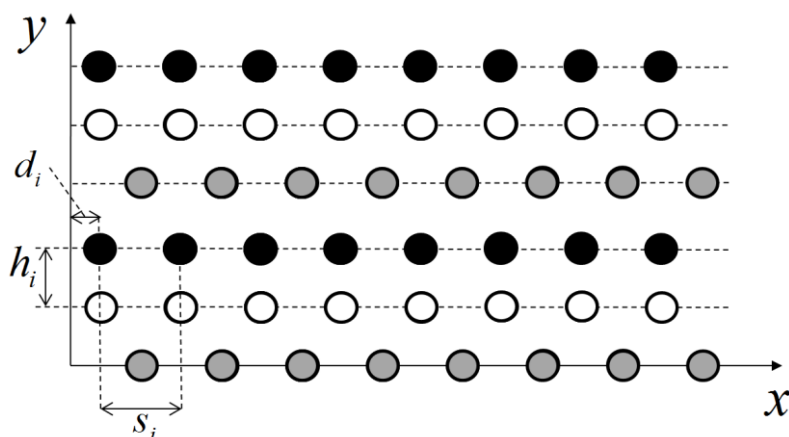
1. әр қабат бірдей атомдардан тұрады, әртүрлі атомдар әртүрлі қабаттарда болуы мүмкін;

2. бір қабаттағы іргелес атомдар арасындағы қашықтықтар бірдей, бірақ әр түрлі қабаттарда олар ерекшеленуі мүмкін;

3. қаралып жатқан жүйеде К параллельді қабаттардан тұратын топ бөлінеді, у өсі бағыты бойынша периодты қайталанып отырады;

4. әр қабаттағы атомдардың саны және қабаттардың жалпы саны потенциалды шектеусіз деп саналады.

1 Суретте үш қабаттан тұратын, қайталанып отыратын атомдар тобы моделінің мысалы көрсетілген. Әрбір қабат өзіне сай атомдардан тұрады.



Сурет. 1 Материалдың екі өлшемді моделі

Берілген модельде атомдардың орналасуы келесі параметрлермен анықталады:

$h_i, i=1, \dots, K$  -  $i$  нөмірлі қабат және алдыңғы қабат арасындағы қашықтық;

$d_i, i=1, \dots, K$  -  $i$  қабаттағы нөлдік белгіге қатысты оң абсциссалы бірінші атомның жылжуы;

$s_i, i=1, \dots, K$  -  $i$  қабаттағы атомдар арасындағы қашықтық.

$d_1 = 0, d_i < s_i, i=2, \dots, K$

Бұл параметрлердің мәндер жиынтығы конфигурация деп аталады. Материалдың бос қалпында болу жағдайы төменде қарастырылады. Материалдың моделденетін фрагментіне енетін, атомдар әсер етуінің минималды қуатына өзара сәйкес келетін конфигурацияны анықтау қажет.

### Оңтайландыру есебінің қойылымы

Материалдың құрылымы  $k$  ұзындықты атомдар түрлерінің векторы арқылы және ұзындығы  $3k$  болатын  $z = (h_1, d_1, s_1, \dots, h_k, d_k, s_k)$  геометриялық сипаттамаларының векторымен толығымен анықталады. Жалпылықты шектемей, айталық  $d_1 = 0, d_i < s_i, i \in \overline{1, k}$ .

Топ деп барлық түрлі қабаттардағы атомдардың  $i$ -қабатындағы  $n_i + 1$  атомдарының жиынтығын атайды.  $i$  тобының шеткі сол атомының абсциссасы  $a_i = d_i$  формуласы бойынша есептеледі, ал шеткі оң атомы -  $b_i = d_i + n_i s_i$  формуласымен есептеледі.

Егер барлық  $i, j \in \overline{1, k}$  үшін  $b_i - a_i = b_j - a_j$  шарты орындалса, топты периодты деп атайды. Периодтылық шарты барлық  $i, j \in \overline{1, k}$  үшін келесі түрде жазылуы мүмкін  $s_i n_i = s_j n_j$ . Бұл шарт келесі формулаға эквивалентті  $s_i n_i = s_1 n_1, i = \overline{1, n}$ .

Жазық кристалдың геометриялық құрылымын есептеу екі кезеңде жүзеге асырылады. Біріншіден,  $k$  қабаттары бар  $[0, L]$  аралықта  $x$ -өсімен шектелген материалдың фрагменті қарастырылады. Содан соң, осы материал фрагменті атомдарының энергиясын минимизациялау жүргізіледі,  $z^*$  параметрлерінің векторлары алынады. Сонда атомдардың периодты тобының анықтамасы шығады. Ол үшін оңтайландыру есебі шешіледі:

$$n_1^* = \min\{n \in \mathbb{N}: \sum_{i=1}^k \min(\frac{s_1^* n}{s_i^*} - \lfloor \frac{s_1^* n}{s_i^*} \rfloor, \frac{s_1^* n}{s_i^*} - \lfloor \frac{s_1^* n}{s_i^*} \rfloor)\} \leq \delta\}, \quad (1)$$

мұнда  $\delta$  – периодтылық анықтамасының дәлдігін анықтаудың жеткілікті аз саны. (1) алынған  $n_1^*$  мәнін қолдану арқылы  $n_i^*$  мәндері формула бойынша есептеледі

$$n_i^* = \lfloor \frac{s_1^* n_1^*}{s_i^*} \rfloor, i \in \overline{1, k}.$$

Екінші кезеңде, бірінші кезеңде анықталған периодтық топтағы энергия минимизацияланады. Периодтық топтың энергиясы оған кіретін атомдар энергиясының қосындысына тең деп есептеледі:

$$E = \sum_{i=1}^k \sum_{j=1}^{n_i} E_{ij},$$

мұндағы  $E_{ij}$  -  $i$ -ші нөмірлі қабатта орналасқан  $j$ -ші атомның энергиясы.

Кристалдардың әр түрлі қасиеттерін және құрылымын сипаттау үшін екі негізгі тәсіл қолданылады. Бірінші тәсіл кванттық-механикалық байланыс теңдеулерін шешуге негізделген. Мұндай тәсіл жүйелерді жоғары дәлдікпен модельдеуге мүмкіндік береді, бірақ есептеу ресурстарын аса талап етеді. Басқа тәсіл «күштік өріс» идеясына негізделген: атомдар тобы энергиясы өзара әрекеттесу потенциалымен жуықтатылады. Бұл тәсіл атомдардың үлкен жиынтығын модельдеу үшін қажет.

Көпбөлшекті өзара әрекеттесу потенциалы идеялары алдымен Абелдің [3] жұмыстарында пайда болады, онда байланыс қуаты үшін жалпы көрініс пайда болады, бұл жақын көршілерінің жұптық өзара әрекеттесуінің сомасы. Абел теориясының одан әрі дамуын Терсофф [4], Бреннер [5] және олардың ізбасарларының [6] жұмыстарынан көруге болады.

**Қорытынды.** Сонымен бұл жұмыста жазық кристалдың екі өлшемді моделі талданды. Заманауи наноматериалдарды қолданудың қажеттілігі көрсетілген. Бұл модель материалдың механикалық қасиеттерін анықтау үшін қолдануы және қабатталған құрылымдардың кең ауқымына кеңейтілуі мүмкін.

#### Пайдаланылған әдебиеттер тізімі

1. Ferrari A.C. et al. Science and technology roadmap for graphene, related two-dimensional crystals and hybrid systems // *Nanoscale*. 2015. V. 7. № 11. P. 4598–4810.
2. Лурье С.А., Посыпкин М.А., Соляев Ю.О. Метод идентификации масштабных параметров градиентной теории упругости на основе численных экспериментов для плоских композитных структур // *Internat. J. of Open Informat. Technol.* 2015. Т. 3. № 6. С. 1–6.
3. Abell G. C. Empirical chemical pseudopotential theory of molecular and metallic bonding // *Physical Review B*. – 1985. – Т. 31. – №. 10. – С. 6184.
4. Tersoff J. Modeling solid-state chemistry: Interatomic potentials for multicomponent systems // *Physical Review B*. – 1989. – Т. 39. – №. 8. – С. 5566.



5. Brenner D. W. Empirical potential for hydrocarbons for use in simulating the chemical vapor deposition of diamond films //Physical review B. – 1990. – Т. 42. – №. 15. – С. 9458.

6. Stuart S. J., Tutein A. B., Harrison J. A. A reactive potential for hydrocarbons with intermolecular interactions //The Journal of chemical physics. – 2000. – Т. 112. – №. 14. – С. 6472-6486.

## **ҰЛЫ ДАЛАНЫҢ БАЙЫРҒЫ КӨНЕ ЕСЕПТЕРІ – ҚАЗАҚ ХАЛҚЫНЫҢ ҚАСИЕТТІ РУХАНИ МҰРАСЫ**

**С.Н. Байбеков<sup>1</sup>, С.А. Алтынбек<sup>2</sup>, А.У. Елеусинова<sup>3</sup>  
А.С. Тургинбаева<sup>4</sup>**

*<sup>1</sup>Ақпараттық және есептеу технологиялары институтының профессоры,  
т.ғ.д.,*

*<sup>2</sup>Ақпараттық және есептеу технологиялары институтының аға ғылыми  
қызметкері, PhD,*

*<sup>3</sup>Ақпараттық және есептеу технологиялары институтының аға ғылыми  
қызметкері,*

*<sup>4</sup>Ақпараттық және есептеу технологиялары институтының аға ғылыми  
қызметкері, Алматы, Қазақстан,*

*e-mail: baibekovsn@mail.ru, serik\_aa@bk.ru, altyn\_eleusinova@bk.ru,  
ssalua@mail.ru*

Байырғы қазақи көне есептер осы уақытқа дейін әдебиеттану оның ішінде фольклористика ғылымының да, басқа ғылымдардың да зерттеу нысанасынан тысқары қалып келді. Осындай жағдай тек бізге ғана емес, әлемнің барлық еліне тән. Бұған негізгі себеп: филология мен нақты ғылымдардың, атап айтсақ әдебиет пен есептеу ғылымдарының бір бірінен алшақтауы, осы ғылымдар арасындағы сабақтастық пен жалғастықтың төмендеуі.

Қазіргі таңда әртүрлі себептерге байланысты гуманитарлық және нақты ғылымдар арасында пәнаралық байланыс, жалғастық пен сабақтастық тіпті кездеспейді десек те болады. Ал кез келген нақты ғылымның бастапқы қайнары көзі, оның шыққан жері - халықтың әлеуметтік жағдайы, тұрмыс салты, күнделікті өмірі т.с.с. Өткен заман ғұламалары, мәселен, Архимед, Аристотель, әл-Фараби және басқалар өздерінің нақты ғылыми трактаттарын өмірдің белгілі бір жағдайына байланысты әңгіме ретінде жазған. Ал қазіргі таңда нақты ғылымдар ауқымы өте зор болғандықтан, осы бағытта жазылатын ғылыми еңбектер құрғақ фактілер ретінде, цифрлар мен формулалардың жиынтығы ретінде жазылады. Осы себепті гуманитарлық және нақты ғылымдар бір бірінен алшақтай берген және алшақтай береді.

Дегенмен кейбір оппоненттер «есептердің ауыз әдебиетіне, оның ішінде фольклорға қандай қатысы бар?» - деуі мүмкін. Әрине, бұл – мәселеге жеңіл қарау. Себебі фольклор тек аңыз-әпсана, жыр-қисса, ертегі, тұрмыс-салт өлеңі мен әндері, мақал-мәтелдер, жұмбақтар, жаңылтпаштар т.б. ғана емес. Фольклордың ауқымы өте

кең. Мәселен, жұмбақ - дедукция және индукция әдістерімен шешілетін математикалық логика есептері екені баршамызға мәлім. Демек, халықтың байырғы көне есептері де - фольклордың өзіндік бір тармағы. Өкінішке қарай, бұл тақырып осы күнге дейін бізде ғана емес, барлық елдерде де зерттелмеген, тіпті зерттелуі басталмаған проблема. Демек, халықтың көне есептерін зертеу және оны бір жүйеге келтіріп, оның ауыз әдебиетіндегі орнын ғылыми түрде анықтау - жаңа ғылыми бағыт деп санаймыз. Біздің ойымызша ғылымдағы ескерілмегенді ескеріп, орнына келтіріп, кемді толтыру арқылы халқымыздың қасиетті рухани құндылығын одан сайын тереңдетіп қана қоймай, фольклортану ғылымына халық есептерін жаңа бағыт ретінде енгізу мүмкіндігі бар.

Қазіргі таңда белгілі қазақи көне есептер қарапайым арифметикадан бастап геометрияны, жоғары математиканы, тіпті сандар теориясы мен комбинаториканы, физиканы да қамтиды. Ал осы есептерге гуманитарлық ғылымдар нысанасынан қарасақ, онда олар әлеуметтік тұрмыс-жағдайды, халықтың салт- дәстүрін, өмір сүру ерекшеліктерін, философия, тарих, заң жүйесін т.с.с. қамтиды. Оның ішінде Абылай хан, Бәйдібек, Зере ана, Төле би, Қазыбек би, Әйтеке би, Сырым батыр, Бөлтірік шешен және басқа да тарихи тұлғалар аттарымен аталатын есептер өзіндік келбетпен ерекшеленіп тұрады.

Президент Н.Назарбаев **«Болашаққа бағдар: рухани жаңғыру»** атты өзекті мақаласында: *«...бұрынғыдай тарихи тәжірибе мен ұлттық дәстүрге шекеден қарамау керек. Керісінше, замана сынынан сүрінбей өткен озық дәстүрлерді жаңғырудың маңызды алғышарттарына айналдыра білу керек. Егер жаңғыру елдің ұлттық-рухани тамырынан нәр алмаса, ол адасуға бастайды»* деп көрсетті. Өте орынды және нақты табылып айталған ой. Енді осы бағытта қазақ халқының көне есептеріне назар аударайық.

Қазақ халқының санау жүйесі басқа халықтармен салыстырғанда өзіндік сипаттармен ерекшеленеді. Мәселен, ондық жүйедегі сандар барлық Еуропа және басқа елдер тілінде «бір ондық», «екі ондық», «үш ондық», сол сияқты «сегіз ондық», «тоғыз ондық» деп аталады. Ал қазақ тілінде олардың әрқайысының өзіндік бөлек атауы бар: он, жиырма, отыз, қырық... сексен, тоқсан. Осы атаулар қазақ халқында есептеу ғылымы көне заманның өзінде-ақ кеңінен дамығандығын көрсетеді. Әрине, бұған әртүрлі тұрмыс-салт ерекшелігі әсер еткен. Мәселен, қазіргі таңда кеңінен тарап жүрген «фотографиялық жад» (фотографическая память), «ойша ұшқыр есептеу» (ментальная арифметика) деген әдістер біздің халықта көне заманның өзінде де болған. Бұған дәлелдің бірі - осы күнге дейін сақталып келген қолдың саусақ буындары арқылы қосу, алу, көбейту және бөлу тәсілдері. Екінші дәлел – қазақта малшы қауымы қазіргі күннің өзінде де қой санағанда оларды он-оннан бейнелей топтастырып (шын майталмандар жиырмадан топтастырып), мың шақты қойды 1-2 минут ішінде лезде санап шығып, нәтижесін «компьютерден» де тез есептейді. Бұндай тәсіл, әрине, фоторграфиялық жады мен лездік ұшқыр есептеу дүниесінің бірін бірі толықтыру арқасында іске аса алады. Мәселен, әйгілі шопан, 2 мәрте Социалистік Еңбек Ері Жазылбек Қуанышбаев малды жиырмадан бейнелей топтастырып санау тәсілінің майталман шебері болғанын халық осы күнге дейін аңыз қылып айтады...

Осы топтастырып санау әдісіне арналған талай көне есептер халықтың көне сандығында әлі де беті ашылмай жатқаны рас.

Халықтың байырғы көне есептері тұрмыс-салт өлеңдері сияқты алуан тақырыпты қамтиды. Тіпті кейбір есептердің өзіндік адуын «мінездері» болады. Кейбірулері – жайдары, кейбіреулері асау мінезді, табан астында жалт береді. Мысалы, мына бір көне есеп (өте қысқартылған түрінде келтірейік):

*«Ертеде бақуатты бір жігіт қарақұрт шағып кенеттен қайтыс болардың алдында бауыр-жолдастарына және артында қалған айы-күні жетіп отырған екіқабат жас келіншегіне аманат айтқан екен: «Менің артымда қалған дүние мүлікті бірдей үш бөлікке бөліп, егер әйелім қыз туса, онда бір бөлікті қыздың еншісіне, ал қалған екі бөлікті әйеліме беріңдер. Ал егер әйелім ұл туса, онда екі бөлікті ұлдың еншісіне, қалған бір бөлікті әйеліме беріңдер» деп дүниеден өтіпті.*

*Көп ұзамай әйелі егіз ұл мен қыз туыпты. Енді еншіні қалай бөлу керек?..».*

Міне келтірілген есеп табан астында тосын мінез көрсете жалт беріп, күтпеген жауапты талап етіп тұр. Тіпті егер әйелі егіз екі ұл, немесе егіз екі қыз туса, сол сияқты екі ұл мен бір қыз туса, немесе екі қыз бен бір ұл туса, онда енші қалай бөлінбек деген сауалдар да туындап тұр...

Келесі есеп даналығы және байсалдығымен ерекшеленеді (қысқартылған түрінде келтіріледі):

*«Ерте заманда халыққа қадірлі, елге сыйлы Зере атты әйгілі әйел өтіпті. Сол әйелде ертеден енеден келінге, тек қана енеден келінге (қызға емес – қыз жат жұрттық) беріліп келе жатқан асыл тастардан жасалған көне және өте қымбат алқа болыпты. Сол кісі қартайған шағында өз кіндігінен тараған жеті келінін шақырып алып, мына алқаны біреуіңе тапсыратын кез келді. Егер біреуіңді таңдап соған аманат қылып тапсырсам, онда қалған алтауың өкпелейсіңдер. Демек, олай жасауға болмайды. Егер екеуіңді таңдап алып, алқаны екіге бөлсем, онда бір моншақ артылып қалады, Демек, олай да болмайды. Сол сияқты алқаны үшке, төртке, беске және алтыға бөлсем, онда әрбір бөліс кезінде бір моншақ артылып қалады, яғни бөліске түспейді. Егер алқаны жетіге бөлсем, онда алқа моншақтары қалдықсыз теңдей бөлінеді. Бірақ алқаны ешелер өсиеті бойынша ешқандай бөліске салмай бүтіндей бір келінге тапсыруым керек. Ендеше осы алқада қанша моншақ бар? Кім шешеді – сол келін осы алқаға лайық, алқа соған беріледі. Қалғандарың бұған ренжімейсіңдер...» деген екен.*

*Осы есепті бірінші болып шешкен келін, көп ұзамай ақылдылығымен, даналығымен аты елге жайылып, Алқа-Тұмар келін, Тұмар келін, Тұмар ана аталып кетіпті...*

*Ел арасында сол алқа әлі де сақтаулы деген сөз бар...»*

Қандай керемет есеп! Есептің даналық ерекшелігі айтпай-ақ көрініп тұр...

Осы жерде айта кету керек, қазіргі заман ғылым дүниесінде осындай есептерді шешу тәсілін «Қытайдың ұлы тәсілі (Великая китайская теорема об остатках)» деп атайды. Өкінішті-ақ... Ұлы Даладан жан-жаққа шашыраған қасиетті көне дүниелердің бірі болар...

Көптеген ежелгі есептер Ұлы Даладағы әлеуметтік тұрмыс-салтты дәлме дәл сипаттауымен ғажайып ерекшеленеді. Мәселен, мына бір есеп (қысқартылған түрінде келтіріледі):

*«Ертеде Ұлы Далада дүркіреген ас беріліпті. Үлен жиын алдында ас иесі «Асқа келген әр төреге төрт табақ, әр қожасаға қос табақ, төрт қараға бір табақ тартылсын» депті. Ас иесінің тапсырмасы мүлтіксіз орындалып, ас біткен соң қорытынды жасаса, асқа жүз қонақ келіп жүз табақ тартылыпты. Ендеше, асқа қанша төре, қанша қожаса және қанша қарапайым адам қатысты?».*

Бұл есептің көркемділігі тыңдаушы мен оқырманға оқиғаны (асты) елестетіп-ақ тұрғандай... Дегенменде, аңғарып қарасақ, бұл есептің артына жасырын жұмбақ ілескен: сонда қалай, төре бір өзі төрт табақ дәм жейтін мешкей болған ба? Әрине, олай емес... Төре мен қожалар - нөкерлерін ертіп жүретін Ұлы Дала ақсүйектері, билік және

дін иелері. Өз дәрежесіне сәйкес, мәселен, төре 15-16, ал қожа 7-8 нөкерін ертіп жүретін болған...

Көне есептердің осындай сипаттарын көптеп келтіруге болады және оларды анықтап бір жүйеге түсіріп, олардың ауыз әдебиетінде алатын орнын анықтау – бүгінгі ғылым алдына қойылар басты міндетінің бірі деп есептейміз. Бұндай мәселелер гуманитарлық және нақты ғылымдар арасындағы пәнаралық байланыс, жалғастық пен сабақтастықтың арқасында орындалмақ. Атап айтқанда, көптеген байырғы көне есептер халық тұрмысындағы болар жағдайды пайымдау, істелмек нәрсені алдын ала мейлінше ұқыптылықпен бағдарлау және оларды нақты есептей отыра жобалау, олардың тиімділігіне алдын ала көз жеткізу мақсатында өмірге келген. Демек, осы жағдайға тиянақты зер салсақ, ғылымның пән аралық байланыс тәсілі осы аталып отырған ғылыми мәселені тиянақты атқаруға толықтай мүмкіндік береді.

Есептеу ғылымы қалыптаспаған ортада есеп үлгілерінің болуы мүмкін еместігі баршамызға мәлім. Ендеше, байырғы көне есептер тек қана *дәстүрлі ортада*, яғни логикалық ойлау қабілеті жоғары дамып, есеп құрастыру және оны ауызша шығару дәстүрі орныққан және де шығармашылық қабілеті ерекше өркендеп, есте сақтау қабілеті кемелденген ортада ғана пайда болып дами алатыны түсінікті құбылыс. Демек, байырғы Ұлы Дала есептері ерте заманның өзінде-ақ қазақ халқының логикалық ойлау жүйесі мен шығармашылық қабылеттері ерекше дамыған ел болғанын көрсететін қасиетті тарихи құндылық және рухани мұра.

Егер байырғы қазақи көне есептер ұрпақтан ұрпаққа, ғасырдан ғасырға қағаз бетіне түсірілмей-ақ халық жадында сақталып келе жатқанын еске алсақ, онда олар халық ауыз әдебиетінің ерекше танымдық нысаны екені түсінікті болады. Олардың мыңдаған жылдар бойы халық жадында мәңгілік орын алған рухани құндылық екені айқындала түседі.

Міне осындай себептермен Президент Н.Назарбаев *«Болашаққа бағдар: рухани жаңғыру»* атты өзекті мақаласында: *«...Ұлттық жаңғыру дегеннің өзі ұлттық сананың кемелденуін білдіреді. Оның екі қыры бар. Біріншіден, ұлттық сана-сезімнің көкжиегін кеңейту. Екіншіден, ...құндылықтар жүйесінде білімді бәрінен биік қоятын ұлт қана табысқа жетеді»* деп жазды.

Ендеше, еліміздің әрбір азаматы үшін ұлттық сана-сезімнің көкжиегін кеңейтуге, қасиетті ұлттық рухани құндылықтың кемелденуіне бір адамдай атсалысу үлкен де ұлағатты абзал борыш.

Бұл мақала Қазақстан Республикасы Оқу және Білім министрлігі Білім комитетінің «Мәңгілік Ел» ғылыми негіздері» атты ғылымды дамытудың басым бағытының Бағдарламалық-нысаналы қаржыландыру шеңберінде №BR05236075 *«Ұлы Даланың байырғы көне есептерін қазақ халқының қасиетті рухани мұрасы және фольклорлық құндылығы ретінде зерттеу, талдау және жаңа жүйеге келтіру»* атты ғылыми-техникалық бағдарлама бойынша жазылды.

## **КРИПТОГРАФИЯЛЫҚ ЖҮЙЕЛЕРДЕ КІЛТТЕРДІ БАСҚАРУ**

**Капалова Н., Хаумен А., Абишева А.**

*ҚР Білім және ғылым министрлігі Ғылым комитетінің  
«Ақпараттық және есептеуіш технологиялар институты», Қазақстан  
e-mail: kapalova@ipic.kz, haumen.armanbek@gmail.com, ak\_maral@mail.ru*

***Аннотация.** Бұл мақалада позициялық емес полиномдық санау жүйелерінің негізінде құрылған ақпаратты криптографиялық қорғау алгоритмдеріндегі кілттерді басқару сұрақтары қарастырылған. Криптографиялық кілттерді генерациялау, жинақтау және тарату мәселелері талқыланып, оларды шешудің жолдары, құрылымы мен қасиеттеріне шолу жасалынады. Сонымен қатар позициялық емес полиномдық санау жүйелеріне негізделіп құрылған алгоритмнің кілттік құрылымы да қарастырылады.*

Ақпараттық қауіпсіздікті қамтамасыз ету – ақпараттық технологияларды дамытудың басым бағыттарының бірі. Осы салада шешілген мәселелердің ауқымы сандық және сапалық тұрғыда үнемі кеңеюде. Компьютерлік жүйелерде ақпаратты қорғау үшін қолданылатын негізгі құралдардың бірі криптографиялық түрлендірулер болып табылады. Заманауи криптографияның төрт басты бөлімі бар: симметриялы криптожүйелер, ашық кілтті жүйелері, электронды қолтаңба жүйесі, кілттерді басқару.

Қазіргі уақытта криптожүйелер кілттерді пайдалануға негізделген. Кілттерді басқару ақпарат алмасудың құпиялығын, деректердің сәйкестілігін және тұтастығын қамтамасыз ету үшін шешуші рөл атқарады. Әдетте, кілттерді басқару криптографиялық қосымшалардың ең осал тұсы болып табылады. Криптографиялық технологияны пайдалану қарапайым, бірақ кілттерді сақтау, кілттерді пайдалану және олардың өзара алмасуын қамтамасыз ету әлдеқайда қиын болып табылады. Сенімділігі төмен кілттік басқару жүйесі өте жақсы ұйымдастырылған жүйелердің сапасын төмендетеді, өйткені алгоритмнің бар қауіпсіздігі кілттердің қауіпсіздігіне шоғырланған.

Алайда заманауи криптографияның алдында тұрған негізгі мәселелер, бұл алгоритмнің тұрақтылығын арттыру және кілттер мен мәліметтер блогының мөлшерін азайту. Бұл мәселені шешудің ең айқын жолы-криптографиялық алгоритмдердегі ақпараттар блогын тек сандар түрінде (немесе ақырлы өрістің элементтерін) көрмей, оларды басқа да үлкен қиындықты алгебралық объект түрінде көру. Осындай типтегі объектілер қатарына позициялық емес полиномды санау жүйелері (ПЕПСЖ) жатады [1-3].

Ақпараттық және есептеуіш технологиялар институтының ақпараттық қауіпсіздік зертханасында позициялық емес санау жүйелеріне негізделген ақпаратты шифрлеу алгоритмдері құрастырылып, сынақтан өтуде [4]. Практикалық іске асыру және теориялық зерттеулер нәтижелері көрсеткендей, осы ПЕПСЖ-сін дәстүрлі емес алгоритмдер мен кодтау әдістерін құрастыруда және шифрлауда, электрондық сандық қолтаңбаны (ЭСК) қалыптастыруда пайдалану, осы криптографиялық рәсімдердің айтарлықтай сенімділігін арттыруға, ЭСК-ның ұзындығын азайтуға мүмкіндік береді. Ұсынылған алгоритмдер криптографиялық тұрғыда жақсы қасиеттерге ие екені арнайы жүргізілген сынақтардың қорытындысы дәлеледеп берді [5]. Ендігі кезекте осы шифрлеу алгоритмдеріне негізделген криптографиялық жүйеге қажетті кілттерді басқару мәселесі туындап отыр.

Кілттерді басқару процедурасына **кілттерді жасау, оларды жинақтау және тарату** қызметтері жатады [6]. Осыларға шолу жасап, міндеттері мен функцияларын сипаттап өтейік. Жақсы ұйымдастырылған кілттерді басқару жүйесінің маңызды сипаты бірнеше кілттердің қауіпсіздігін қамтамасыз ету мәселесіне көптеген кілттердің қауіпсіздігін қамтамасыз етудің күрделі мәселелерін біріктіру болып табылады. Сақталатын ақпараттың қауіпсіздігін қамтамасыз ету үшін кілттерді қолданған жағдайда, субъекті бір пайдаланушы болуы мүмкін және ол дәйекті уақыт

аралығында деректермен жұмыс істейді. Байланыс желілеріндегі кілттерді басқару кем дегенде екі субъектіні қамтиды – хабарды жіберуші және алушы.

Негізгі кілттерді басқарудың мақсаты – қауіптерді бейтараптандыру. Ол қауіптерге мыналарды жатқызуға болады:

- Жеке кілттердің құпиялылығын бұзу;

- Ашық немесе жабық кілттердің шынайылығын және аутентификациялығын бұзу. Бұл жерде шынайылық ретінде осы кілтті қолданатын желінің конфиденциалдығын қамтамасыз ету үшін корреспонденттің түпнұсқалылығын білу немесе тексеру мүмкіндігін түсінеміз;

- Ашық немесе жабық кілттерді рұқсатсыз пайдалану, мысалы, кілтті пайдалану мерзімі аяқталған кезде.

Кілттерді басқару функциялары мыналарды қамтиды:

- ЭЦҚ кілттерін қалыптастыру;

- Ашық кілтті тіркеу және тіркеу туралы куәлікті беру;

- Ашық кілтті куәліктің қолданылу мерзімін белгілеу;

- Ашық кілттерге қол жеткізуді ұйымдастыру;

- Сертификат мәртебесін анықтау;

- Қайтарып алудың себебін көрсететін куәліктерді қайтарып алу;

- Сертификаты тоқтата тұру / ұзарту;

- Сертификаттарды мұрағаттау;

- Сертификаты және қауіпсіздік саясатын пайдалануды басқару [7].

Кілттерді **жасаудың** қолданыстағы әдістерін аппараттық және бағдарламалық деп бөлуге болады. Бұл жағдайда басты талап – кілттердің үлестірімінің бірқалыптылығы. Кілттерді аппараттық жасау кезінде кездейсоқ физикалық процесс жүретін электрондық құрылғылар – шу генераторлары қолданылады.

Бағдарламалық іске асыруда - псевдокездейсоқ тізбек генераторлары қолданылады. Псевдокездейсоқ сандық генераторларды таңдағанда арнайы критерийлер орындалуы қажет [8].

Кілттерді **жинақтауды** ұйымдастыру – оларды сақтау, есепке алу және жою әрекеттерімен байланысты. Күрделі ақпараттық жүйеде бір пайдаланушы негізгі ақпараттың үлкен көлемімен жұмыс істей алады, соның әсерінен кейде негізгі ақпарат үшін шағын дерекқорларды ұйымдастыруды талап етеді. Мұндай жүйелер пайдаланылатын кілттерді қабылдау, сақтау, есепке алу және жоюға жауап береді. Пайдаланылған кілттер туралы ақпарат шифрланған түрде сақталу керек. Кілт туралы мәліметтерді шифрлайтын кілттер мастер-кілттер деп аталады және әдетте олар компьютер жүйесінде сақталмайды, оларды түрлендіру үшін криптографиялық алгоритмдер қолданылады. Пайдаланылатын кілттердің саны абоненттер санына, берілген ақпараттың көлемі мен шифрлау алгоритмінің ерекшелігіне байланысты. Бұл жағдайда сеанстық кілттер жойылуы қажет.

Кілттерді басқарудың үшінші мәселесі кілттерді жаңарту туралы сұрақпен тікелей байланысты – кілттерді тарату. Кілттерді тарату криптографияның негізгі міндеттерінің бірі болып табылады. Мәселенің ауқымын түсіну үшін, жеке мәліметтерді бір-бірімен алмастыратын  $n$  пайдаланушыларға қызмет көрсету кезінде  $n(n - 1)/2$  түрлі құпия кілттер қажет екенін көрсету қажет. Сонымен  $n$  өсуімен көптеген кілттерді басқару мәселесі туындайды. Оны шешудің бірнеше жолы бар. Ең қолайлысын анықтау келесі жағдайларға байланысты [9]:

- *Физикалық үлестіру.* Сенімді курьерлер немесе қарулы күзетшілердің көмегімен кілттерді дәстүрлі физикалық құралдармен жіберуге болады. Бұл процедура симметриялық немесе асимметриялық криптожүйелерде де қолданылады. Кілт

жасаушы қолданушыға асимметриялы құпия (немесе асимметриялық ашық кілтті) кілтті физикалық қауіпсіз түрде жібереді деп алдын ала белгіленеді.

- *Орталық кілт берушімен өзара әрекеттесу барысында қатысушылар ортақ кілт алады, яғни - «абоненттік шифрлау» схемасы.* Мұндай жүйеде кілттерді шығару және тарату орталығы жіберілген хабарламалардың шынайы және дұрыс екенінің кепілі болып табылады. Өйткені ол пайдаланушыларды кілттермен қамтамасыз етіп қана қоймай, сонымен қатар кілттерді жасауда және жеткізуде олардың құпиялығына жауап береді. Егер орталыққа зиян келтіріліп сенім жоғалса, кілттерге қатысты келесі сұраныстарды қамтамасыз етуде қиындық туады және бұрын берілген кілттердің қауіпсіздігі криптожүйеге байланысты болады.

- *Кілттерді сертификациялау орталығы арқылы пайдаланушылардың ашық кілттеріне қол жеткізу және құпия кілттерді беру.* Мұнда да симметриялық және асимметриялық криптожүйелер қолданылады. Пайдаланушы жұмыстың ең басында әр кілттерді тарату орталығымен қауіпсіз жолмен өзара байланысуы керек, сонда бұл бастапқы кілт алмасу жағдайында мәселенің бірі болып табылады. Егер ұзақ мерзімді құпия кілттер пайдаланушылар арасында және қандай да бір кілттерді тарату орталығы арасында таратылса, онда арнайы криптографиялық хаттамалар қолданылады. Таратудың бұл әдісінде пайдаланушылар және орталық «онлайн» режимінде жұмыс істеу ескеріледі.

- *Сенім желісі.* Асимметриялық криптожүйелерде қолданылады. Пайдаланушылар өз кілттерін таратады және басқа пайдаланушылардың кілттерін бақылайды. Кілттерді алмасудың бұл ресми емес түрінде сенім жатыр. Бір шешім, яғни әрбір пайдаланушыға сенім орталығымен байланыса алатын бір кілт бекітіледі. Бұл жағдайда  $n$  пайдаланушыларға арналған жүйе тек  $n$  кілтті қажет етеді. Екі пайдаланушы құпия мәлімет алмасқысы келгенде, олар осы хабарламаны жіберу үшін ғана пайдаланылатын кілт жасайды. Бұл *сеанс кілті* деп аталады. Сеанс кілті сенім орталығының қатысуымен хаттамалардың бірі арқылы жасалады.

- *Кілттермен алмасу хаттамасы.* Осы уақытқа дейін ортақ құпия кілті болмаған өзара қатысушылар арасында қорғалмаған байланыс арналары арқылы құпия кілт жасалады және алмасады. Ашық кілтті криптожүйелерді қолдана отырып, делдалдарға сенбейтін және кездесуге қабілетсіз серіктестер кілт айырбастау хаттамасына сәйкес желіде ортақ құпия кілтпен «онлайн» режимде келіссөз жүргізе алады. Бұл ашық кілтті шифрлау технологиясының ең көп қолданылатын түрі. Алдымен, тараптар алдын ала құпия кілтке келіседі. Содан кейін керекті ақпаратты шифрлау үшін келісілген кілтпен симметриялық шифр қолданылады.

Кілттерді орталықтан таратуды қолданатын әдістердің *кемшілігі* – орталық кімге және қандай кілттер тағайындалғанын біледі, ол ақпараттық жүйеде жүрген барлық хабарламаларды оқуға мүмкіндік береді. Кілттерді тікелей айырбастау кезінде субъектілердің түпнұсқалығын аутентификациялау проблемасы бар.

Сипатталған жүйелердегі кемшіліктерді есере отырып, заманауи талаптарға сай кілттерді басқару жүйесінің нұсқасын құру қарастырылды. Сонымен қатар, ұсынылған шифрлеу алгоритмдерінің ерекшеліктері ескеріледі.

Құрастырылған алгоритмнің бір ерекшелігі – бұл жүйеде кілт ретінде тек кездейсоқ сандар ғана емес, сонымен қатар жұмыс негізі болатын көпмүшеліктер де кілт ретінде алынады. Яғни, негізгі кілтке осы көпмүшеліктердің  $GF(2)$  өрісіндегі балама түрі де қосылады. Осыған байланысты бұл алгоритмдегі кілттер басқа алгоритмдерден қарағанда құрамдас болғандықтан толық кілт деп аталады. Бұл жүйеге арналған кілттерді басқаруда негізгі кілттерді генерациялаумен қоса осы аталған көпмүшеліктерді таңдау да қарастырылады.

Кілттерді басқаруды іске асыру симметриялық және асимметриялық біріктірілген криптожүйенің шеңберінде жүзеге асырылады. Осындай тәсілмен ашық мәтінді шифрлау және жіберу үшін ПЕПСЖ негізделген симметриялы криптожүйе пайдаланылады, ал симметриялық криптожүйенің құпия кілтін (толық кілтті) шифрлау, яғни толық кілтті шифрлау және жіберу үшін асимметриялық криптожүйесі қолданылады. Толық кілтпен шифрленген ақпаратпен бірге ашық кілттер жүйесінде шифрленген.

Алдағы уақытта позициялық емес санау жүйелеріне негізделген ақпаратты шифрлеу алгоритмінің кілттерін басқару жүйесін жобалау және құрастыру жұмыстары жалғасын табатын болады.

#### **Пайдаланылған әдебиеттер:**

1. Амербаев В.М., Бияшев Р.Г., Нысанбаева С.Е. Применение непозиционных систем счисления при криптографической защите информации, // Изв. Нац. Акад. наук РК. Сер. физ.- мат. наук. – Алматы: Ғылым, 2005. – № 3. – С. 84-89.

2. Бияшев Р.Г., Нысанбаева С.Е. Формирование электронной цифровой подписи с проверяющими функциями // Комплексная защита информации: Матер. XI Междунар. конф. (20-23 марта 2007 г., Новополюцк, Республика Беларусь). – Минск: Амалфея, 2007. – С. 51-54.

3. Капалова Н.А., Нысанбаева С.Е. Алгоритм открытого распределения ключей на базе непозиционной полиномиальной системы счисления // Вестник КазНУ. Сер. мат., мех., информат. - 2007. - №3 (54), - С. 82-87.

4. Biyashev R., Nyssanbayeva S., Kapalova N., Haumen A. Modified symmetric block encryption-decryption algorithm based on modular arithmetic // Proceedings of the International Conference on Wireless Communications, Network Security and Signal Processing (WCNSSP 2016). – Chiang Mai, Thailand, 2016. –P. 263-265. (Web of Science, Thomson Reuters)

5. Капалова Н.А., Хаумен А. Алгоритм шифрования на SP-сети // Материалы научной конференции Института информационных и вычислительных технологий МОН РК «Современные проблемы информатики и вычислительных технологий». – Алматы, 2017. – С. 113-118.

6. Фомина И.А. Управление ключами в криптографических системах. Вестник Нижегородского университета им. Н.И. Лобачевского. 2010, №4(1), стр.165-169.

7. Аристархов И.В. дисс. Управление сертификатами ключей проверки электронной подписи, Москва, 2012

8. Кнут Д.Э. Искусство программирования. Т.2. Получисленные алгоритмы. М.:Издательский дом «Вильямс», 2004. 832 с.

9. Шнайер Б. Прикладная криптография. Протоколы, алгоритмы, исходные тексты на языке Си. М.: Из-во «Триумф», 2003. 816с.

## **ЖҮЙКЕ ЖҮЙЕСІНІҢ ҚҰРЫЛЫМЫ МЕН ҚАСИЕТТЕРІН ИМИТАЦИЯЛАЙТЫН НЕЙРОНДЫ ЖЕЛІЛЕР**

**Қожахмет Б.<sup>1,2</sup>, Калижанова А.У.<sup>1,2</sup>**

*ҚР БҒМ ҒК, Ақпараттық және есептеуіш технологиялар институты,  
e-mail: banu\_kozhakhmet@mail.ru*

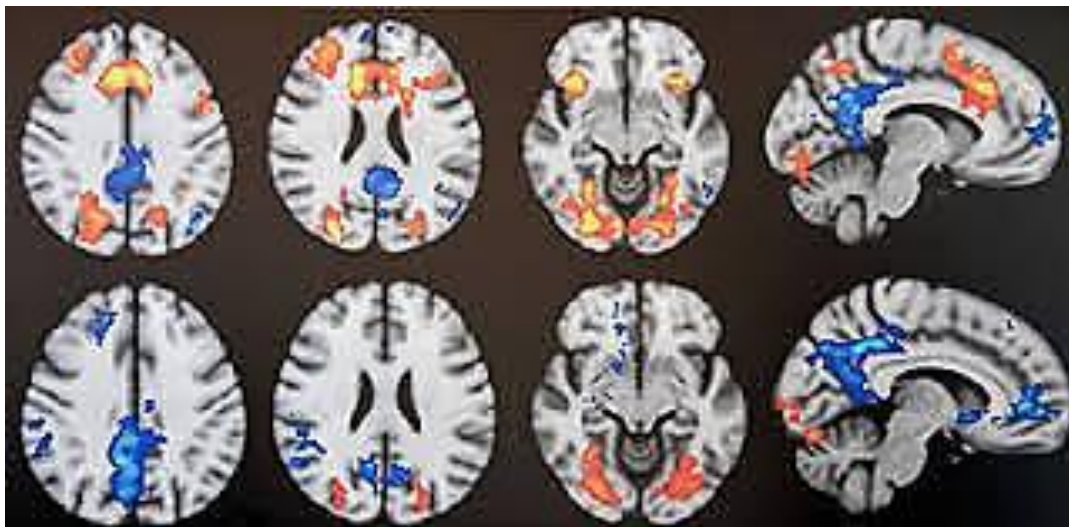


***Аңдатпа:** Мақалада адамның миына байланысты аурулардың диагнозын анықтау кезінде қазіргі уақытта мәселелерді шешуге белгіленген, мидың бейнелеуі, бағыттары мен нақты әдісін артықшылықтары мен кемшіліктері тізілген әдістердің даму перспективалары әдістерін зерттеу, моделдеу және деректерді өңдеу әдістері қарастырылған.*

Нейронды желілер – күрделі тәуелділіктерді жүргізуге арналған қуатты модельдеу әдісі. Нейронды желілер табиғатына сәйкес, әдетте, сызықты емес. Көптеген жылдар бойы сызықты модельдеу облыстардың көпшілігінде модельдеудің негізгі әдісі болды, себебі ол үшін оптимизация процедуралары жақсы құрастырылған. Сызықты аппроксимация қанағаттандырылмаған есептерде сызықты әдістер нашар жұмыс істейді. Сонымен бірге, нейронды желілер айнымалылар саны көп жағдайда сызықты тәуелділіктерді модельдеуге мүмкіндік бермейтін «өлшем қарғысына» қарсы тұра алады.

Соңғы жылдары біз көптеген облыстарда – бизнесте, медицинада, техникада, геологияда, физикада кеңінен қолданылып жүрген нейронды желілерге деген қызығушылықтың көбеюін байқап жүрміз. Нейронды желілер болжауды, классификацияны немесе басқаруды қажет ететін есептерді шешу керек тәжірибелердің барлығына ендірілді.

Нейронды желілер интуитивті тұрғыдан дұрыс болып табылады, себебі олар жүйке жүйелерінің примитивті биологиялық моделінде негізделген. Болашақта мұндай нейро-биологиялық модельдердің өсуі шынымен де ойлай алатын компьютерлердің пайда болуына алып келуі мүмкін[1].



**1-ші сурет- Функционалды магнитті-резонансты томография**

Өткен онжылдық ішінде нейронды желілер айтарлықтай алға жылжыды, алайда У. Маккаллоктң көптеген тұжырымдары бүгінге дейін өзекті болып қалды. Нейрон модельдерінің түрлілігіне қарамастан, У. Маккаллок және У. Питтспен ұсынылған олардың қозғалыс қағидалары өзгеріссіз қалды. У. Маккаллоктң моделінің кемшілігі нейронның өз моделі болып табылады – белсенді функцияның «шекаралық» түрі. Функцияның шекаралық түрі нейронды желіге оқыту және қойылған есепті баптау барысында жеткілікті илгіштікті бермейді[3].

Нейронды желілер теориясының дамуы америкалық нейрофизиолог Френсис Розенблаттың (Rosenblatt F.) жұмыстарында жалғасын тапты. 1958 жылы ол нейронды желінің моделін ұсынды. Ф. Розенблат У. Маккаллоктың және У. Питтстің моделіне оны оқытуға мүмкін ететін модификацияға байланыс қабілеттілігін енгізді. Бұл модель персептрон деп аталды. Алғашында персептрон процессорлы элементтің қатты шекаралы функциялы, бинарлы немесе көпбелгілі кірістері бар бір қабатты құрылымынан тұратын. Одан кейін бұл модель жеткілікті түрде дамытылды. Алғашында үлкен энтузиазммен қабылданған персептрон ірі ғылыми беделділер жағынан интенсивті шабуылға ұшырады[4].

1982 жылы америкалық биофизик Дж. Хопфилд (Hopfield J.) кейінірек оның есімімен аталған нейронды желінің негізгі нұсқасын ұсынды. Джн Хопфилдтің желісі қабаттар арасында кері байланысқа ие, бұл оның жалпылау қасиеттерін ұлғайтады. Мұндай желілер үлгілерді танып білуде кең қолданысқа ие болды. Келесі жылдары кері тарату, екі бағытты ассоциативті жады, өздігінен ұйымдастырылатын карталар және т.б. көптеген эффективті алгоритмдер мен нейрожелілер архитектуралары табылды. Кез келген нейронды желіні құру есептеудің үлкен көлемін қажет етеді (желіні оқыту, әдетте, итерационды процесс болып табылады). Сондықтан компьютерлердің есептеу қуаттылығының артуымен ғана нейрожелілерді тәжірибеде қолдануға мүмкіндік туды. Бұл нейрожелілік деректерді өңдеу принциптерін қолданатын бағдарламалардың кеңінен таралуына мүмкіндік берді.

Нейрожелілер мен нейромодельдеу пайда болып, өскен негізгі идеялар келесілер болып табылады:

1. Нейронды желі тірі ағзалардың жүйке жүйесінің құрылымы мен қасиеттерін имитациялайды: ол өзара байланысқан қарапайым бір типті есептеу элементтерінің (нейрондар) көп мөлшерінен тұрады және әр бөлек нейронның мүмкіндіктерімен салыстырғанда иілгіш және күрделі мінез-құлыққа ие.

2. Нейрожелі кірісте кіріс сигналдардың жиынын қабылдайды және оларға сәйкес қандай да бір есептің шешімі болып табылатын жауапты (шығыс сигналдарын) шығарады.

Әр нейрон кіші көлемді жұмыс орындайды – мысалы, оған келген кейбір салмақты коэффициенттері бар сигналдарды суммалайды және бұл кіріс деректерінің суммасын сызықты емес түрлендіреді. Басқа таралған нұсқасы кейбір есте сақталған эталоннан жоғары шығыс сигналын шығаратын нейрон-детектор болып табылады.

Жасанды нейронды желілер, биологиялық тәріздес, көптеген байланыстарға ие параллель функцияланатын қарапайым процессорлардың үлкен көлеміне ие есептеу жүйесі болып табылады. Мұндай желілерді құрастыру барысында әдетте биологиялық аналогтардан ерекшеленетін айтарлықтай қарапайымдылау қатары жасалғанымен, жасанды нейронды желілер таңқаларлық қасиеттер көлемін ұсынады.

Нейронды желілер өз мінезін оларды қоршайтын ортаның күйіне байланысты өзгерте алады. Кіріс сигналдарының анализінен кейін (талап етілетін шығыс сигналдарымен қоса) олар дұрыс реакцияны қамтамасыз ету үшін өздігінен бапталып, оқытылады. Оқытылған желі оған түрлі кедергілерге ие бейнені дұрыс «көруге» мүмкіндік беретін кіріс деректердің кейбір ақаулықтарына төзімді болуы мүмкін.

Өткен ғасырдың 50-ші жылдарында зерттеушілер тобы биологиялық және физиологиялық әдістерді біріктіре отырып, алғашқы жасанды нейронды желілерді құрастырды. Сол кезде жасанды интеллектке кілт табылды деп есептелінді. Бірақ бұл желілер жасанды көзқарас облысынан ауа-райын болжау және деректер анализі сияқты кейбір есептерді эффективті түрде шешкенімен, мұндай ой өзгертілді. Желілер өздері сәтті шығарып жатқан есептерге ұқсас есептерді шешуге қабілетсіз болды. Осы

уақыттан бастап интенсивті анализ периоды басталды. Теориялар құрылып, теоремалар қатары дәлелденді. Бірақ сол кездің өзінде күрделі математикалық әдістердің қолданылуынсыз ірі жетістіктерге жету қиын екені түсінікті болды.

70-жылдардан бастап ғылыми журналдарда жасанды нейронды желілерге қатысты мақалалар жариялана бастады. Біраз уақыттан кейін негізінде көптеген желілер құрылған жақсы теориялық фундамент қалыптасты. Соңғы екі онжылдықта құрастырылған теория іргелі есептердің шешімі үшін кең қолданыла бастады. Жасанды нейронды желілерді конструкциялау үшін іргелі бағдарламалық камтаманы құрастырумен айналысатын фирмалар пайда болды. Сонымен бірге, 90-шы жылдар жасанды нейронды желілердің бизнеске келуімен ерекше болды. Бұл облыста, өнімге сұранысты болжаудан банк клиенттерінің төлеу қабілеттіліктерін анализдеуге дейін, олар көптеген есептерді шешу барысында өздерінің эффективтілігін көрсетті.

Бүгінгі таңда түрлі есептерді шешуге бағытталған, көптеген функциялау қағидаларына ие нейронды желі конфигурацияларының үлкен көлемі бар.

Оқуға қабілеттілік мидың негізгі қасиеті болып табылады. Жасанды нейронды желілер үшін оқыту ретінде қойылған есептің эффективті шешімі үшін синаптикалық байланыстардың желі архитектурасын баптау болып есептелінеді. Әдетте нейронды желіні оқыту кейбір таңдалымда жүзеге асырылады. Қандай да бір алгоритм бойынша жүретін оқыту процесінің мөлшері бойынша желі кіріс сигналдарға жақсырақ (дұрысырақ) жауап қайтаруы тиіс.

Түрлі есептерді шешуге бағытталған оқыту алгоритмдерінің үлкен көлемі бар.

Көпөлшемді статистика қағидаларына негізделген математикалық модельдеу әдістері диагностика, болжау және тәжірибелі медицинада ішкі ауру түрлерінің терапиясы үшін қолданылуы қиын. Бұл мақсаттар үшін факторлы, кластерлі, регрессионды және дискриминантты анализдің қолданылуы жоғары деңгейдегі ғылыми қызығушылықты тудырады.

Тәжірибелік мақсатта дәрігерлер өз науқастарының диагнозын қоюы, болжауы және емдеуі үшін табылған математикалық модельдерді қолданбайды. Бұл модельдердің тәжірибеде қолданылуының қиындылығымен түсіндіріледі[5].

Нейрокомпьютерлердің немесе адам миының нейрондарының жұмысын имитациялайтын бағдарламалардың – нейронды желілердің медицинада қолданылуы дәрігерлердің жұмысын жеңілдетіп, жылдамдатады. Мұндай жүйелер мысалдарға негізделіп, өздігінен оқуға және есептердің белгілі бір кластарын шешуге қабілетті. Нейрокомпьютерлердің бұл қасиеті, әсіресе, көптеген есептер формализацияға ұшырамайтын, ал кіріс деректері жиі толық емес болатын медицинада өзекті.

Оқыту барысында нейронды желі медициналық есептерлі шешу «тәжірибесіне» ие болғандықтан, деректер базасы білім базасына модифицияланады. Тәжірибенің болуы дұрыс шешім мүмкіндігін мұндай жағдай бұрын кездеспесе де, алдын ала қарастырады. Яғни, дәрігердің жұмысы сияқты аналогия бақыланады.

Біздің жағдайымызда оқытушы таңдалым ретінде науқас аурулары жайлы 180 мәлімет қолданылды. Нейронды желінің тәжірибеде қолданылуы миокард инфаркті бар науқастардың рискстратификациясы барысында оның жоғары эффективтілігін көрсетті. 90% жағдайда диагнозды қоюда және осыған ұқсас науқастардың асқынуларын болжауда негізгі болып табылған 15 параметр анықталды.

Нейрожелілік мамандық жүйелердің артықшылықтары көрініп тұр. Дәрігерге (қолданушыға) кіріс деректері мен ақырлы шешім арасында ешқандай байланыс орнатудың қажеті жоқ. Дәрігер ақпараттың классикалық статистикалық түрлендірілуі мен математикалық модельдің тексерісіне уақыт кетірмейді. Нейрожелілер категориялық болып табылмайды. Нейрожелілер шешімді дәрігерге алынған нәтижені

(диагнозды) түрлендіру мүмкіндігінде сенімділік деңгейімен шығарады. Ол шешімді өте жылдам (секунд ішінде) береді. Бұл нейрожеліні диалогтық режимде түрлі динамикалық жүйелерде қолдануға мүмкіндік береді. Нейронды бағдарламалардың мұндай ерекшеліктерінің барлығы миокард инфаркті бар науқастардың рискстратификациясы үшін ортақ медициналық мамандандырылған жүйені құруды жеңілдетуге мүмкіндік береді.

**Қолданылған әдебиеттер:**

1. Боровикова В.П. Нейрондық желілер. STATISTIC A NeuralNetworks: Қазіргі заманғы мәліметтерді сараптау технологиясымен Методологиясы методологиясы/ Ред. 2-ші шығ., 2008. 392 б.
2. Галушин А.И. Нейронные сети: основы теории. М.: Горячая линия-Телеком, 2010. 496 б.
3. Патрик Э. Основы теории распознавания образов / ағл.тілінен аударылған.; ред. Б. Р. Левина. М.: Сов.радио, 1980. 408 б.
4. РозенблаттФ. Аналитические методы изучения нейронных сетей // Шет ел радиоэлектроникасы. 1965. №5. б. 40-45.
5. Хайкин С. Нейронные сети: толық курс 2-ші шығарылым: ағл.тілінен ауд. М.: Вильяме, 2006. 1104 б.

## **СӨЙЛЕУДІ ТАНУДА MFCC ЖӘНЕ DTW АЛГОРИТМДЕРІНІҢ ЕРЕКШЕЛІГІ**

**О.Ж. Мамырбаев<sup>1</sup>, Н.О. Мекебаев<sup>2</sup>, М. Тұрдалыұлы<sup>2</sup>**

*<sup>1</sup>PhD, қау.профессор, ҚР БҒМ ҒК Ақпараттық және есептеуіш  
технологиялар институты, Алматы қ., Қазақстан*

*<sup>2</sup>PhD докторант, Әл-Фараби атындағы Қазақ ұлттық университеті  
Алматы қ., Қазақстан*

*<sup>1</sup>morkenj@mail.ru, <sup>2</sup>nurbapa@mail.ru, <sup>3</sup>mkt\_001@mail.ru*

**Аңдатпа.** Сөйлеуді тану қауіпсіздік жүйелерінде, денсаулық сақтауда, телефонияда әскери қызмет көрсетуде және мүгедектерге арналған жабдықта кең ауқымда қолданылады. Сөйлеу әрдайым әр түрлі сигнал болып табылады. Сонымен, автоматты түрде сөйлеуді тану жүйесі үшін цифрлық өңдеудің тиісті алгоритмі таңдалуы керек. Сөйлеу үлгісінен талап етілетін ақпаратты алу үшін оның ерекшеліктерін алу керек. Сөйлеуді тану мақсатында бұл шешімдерді қабылдау үшін талданады. Осы мақалада MATLAB ортасында сөйлеуді тану жүйесін енгізу түсіндіріледі. Mel-Frequency Cepstral коэффициенттері (MFCC) және динамикалық уақытты орау (DTW) - бұл функцияның экстракциясы мен үлгіні сәйкестендіру үшін бейімделген екі алгоритм қарастырылады. Нәтижелер бір уақытта және үздіксіз тестілеу кезеңдерінде жүргізіледі.

**Түйінді сөздер:** Сөйлеуді тану, дауыс, Фурье қатары, MFCC, DTW, FIR, FFT;

## 1. Кіріспе

Бүгінде сөйлеуді тану жүйесі тану формаларын тану принциптеріне негізделген. Бүгінгі күні сөйлеуді тану міндеті - бұл өзекті мәселе. Сөйлеуді тану жылдамдығы белгісіз үлгі мен дерекқор үлгісінің арасындағы сәйкес келетін дәлдікке байланысты. Үлгі сәйкестігі әдісі сөйлеуді танудың соңғы нәтижесін анықтайды.

Сөйлеуді тану - автоматты сөйлеу сигнал ақпарат негізінде адамды айтқан сөздерді тану процесі. Сол сөздің бірнеше ұзақтылығы, әр түрлі жиіліктерде болуы мүмкін және сол ұзындығы бірдей сөздер байланысты, сөздер әр түрлі бөліктері әр түрлі қарқыны айқын фактісі ортасында түрліше болады. Сөйлеу үлгілерін (векторлардың дәйектілігі ретінде ұсынылған) арасындағы қашықтықты алу үшін уақытты туралау керек. Уақыт динамикалық теңестіру тұжырымдамасы салыстыру спектрлік сөйлеу реттілігі байланысты проблемаларды шешу үшін пайдаланылды.

Сөйлеуді тану - ауызша сөзді компьютерлік бағдарламаға енгізу ретінде қабылдау. Адамдар айтқан дыбыстар, сөздер немесе сөз тіркестерінің электр сигналдарына айналатын технологиясы және бұл сигналдар мән берілген кодтау үлгісіне айналады. Сөйлеуді тану - сөйлеу сигналын орфографиялық көріністе айналдыру міндеті. Сөйлеуді тану жүйесі сөйлеу түріне қарай жіктелуі мүмкін:

- ✓ Сөйлеушіге тәуелді: пайдаланушыны жүйені дауысына сәйкес оқытуды талап ететін жүйелер.

- ✓ Сөйлеушіге тәуелді емес: пайдаланушыны жүйені оқытуды талап етпейтін жүйелер, яғни олар кез-келген сөйлеушімен жұмыс істеуге арналған.

- ✓ Оқшауланған сөз тану: жүйе бір уақытта бір сөзді қабылдайды, олар әрдайым үздіксіз сөйлеуге мүмкіндік береді.

- ✓ Қосылған сөздік жүйесі: сіз қысқа кідіріспен, яғни жоспарланған сөзбен, әр сөзді баяу және анық айтуға болады.

- ✓ Үздіксіз сөйлеу: пайдаланушыға табиғи түрде сөйлеуге мүмкіндік береді, сонымен қатар компьютерлік диктация деп аталады.

## 2. Сөйлеуді танудың алгоритмдері

Бүгінде сөйлеуді тану жүйесі тану формаларын тану принциптеріне негізделген. Сөйлеуді тану әдістерінің классификациясы:

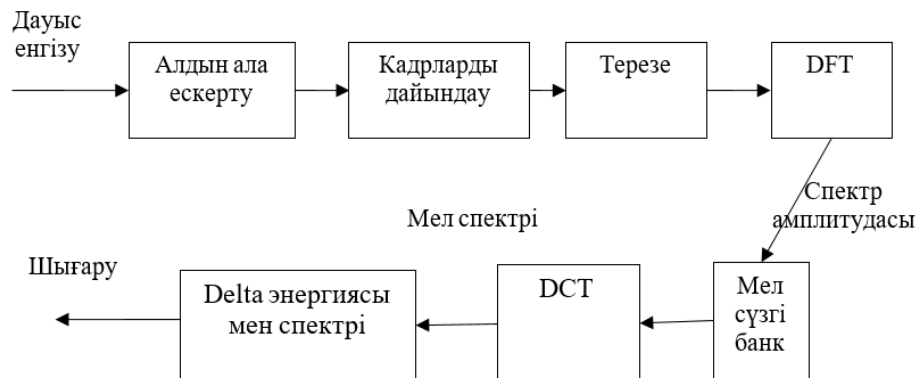
1. Сөйлеу сигналдарының ақпараттық белгілерін анықтау үшін MFCC (Мел-жиіліктегі кепстралды коэффициенттер) алгоритмдері қолданылады.

2. Динамикалық программалау - уақытша динамикалық алгоритмдер (Dynamic Time Warping). Контекстке тәуелді жіктеуде ол жүзеге асырылғанда, жеке лексикалық элементтер, фонемалар мен аллофондар сөйлеу ағынынан таңдап алынады, олардан кейін слог және морфемалармен біріктіріледі.

MFCC (Мел-жиіліктегі кепстралды коэффициенттер) кодтарын DTW алгоритмімен бірге пайдалану сөйлеуді тану технологиясында салыстырмалы түрде жетілдірілген және тану дәлдігі де жоғары болады. Бірақ адамдар сөйлеуді тану технологиясын дамытуға және сөйлеуді тану технологиясының ауқымын кеңейтуге барынша жоғары талаптар қойып отырады. Dynamic Time Warping (DTW) әдісінде қолданылатын тән үлгі ретінде MEL-Frequency Cepstral (MFCC) коэффициенттері қолданылады. Үлгі жасау кезінде және сөйлеу сигналының алынған сипаттамаларын сәйкестендіру процесінде алгоритм жақсарды. Осылайша, DTW алгоритмі оның тиімділігі мен дәлдігін әлде қайда жақсарта алады. Сөйлеуді тану үшін тікелей қолданылатын классикалық Мел-жиіліктегі кепстралды коэффициенттер жиілік әдісімен салыстырғанда бұл әдіс ауызша эксперименттерді тану жылдамдығын арттырады.

## 2.2 MFCC (Мел-жиіліктегі кепстралды коэффициенттер) алгоритмдері

MFCC (Мел-жиіліктегі кепстралды коэффициенттер) - тану әдісінде қолданылатын спектрлік негізделген ең танымал параметр. Мультифункционалды алгоритмді іске асыруда азғантай күрделіліктің артықшылығына байланысты, MFCC - тің Мел шкаласының жиіліктеріне сәйкес келетін белгілі коэффициенттері деректер базасындағы ауызша сөз үлгілерінен алынған.



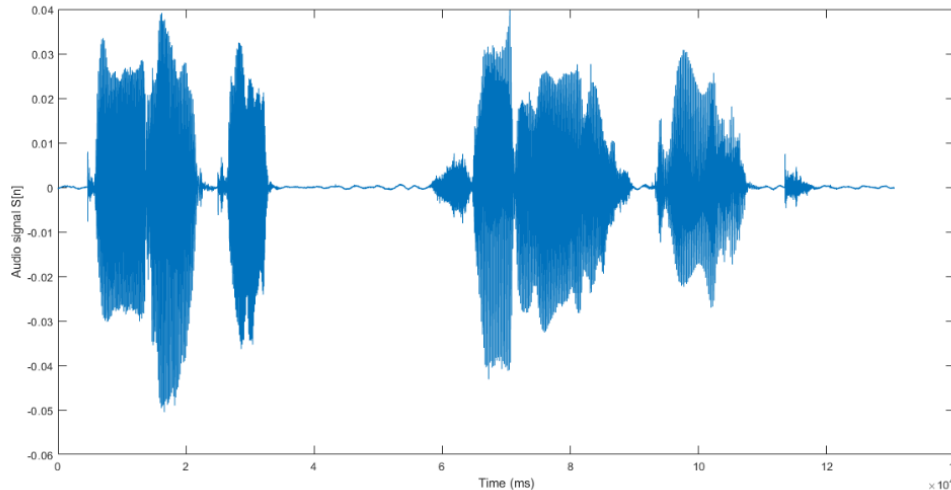
Сурет 1. MFCC блок схемасы

MFCC (Мел-жиіліктегі кепстралды коэффициенттер) адамның есту қабылдауына негізделген, ол жиіліктерді 1КГц-ден артық қабылдай алмайды. MFCC алгоритмімен алынған ерекшеліктер адамның құлағының ішіндегі иірім жиілігімен сыни өткізу қабілетін анықтайтын жолағында белгілі өзгерістерге ұқсас. Берілгендерді шығарып алу үрдісіндегі MFCC үшін дауыстық үлгі 2 суретте көрсетілген.



Сурет 2. MFCC ағынының диаграммасы

Функцияны шығармас бұрын, дауыс үлгісі аналогты цифрлық түрлендіруге, содан кейін алдын-ала басу және сүзу арқылы өтеді. Тегістеуді болдырмау үшін жеткілікті іріктеу жылдамдығына ие болу маңызды. Nyquist үлгісінің сынамалары бойынша теоремаға сәйкес, ең жоғарғы жиіліктегі  $f$  сигналының ең төменгі жиілігінің жиілігі  $2$  фГц болуы керек. 3-суретте «Төре көп, Сары Нияз біреу-ақ» сөзі үшін сандық дауыстың дауыс үлгісі көрсетілген.



**Сурет 3. «Төре көп, Сары Нияз біреу-ақ» дауысты дыбысы бар сөзі үшін сандық дауыс беру үлгісі**

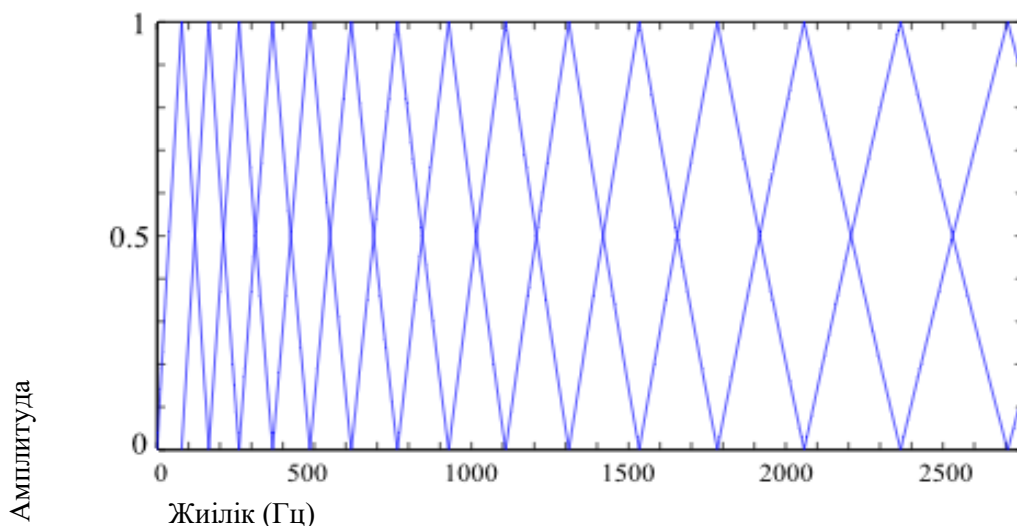
Алдын ала басымдық кезеңі төменгі жиіліктерге қатысты жоғары жиіліктің шамасын арттырады. Осы мақсатта пайдаланылатын FIR сүзгісі және оның тиісті дискретті шығарылуы тиісінше 1 теңдеуде және 2 теңдеуде берілген.

$$F(x) = 1 - kx^{-1} \quad 0 < k < 1 \quad (1)$$

$$y[m] = s[m] - k \cdot s[m - 1] \quad 0 < k < 1 \quad (2)$$

Мұнда  $y[m]$  - шығу және  $s[m]$  FIR фильтрінің сигнал кірісі.

Фурье түрлендірудің жылдам форматы (FFT) жиіліктік доменге сигнал түрлендіретін әр кадрға қолданылады. Біз әдетте 613 нүкте FFT жасап, алғашқы 258-ді коэффициенттерін сақтаймыз. Осылайша әр кадрға арналған спектр алынады. Бірақ ерекшеліктерді сәйкестендіру кезеңі үшін қажет емес көптеген ақпарат алынады. Мүмкіндіктің сәйкестігі алгоритмі екі аралық жиіліктің айырмашылығын анықтай алмайды. Осы себепті спектральды контейнерлердің топтамаларын аламыз және оларды әртүрлі жиілік аймақтарында қанша энергия бар екенін түсіну үшін жинақтаймыз. Бұл әрбір кадрды MEL сүзгі банктерімен көбейту арқылы жүзеге асырылуы мүмкін 4-суретте көрсетілген.



Сурет 4. Мел аралық сүзгі банктер

### 2.3 Дауысты тану кезінде DTW алгоритмін пайдалану

DTW алгоритмін қолданатын сөздерді анықтау

Сөзді анықтау сандық пішіндерді салыстыру арқылы жасалуы мүмкін немесе сигналдар спектрограммасын салыстыру арқылы. Салыстыру үрдісі екі жағдайда да өтеуі тиіс әр түрлі ұзындықтағы реттілігі және бейсызық сипат дыбыс арқылы. DWT алгоритмі осы мәселелерді табу арқылы шешеді, екі жол арасындағы оңтайлы қашықтыққа сәйкес келетін деформация әртүрлі ұзындықтар болуы мүмкін.

Алгоритмнің 2 ерекшелігі бар:

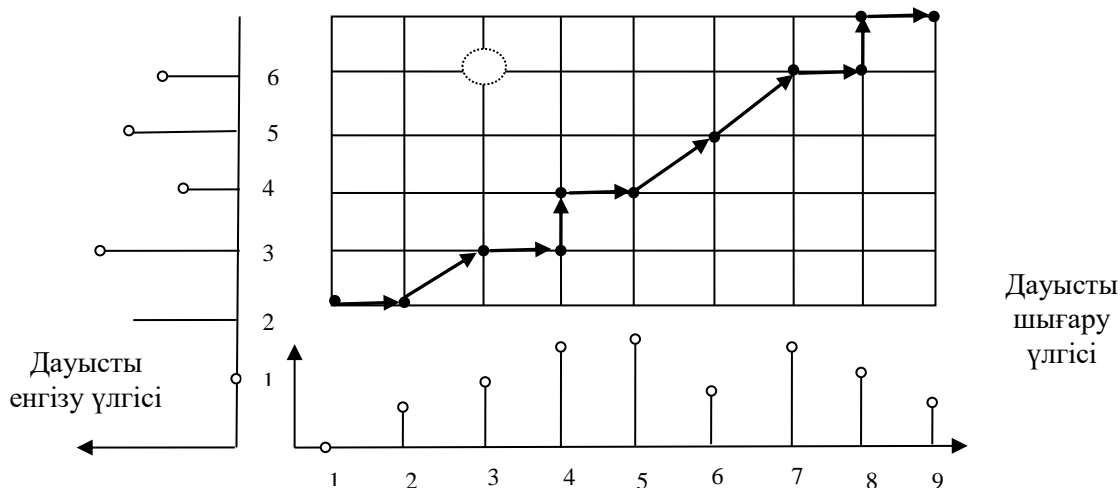
1. Сандық толқындардың тікелей салыстыру. Бұл жағдайда әр сандық жүйелілік үшін жаңа ретін құрайды, олардың өлшемдері айтарлықтай аз. Сандық тізбектегі бірнеше мың сандық мәндер болуы мүмкін, ал кейіннен бірнеше жүз мән болуы мүмкін. Сандық мәндердің санын азайту, бұрыштық нүктелердің арасында оларды жою арқылы жүзеге асады. Бұл сандық жүйенің ұзындығын қысқарту үдерісі оның үлгісін өзгертпеуі керек. Әрине, процесс тану дәлдігін төмендетеді. Алайда, жылдамдықтың, дәлдіктің, сөздіктегі сөздердің көбеюіне байланысты көбейді.

2. Спектрограммалардың сигналдарын ұсыну және екі спектрограммды салыстыру үшін DTW алгоритмін қолдану. Әдіс сандық сигналды қабаттасатын бірнеше аралықтарға бөлуден тұрады. Әрбір импульс үшін нақты сандардың интервалдары (дыбыс жиілігі) жылдам Фурье түрлендіруін есептейді және дыбыс спектрограммасының матрицасында сақталады. Параметрлер барлық есептеу операциялары үшін бірдей болады: импульстік ұзындығы, Фурье түрлендіруіндегі екі қатардағы импульстардың ұзындығы. Фурье түрлендіруі симметриялы түрде орталыққа байланысты, ал бір жағынан күрделі сандар екінші жағынан сандармен байланысты.

Осыған байланысты, симметрияның бірінші бөлігінің мәндері ғана сақталуы мүмкін, осылайша, спектрограмма күрделі сандардың матрицасын, мұндай матрицаның жолдарының саны трансформацияның жартысына тең, Фурье және бағандардың саны дыбыс ұзақтығына байланысты анықталады. Спектрограмма DTW ұштасқаннан кейін нақты сандар матрицасына қолданылатын болады, мұндай матрица энергия матрицасы деп аталады.



Бұл кезеңде алдыңғы қадамда есептелген сөздердің ерекшеліктері анықтамалық үлгілермен салыстырылады. DTW алгоритмі сөздердің және анықтамалық үлгілердің ерекшеліктері арасындағы ең аз қашықтықты есептеу үшін енгізіледі. Әр үлгімен есептелген баллдар арасында ең аз мәнге сәйкес келетін сөз анықталады. DTW екі рет сериялы арасындағы оңтайлы теңдеулерді табады, егер бір уақыттық серия оны уақыт осі бойымен созу немесе азайту арқылы сызықты емес «соғылған» болуы мүмкін. Екі уақыт аралығы арасындағы сәйкестіктің ұзақтығы қашықтық факторы бойынша өлшенеді. Екі дыбыс үлгілері үшін динамикалық уақытты орау 5-суретте көрсетілген.



Сурет 5. Екі дауысты динамикалық уақытты белгілеу үлгісі

$(i, j)$  бұл аралықтағы  $d(a_i, b_j)$  нүктелер арасындағы  $a_i, b_j$  қашықта  $n$  және  $m$  тізбегінің матрицасы құрылады.

$$D(i, j) = \min[D(i - 1, j - 1), D(i - 1, j), D(i, j - 1)] + d(i, j) \quad (3)$$

Ең аз қашықтыққа сәйкес келетін үлгі - табылған сөз болады.

### Қорытынды

MFCC және DTW көмегімен оқшауланған сөз табу жүйесі MATLAB ортасында жасалады. Жүйе бес бөлек сөздің үлгілерін сақтау арқылы оқытылады. Нәтижелер көрсеткендей, оқу үлгеріміндегі әрбір сөз үшін он шақты үлгіні сақтау бес үлгімен салыстырғанда жақсы нәтижелер береді. Оқшауланған сөздерді табудың тиімділігі екі буын сөзі үшін бір буын сөзімен салыстырғанда 100% құрайды. Жоғарыда келтірілген нәтижелерден бірдей сөздер арасындағы DTW қашықтығы 100-ден аз, ал әртүрлі сөздер арасында 290-нан асатынына көз жеткізе аламыз. Сондықтан, 150-ден шектік мәнді белгілеу пайдаланушының сөзін басқа сөзден оңай сүзуге болады, олардың үлгілері оқу кезеңінде сақталған. Мақала ЖТН-АР05131207 «Терең нейрондық желілерді пайдаланатын мультитілдік автоматты сөйлеуді тану технологиясын құру» жоба негізінде дайындалады.

### Әдебиеттер тізімі

1. Фланаган Дж.Л. Анализ, синтез и восприятие речи / пер. с англ. А. А. Пирогова. М.: Связь, 1968.
2. Кузнецов В., Отт А. Автоматический синтез речи. Таллинн: Валгус, 1989.
3. М.А.Муқет, “Speech Recognition using Digital Signal Processor”, unpublished.
4. A. Bala, Abhijit kumar, Nidhika Birla, “Voice Command Recognition System Based On MFCC And DTW”, International Journal of Engineering Science and Technology, Vol. 2, No. 12, 2010.
5. M.R. Hasan, M. Jamil, M.G. Rabbani and M.S. Rahman, “Speaker Identification Using MEL Frequency Cepstral Coefficient”, Proceedings of 3rd International conference on Electrical and Computer Engineering (ICECE), December,28-30, 2004, Dhaka, Bangladesh.

## АРАҚАШЫҚТЫҚ МЕТРИКАСЫ АУЫСПАЛЫ БОЛЫП КЕЛЕТІН АЛГОРИТМДЕР

Л.Ш. Черикбаева<sup>1</sup>, Б. Калыбек уулы<sup>2</sup>

<sup>1</sup>әл-Фарабиатындағы Қазақ Ұлттық Университеті,  
Алматы қаласы, Қазақстан Республикасы,

<sup>2</sup>Ақпараттық технологиялар және автоматтандыру институты ҚР ҰҒА,  
Бішкек, Қырғызстан  
e-mail: lyailya\_sh@mail.ru

*Аңдатпа.* Бұл жұмыста объектілер арасындағы арақашықтықты түрлі метрикалар арқылы есептеу нәтижелері өзгешеленетін, K-means базалық алгоритмдер негізіндегі кластерлеу нәтижелер тобы (ансамблі) қарастырылған. Сонымен қатар арақашықтық метрикалары қарастырылып, ансамбль жұмысы бағаланған.

Кейінгі жылдары ақпараттарды өңдеу – деректерді интеллектуалды түрде талдау (Data Mining) бағытын қарқынды түрде зерттеу жүргізілуде және де осы бағытқа қызығушылық артып келеді. Деректерді талдаудың классикалық тәсілдерінен айырмашылығы, бұл жерде адамның мінез-құлқын модельдеуге, жалпылама интеллектуалдық мәселелерді шешуге, үлгілерді анықтауға басты назар аударылады. Кластерлерді талдаудың негізгі мақсаты – топ ішінде бір-бірімен ұқсас болып келетін және басқа топтардан айырмашылықтары бар объектілер топтарының салыстырмалы түрде аз бөлігін бөліпалу. Талдаудың бұл түрі классификациялау есептерін шешуде ақпараттық жүйелерде, сонымен қатар деректер қорымен жұмыс жасау, интернет-құжаттарды талдау, кескіндерді сегментациялау және т.б. деректердегі заңдылықтарды анықтауда қолданылады [1]. Қазіргі уақытта әлеуметтік, корпоративтік, ғылыми және басқа да ақпараттық орталарда үлкен көлемді деректер көптеп жинақталуда. Өртүрлі аймақтарда үлкен деректердің қарқынды түрде қолданылуы, зерттеушілердің үлкен көлемді ақпараттарды талдау және өңдеу құрылғылары мен әдістерін дамытуға аса жоғары қызығушылығын арттырып отыр. Деректерді талдауда болашағы бар кемел бағыттардың бірі кластрелік талдау болып табылады [2].

Үлкен көлемді деректер (BigData) веб-ресурстардан ақпараттарды алу, деректерді қарқынды пайдаланатын аудандардағы жалпы үлгілерді анықтау және т.б. осы сияқты әртүрлі процестерде қолданылады. Мұндай деректерге жіктеу, құрылымдау, мұқият талдау жасау қажет. Бұл жағдайда кластерлік талдау көптеген ғылыми зерттеулердің негізі болып табылады. Кластерлеу – бұл белгілі бір қасиеттері бар тәуелсіз бірлік ретінде қарастырылатын біртекті элементтердің белгілі бір біріктірілуін бөліп алу арқылы сегменттеу. Кластерлеу процедурасы нәтижесінде «кластерлер» пайда болады, яғни бір-біріне өте ұқсас топтар пайда болады.

Кластерлеу сапасының критеріі, әдетте, топ ішіндегі объектілердің шашырап орналасуына және олардың арасында қашықтыққа байланысты белгілі бір функционалдылық деп қарастырылады [3].

Алгоритмдер ансамблінің (тобының) қалыптасуына және ансамблге қатысушылар пікірлері негізінде олардың көмегімен топтық шешімдерді құрумен байланысты кластерлік мәселелердегі шешімдердің орнықтылығы жақсаруы мүмкін. Бұл жұмыста үлкен деректерді талдау үшін әртүрлі қашықтық метрикасына негізделген кластерлік алгоритмдер ансамблінің құрылысына ерекше көңіл бөлінеді.

Ансамблдік тәсіл кластерлеу сапасын жақсартуға мүмкіндік береді. Кластерлік талдаудың топтық шешімдерін құру тәсілінің бірнеше негізгі бағыты бар: консенсустік үлестіруге, коассоциативті матрицаға негізделген, теоретикалық-графтық әдіс, орташа жұпты қайырымашылықтар матрицасы және т.б. Кластерлеу әдісін таңдау деректердің көлеміне және бір уақытта деректердің бірнеше түрін өңдеуге және талдауға тәуелді болады [7]. Практикада кластерге бөлу K-means әдісімен жүргізілетін, гибридті тәсілдер жиі пайдаланылады. Бұл әмбебап, танымал әдістердің бірі.

$$V = \sum_i^k \sum_{x_j \in S_i} (X_j - \mu_i)^2, \quad (1)$$

мұндағы,  $k$  – кластерлер саны,  $S_i$  – алынған кластерлер,  $i=1,2,\dots, k$  және  $\mu_i$  – векторлар ортасы. Алгоритмдерді салыстыру туралы деректер 1-кестеде көрсетілген.

### 1 – кесте. Алгоритмдерді салыстыру кестесі

Кластерлеу алгоритмі	Кіріс деректері	Нәтиже
Иерархиялық	Иерархияны қысқарту үшін кластерлердің саны немесе қашықтық шегі	Кластерлердің бинарлы ағашы
K-means	Кластерлер саны	Кластерлер орталығы
C-means	Кластерлер саны, айқындық дәрежесі	Кластерлер орталығы, тәуелділік матрицасы
Байланысқан компоненттерін ерекшелеу	R арақашықтық шегі	Кластерлердің ағаш тәріздес құрылымы

Кластерлік талдауда объектілер арасындағы арақашықтықтарды анықтау үшін  $x$  және  $x'$  объектілері арасындағы әртүрлі арақашықтық метрикасы қолданылады. Кластерлік талдауда келесі метрикалар ең танымал болып саналады:

1. Евклид арақашықтығы:  $p(x, x') = \sqrt{\sum_i^n (x_i - x'_i)^2}$  (2)

2. Манхеттен арақашықтығы:  $p(x, x') = \sum_i^n |x_i - x'_i|$  (3)

3. Чебышев арақашықтығы:  $p(x, x') = \max(|x_i - x'_i|)$  (4)

Метрику дұрыс таңдау кластерлеу сапасына тікелей әсер етеді. Қазіргі уақытта кластерлік талдауда топтық әдістер көптеп қолдануда. Осыған дейін айтып кеткендей, кластерлік талдау алгоритмдері әмбебап емес, әр алгоритмнің өз қолдану аясы бар (1-кесте). Алгоритмдердің топтық тәсілі бастапқы алгоритмдердің таңдап алынған параметрлерінен соңғы шешім тәуелділігін төмендетеді және нақты шешім алуға мүмкіндік береді [4]. Топтық алгоритмдер алудың келесі негізгі әдістері бар (1-сурет) [5].

1. Консенустық бөлуді табу, яғни қандайда бір тиімді критеріі бойынша, бірнеше алынған шешімдердің келісе отырып, кластерлерге бөлуі;
2. Ұқсастықтар/айырмашылықтардың келісім матрицасын (co-occurrence matrix) есептеу.



1 – сурет. Кластерлеудің алгоритмдер ансамблі

Алгоритмдер ансамблінің қолданылу мысалы ретінде параметрлік емес MeanSC алгоритмі негізінде құрылған ансамблді алуға болады. Бұл көп арналы бейнелерді кластерлеу көрсеткішін жақсартуға мүмкіндік береді [6]. Сонымен қатар, K-means және SVM (Support Vector Machines) алгоритмдері негізіндегі кластерлеу алгоритмдерінің ансамблін қолдану диагностика уақытын қысқартуға, жүрек ақауларын анықтаудың дәлдігін жақсартуға мүмкіндік берді [8]. Осылайша, әртүрлі алгоритмдер жиынтығы мен ансамблді олардың артықшылықтары мен ерекшеліктеріне сәйкес қолдану белгілі бір нақты аймақ үшін ең қолайлы кластерлеу схемасын жасауға мүмкіндік береді. Осыдан бұрын да кластерлеудің нәтижесіне әсер ететін маңызды фактор - бұл объектілер арасындағы қашықтықтардың белгілі бір метрикасын таңдау болып табылатындығын айтып өткенбіз. Осы екі тәсілдерді біріктіру арқылы кластерлік талдаудың тиімділігін едәуір жақсартуға болады [16].

Кластерлеу алгоритмдерінің тобы ретінде K-means алгоритмін қарастырайық. K-means алгоритмі бірінші қадамда арақашықтық метрикасын пайдалана отырып, деректерді кластерлерге бөледі. Мұнда келесі метрикалар қолданылды: Евклид арақашықтығы, Манхэттен арақашықтығы, Чебышев арақашықтығы.

Содан соң (5) – формуласы бойынша алгоритмдердің топтық пікірлерінің салмағы мен дәлдігі есептелінеді:

$$\omega_l = \frac{Acc_l}{\sum_{l=1}^L Acc_l}, \quad (5)$$

мұндағы,  $A_{CC_i}$  - l алгоритмінің дәлдігі, ал L-алгоритмдер тобындағы алгоритмдер саны. Әрбір алынған кластерлік бөлінулер үшін бинарлы  $n \times n$  өлшемді матрица құрылады, мұндағы  $n$ -объектілер саны. Кластерлерге бөлудің жақсы нәтижесін алу үшін, ансамбльде L бөлінуіне ұқсастық/айырмашылықтың бинарлы матрицасын құру:

$$H_i = \{h_i(i, j)\}, \quad (6)$$

мұндағы  $i$  және  $j$  элементтері бір кластерге түссе, онда  $h_i(i, j)$  нөлге тең, кері жағдайда 1-ге тең.

Содан соң әр элементі мүмкін болатын матрица элементтерінің салмақталған қосындысын беретін айырмашылықтар матрицасы есептеледі. Бинарлық бөлінулердің келісілген матрицасын құру үшін:

$$H' = \{h'(i, j)\}, \quad (7)$$

$$h'(i, j) = \sum_{i=1}^L w_l h_l(i, j), \quad (8)$$

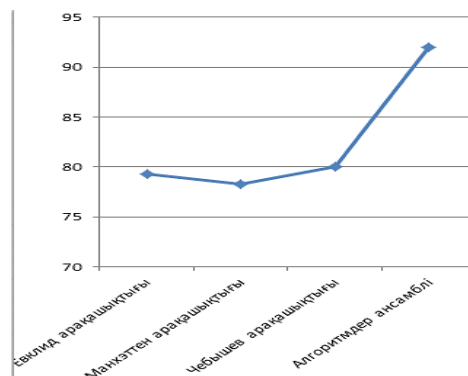
мұндағы  $w_l$ -алгоритм салмағы.

Келісілген матрица бойынша бөлу нәтижесін жақсарту мақсатында жақын көрші алгоритмі қолданылды. Алынған матрица иерархиялық агломерациялық кластерлеу алгоритмі үшін кіріс деректері ретінде қолданылады. Осыдан кейін қарапайым тәсілдер көмегімен, қолайлы кластерлік шешімді таңдап алуға болады. Топтық алгоритмді құру процедурасы 2-суретте көрсетілген:



2 – сурет. Кластерлеу алгоритмдерінің тобы

Кестеден Чебышев арақашықтығын қолданған алгоритмнің салмағы жоғары екендігін көреміз.



3– сурет. Алгоритмдердің дәлдігін салыстыру

3-суреттен әртүрлі метрикалы стандартты K-means алгоритмінің дәлдігі берілген ансамбль дәлдігінен төмен екендігін көруге болады. Кластерлік талдау деректерді өңдеудің негізгі бағыттарының бірі болып табылады, бұл қазіргі уақытта бар тиімді әдістердің бірі. Кластерлеу алгоритмдерінің ансамблін пайдалану арқылы деректерді топтарға бөлудің сенімділігін арттырады [19-20]. Бұл әдіс әр түрлі салаларда қолданылуы мүмкін. Мақалада қарастырылған кластерлік алгоритмдер ансамблі K-means алгоритмдері үшін арақашықтық метрикасының кемшіліктерін түзете отырып, кластерлерге бөлудің сенімділігін арттырады.

#### **Әдебиеттер**

1. В.Б. Бериков, Г.С. Лбов. Современные тенденции в кластерном анализе. [http://biocomparison.ucoz.ru/\\_ld/0/49\\_berikov\\_lbov.pdf](http://biocomparison.ucoz.ru/_ld/0/49_berikov_lbov.pdf)
2. П.В. Бочкарёв. Разработка ансамбля алгоритмов кластеризации на основе изменяющихся метрик расстояний.
3. В.Б. Бериков. Коллектив алгоритмов с весами в кластерном анализе разнородных данных. // Вестник Томского Государственного Университета. 2013. № 2(23). Стр. 22-31.
4. Berikov V.B. Weighted ensemble of algorithms for complex data clustering // Pattern Recognition Letters. 2014. Vol. 38. P. 99-106.
5. В.Б. Бериков. Классификация данных с применением коллектива алгоритмов кластерного анализа // Знания-ОнтологииТеории (ЗОНТ-2015), 2015, С. 29-38
6. И.А. Пестунов, В. Б. Бериков, Ю. Н. Синявский. Сегментация многоспектральных изображений на основе ансамбля не параметрических алгоритмов кластеризации // Вестник СибГАУ, 2010, №5(31), С. 56-64.

## **МОДУЛЬ БОЙЫНША ДӘРЕЖЕГЕ ШЫҒАРУ ОПЕРАЦИЯСЫН ҚОЛДАНУ АРҚЫЛЫ АҚПАРАТТЫ ШИФРЛАУ ӘДІСІ**

**Хомпыш А.**

*ҚР БжҒМ ҒК «Ақпараттық және есептеуіш технологиялар институты»  
e-mail: ardabek@mail.ru*

***Аңдатпа.** Мақалада Галуа өрісіндегі модуль бойынша дәрежеге шығару операциясын қолдана отырып ақпаратты криптографиялық қорғаудың шифрлеу алгоритмі ұсынылған.*

Ақпараттық технологияларды дамытудың қазіргі кезеңінде қорғау функциясы ақпараттық өңдеу кешендерінің ажырамас бөлігі. Ақпаратты қорғау проблемасын шешудің ең тиімді әдістерінің бірі криптографиялық қорғау жүйелерін пайдалану болып табылады. Қазіргі есептердің бірі үлкен көлемдегі ақпаратты шифрлеуде жоғары жылдамдықты қамтамасыз ететін шифрлеу алгоритмдері мен әдістерін құру. Мұндай есептерді шешу үшін түрлі математикалық приметивтер қолданылады. Соның бірі кеңейтілген Галуа  $GF(p^r)$  өрісі [1-3].

Галуа өрісіндегі аддитивті және мультипликативтік операциялар модуль бойынша (қосу, көбейту, дәрежеге шығару) және олардың әр түрлі комбинациясын

қолдану ақпаратты қорғау деңгейін жоғарлатуға мүмкіндік береді. Ақырлы Галуа өрісіндегі жұмыс істейтін модульді арифметика кең қолданысқа ие екендігін айта кеткен жөн. Бұл,  $GF(p^v)$  Галуа өрісінің характеристикасының модуль бойынша дискретті логарифмді және түбірді есептеу аппараттық және уақыт шығындарын қажет ететіндігіне байланысты. Бұл мақалада кеңейтілген Галуа  $GF(p^v)$  өрісінде модуль бойынша дәрежеге шығару операциясын қолдану арқылы ақпараттарды шифрлау әдісі ұсынылған [4].

Мұндай криптографиялық ақпараттарды қорғау жүйесінде осы өрістеги мультипликативті топтағы барлық элементтерді тудыратын  $\pi(z)$  келтірілмейтін көпмүшелік таңдап алынады [5]. Сондықтан мұндай тудырушы көпмүшеліктің дәрежесі  $v$ -ға тең.

$$\text{ord}\pi(z) = v$$

Мұндай жағыдайда екілік жүйеде берілген ашықмәтін  $v$  блоктарға бөлінеді және әрбір мұндай екілік жүйедегі блоктар формулаға сәйкес көпмүшеліктер түрінде жазылады:

$$A(z) = a_{v-1}z^{v-1} + a_{v-2}z^{v-2} + a_{v-3}z^{v-3} + \dots + a_2z^2 + a_1z^1 + a_0z^0, \quad (1)$$

мұндағы  $a_{v-1}, a_{v-2}, \dots, a_2, a_1, a_0$  —  $A$  блогының екілік жүйедегі мәні, яғни  $a_i \in \{0,1\}$ ,  $i = 0, 1, \dots, v-1$  [2].

Мұндай  $A(z)$  көпмүшелігінің дәрежесінің мәні  $\pi(z)$  тудырушы көпмүшеліктің дәрежесінің мәндерінен аспайды

$$\text{ord}A(z) < \text{ord}\pi(z).$$

Алынған әрбір  $v$  блок екілік жүйедегі сандардан құралған және кеңейтілген Галуа өрісінің элементі ретінде қарастырылады. Кеңейтілген Галуа өрісінде дәрежеге шығару операциясына негізделген шифрлеу процедурасын жүзеге асыру үшін  $K$  кілтінің мәнін анықтаймыз. Оны кездейсоқ тізбектер генераторының (КТГ) көмегімен аламыз[5]:

- тізбекті түзу;
- алынған тізбекті  $v$  ұзындыққа сәйкес бөлшектеу;
- екілік жүйедегі тізбектерді ондық жүйеге ауыстыру.

$$K = \{k_0, k_1, k_2, \dots\}, \quad (2)$$

мұндағы  $k_0, k_1, k_2, \dots$   $v$  ұзындыққа сәйкес алынған бөліктердің ондық жүйедегі мәні.

Модуль бойынша дәрежеге шығару операциясын қолдану арқылы деректерді шифрлеу формуласы келсідей анықталынады:

$$\beta_i(z) = A(z)^{k_i} \text{ mod } \pi(z), \quad (3)$$

мұндағы  $\beta_i$  - шифрмәтін,  $i=0, 1, \dots$

Хабарламаны дешифрлау үшін (4) формула бойынша есептейміз.

$$A(z) \equiv \sqrt[k_i]{\beta_i(z)} \text{ mod } \pi(z) \quad (4)$$

Мұнда (5) теңдеуді қанағаттандыратын  $k_i$ -дің кері элементін табу арқылы ашық мәтінді есептейміз:

$$k_i \cdot (k_i)^{-1} \equiv 1 \pmod{\pi^v - 1} \quad (5)$$

Онда (5) формула негізінде дешифрлау формуласын келесідей есептеуге болады:

$$A_i(z) \equiv \beta_i(z)^{k_i^{-1}} \pmod{\pi(z)}. \quad (6)$$

Ұсынылған шифрлеу алгоритмде кілттерді генерациялауға қолданылатын осындай топтың реті  $\pi(z)$  келтірілмейтін көпмүшелікпен анықталынады. Таңдап алынатын кілт КТГ-нан алынғандықтан олардың барлығы нөлдік күйде болуы мүмкін, яғни  $K=0$  ондай жағыдайда кілттерді қайтадан генерациялаймыз.

Бұл процедураның қиындығы кілтті білмейтін пайдаланушылар ең алдымен бастапқы деректерді анықтау үшін шифрланған блоктың символдарына байланысты дискретті логарифмді есептеу қажет. Алайда, мультипликативтік операцияларға байланысты дәрежеге шығару және оның керісін есептеу ең қиын және көп уақытты талап етеді.

Мысалы:  $GF(2^3)$  өрісін қарастырайық. Осы өрісте жататын келесі келтірілмейтін көпмүшелікті таңдап алайық.

$$\pi(z) = z^3 + z + 1$$

Бізге келесі түрде екілік тізбектен құралған ашықмәтін берілсін:

$$A = 110111000111$$

Модуль бойынша дәрежеге шығару негізінде ақпараттарды шифрлеу процессіне сәйкес ашықмәтіндерді  $v$  блоктарға бөлеміз.

Біздің жағыдайда:

$$\begin{aligned} A_0 &= 110; & A_0(z) &= z^2 + z; \\ A_1 &= 111; & A_1(z) &= z^2 + z + 1; \\ A_2 &= 000; & A_2(z) &= 0; \\ A_3 &= 111; & A_3(z) &= z^2 + z + 1; \end{aligned}$$

Кездейсоқ сандар тізбегінің генераторынан алынған кілт мынаған тең болсын:

$$K = (3, 5, 6, 6).$$

Шифрлеу процесін  $v$  бойынша бөлініп алынған бір блокты ғана сипаттайық. Онда бөлініп алынған бірінші блокты (3) формулаға сәйкес шифрлейміз:

$$\beta_0(z) = A_0(z)^{k_1} \pmod{\pi(z)} = (z^2 + z)^3 \pmod{z^3 + z + 1} = z^2 + z + 1$$



Енді дешифрлау үшін (5) формула бойынша  $k_0$  -дің кері элементін табамыз, біздің жағында 3-тің кері элементі модуль 7 бойынша 5-ке тең.

Онда (6) формулаға сәйкес дешифрлейміз:

$$A_0(z) = \beta_0(z)^{k_0^{-1}} \bmod \pi(z) = (z^2 + z + 1)^5 \bmod z^3 + z + 1 = z^2 + z.$$

Алдағы жұмыстарда осы ұсынған шифрлеу алгоритмі жан жақты зерттеліп, криптоұрақтылығы талданады. Алгоритмнің тиімділігін жетілдіру, бағдарламалық жүзеге асыру және басқада алгоритмдермен салыстыру жұмыстары жоспарланған.

#### **Қолданылған әдебиеттер**

1. Акушкин И.Я., Юдицкий Д.И. Машинная арифметика в остаточных классах, Из. «Советское радио», 1968 г., стр.430.

2. Бияшев Р.Г. Разработка и исследование методов сквозного повышения достоверности в системах обмена данными распределенных АСУ: Дис. ... докт. тех. наук. – М., 1985. – 328 с.

3. Капалова Н.А. Алгоритм шифрования на базе непозиционных полиномиальных систем счисления с использованием системы Эль-Гамала // Инфокоммуникационные технологии в науке, производстве и образовании: Пятая Межд.науч.-кон.-Ч. П.-г.Ставрополь, Северо-Кавказ.гос.тех.ун-т, 2-6 мая, 2012.- С.225-227.

4. Чипига А.А. Криптографическая защита данных в информационных технологиях на базе непозиционных полиномиальных систем / И.А. Калмыков, А.А. Чипига, А.В. Барильская, О.А. Кихтенко // «Известия ЮФУ. Технические науки». - Таганрог, - 2009. - С.210-220.

5. Чипига А.А. Нейроалгоритм для вычисления коэффициентов обобщенной полиадической системы в  $GR(p^v)$  [Текст] / А.А. Чипига, И.А. Калмыков, Ю.О. Щелкунова, А.А. Шилов // Материалы III Санкт-Петербургской межрегиональной Конференции «Информационная безопасность регионов России(ИБРР-2003)», Санкт-Петербург, - 2003- С. 55.

## **ОБ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЕ И МАТЕМАТИЧЕСКОЙ МОДЕЛИ ЭЛЕКТРОЭНЕРГЕТИЧЕСКИХ СИСТЕМ**

**Абдилдаева А.А., Галиева Ф.М., Базарбекова М.О., Даулетбек Е.Т.**

*Институт информационных и вычислительных технологий КН МОН РК,  
Казахстан*

*e-mail: abass\_81@mail.ru, bazarbekovameruert@gmail.com,  
dauletbekyergali2011@gmail.com*

**Аннотация.** Электроэнергетические объекты представляют собой сложную многосвязную систему автоматизированного управления. Из года в год возрастающая энергоемкость малых и больших промышленных предприятий и городов делает актуальным создание безотказно и безаварийно функционирующих

электро-энергетических систем (ЭЭС). В данной статье рассмотрена современная математическая модель ЭЭС, а также практическое исследование производительности информационно-интеллектуальных систем (ИИС) для ЭЭС. Предложена оптимальная система с целью увеличения производительности для ЭЭС.

### Введение

На сегодняшний день исследуются и формируются новые концепции развития электроэнергетических систем. Технологическая инфраструктура современных электроэнергетических систем является сложной и включает в себя множество различных пространственно-распределенных, но взаимосвязанных технических элементов, в режиме реального времени осуществляющих процессы производства, передачи и распределения электрической энергии и реализующих общую стратегическую цель – обеспечить надежное электроснабжение потребителей энергии. Для этого существует ряд методов прогнозирования потребления электроэнергии. Однако все эти методы могут обработать информацию, которая поступает от каждой точки потребления с помощью счетчиков. В свою очередь, счетчики имеют возможность передавать информацию о мгновенных значениях потребленных мощностей в центры обработки данных. Казахстан не имеет подобной развитой инфраструктуры, что делает невозможным использование иностранных методик, поэтому возникает необходимость создания интеллектуальной системы для ЭЭС на основе математических моделей. Существующие методики мирового масштаба имеют разветвленную систему датчиков и счетчиков.

### Математическая модель многомашинной энергосистемы

Рассмотрим классическую математическую модель многомашинной системы. На рисунке 1 приведена электрическая схема системы из  $n$  машин. Все напряжения измеряются относительно узла 0 (нейтраль). Узлы  $1, 2, \dots, n$  –шины соответствующих машин или точки подключения э.д.с. за переходными реактивными сопротивлениями. Различные узлы соединены между собой и с узлом 0 пассивными элементами, представленными полными сопротивлениями. Исходные значения  $E_1, E_2, \dots, E_n$  определяются из условий предаварийного режима. Величины  $E_i (i=1, 2, \dots, n)$  в переходном режиме принимаются постоянными.

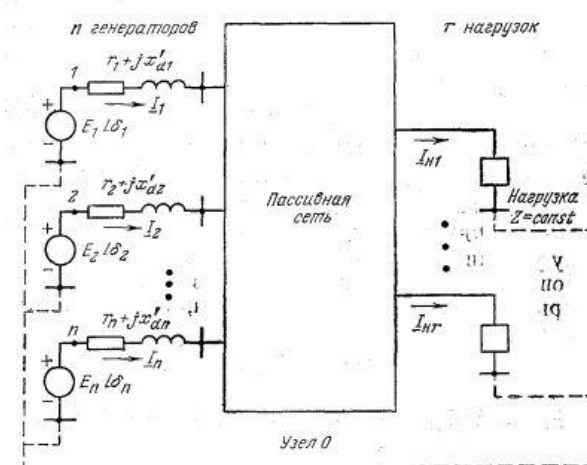


Рисунок 1. Классическая модель многомашинной системы

Матрица проводимостей  $n$ -полюсника со стороны выводов генераторов определяется выражением

$$I = YE$$

где  $Y$  имеет диагональные элементы  $Y_{ii}$  и недиагональные элементы  $Y_{ij}$ . По определению  $Y_{ii} = Y_{ii} < \theta_{ii} = G_{ii} = jB_{ii}$  равна собственной проводимости узла  $i$ , а  $Y_{ij} = Y_{ij} < \theta_{ij} = G_{ij} = jB_{ij}$  равна взаимной проводимости между узлами  $i$  и  $j$  со знаком минус.

Мощность, притекающая в сеть в узле  $i$  и равная электрической мощности  $i$ -й машины, определяется из выражения

$$P_i = \operatorname{Re} E_i I_i^*$$

$$P_{\Sigma i} = E_i^2 G_{ii} + \sum_{\substack{j=1 \\ j \neq i}}^n E_i E_j Y_{ij} \cos(\theta_{ij} - \delta_i + \delta_j) =$$

$$E_i^2 G_{ii} + \sum_{\substack{j=1 \\ j \neq i}}^n E_i E_j Y_{ij} [B_{ij} \sin(\delta_i - \delta_j) + G_{ij} \cos(\delta_i - \delta_j)]$$

$$i = 1, \dots, n$$

Уравнения движения имеют вид [1]:

$$\frac{\tau_{J,i}}{\omega_{ном}} \frac{d\omega}{dt} + D_i \omega = P_{MXi} - \left[ E_i^2 G_{ii} + \sum_{\substack{j=1 \\ j \neq i}}^n E_i E_j Y_{ij} \cos(\theta_{ij} - \delta_i + \delta_j) \right]$$

$$\frac{d\delta_i}{dt} = \omega_i - \omega_{ном}, i = 1, 2, \dots, n$$

Необходимо отметить, что в момент, предшествующий аварийному возмущению ( $t=0$ ),

$$P_{MXi,0} = P_{\Sigma i,0}, \text{ т.е. } P_{MXi,0} = E_i^2 G_{ii,0} + \sum_{\substack{j=1 \\ j \neq i}}^n E_i E_j Y_{ij,0} \cos(\theta_{ij,0} - \delta_{i,0} + \delta_{j,0}).$$

При некоторых стандартных предположениях динамика  $n$  взаимосвязанных генераторов через сеть передачи может быть описана классической моделью с динамикой распада потока [2]. Сеть была сведена к представлению внутренней шины, предполагая, что нагрузки являются постоянными импедансами и учитывают наличие проводимости переноса. Динамическая модель  $i$ -й машины представлена классической моделью третьего порядка

$$\dot{\delta}_i = \omega_i - \omega_s$$

$$\dot{\omega}_i = \frac{\omega_s}{2H_i} (P_{mi} - D_i(\omega_i - \omega_s) - E_{qi}' I_{qi})$$

$$\dot{E}_{qi}' = \frac{1}{T_{di}} (E_{fi}' - E_{qi}' - (X_{di}' - X_{di}') I_{di})$$

где

$$I_{q_i} = G_{ii} E_{q_i} + \sum_{j=1, j \neq i}^n E_{q_j} \{G_{ij} \cos(\delta_j - \delta_i) - B_{ij} \sin(\delta_j - \delta_i)\}$$

$$I_{d_i} = -B_{ii} E_{q_i} - \sum_{j=1, j \neq i}^n E_{q_j} \{G_{ij} \sin(\delta_j - \delta_i) + B_{ij} \cos(\delta_j - \delta_i)\}$$

$I_{q_i}(t)$  и  $I_{d_i}(t)$  представляют токи в d-q системе отсчета i-го генератора,  $E_{q_i}(t)$  - переходная ЭДС в квадратурной оси,  $E_{f_i}(t)$  - является эквивалентной ЭДС в катушке возбуждения,  $X_{d_i}$  и  $X'_{d_i}$  - реактивное сопротивление прямой оси и переходное реактивное сопротивление прямой оси соответственно;  $P_m$  - механическая входная мощность, принимаемая постоянной,  $D_i$  - коэффициент затухания;  $H_i$  - представляет собой постоянную инерции в секундах;  $T'_{d_i}$  - постоянная времени короткого замыкания на прямой оси, в секундах;  $\delta_i(t)$  - угол ротора в радианах;  $\omega_i(t)$  - представляет относительную скорость,  $\omega_s = 2\pi f_s$  - синхронная скорость машины, в рад / с;  $G_{ij}$  и  $B_{ij}$  представляют собой i-й элемент строки и j-го столбца матрицы узловых проводимости и матрицы узловых восприимчивости соответственно, которые являются симметричными, во внутренних узлах после устранения всех физических шин. Рассмотрим  $E_{f_i}(t)$  как входной сигнал системы.

### Интеллектуальная система для ЭЭС

Современные методики распространены и использованы для новых ИИС, как GRID, SCADA. В этих системах сложилась определенная технология разработки интеллектуальных информационных систем, которая включает следующие шесть этапов: идентификация, концептуализация, формализация, конструирование и реализация, тестирование и опытная эксплуатация. Классифицируя эти системы, технология Supervisory Control and Data Acquisition (SCADA) - программный пакет, предназначенный для разработки или обеспечения работы в реальном времени систем сбора, обработки, отображения и архивирования информации об объекте мониторинга или управления.

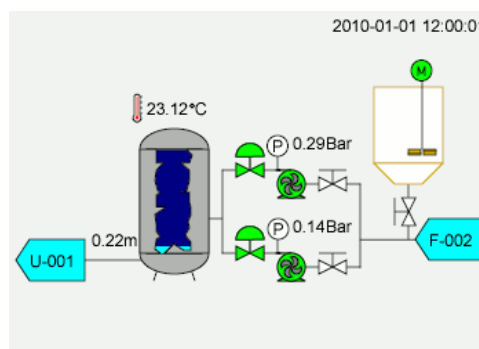
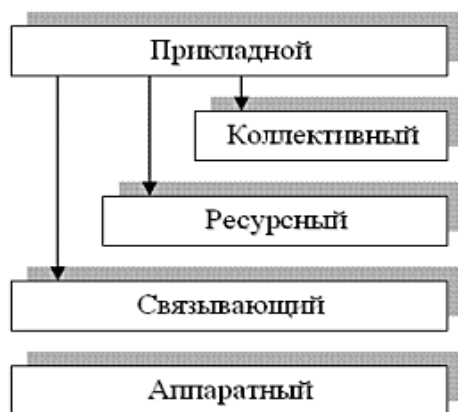


Рисунок 2 Интегрированные SCADA-системы

Данное программное обеспечение устанавливается на компьютеры и, для связи с объектом, использует драйверы ввода-вывода или OPC / DDE-серверы. Этот метод вполне подходящий под отечественные требования. Программный код может быть как

написан на языке программирования, так и сгенерирован в среде автоматизированного проектирования. Иногда SCADA-системы комплектуются дополнительным ПО для программирования промышленных контроллеров. Такие SCADA-системы называются интегрированными как показано на рисунке 2 [5].

Общая структура глобальной системы GRID описывается в виде стека (набора уровней или слоев) протоколов как на рисунке 3. В такой модели каждый уровень предназначен для решения узкого круга задач и используется для предоставления услуг для более высоких уровней [6]. Верхние уровни ближе к пользователю и работают с наиболее абстрактными объектами, тогда как нижние уровни сильно зависят от физической реализации грид-ресурсов.



**Рисунок 3. Стек протоколов**

Итак, стек грид-протоколов могут непосредственно работать с ресурсами; обеспечивать обмен данными между компонентами базового уровня и протоколы аутентификации; также взаимодействуют с ресурсами, используя унифицированный интерфейс и не различая архитектурные особенности конкретного ресурса; отвечает за координацию использования имеющихся ресурсов; описывает пользовательские приложения, работающие в среде виртуальной организации; приложения функционируют, используя протоколы, определенные на нижележащих уровнях.

В Казахстане большинство этих функций выполняется в «ручном» режиме, что связано с наличием многих ошибок из-за недостаточной квалификации обслуживающего персонала, следствием которых являются высокие экономические и социальные потери.

Изучение слабых мест существующих подходов с целью разработки универсальной системы, описывающая нестабильные быстро меняющиеся процессы и распознающая скрытую информацию, обеспечивая адаптационные свойства модели, представляет большую ценность как для отечественной и мировой науки. Надежное и эффективное управление режимом электроснабжения на всех уровнях диспетчерского управления в новых модулях для решения конкретных задач сложного процесса оперативно-диспетчерского управления тепло- и электроэнергетика Казахстана. Решить эту задачу возможно на основе использования концепции интеллектуальных систем управления, разрабатываемых в последние годы и применяемых в зарубежных энергетических компаниях.

С помощью сделанного сравнительного анализа информационно-интеллектуальных системы можем узнать нужные требования отечественного ЭЭС и

построить нужную комплексную модель. Основой разрабатываемой системы является комплекс, который базируется на объединении генетического, пространственно-временного и функционального подходов.

Кроме комплексного подхода будет использоваться анализ полной (зеркальной и инверсионной) симметрии, геометрически отражающей законы сохранения. Базируясь на ограниченном объеме информации реального времени, но учитывая наиболее вероятные внутрисистемные и межсистемные связи, система сможет позволять эффективно прогнозировать развитие событий, связанных с процессами генерации, передачи и потребления электроэнергии.

Система сможет обеспечивать контроль дистанционного управления и мониторинг сигналов защит, управлять каналами связи с диспетчерским центром, вести замер показателей качества электроэнергии, контролировать метки времени, вести постоянный обмен с системой коммерческого учета электроэнергии, блокировать аппаратуру в случае неправильных действий персонала, а также вести мониторинг параметров основного и вспомогательного оборудования, длительно работающего в режиме постоянной готовности к локализации аварии.

Использование разрабатываемой системы для ЭЭС приведет к увеличению надежности энергообъектов, и даст нам высший уровень управления и организацию. Модель работы ИИС показан на рисунке 4.

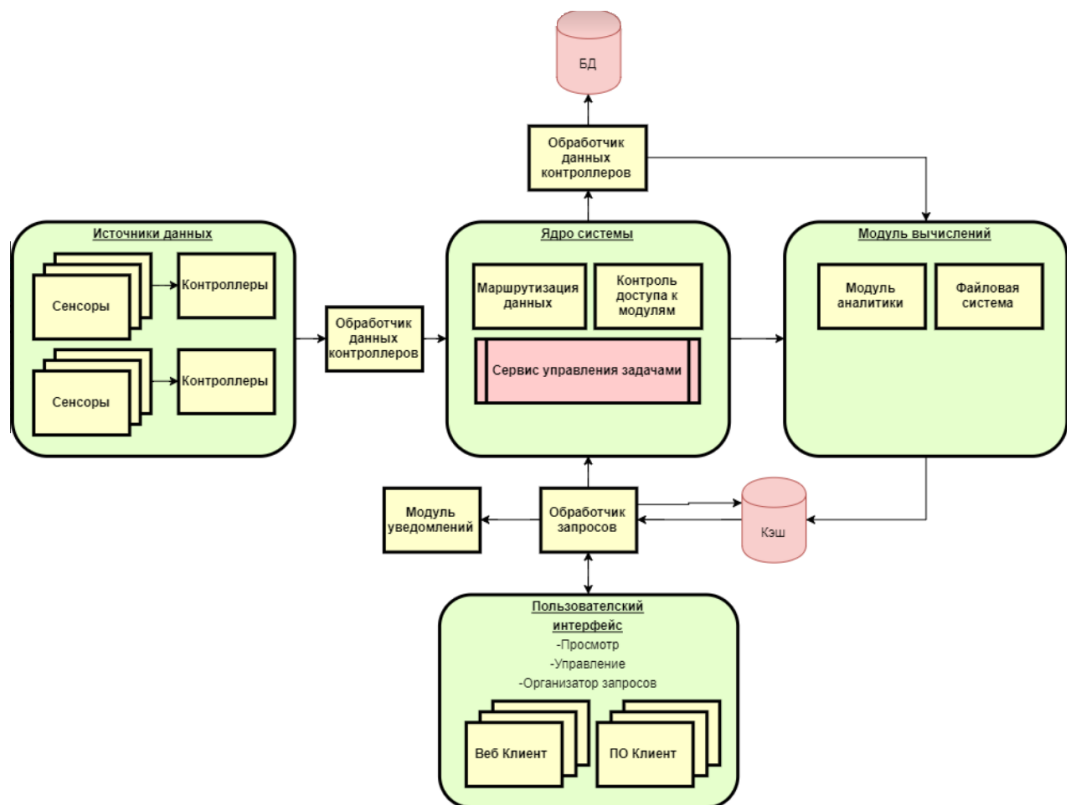


Рисунок 4. Модель системы

### Заключение

Интенсивное развитие средств вычислительной техники, современных технологий связи, а также появление новейших алгоритмов и методик анализа и

синтеза многоуровневой информации способствовали созданию информационной интеллектуальной системы для отечественных электроэнергетических станций.

Отечественная информационно-интеллектуальная система – это прежде всего знание технологии и местных условий, включающее знание отечественной нормативно-технической документации. Концепция интеллектуальности системы заключается в запрограммированных алгоритмах, позволяющих вести автоматическое управление системой в зависимости от складывающейся ситуации. Система автономно будет предпринять действия, предотвращающие возникновение или развитие аварийной ситуации, при выходе определенных параметров за установленные пределы, либо оптимизировать работу для достижения наибольшей экономии и т.д. Система позволит полностью отслеживать и фиксировать в журнале события, происходящие в системе. Кроме того в журнале фиксироваться действия персонала, либо бездействие. Это позволит с достаточной точностью восстановить ход событий, а также полностью отказаться от ведения обычного журнала, полностью заменив его электронной версией. Таким образом, система должна стать «автопилотом» оператора энергосистемы, а журнал событий, переданный в базы высшего уровня «черным ящиком».

#### **Список использованной литературы**

1. П. Андерсон, А. Фуад. Управление энергосистемами и устойчивость / пер. с англ. А. М. Машанского, Р. С. Рабиновича. - М. : Энергия, 1980.
2. M.A. Pai, Energy function analysis for power system stability, in: Kluwer International Series in Engineering and Computer Science, Power Electronics and Power Systems, Springer-Verlag, New York, August 1989.
3. Поспелов, Г.С. Искусственный интеллект – основа новой информационной технологии / Г.С. Поспелов. – М. : Наука, 2008.
4. «Развитие методов оценивания состояния ЭЭС на основе интеграции данных СКАДА», 2017
5. FP7-ENERGY-2008-Russia - Intelligent Coordination of Operation and Emergency Control of EU and Russia Power Grids (ICOEUR)
6. Международный научно-практический семинар «Современные программные средства для расчётов нормальных и аварийных режимов, анализа надежности, оценивания состояния, проектирования и автоматизации оперативно-диспетчерского управления электроэнергетических систем». - Иркутск: ИДУЭС. 2002, 2003, 2005, 2006, 2007, 2008гг.

## **АНАЛИЗ СОВРЕМЕННЫХ МЕТОДОВ МОДЕЛИРОВАНИЯ ПОТРЕБЛЕНИЯ ЭЛЕКТРИЧЕСКОЙ ЭНЕРГИИ**

**<sup>1</sup>Абдилдаева А.А., <sup>2</sup>Дрозденко А.А., <sup>2</sup>Коплык И.В., <sup>2</sup>Маринич Т.А.**

<sup>1</sup> *Институт информационных и вычислительных технологий КН МОН РК,  
Казахстан*

<sup>2</sup> *Сумский государственный университет, Украина  
e-mail: aleksey.drozdenco@ms.sumdu.edu.ua*

***Аннотация.** Проведен обзор современных отечественных и зарубежных методов формирования математической модели электроэнергетических систем и разработки интеллектуальной информационной системы. Выявлены основные недостатки существующих подходов к управлению энергетическими системами. Определены основные значащие факторы, которые влияют на повышение потребления электроэнергии потребителями. Выработаны перечень основных задач, которые необходимо разработать с целью создания алгоритмов прогнозирования потребления электроэнергии. Проведено исследование детерминированных и стохастических компонент временных рядов.*

В настоящее время совершенствование системы управления процессами потребления и сбережения электроэнергии отдельными объектами (учреждениями) и отраслями экономики в целом, предусматривает создание регионально-отраслевой организационно-экономической модели. Разрабатываемый подход развития системы управления энергосбережения на примере образовательных учреждений создает методические основы и алгоритмическую базу сбора, обработки, анализа информации по использованию электроэнергии, принятия управленческих решений и проведения электроэнергетического аудита, а также обеспечивает функции контроля электропотребления, совершенствование системы лимитирования электроэнергии; автоматизированное создание электроэнергетических паспортов объектов [6, 9]. Настоятельная необходимость создания эффективной системы управления энергосбережения на примере образовательных учреждений и недостаточная степень разработки теоретико-методологических подходов к оценке уровня эффективности использования электроэнергии, свидетельствуют об объективной необходимости проведения дальнейшего углубления указанного научного исследования. В настоящее время в Казахстане и Украине подобные комплексные системы управления отраслевого уровня отсутствуют.

В ходе проведения аналитического обзора сформированы основные положения теории моделирования систем управления эффективностью и прогнозирования использования электрической энергии потребителями, которые базируются на закономерности динамики временных рядов внутренних (техничко-экономических, структурных, режимных) и внешних (метеорологических, экологических, энергетических, макроэкономических) факторов, характеризующих систему «генерация – климатические условия – энергопотребление» [4, 5]. Установлено, что динамика временных рядов этих факторов характеризуется слабой прогнозируемостью. Вместе с тем, несмотря на этот негативный фактор, между ними существует взаимосвязь (коинтеграция), которая приводит к некоторым общим, взаимосвязанным изменениям. Именно эта особенность динамики временных рядов внутренних и внешних факторов, позволяет применить при их исследовании, теорию Р. Энгла и К. Грэнджера.

Характерной особенностью системы «объект – энергопотребления» является стационарный характер временных рядов факторов ее функционирования. Иначе говоря, функции распределения стационарных динамических рядов не изменяются при сдвиге времени. Эта особенность динамики временных рядов системы «объект – энергопотребления» позволяет применить при их исследовании [8], теорию авторегрессионного моделирования.

В рамках структурного исследования проведены сбор и подготовка статистических данных, необходимых для дальнейшего тестирования различных



методик моделирования, проверки моделей на адекватность и оценки качества полученных прогнозов.

Установлено, что важным этапом является сочетание различных авторегрессионных подходов, структурного моделирования и нейросетевого прогнозирования [1, 7] для дальнейшей математической и компьютерной реализации адаптивной модели энергопотребления с элементами искусственного интеллекта.

Проведенный анализ показывает, что для поддержки системы управления и принятия решений по вопросам использования и экономии электроэнергии необходима детализация междисциплинарных, системных подходов, что, в свою очередь, обеспечит функционирование системы энергосбережения на всех уровнях, включая вопросы: анализа и сравнения объектов с другими, оптимизации расходов бюджетных и собственных средств учреждений образования на оплату электроэнергии, разработку форм отчетности и методических и сопроводительных документов, создание автоматизированного программного комплекса для внедрения системы сбора и анализа информации по использованию объектами электроэнергии, а внедрение пилотных систем управления эффективностью использования электрической энергии для отдельных потребителей системы.

В результате аналитического обзора были выбраны теоретические основы моделирования систем управления эффективностью и прогнозирования использования электрической энергии потребителями, базирующиеся на сочетании элементов, обобщенных авторегрессионных моделей скользящей средней (ARIMA), структурных и коинтеграционных моделей (теория Р. Энглу и К. Грэнджера) [3]. На основе выбранных теоретических моделей предполагается разработка научно-методического обеспечения (инструментария) создание многоуровневой (в пилотном случае – двухуровневой) системы управления процессами эффективного потребления электроэнергии: универсальных экономико-математических моделей процессов электропотребления.

### **Литература**

1. Multifactor-influenced energy consumption forecasting using enhanced back-propagation neural network / Y. Zeng, Y. Zeng, B. Choi, L. Wang // *Energy*. – 2017. – Vol. 127 – P. 381-396. (<http://www.sciencedirect.com/science/article/pii/S0360544217304759>).
2. Time series forecasting for building energy consumption using weighted Support Vector Regression with differential evolution optimization technique / [F. Zhang, C. Deb, S. Lee and oth.]. // *Energy and Buildings*. – 2016. – Vol. 126. – P. 94-103. (<http://www.sciencedirect.com/science/article/pii/S0378778816303899>).
3. Sen P. Application of ARIMA for forecasting energy consumption and GHG emission: A case study of an Indian pig iron manufacturing organization / P. Sen, M. Roy, P. Pal. // *Energy*. – 2016. – Vol. 116, P.1. – P. 1031-1038. (<http://www.sciencedirect.com/science/article/pii/S0360544216315043>).
4. A review on time series forecasting techniques for building energy consumption / [C. Deb, F. Zhang, J. Yang and oth.] // *Renewable and Sustainable Energy Reviews*. – 2017. – Vol. 74. – P. 902-924. (<https://ideas.repec.org/a/eee/rensus/v74y2017icp902-924.html>).
5. Building electrical energy consumption forecasting analysis using conventional and artificial intelligence methods: A review / [M. Mat Dauta, M. Yusri Hassan, H. Abdullaha and oth.]. // *Renewable and Sustainable Energy Reviews*. – 2017. – Vol. 70. – P. 1108-1118. (<https://ideas.repec.org/a/eee/rensus/v70y2017icp1108-1118.html>).
6. Старкова Г. Методы и модели прогнозирования электропотребления на региональном уровне / Г. Старкова // *International Journal «Information Theories and*

Applications», vol. 19, Number 4, 2012. pp. 378-383. (<http://www.foibg.com/ijita/vol19/ijita19-4-p10.pdf>).

7. Нечеткие модели и нейронные сети в анализе и управлении экономическими объектами: монография / [Ю.Г. Лысенко, Е.Е. Бизянов, А.Г. Хмелев, А.Ю. Минц и др.]; под ред. Ю.Г. Лысенко. – Донецк: Юго-Восток, 2012. – 388 с.

8. Калінчик В. П. Методологія оперативного управління споживанням електричної енергії / В. П. Калінчик // Енергетика. – 2013. – № 1. – С. 49-53. ([http://nbuv.gov.ua/UJRN/eete\\_2013\\_1\\_10](http://nbuv.gov.ua/UJRN/eete_2013_1_10)).

9. Калінчик В. П. Аналіз показників нерівномірності графіків навантаження промислових підприємств / В. П. Калінчик, В. П. Розен, О. В. Скачок // Вісник КрНУ імені Михайла Остроградського. – 2014 – Випуск 2 (85). – С. 67-72. ([http://nbuv.gov.ua/UJRN/Vkdpu\\_2014\\_2\\_12](http://nbuv.gov.ua/UJRN/Vkdpu_2014_2_12)).

## АЛГЕБРАИЧЕСКИЕ СВОЙСТВА СТЕПЕНЕЙ ПОЗИТИВНЫХ ПРЕДПОРЯДКОВ

**Абешев К.Ш., Бадаев С.А., Кабылжанова Д.К., Калмурзаев Б.С.,  
Серикжан Р., Сыдыкбек А.Ж.**

*Институт информационных и вычислительных технологий КН МОН РК,  
Казахстан*

*e-mail: [kuanqk@gmail.com](mailto:kuanqk@gmail.com), [sbadaev@gmail.com](mailto:sbadaev@gmail.com), [dkabylzhanova@gmail.com](mailto:dkabylzhanova@gmail.com),  
[birzhan.kalmurzayev@gmail.com](mailto:birzhan.kalmurzayev@gmail.com), [r.serikzhan@gmail.com](mailto:r.serikzhan@gmail.com),  
[aibeksydykbek@gmail.com](mailto:aibeksydykbek@gmail.com)*

***Аннотация.** Доклад посвящен результатам исследований алгебраических свойств структуры степеней позитивных предпорядков, выполненных в соответствии с календарным планом по проекту АР05131579 «Позитивные предпорядки и вычислимая сводимость на них как математические модели баз данных». В структуре степеней позитивных предпорядков относительно вычислимой сводимости выделены нерасщепляемые степени, показано существование цепей и антицепей разного рода, установлено, что каждая степень состоит либо из одного, либо из бесконечно многих типов рекурсивного изоморфизма.*

Позитивный предпорядок - это вычислимо перечислимое, рефлексивное и транзитивное бинарное отношение. На позитивных предпорядках естественным образом вводится отношение вычислимой сводимости [1,2] и соответствующее понятие степени позитивного предпорядка. В самой общей ситуации вычислимая сводимость может быть определена для бинарных отношений следующим образом:

Бинарное отношение  $P$  на множестве натуральных чисел  $\omega$  вычислимо сводится к бинарному отношению  $Q$  на  $\omega$  (символически,  $P \leq_c Q$ ), если существует вычислимая функция  $f$  такая, что

$$x P y \Leftrightarrow f(x) Q f(y),$$

для любых  $x, y \in \omega$ .

Понятия вычислимого изоморфизма [3] и вычислимой эквивалентности определяются обычным образом: бинарное отношение  $P$  на  $\omega$  вычислимо изоморфно бинарному отношению  $Q$  на  $\omega$ , если существует вычислимая перестановка  $p$ , осуществляющая сведение  $P \leq_c Q$ .

Отношение  $\leq_c$  является предпорядком, который с точки зрения теории нумераций является мономорфизмом нумерованных множеств, для которых  $R$  и  $Q$  являются нумерационными эквивалентностями.

Отношение  $P$  вычислимо эквивалентно отношению  $Q$  ( $P \equiv_c Q$ ) тогда и только тогда, когда  $P \leq_c Q$  и  $Q \leq_c P$ . Множество отношений, вычислимо эквивалентных позитивному предпорядку  $P$ , называется  $\leq_c$ -степенью предпорядка  $P$ , а совокупность  $\leq_c$ -степеней всех позитивных предпорядков относительно частичного порядка, индуцированного вычислимой сводимостью  $\leq_c$ , является алгебраической структурой, которая содержит всю информацию о позитивных предпорядках в целом. Исследование алгебраических свойств этой структуры было запланированы в проекте в течение первых трех кварталов 2018 года.

Заметим, что в частном случае позитивных эквивалентностей исследования структуры степеней относительно вычислимой сводимости проводилось исполнителями проекта по гранту КН МОН РК ГФ 4/3952 «Отношения эквивалентности, предпорядоченные структуры и алгоритмические сводимости на них, как математическая модель баз данных». По тематике названного проекта опубликованы десятки статей, из которых для текущего проекта, как мы считаем, наиболее важными являются работы американских и итальянских авторов [5,6]. Отметим, что совокупность  $\leq_c$ -степеней позитивных эквивалентностей образует начальный сегмент в структуре  $\leq_c$ -степеней всех позитивных предпорядков. Поэтому продолжение исследований структуры  $\leq_c$ -степеней позитивных эквивалентностей лежит также в русле исследований структуры  $\leq_c$ -степеней всех позитивных предпорядков.

В ходе выполнения проекта ГФ 4/3952 было замечено, что некоторые из разработанных для позитивных эквивалентностей методов могут быть использованы и для позитивных предпорядков. В частности, количество типов вычислимого изоморфизма внутри произвольной  $\leq_c$ -степени в обоих случаях идентично. Этот результат был получен исполнителем Д.К. Кабылжановой (Теорема 8 ниже) и по нему была подготовлена статья в журнал Алгебра и логика. Особо отметим, что статья Д.К. Кабылжановой еще не вышла из печати и не была включена в перечень опубликованных работ по проекту ГФ 4/3952.

В первом полугодии 2018 г. наши исследования алгебраических свойств структуры степеней позитивных предпорядков были в основном сосредоточены на частном, но очень важном классе слабо предполных позитивных предпорядков.

Позитивный предпорядок  $P$  называется слабо предполным [4,7], если для любой всюду определенной функции  $\varphi_e$  существует элемент  $x_e$  такой, что  $\varphi_e(x_e)Px_e$ . Иными словами, на слабо предполных позитивных предпорядков невозможно проводить диагональные процедуры, являющиеся мощным инструментом исследований в теории алгоритмов.

Интерес к исследованию слабо предполных позитивных предпорядков вызван еще и тем, что  $\leq_c$ -степени слабо предполных позитивных предпорядков не расщепляются на две несравнимые  $\leq_c$ -степени, для которых она является точной верхней гранью.

Более точно это свойство можно сформулировать на основе следующего понятия. Прямой суммой позитивных предпорядков  $R$  и  $Q$  называется предпорядок  $R \oplus Q$ , для которого  $xR \oplus Qu$  тогда и только тогда, когда  $x = 2n, y = 2m$  и  $nRm$ , либо  $x = 2n + 1, y = 2m + 1$  и  $nQt$ .

**Теорема 1.** Если  $\leq_c$  – степени позитивных предпорядков  $R$  и  $Q$  не сравнимы, то  $\leq_c$  – степень позитивного предпорядка  $R \oplus Q$  не содержит слабо предполных позитивных предпорядков.

Степени, обладающие приведенным в Теореме 1 свойством, в западной литературе называются *join irreducible*. Это алгебраическое свойство является определимым в языке логики первого порядка. Поэтому характеристика *join irreducible*  $\leq_c$  – степеней является интересной и важной математической проблемой. Теорема 1 является частичным решением этой проблемы.

Позитивный предпорядок  $P$  назовем конечным, если он имеет конечное число классов эквивалентности. Как известно [1],  $\leq_c$  – степени конечных позитивных эквивалентностей образуют строго возрастающую цепь, каждый элемент которой  $\leq_c$  – сводится к степени любой бесконечной позитивной эквивалентности. Пределом этой цепи является  $\leq_c$  – степень тождественной эквивалентности  $Id$ . Степени бесконечных позитивных эквивалентностей делятся на два класса: расположенные выше  $\leq_c$  – степени  $Id$  и несравнимые с  $\leq_c$  – степенью  $Id$ . В соответствии с этим бесконечные позитивные эквивалентности делятся на светлые (*light*) и темные (*dark*). Термины *light, dark* предложены зарубежными исполнителями проекта U.Andrews, A.Sorbi в работе [8]. Понятия светлых и темных позитивных эквивалентностей мы обобщаем естественным образом на позитивные предпорядки следующим образом.

Позитивный предпорядок  $P$  назовем светлым, если найдется позитивный предпорядок  $Q$  такой, что

1.  $\{(x, y): xQy\} = Id$ .
2.  $Q \leq_c P$ .

Если позитивный предпорядок  $P$  не является ни светлым, ни конечным, то назовем  $P$  темным позитивным предпорядком.

**Теорема 2** (критерий темноты предпорядка).

Позитивный предпорядок  $P$  является темным тогда и только тогда, когда эквивалентность  $E = \{(x, y): xPy\}$  является темной.

В [8] показано, что структура  $\leq_c$  – степеней позитивных эквивалентностей не является ни верхней полурешеткой, ни нижней полурешеткой. В частично упорядоченном множестве  $\leq_c$  – степеней всех позитивных предпорядков операции взятия точной верхней и взятия точной нижней граней также определены не для всех пар элементов. Иными словами, структура  $\leq_c$  – степеней позитивных предпорядков также не является ни верхней полурешеткой (а верхними полурешетками являются структуры относительно большинства известных алгоритмических сводимостей), ни нижней полурешеткой. Следующей по важности характеристикой бесконечных частично упорядоченных множеств является наличие в них бесконечных цепей и антицепей. Существование антицепей в структуре  $\leq_c$  – степеней позитивных предпорядков установлено нами даже в начальном сегменте их сегменте  $\leq_c$  – степеней позитивных эквивалентностей, а, именно, справедлива следующая теорема.

**Теорема 3.** Существует бесконечная  $\leq_c$  – антицепь слабо предполных позитивных эквивалентностей.

Мы уже указывали на тривиальную бесконечную возрастающую цепь конечных позитивных эквивалентностей. Более того, среди  $\leq_c$  – степеней позитивных

эквивалентностей существуют бесконечные возрастающие цепи  $\leq_c$ -степеней слабо предполных позитивных эквивалентностей [9]. Естественен вопрос: а существуют ли бесконечные возрастающие цепи  $\leq_c$ -степеней светлых (темных) позитивных предпорядков?

**Теорема 4.** Для любого позитивного предпорядка  $P$  найдется слабо предполный позитивный предпорядок  $Q$  такой, что  $P \leq_c Q$ .

**Следствие.** Существуют бесконечные возрастающие цепи  $\leq_c$ -степеней светлых позитивных предпорядков.

Для темных позитивных предпорядков справедлив более сильный результат:

**Теорема 5.** Для любого темного позитивного предпорядка  $P$  найдется слабо предполный позитивный предпорядок  $Q$  такой, что  $P <_c Q$ .

Для темных позитивных предпорядков справедлив более сильный вариант Теоремы 3:

**Теорема 6.** Для любого конечного позитивного предпорядка  $P$  найдется бесконечно много попарно  $\leq_c$ -несравнимых слабо предполных минимальных темных позитивных предпорядков  $P_d$  таких, что  $P \leq_c P_d$ .

Здесь под минимальным понимается позитивный предпорядок, индуцирующий минимальный элемент в структуре степеней позитивных предпорядков. Заметим, что минимальных элементов в структуре  $\leq_c$ -степеней бесконечно много. Антиподом минимальных позитивных предпорядков являются универсальные позитивные предпорядки, к которым  $\leq_c$  сводятся любые позитивные предпорядки.

**Теорема 7.** Пусть  $P$  – не универсальный позитивный предпорядок. Тогда существует слабо предполный не универсальный позитивный предпорядок  $Q$  такой, что  $P \leq_c Q$ .

Таким образом, наибольшие элементы структуры  $\leq_c$ -степеней позитивных предпорядков являются предельными снизу.

Понятие  $\leq_c$ -эквивалентности позволяет изучать позитивные эквивалентности с точностью до вычислимой реализуемости. А понятие вычислимого изоморфизма изучает их с точностью до кодирования. Естественным является вопрос о взаимосвязи этих двух понятий. С практической точки зрения этот вопрос можно переформулировать так: сколькими принципиально разными способами можно эффективно закодировать одну и ту же базу данных. Ответ содержится в следующей теореме.

**Теорема 8.** Каждая  $\leq_c$ -степень позитивного предпорядка содержит либо один, либо бесконечно много типов вычислимого изоморфизма.

*Приведенные выше результаты подготовлены в виде статей или приняты к печати в научные журналы [10-12], а также будут также доложены на двух международных конференциях:*

*Logic Colloquium 2018, July 23-28, Udine, Italy,*

*Мальцевские чтения 2018, 19-23 ноября, Новосибирск, Россия.*

## **Литература**

1. Gao S. and Gerdes P. Computably enumerable equivalence relations // *Studia Logica*, –2001. – Vol. 67. – P. 27–59.
2. Bernardi C., Sorbi A. Classifying positive equivalence relations // *Journal of Symbolic Logic*. – 1983. – Vol. 48, № 3. – P. 529–538.
3. Lachlan A. H. A note on positive equivalence relations // *Z. Math. Logik Grundlagen Math*. – 1987. - Vol. 33. - P. 43–46.

4. Бадаев С. А. О слабо предполных позитивных эквивалентностях // Сибирский математический журнал. - 1991. - Т. 32, № 2. - С. 166–169.
5. Andrews U., Lempp S., Miller J.S., Ng K.M., San Mauro L., and Sorbi A. Universal computably enumerable equivalence relations // Journal of Symbolic Logic. – 2014. – Vol. 79, №1. – P. 60–88.
6. F. Montagna and A. Sorbi. Universal recursion theoretic properties of r.e. preordered structures // Journal of Symbolic Logic. – 1985. –Vol.50, № 2. –P.397-406.
7. S. Badaev, A. Sorbi. Weakly precomplete computably enumerable equivalence relations // Mathematical Logic Quarterly. – Vol.62, №1-2. – P.111-127.
8. Andrews U., Sorbi A. Joins and meets in the structure of computably enumerable equivalence relations // Journal of Symbolic Logic (принята к печати).
9. Badaev S. A chain of weekly precomplete computably enumerable equivalence relations // Bulletin of Symbolic Logic. – 2017. - Vol. 23, no.2. - P. 230-231.
10. Andrews U., Badaev S. On isomorphism classes of computably enumerable equivalence relations // Journal of Symbolic Logic (подготовлена к печати).
11. Кабылжанова Д.К. О позитивных предпорядках // Алгебра и логика (принята к печати).
12. Баженов Н.А., Калмурзаев Б.С. О темных вычислимо перечислимых отношениях эквивалентности // Сибирский математический журнал (сдана в печать).

## **ПОТЕНЦИАЛ ВОЗОБНОВЛЯЕМОЙ ЭНЕРГЕТИКИ И ОБЗОР ИМЕЮЩИХСЯ ЭНЕРГО И ТЕПЛО СБЕРЕГАЮЩИХ ТЕХНОЛОГИЙ В КАЗАХСТАНЕ**

**Амиргалиев Е.Н., Иржанова А.А.**

*Институт информационных и вычислительных технологий КН МОН РК,  
Казахстан  
amir\_ed@mail.ru, irzhanovaai@gmail.com*

**Аннотация.** В настоящей работе изучены существующие технологии энергосбережения, а так же потенциал возобновляемой энергетики. Проведен обзор имеющихся энергосберегающих и тепло сберегающих технологий. Казахстан имеет благоприятные географические факторы и потенциал нетрадиционных источников возобновляемой энергии для развития альтернативной энергетики. Внедрение энергосберегающих технологий в хозяйственную деятельность как предприятий, так и частных лиц на бытовом уровне, является одним из важных шагов в решении многих экологических проблем – изменения климата, загрязнения атмосферы, истощения ископаемых ресурсов и др.

**Ключевые слова:** возобновляемая энергия, альтернативные источники энергии, энергосбережение, тепло сбережение, умный дом.

На сегодняшний день энергосбережение – это одна из приоритетных задач. Связано эта ситуация с дефицитом основных энергоресурсов, возрастающей стоимостью их добычи, а также с глобальными экологическими проблемами. «Экономия энергии – это эффективное использование энергоресурсов за счет

применения инновационных решений, которые осуществимы технически, обоснованы экономически, приемлемы с экологической и социальной точек зрения, не изменяют привычного образа жизни». Это определение было сформулировано на Международной энергетической конференции ООН.

Энергосбережение в любой сфере сводится по существу к снижению бесполезных потерь энергии. Анализ потерь в сфере производства, распределения и потребления электроэнергии показывает, что большая часть потерь – до 90% – приходится на сферу энергопотребления, тогда как потери при передаче электроэнергии составляют лишь 9–10%. Поэтому основные усилия по энергосбережению сконцентрированы именно в сфере потребления электроэнергии. Основная роль в увеличении эффективности использования энергии принадлежит современным энергосберегающим технологиям. Энергосберегающая технология – новый или усовершенствованный технологический процесс, характеризующийся более высоким коэффициентом полезного использования топливно-энергетических ресурсов (ТЭР). Внедрение энергосберегающих технологий в хозяйственную деятельность как предприятий, так и частных лиц на бытовом уровне, является одним из важных шагов в решении многих экологических проблем – изменения климата, загрязнения атмосферы (например, выбросами от ТЭЦ), истощения ископаемых ресурсов и др. [1]

Примером домов, которые в будущем позволят человеку жить в гармонии с природой, в то же время не лишая себя привычного комфорта, являются так называемые жилища нулевой энергии (zero energy house) или пассивные дома (passive house), объединяемые общим термином "энергоэффективные дома". "Энергоэффективным" будет считаться такой дом, в котором комфортная температура поддерживается зимой без применения системы отопления, а летом без применения системы кондиционирования. Чтобы дом был энергоэффективным, при его строительстве должно быть сделано следующее:

- применение современной тепловой изоляции трубопроводов отопления и горячего водоснабжения;
- индивидуальный источник тепло энергоснабжения (индивидуальная котельная или источник когенерации энергии);
- тепловые насосы, использующие тепло земли, тепло вытяжного вентиляционного воздуха и тепло сточных вод;
- солнечные коллекторы в системе горячего водоснабжения и в системе охлаждения помещения;
- поквартирные системы отопления с теплосчетчиками и с индивидуальным регулированием теплового режима помещений;
- система механической вытяжной вентиляции с индивидуальным регулированием и утилизацией тепла вытяжного воздуха;
- поквартирные контроллеры, оптимизирующие потребление тепла на отопление и вентиляцию квартир;
- ограждающие конструкции с повышенной теплозащитой и заданными показателями теплоустойчивости;
- утилизация тепла солнечной радиации в тепловом балансе здания на основе оптимального выбора свето-прозрачных ограждающих конструкций;
- устройства, использующие рассеянную солнечную радиацию для повышения освещенности помещений и снижения энергопотребления на освещение;
- выбор конструкций солнцезащитных устройств с учетом ориентации и посезонной облученности фасадов;

– использование тепла обратной воды системы теплоснабжения для напольного отопления в ванных комнатах;

– система управления тепло энергоснабжением, микроклиматом помещений и инженерным оборудованием здания на основе математической модели здания как единой теплоэнергетической системы. [2]

«Умный дом» (Smart house) — чрезвычайно популярные сегодня автоматизированные системы управления, внедряемые как в жилые дома, так и в офисы. В них используются датчики движения, отслеживающие присутствие человека, в зависимости от чего происходит включение/выключение энергообеспечения помещения.

Умный дом — новейшие технические решения по энергетическому сбережению. Новейшие технологии энергосбережения в последнее время включают освоение возобновляемых источников энергии, к которым относится солнечная энергия. Однако широкого распространения они пока не получили из-за высокой цены. Тем не менее, в мире уже активно ведется строительство зданий по двум направлениям: активные и пассивные системы.

При пассивных – для накопления тепла используются свойства самого здания, а для при активных – в пределах здания создаются устройства, преобразующие солнечную энергию в тепловую. [3]

Разумеется, современные технологии энергосбережения не стоят на месте. Их развитие обусловлено высокой потребностью людей в энергосберегающих мерах в условиях дефицита и истощения энергоресурсов. Очевидно, что вскоре производители предложат еще более эффективные решения для сбережения энергии.

Общая мощность солнечной радиации, достигающая на долю нашей планеты, составляет  $1.7 \cdot 10^{14}$  кВт, что более чем в 14 000 раз больше современного мирового энергопотребления. Если суммировать всю солнечную энергию, которую наша планета получает за один год, то она составит  $10^{18}$  кВт/ч, что примерно в 10 раз больше энергии всех разведанных и прогнозируемых ископаемых топлив, включая и расщепляющиеся вещества. Прогнозы развития мировой энергетики предсказывают все возрастающую долю возобновляемых источников в будущем. Доля возобновляемых источников энергии, среди которых солнечная энергия играет основную роль, будет неуклонно расти и, по прогнозам, к концу XXI века может составить более 65%.

В развитых странах темпы роста основных отраслей возобновляемой энергетики (ВЭ) весьма высоки и составляют до 65% в год. Возобновляемая энергетика основана главным образом на нескольких формах: ветроэнергетика, солнечные тепловые низко потенциальные системы, фото-энергетика, биогаз, малая гидроэнергетика, геотермальная энергетика. [4]

В последние годы рост потребления электроэнергии в Казахстане был устойчивым. Однако глобальные проблемы с энергетикой, которые могут привести к деформированию экономики многих стран, в том числе и Казахстана, не за горами. Из страны, экспортирующей энергоносители, Казахстан может трансформироваться в государство, импортирующее энергетические установки возобновляемого типа. Если учесть, что неисчерпаемым источникам возобновляемой энергетики не свойственны проблемы истощения запасов, то можно говорить об отсутствии основной причины, вызывающей повышение цен на производимую электроэнергию.

Казахстан имеет благоприятные географические факторы и потенциал нетрадиционных источников возобновляемой энергии для развития альтернативной энергетики. По потенциалу солнечной энергии Казахстан занимает ведущее место в



мире. Как большая часть субрегиона Центральной Азии, Казахстан занимает территорию свыше 2,7 млн. км<sup>2</sup> с благоприятными географическими и климатическими условиями для развития возобновляемых источников энергии, таких, как солнечная и ветроэнергетика. Количество солнечных дней составляет до 300 дней в году.

Природные условия Казахстана благоприятны также для развития фото-энергетики и низко потенциального тепла. Количество солнечных часов составляет 2200-3000 часов в год, а суммарная дневная радиация при реальных условиях облачности в республике составляет 3,8-5,2 кВт•ч/м<sup>2</sup>, что делает возможным использование панелей солнечных батарей, портативных систем фото-электроисточников и солнечных нагревателей воды в сельской местности, на животноводческих отгонах. [5]

В то же время по энергообеспеченности на душу населения, по данным на 2002 г., ряд областей республики (Акмолинская, Алматинская, Жамбылская, Западно-Казахстанская, Северо-Казахстанская и Южно-Казахстанская) попадают в разряд районов, которые можно отнести к энергодефицитным (приходится менее 2 кВт•ч в сутки на человека, тогда как в среднем по стране этот показатель в 5 раз выше). Для повышения энергообеспечения населения можно использовать автономные фотоэнергетические установки.

Основные мотивы разработки возобновляемых ресурсов в Казахстане заключаются в замещение импорта электроэнергии, особенно в южных регионах, экологически чистыми и конкурентоспособными возобновляемыми энергетическими ресурсами в расширении доступа к электроэнергии населения отдаленных населенных пунктов и кочевий. Это возможно сделать посредством применения бытовых солнечных батарей на фотоэлементах.

ВЭ обеспечивает также защиту экосистемы страны путем уменьшения зависимости от выработки электроэнергии на основе угля (составляющей в настоящее время около 85%), которая оказывает негативное воздействие на окружающую среду. Снижение потерь на линиях электропередачи и усовершенствование стабильности и надежности через монтаж распределительных и терминальных станций, генерирующих электроэнергию с использованием возобновляемых энергетических ресурсов. [6]

В стране имеется хороший опыт работ в области низко потенциальных источников тепла, опреснителей и осушителей. Например, ТОО «Еркин и К» внедрило в производство серию изобретений по солнечным коллекторам и опреснителям, установило более 160 единиц таких устройств с мощностью от 100 до 20 тыс. литров горячей воды в день. Установки позволяют получать горячую воду с температурой до 100°C, максимальное давление в установках - до 20 атмосфер, срок службы - более 10 лет. Установки имеют небольшой вес, надежны в эксплуатации, недороги.

### **Заключение**

Для развития возобновляемой и альтернативной энергетики в Республике Казахстан необходимо, обеспечить разработку технологий в области ВЭ и их промышленное внедрение на существующих предприятиях; создание новых производств в области ВЭ, необходимой инновационной инфраструктуры, обеспечивающей развитие, интеграцию и специализацию участников выбранных направлений ВЭ, определение правовых и законодательных мероприятий, направленных на продвижение продукции и услуг, и упрощение регистрации организаций, специализирующихся в области ВЭ, субсидирование расходов на

внедрение технологий ВЭ, в том числе в виде льготного налогообложения; создать конкурентоспособность казахстанской энергетике.

#### Список использованной литературы

- [1] Матросов Ю.А. Энергосбережение в зданиях. Проблема и пути её решения. – М.:НИИСФ, 2008.
- [2] Рахматова М. У., Низомадинов И. М. Энергосберегающие технологии и способы энергосбережения // Молодой ученый. — 2015. — №23. — С. 212-214.
- [3] Пассивный дом (Passive House). <http://www.escoteco.ru>
- [4] Мендебаев Т. Нетрадиционная энергетика. Выбор за нами // Наука и высшая школа Казахстана, 2007.-С.15 сентября
- [5] Токмолдин С.Ж. Развитие солнечной фотоэнергетики в Республике Казахстан // Энергетика и топливные ресурсы Казахстана. — — № 8. — С. 43-47.
- [6] Бутырина Е. Приближающаяся угроза топливного голода и загрязнения окружающей среды требует увеличения мер по развитию энергосбережения и альтернативных источников энергии // Панорама. — 2009. — № 10. — 20 марта. — С. 12.

## ОБ ОДНОМ ПОДХОДЕ К РЕШЕНИЮ ДВУХТОЧЕЧНОЙ ЗАДАЧИ С ПАРАМЕТРОМ ДЛЯ СИСТЕМЫ ИНТЕГРО-ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ

Асанова А.Т.<sup>1,2</sup>, Бакирова Э.Б.<sup>1,2</sup>, Темешева С.М.<sup>1,2</sup>

<sup>1</sup>Институт математики и математического моделирования МОН РК

<sup>2</sup>Институт информационных и вычислительных технологий МОН РК  
e-mail: [assanova@math.kz](mailto:assanova@math.kz)

***Аннотация.** Исследуется линейная двухточечная задача с параметром для системы интегро-дифференциальных уравнений Фредгольма с вырожденным ядром. Методом параметризации рассматриваемая задача сведена к эквивалентной задаче, состоящей из специальной задачи Коши для системы интегро-дифференциальных уравнений Фредгольма, краевых условий и условий непрерывности решения в точках разбиения. Установлены условия однозначной разрешимости рассматриваемой задачи в терминах разрешимости системы алгебраических уравнений относительно параметров.*

Математические модели, базирующиеся на дифференциальных и интегро-дифференциальных уравнениях, широко используются при описании различных процессов и объектов. Важное прикладное значение имеют задачи, в которых требуется определить неизвестные параметры математической модели по дополнительной информации о решении дифференциального или интегро-дифференциального уравнения, получаемой из эксперимента [1-8]. Задачи такого типа также трактуют как задачу управления или обратной задачей для дифференциальных и интегро-дифференциальных уравнений [9-13]. В связи с этим вопросы

существования, единственности и устойчивости решения задач с параметрами весьма существенны для разработки численных методов идентификации параметров математических моделей, описываемых дифференциальными и интегро-дифференциальными уравнениями [1-13].

В предлагаемой работе исследуется двухточечная краевая задача с параметром для интегро-дифференциального уравнения Фредгольма с вырожденным ядром. С помощью введения дополнительных параметров [14] как значений искомого решения в некоторых точках отрезка  $[0, T]$ , где рассматривается задача, исходная задача сводится к эквивалентной задаче, состоящей из специальной задачи Коши для системы интегро-дифференциальных уравнений Фредгольма, краевых условий и условий непрерывности решения в точках разбиения. Используя интегральное уравнение, эквивалентное специальной задаче Коши для системы интегро-дифференциальных уравнений Фредгольма, и свойство вырожденности ядра интегрального слагаемого, получено представление решения специальной задачи Коши через введенные параметры при предположении обратимости некоторой матрицы. С учетом этого представления из краевых условий и условий непрерывности решения построена система алгебраических уравнений относительно параметров. Установлены условия однозначной разрешимости рассматриваемой задачи в терминах разрешимости системы алгебраических уравнений относительно параметров.

Рассматривается линейная двухточечная краевая задача с параметром для интегро-дифференциального уравнения Фредгольма

$$\frac{dx}{dt} = A(t)x + \sum_{k=1}^m \int_0^T \varphi_k(t) \psi_k(s) ds + B(t)\mu + f(t), \quad x, \mu \in R^n, \quad t \in (0, T), \quad (1)$$

$$B_0\mu + B_1x(0) + B_2x(T) = x^0, \quad x^0 \in R^n, \quad (2)$$

$$C_0\mu + C_1x(0) + C_2x(T) = x^1, \quad x^1 \in R^n, \quad (3)$$

где  $(n \times n)$ -матрицы  $A(t), B(t), \varphi_k(t), \psi_k(s), k = \overline{1, m}$ , и  $n$ -вектор  $f(t)$  непрерывны на  $[0, T]$ ,  $B_i, C_i$  - постоянные  $(n \times n)$ -матрицы,  $i = 0, 1, 2$ ,  $\|x\| = \max_{i=1, n} |x_i|$ .

Решением задачи (1)--(3) является пара  $(x^*(t), \mu^*)$ , где непрерывная на  $[0, T]$  и непрерывно дифференцируемая на  $(0, T)$  функция  $x^*(t)$ , удовлетворяет интегро-дифференциальному уравнению (1) при  $\mu = \mu^*$  и краевым условиям (2), (3).

Двухточечная краевая задачи для линейной системы интегро-дифференциальных уравнений Фредгольма рассматривалась в работах [15-20]. Установлены условия однозначной разрешимости исследуемой задачи в терминах фундаментальной матрицы дифференциальной части [15], в терминах аппроксимирующих краевых задач для нагруженных дифференциальных уравнений [16], в терминах регулярных разбиений [17-19]. В работах [20-21] изучалась двухточечная краевая задача с параметром для системы обыкновенных дифференциальных уравнений методом параметризации. Были установлены критерии корректной разрешимости рассматриваемой задачи в терминах исходных данных и предложены алгоритмы нахождения ее приближенных решений. Результаты для случая  $B_0 = B_2 = 0, C_0 = C_1 = 0, B_1 = C_2 = 1$  анонсированы в [22].

Для решения задачи с параметром (1)—(3) используется подход, разработанный в [23], на основе алгоритмов метода параметризации и численных методов решения задач Коши. Схема метода. Берутся точки  $t_0=0 < t_1 < \dots < t_{N-1} < t_N=T$  и осуществляется разбиение интервала  $[0, T]$  на  $N$  подинтервалов:  $[0, T] = \bigcup_{r=1}^N [t_{r-1}, t_r)$ , которое обозначим через  $\Delta_N$  [22]. Сужение функции  $x(t)$  на  $r$ -ый интервал  $[t_{r-1}, t_r)$  обозначим через  $x_r(t)$ , т.е.  $x_r(t) = x(t)$  для  $t \in [t_{r-1}, t_r)$ ,  $r = \overline{1, N}$ .

Вводятся следующие пространства:

$C([0, T], R^r)$  - пространство непрерывных на  $[0, T]$  функций  $x: [0, T] \rightarrow R^r$  с нормой  $\|x\|_1 = \max_{t \in [0, T]} \|x(t)\|$ ,

$C([0, T], \Delta_N, R^{nN})$  - пространство систем функций  $x[t] = (x_1(t), x_2(t), \dots, x_N(t))$ , где  $x_r: [t_{r-1}, t_r) \rightarrow R^r$  непрерывны на  $[t_{r-1}, t_r)$  и имеют конечные левосторонние пределы  $\lim_{t \rightarrow t_r-0} x_r(t)$  для всех  $r = \overline{1, N}$ , с нормой  $\|x[\cdot]\|_2 = \max_{r=\overline{1, N}} \sup_{t \in [t_{r-1}, t_r)} \|x_r(t)\|$ .

Вводятся дополнительные параметры  $\lambda_r = x_r(t_{r-1})$ ,  $r = \overline{1, N}$ , обозначим через  $\lambda_{N+1} = \mu$ , затем на каждом  $r$ -ом интервале  $[t_{r-1}, t_r)$  производится следующая замена функции  $x_r(t)$ :

$$x_r(t) = u_r(t) + \lambda_r, \quad t \in [t_{r-1}, t_r), \quad r = \overline{1, N}, \quad \text{и} \quad \mu = \lambda_{N+1}.$$

В результате получим линейную краевую задачу с параметрами

$$\frac{du}{dt} = A(t)(u_r + \lambda_r) + \sum_{j=1}^N \sum_{k=1}^m \int_{t_{j-1}}^{t_j} \varphi_k(t) \psi_k(s) (u_j(s) + \lambda_j) ds + B(t) \lambda_{N+1} + f(t), \quad t \in [t_{r-1}, t_r), \quad (4)$$

$$u_r(t_{r-1}) = 0, \quad r = \overline{1, N}, \quad (5)$$

$$B_0 \lambda_{N+1} + B_1 \lambda_1 + B_2 \lim_{t \rightarrow T-0} u_N(t) + B_2 \lambda_N = x^0, \quad (6)$$

$$C_0 \lambda_{N+1} + C_1 \lambda_1 + C_2 \lim_{t \rightarrow T-0} u_N(t) + C_2 \lambda_N = x^1, \quad (7)$$

$$\lambda_s + \lim_{t \rightarrow t_s-0} u_s(t) = \lambda_{s+1}, \quad s = \overline{1, N-1}, \quad (8)$$

где соотношения (8) является условием склеивания решения в точках разбиения отрезка  $[0, T]$ . Заметим, что условия (8) и интегро-дифференциальные уравнения (4) также обеспечивают непрерывность производных решения в этих точках.

Решением задачи с параметрами (4)—(8) является пара  $(u^*[t], \mathcal{L}^*)$ , где функция  $u^*[t] = (u_1^*(t), u_2^*(t), \dots, u_N^*(t)) \in C([0, T], \Delta_N, R^{nN})$  с непрерывно дифференцируемыми на  $[t_{r-1}, t_r)$  компонентами  $u_r^*(t)$ , и  $\mathcal{L}^* = (\lambda_1^*, \lambda_2^*, \dots, \lambda_N^*, \lambda_{N+1}^*) \in R^{(N+1)}$ , удовлетворяет интегро-дифференциальному уравнению с параметрами (4), начальным условиям (5), соотношениям (6)--(8) при  $\lambda_j = \lambda_j^*$ ,  $j = \overline{1, N+1}$ .

Если пара  $(x^*(t), \mu^*)$  является решением задачи (1)--(3), то пара  $(u^*[t], \lambda^*)$  с элементами  $u^*[t] = (u_1^*(t), u_2^*(t), \dots, u_N^*(t)) \in C([0, T], \Delta_N, R^N)$ ,  $\lambda^* = (\lambda_1^*, \lambda_2^*, \dots, \lambda_N^*, \lambda_{N+1}^*) \in R^{(N+1)}$ , где  $\lambda_r^* = x_r^*(t_{r-1})$ ,  $u_r^*(t) = x_r^*(t) + x_r^*(t_{r-1})$ ,  $t \in [t_{r-1}, t_r)$ ,  $r = \overline{1, N}$ ,  $\lambda_{N+1}^* = \mu^* \in R^n$ , будет решением задачи (4)--(8). И обратно, если пара  $(\tilde{u}[t], \tilde{\lambda})$  с элементами  $\tilde{u}[t] = (\tilde{u}_1(t), \tilde{u}_2(t), \dots, \tilde{u}_N(t)) \in C([0, T], \Delta_N, R^N)$ ,  $\tilde{\lambda} = (\tilde{\lambda}_1, \tilde{\lambda}_2, \dots, \tilde{\lambda}_N, \tilde{\lambda}_{N+1}) \in R^{(N+1)}$ , является решением задачи (4)--(8), то пара  $(\tilde{x}(t), \tilde{\mu})$ , определяемая равенствами

$$\tilde{x}(t) = \tilde{u}_r(t) + \tilde{\lambda}_r, \quad t \in [t_{r-1}, t_r), \quad r = \overline{1, N}, \quad \tilde{x}(T) = \lim_{t \rightarrow T-0} \tilde{u}_N(t) + \tilde{\lambda}_N \quad \text{и} \quad \tilde{\mu} = \tilde{\lambda}_{N+1},$$

будет решением исходной краевой задачи с параметром (1)--(3).

Используя фундаментальную матрицу  $X_r(t)$  дифференциального уравнения

$$\frac{dx}{dt} = A(t)x, \quad t \in [t_{r-1}, t_r), \quad r = \overline{1, N},$$

сведем решение специальной задачи Коши для интегро-дифференциального уравнения с параметрами (4), (5) к эквивалентной системе интегральных уравнений

$$\begin{aligned} u_r(t) = & X_r(t) \int_{t_{r-1}}^t X_r^{-1}(\tau) A(\tau) d\tau \lambda_r + X_r(t) \int_{t_{r-1}}^t X_r^{-1}(\tau) \sum_{j=1}^N \sum_{k=1}^m \int_{t_{j-1}}^{t_j} \varphi_k(\tau) \psi_k(s) [u_j(s) + \lambda_j] ds d\tau + \\ & + X_r(t) \int_{t_{r-1}}^t X_r^{-1}(\tau) B(\tau) d\tau \lambda_{N+1} + X_r(t) \int_{t_{r-1}}^t X_r^{-1}(\tau) f(\tau) d\tau, \quad t \in [t_{r-1}, t_r), \end{aligned} \quad (9)$$

Пусть  $\xi_k = \sum_{j=1}^N \int_{t_{j-1}}^{t_j} \psi_k(s) u_j(s) ds$  и перепишем систему интегральных уравнений (9)

в следующем виде

$$\begin{aligned} u_r(t) = & \sum_{k=1}^m X_r(t) \int_{t_{r-1}}^t X_r^{-1}(\tau) \varphi_k(\tau) \xi_k d\tau + \\ & + X_r(t) \int_{t_{r-1}}^t X_r^{-1}(\tau) \left[ A(\tau) \lambda_r + \sum_{j=1}^N \sum_{k=1}^m \int_{t_{j-1}}^{t_j} \varphi_k(\tau) \psi_k(s) \lambda_j ds + B(\tau) \lambda_{N+1} + f(\tau) \right] d\tau, \quad t \in [t_{r-1}, t_r), \\ & r = \overline{1, N}. \end{aligned} \quad (10)$$

Умножив обе части (10) на  $\psi_p(t)$ , интегрируя на  $[t_{r-1}, t_r]$  и суммируя по  $r$ , получим систему линейных алгебраических уравнений относительно  $\xi = (\xi_1, \xi_2, \dots, \xi_m) \in R^m$

$$\xi_p = \sum_{k=1}^m G_{p,k}(\Delta_N) \xi_k + \sum_{r=1}^{N+1} V_{p,r}(\Delta_N) \lambda_r + g_p(f, \Delta_N), \quad p = \overline{1, m}, \quad (11)$$

с  $(n \times n)$  - матрицами

$$\begin{aligned}
 G_{p,k}(\Delta_N) &= \sum_{r=1}^N \int_{t_{r-1}}^{t_r} \psi_p(\tau) X_r(\tau) \int_{t_{r-1}}^{\tau} X_r^{-1}(s) \varphi_k(s) ds d\tau, \\
 V_{p,r}(\Delta_N) &= \int_{t_{r-1}}^{t_r} \psi_p(\tau) X_r(\tau) \int_{t_{r-1}}^{\tau} X_r^{-1}(s) A(s) ds d\tau + \\
 &+ \sum_{j=1}^N \sum_{k=1}^m \int_{t_{j-1}}^{t_j} \psi_p(\tau) X_r(\tau) \int_{t_{j-1}}^{\tau} X_r^{-1}(\tau_1) \varphi_k(\tau_1) d\tau_1 d\tau \int_{t_{r-1}}^{t_r} \psi_k(s) ds, \quad r = \overline{1, N}, \\
 V_{p,N+1}(\Delta_N) &= \sum_{r=1}^N \int_{t_{r-1}}^{t_r} \psi_p(\tau) X_r(\tau) \int_{t_{r-1}}^{\tau} X_r^{-1}(s) B(s) ds d\tau
 \end{aligned}$$

и  $n$  - векторами

$$g_p(f, \Delta_N) = \sum_{r=1}^N \int_{t_{r-1}}^{t_r} \psi_p(\tau) X_r(\tau) \int_{t_{r-1}}^{\tau} X_r^{-1}(s) f(s) ds d\tau, \quad p = \overline{1, m}, \quad k = \overline{1, m}$$

Используя матрицы  $G_{p,k}(\Delta_N)$ ,  $V_{p,r}(\Delta_N)$  составим матрицы  $G(\Delta_N) = (G_{p,k}(\Delta_N))$ ,  $p, k = \overline{1, m}$ ,  $V(\Delta_N) = (V_{p,r}(\Delta_N))$   $p = \overline{1, m}$ ,  $r = \overline{1, N+1}$ .

Систему (11) запишем в виде

$$[I - G(\Delta_N)] \xi = V(\Delta_N) \lambda + g(f, \Delta_N), \quad (12)$$

где  $I$  — единичная матрица размерности  $nm$   
 $g(f, \Delta_N) = (g_1(f, \Delta_N), g_2(f, \Delta_N), \dots, g_m(f, \Delta_N)) \in R^{nm}$ .

Специальная задача Коши (4), (5) эквивалентна системе интегральных уравнений (9). Эта система в силу вырожденности ядра будет эквивалентна системе алгебраических уравнений (11) относительно  $\xi = (\xi_1, \xi_2, \dots, \xi_m) \in R^{nm}$ . Вопросы однозначной разрешимости специальной задачи Коши исследованы в [15, 16]. Установлено, что при достаточно малом шаге  $h > 0$ :  $Nh = T$  разбиения отрезка  $[0, T]$  специальная задача Коши будет однозначно разрешима.

Представим матрицу  $[I - G(\Delta_N)]^{-1}$  в виде  $[I - G(\Delta_N)]^{-1} = (M_{k,p}(\Delta_N))$ ,  $k, p = \overline{1, m}$  где  $M_{k,p}(\Delta_N)$  - квадратная матрица размерности  $n$ .

Тогда согласно (12) элементы вектора  $\xi \in R^{nm}$  определяются равенствами

$$\xi_k = \sum_{j=1}^{N+1} \left( \sum_{p=1}^m M_{k,p}(\Delta_N) V_{p,j}(\Delta_N) \right) \lambda_j + \sum_{p=1}^m M_{k,p}(\Delta_N) g_p(f, \Delta_N), \quad k = \overline{1, m}. \quad (13)$$

В (10) вместо  $\xi_k$  подставляя правую часть (13) получим представление функции  $u_r(t)$  через  $\lambda_r$ ,  $r = \overline{1, N+1}$ :

$$u_r(t) = \sum_{j=1}^N \left\{ \sum_{k=1}^m X_r(t) \int_{t_{r-1}}^t X_r^{-1}(\tau) \varphi_k(\tau) d\tau \left[ \sum_{p=1}^m M_{k,p}(\Delta_N) V_{p,j}(\Delta_N) + \int_{t_{j-1}}^{t_j} \psi_k(s) ds \right] \right\} \lambda_j +$$

$$\begin{aligned}
& + X_r(t) \int_{t_{r-1}}^t X_r^{-1}(\tau) A(\tau) d\tau \lambda_r + X_r(t) \int_{t_{r-1}}^t X_r^{-1}(\tau) \left[ \sum_{k=1}^m \varphi_k(\tau) \sum_{p=1}^m M_{k,p}(\Delta_N) V_{p,N+1}(\Delta_N) + B(\tau) \right] d\tau \lambda_{N+1} + \\
& + X_r(t) \int_{t_{r-1}}^t X_r^{-1}(\tau) \left[ \sum_{k=1}^m \varphi_k(\tau) \sum_{p=1}^m M_{k,p}(\Delta_N) g_p(f, \Delta_N) + f(\tau) \right] d\tau, \quad t \in [t_{r-1}, t_r), \quad r = \overline{1, N}.
\end{aligned} \tag{14}$$

Введем обозначения

$$D_{r,j}(\Delta_N) = \sum_{k=1}^m X_r(t_r) \int_{t_{r-1}}^t X_r^{-1}(\tau) \varphi_k(\tau) d\tau \left[ \sum_{p=1}^m M_{k,p}(\Delta_N) V_{p,j}(\Delta_N) + \int_{t_{j-1}}^t \psi_k(s) ds \right], \quad r \neq j,$$

$$r, j = \overline{1, N},$$

$$D_{r,r}(\Delta_N) = \sum_{k=1}^m X_r(t_r) \int_{t_{r-1}}^t X_r^{-1}(\tau) \varphi_k(\tau) d\tau \left[ \sum_{p=1}^m M_{k,p}(\Delta_N) V_{p,r}(\Delta_N) + \int_{t_{j-1}}^t \psi_k(s) ds \right] + X_r(t_r) \int_{t_{r-1}}^t X_r^{-1}(\tau) A(\tau) d\tau$$

$$D_{r,N+1}(\Delta_N) = \sum_{k=1}^m X_r(t_r) \int_{t_{r-1}}^t X_r^{-1}(\tau) \varphi_k(\tau) d\tau \sum_{p=1}^m M_{k,p}(\Delta_N) V_{p,N+1}(\Delta_N) + X_r(t_r) \int_{t_{r-1}}^t X_r^{-1}(\tau) B(\tau) d\tau,$$

$$F_r(\Delta_N) = \sum_{k=1}^m X_r(t_r) \int_{t_{r-1}}^t X_r^{-1}(\tau) \varphi_k(\tau) d\tau \sum_{p=1}^m M_{k,p}(\Delta_N) g_p(f, \Delta_N) + X_r(t_r) \int_{t_{r-1}}^t X_r^{-1}(\tau) f(\tau) d\tau,$$

$$r = \overline{1, N}.$$

Тогда из (14) получим

$$\lim_{t \rightarrow x_r - 0} u_r(t) = \sum_{j=1}^{N+1} D_{r,j}(\Delta_N) \lambda_j + F_r(\Delta_N), \quad r = \overline{1, N}. \tag{15}$$

Подставив соответствующие правые части (15) в краевые условия (6), (7) и условия склеивания (8), получим систему линейных алгебраических уравнений относительно параметров  $\lambda_r$ ,  $r = \overline{1, N+1}$ :

$$[B_1 + D_{N,1}(\Delta_N)] \lambda_1 + B_2 \sum_{j=2}^{N+1} D_{N,j}(\Delta_N) \lambda_j + B_2 [I + D_{N,N}(\Delta_N)] \lambda_N + [B_0 + D_{N,N+1}(\Delta_N)] \lambda_{N+1} = x^0 - B_2 F_N(\Delta_N) \tag{16}$$

$$[C_1 + D_{N,1}(\Delta_N)] \lambda_1 + C_2 \sum_{j=2}^{N+1} D_{N,j}(\Delta_N) \lambda_j + C_2 [I + D_{N,N}(\Delta_N)] \lambda_N + [C_0 + D_{N,N+1}(\Delta_N)] \lambda_{N+1} = x^1 - C_2 F_N(\Delta_N) \tag{17}$$

$$[I + D_{s,s}(\Delta_N)] \lambda_s - [I - D_{s,s+1}(\Delta_N)] \lambda_{s+1} + \sum_{j=1}^{N+1} D_{s,j}(\Delta_N) \lambda_j = -F_s(\Delta_N), \quad s = \overline{1, N-1}. \tag{18}$$

Матрицу, соответствующую левой части системы уравнений (16)-(18) обозначим через  $Q(\Delta_N)$  и систему запишем в виде

$$Q_*(\Delta_N)\lambda = -F_*(\Delta_N), \quad \lambda \in R^{n(N+1)}, \quad (19)$$

где  $F_*(\Delta_N) = (-x^0 + B_2 F_N(\Delta_N), -x^1 + C_2 F_N(\Delta_N), F_1(\Delta_N), \dots, F_{N-1}(\Delta_N)) \in R^{n(N+1)}$ .

Пусть  $\alpha = \max_{t \in [0, T]} \|A(t)\|$ ,  $\beta = \max_{t \in [0, T]} \|B(t)\|$ ,  $L_N = \max_{r=1, \overline{N}} (t_r - t_{r-1})$ ,

$$\Phi_m = \max_{r=1, \overline{N}} \int_{t_{r-1}}^{t_r} \sum_{k=1}^m \|\varphi_k(t)\| dt, \quad \Psi_m = \max_{p=1, \overline{m}} \int_0^T \|\psi_p(t)\| dt.$$

Справедлива

**Теорема 1.** Пусть при некотором разбиении  $\Delta_N$  матрицы  $I - G(\Delta_N): R^m \rightarrow R^m$  и  $Q_*(\Delta_N): R^{n(N+1)} \rightarrow R^{n(N+1)}$  обратимы.

Тогда задача с параметром (1)–(3) имеет единственное решение – пару  $(x^*(t), \mu^*)$  для произвольных  $f(t) \in C([0, T], R^n)$ ,  $x^0, x^1 \in R^n$  и справедлива оценка

$$\max(\|x^*\|, \|\mu^*\|) \leq K(m, \Delta_N) \max(\|f\|, \|x^0\|, \|x^1\|), \quad (20)$$

где

$$\begin{aligned} K(m, \Delta_N) = & \left\{ e^{\alpha L_N} \Phi_m \left\| [I - G(\Delta_N)]^{-1} \Psi_m (e^{\alpha L_N} - 1 + e^{\alpha L_N} \Phi_m \Psi_m) + \Psi_m \right\| + e^{\alpha L_N} - 1 + \right. \\ & \left. + e^{\alpha L_N} L_N \beta \left\| \Phi_m \left\| [I - G(\Delta_N)]^{-1} \Psi_m e^{\alpha L_N} + 1 \right\| \left\| [Q_*(\Delta_N)]^{-1} \right\| \cdot [1 + \max(\|B_2\|, \|C_2\|)] \right\} \times \\ & \times \max(\|e^{\alpha L_N} L_N \left\| \Phi_m \left\| [I - G(\Delta_N)]^{-1} \Psi_m \right\| + e^{\alpha L_N} L_N \left\| \Phi_m \left\| [I - G(\Delta_N)]^{-1} \Psi_m e^{\alpha L_N} + 1 \right\| \right\}. \end{aligned}$$

Доказательство. Пусть даны  $f(t) \in C([0, T], R^n)$ ,  $x^0, x^1 \in R^n$  и матрица  $I - G(\Delta_N)$  обратима. Тогда, используя обратимость матрицы  $Q_*(\Delta_N)$ , найдем единственное решение системы линейных алгебраических уравнений (19):

$$\lambda^* = -[Q_*(\Delta_N)]^{-1} F_*(\Delta_N). \quad (21)$$

Решая специальную задачу Коши (4), (5) при  $\lambda = \lambda^*$ , определяем систему функций  $u^*[t] = (u_1^*(t), u_2^*(t), \dots, u_N^*(t)) \in C([0, T], \Delta_N, R^{nN})$ . Обратимость матрицы  $I - G(\Delta_N)$  приводит к существованию единственной системы функций  $u^*[t]$  с элементами  $u_r^*(t)$ , определяемых правой частью (14) при  $\lambda = \lambda^* = (\lambda_1^*, \lambda_2^*, \dots, \lambda_N^*, \lambda_{N+1}^*) \in R^{n(N+1)}$ . Несложно установить, что пара  $(x^*(t), \mu^*)$ , определяемая равенствами

$$x^*(t) = u_r^*(t) + \lambda_r^*, \quad t \in [t_{r-1}, t_r], \quad r = \overline{1, N}, \quad x^*(T) = \lim_{t \rightarrow T-0} u_N^*(t) + \lambda_N^* \quad \text{и} \quad \mu^* = \lambda_{N+1}^*,$$

будет решением задачи с параметром (1)–(3). Единственность решения доказывается методом от противного. Покажем справедливость оценки (20).

Из равенств

$$X_r(t) \int_{t_{r-1}}^t X_r^{-1}(\tau) P(\tau) d\tau = \int_{t_{r-1}}^t P(\tau_1) d\tau_1 + \int_{t_{r-1}}^t A(\tau_1) \int_{t_{r-1}}^{\tau_1} P(\tau_2) d\tau_2 d\tau_1 + \int_{t_{r-1}}^t A(\tau_1) \int_{t_{r-1}}^{\tau_1} A(\tau_2) \int_{t_{r-1}}^{\tau_2} P(\tau_3) d\tau_3 d\tau_2 d\tau_1 + \dots$$



,  $t \in [t_{r-1}, t_r]$ , получим оценки

$$\left\| X_r(t_r) \int_{t_{r-1}}^{t_r} X_r^{-1}(\tau) \varphi_k(\tau) d\tau \right\| \leq e^{\alpha(t_r - t_{r-1})} \int_{t_{r-1}}^{t_r} \|\varphi_k(\tau)\| d\tau, \quad r = \overline{1, N}. \quad (22)$$

Используя обозначения  $g_p(f, \Delta_N)$ ,  $F_r(\Delta_N)$  и (21), установим неравенства

$$\begin{aligned} \|g_p(f, \Delta_N)\| &\leq \sum_{r=1}^N \int_{t_{r-1}}^{t_r} \|\psi_p(\tau)\| \cdot \left\| X_r(\tau) \int_{t_{r-1}}^{\tau} X_r^{-1}(s) f(s) ds \right\| d\tau \leq \sum_{r=1}^N \int_{t_{r-1}}^{t_r} \|\psi_p(\tau)\| d\tau \cdot e^{\alpha L_N} \cdot L_N \cdot \|f\|_1 = \\ &= \int_0^T \|\psi_p(\tau)\| d\tau \cdot e^{\alpha L_N} \cdot L_N \cdot \|f\|_1, \quad p = \overline{1, m}, \end{aligned} \quad (23)$$

$$\|F_r(\Delta_N)\| \leq e^{\alpha(t_r - t_{r-1})} \sum_{k=1}^m \int_{t_{r-1}}^{t_r} \|\varphi_k(\tau)\| d\tau \| [I - G(\Delta_N)]^{-1} \| \max_{p=\overline{1, m}} \|g_p(f, \Delta_N)\| + e^{\alpha(t_r - t_{r-1})} \cdot (t_r - t_{r-1}) \|f\|_1.$$

Из неравенства  $\|F_*(\Delta_N)\| \leq \max(\|x^0\|, \|x^1\|, \max_{r=\overline{1, N}} \|F_r(\Delta_N)\|) [1 + \max(\|B_2\|, \|C_2\|)]$ ,

с учетом обозначений  $D_{r,j}(\Delta_N)$  и на основе оценок (21), (22), получим

$$\begin{aligned} \|F_*(\Delta_N)\| &\leq \max \left\{ 1, e^{\alpha L_N} L_N \left[ e^{\alpha L_N} \max_{r=\overline{1, N}} \int_{t_{r-1}}^{t_r} \sum_{k=1}^m \|\varphi_k(\tau)\| d\tau \| [I - G(\Delta_N)]^{-1} \| \max_{p=\overline{1, m}} \int_0^T \|\psi_p(\tau)\| d\tau \right] \right\} \times \\ &\times [1 + \max(\|B_2\|, \|C_2\|)] \max(\|x^0\|, \|x^1\|, \|f\|_1). \end{aligned} \quad (24)$$

Отсюда и из (21) следует оценка

$$\begin{aligned} \|\lambda^*\| &\leq \| [Q_*(\Delta_N)]^{-1} \| \cdot \|F_*(\Delta_N)\| \leq \| [Q_*(\Delta_N)]^{-1} \| \cdot [1 + \max(\|B_2\|, \|C_2\|)] \times \\ &\times \max \left\{ 1, e^{\alpha L_N} L_N \left[ e^{\alpha L_N} \max_{r=\overline{1, N}} \int_{t_{r-1}}^{t_r} \sum_{k=1}^m \|\varphi_k(\tau)\| d\tau \| [I - G(\Delta_N)]^{-1} \| \max_{p=\overline{1, m}} \int_0^T \|\psi_p(\tau)\| d\tau \right] \right\} \max(\|x^0\|, \|x^1\|, \|f\|_1) \end{aligned} \quad (25)$$

Представление (14) и обозначения  $V_{p,r}(\Delta_N)$  позволяют нам установить оценку

$$\begin{aligned} \|u^*[\cdot]\|_2 &\leq \left\{ e^{\alpha L_N} \max_{r=\overline{1, N}} \int_{t_{r-1}}^{t_r} \sum_{k=1}^m \|\varphi_k(\tau)\| d\tau \| [I - G(\Delta_N)]^{-1} \| \max_{p=\overline{1, m}} \int_0^T \|\psi_p(t)\| dt \times \right. \\ &\times \left( e^{\alpha L_N} - 1 + e^{\alpha L_N} \max_{r=\overline{1, N}} \int_{t_{r-1}}^{t_r} \max_{k=\overline{1, m}} \|\varphi_k(t)\| dt \max_{p=\overline{1, m}} \int_0^T \|\psi_p(t)\| dt + \max_{p=\overline{1, m}} \int_0^T \|\psi_p(t)\| dt \right) + e^{\alpha L_N} - 1 + \\ &\left. + e^{\alpha L_N} L_N \left[ \max_{r=\overline{1, N}} \int_{t_{r-1}}^{t_r} \sum_{k=1}^m \|\varphi_k(\tau)\| d\tau \| [I - G(\Delta_N)]^{-1} \| \max_{p=\overline{1, m}} \int_0^T \|\psi_p(t)\| dt \cdot e^{\alpha L_N} + 1 \right] \cdot \beta \right\} \|\lambda^*\| + \end{aligned}$$

$$+e^{\alpha L_N} L_N \left[ \max_{r=1, N} \int_{t_{r-1}}^{t_r} \sum_{k=1}^m |\varphi_k(\tau)| d\tau \| [I - G(\Delta_N)]^{-1} \| \max_{p=1, m} \int_0^T \|\psi_p(t)\| dt \cdot e^{\alpha L_N} + 1 \right] \cdot \|f\|_1. \quad (26)$$

Используя оценки (25), (26) и соотношения  $\|x^*\|_1 \leq \max_{r=1, N} \|\lambda_r^*\| + \|u^*[\cdot]\|_2$ ,  $\|\mu^*\| \leq \|\lambda_{N+1}^*\|$ , получим оценку (20). Теорема 1 доказана.

Таким образом, Теорема 1 устанавливает условия однозначной разрешимости задачи с параметром (1)–(3) в терминах матриц  $I - G(\Delta_N)$  и  $Q(\Delta_N)$ , соответствующих систем линейных алгебраических уравнений (12) и (19). Оценка (20) дает нам непрерывную зависимость решения задачи с параметром (1)–(3) от правой части системы (1) – функции  $f(t)$ , данных граничных условий  $x^0$ ,  $x^1$ . Другими словами, условия Теоремы 1 обеспечивают корректную разрешимость задачи с параметром (1)–(3) с константой  $K(m, \Delta_N)$ .

Следующим этапом исследования является построение алгоритмов нахождения численного краевой задачи с параметром для интегро-дифференциального уравнения Фредгольма (1)–(3). При численном решении задачи с параметром (1)–(3) будут также использоваться метод Рунге-Кутты 4-го порядка, формула Симпсона и методы нахождения решения систем алгебраических уравнений (12) и (19).

*Работа выполнена в рамках проекта № AP05132455 по грантовому финансированию МОН РК на 2018-2020 гг.*

### Литература

1. Nesterenko, O.B. Iteration method for the solution of integro-differential equations with constraints / O.B. Nesterenko // *Nonlinear Oscillations*. 2007. - Vol. 10, No. 3. - P. 339-350.
2. Luchka, A. Yu., Nesterenko, O.B. Projection method for the solution of integro-differential equations with restrictions and control / A. Yu. Luchka, O.B. Nesterenko // *Nonlinear Oscillations*. 2008. –Vol. 11, No. 2. - P. 219-228.
3. Luchka, A. Yu., Nesterenko, O.B. Construction of solution of integro-differential equations with restrictions and control by projection-iterative method / A. Yu. Luchka, O.B. Nesterenko // *Nonlinear Oscillations*. 2009. –Vol. 12, No. 1. - P. 85-93.
4. Nesterenko, O.B. Modified projection-iterative method for weakly nonlinear integro differential equations with parameters / O.B. Nesterenko // *Journal of Mathematical Sciences*. 2014. - Vol. 198, No. 3. - P. 328-335.
5. Dauylbaev, M.K. The asymptotic behavior of solutions to singularly perturbed nonlinear integro-differential equations / M.K. Dauylbaev // *Siberian Mathematical Journal*. 2000. - Vol. 41, No. 1. - P. 49-60.
6. Dauylbayev, M.K., Atakhan, N. The initial jumps of solutions and integral terms in singular BVP of linear higher order integro-differential equations / M.K. Dauylbaev, N. Atakhan // *Miskolc Math. Notes*. Hungary, 2015. - Vol. 16, No. 2. - P. 747-761.
7. Dauylbaev, M.K., Mirzakulova, A.E. Asymptotic behavior of solutions of singular integro-differential equations / M.K. Dauylbaev, A.E. Mirzakulova // *Journal of Discontinuity, Nonlinearity, and Complexity*. - L&H Scientific Publishing: USA. 2016. – Vol. 5, No. 2. - P. 147-154.

8. Dauylbaev, M.K., Mirzakulova, A.E. Boundary-value problems with initial jumps for singularly perturbed integrodifferential equations / M.K. Dauylbaev, A.E. Mirzakulova // *Journal of Mathematical Sciences*. 2017. – Vol. 222, No. 3. - P. 214-225.
9. Akhmetov, M.U., Zafer, A., Sejlilova, R.D. The control of boundary value problems for quasilinear impulsive integro-differential equations / M.U. Akhmetov, A. Zafer, R.D. Sejlilova // *Nonlinear Analysis: Theory, Methods & Applications*. 2002. –Vol. 48, No. 2. – P.271-286.
10. Akhmet, M.U., Kirane, M., Tleubergenova, M.A. and Weber, G.W. Control and optimal response problems for quasi-linear impulsive integro-differential equations / M.U. Akhmet, M. Kirane, M.A. Tleubergenova, and G.W. Weber // *European Journal of Operational Research*. 2006. – Vol. 169, - P. 1128-1147.
11. Alimhan, K., Otsuka, N., Adamov, A.A., and Kalimoldayev, M.N. Global Practical Output Tracking of Inherently Nonlinear Systems Using Continuously Differentiable Controllers / K. Alimhan, N. Otsuka, A.A. Adamov, and M.N. Kalimoldayev // *Mathematical Problems in Engineering*. 2015. – Vol. 2015, No. 932097. – P. 1-10.
12. Romanov, V. G., *Inverse Problems of Mathematical Physics* / V. G. Romanov. - Utrecht: VNU Science Press, 1987. – 239 p.
13. Neto, F.D.M., Neto, A.J.S. An introduction to inverse problems with applications / F.D.M. Neto, A.J.S. Neto. – Heidelberg: Springer, 2013. — 255 p.
14. Dzhumabayev, D.S. Criteria for the unique solvability of a linear boundary-value problem for an ordinary differential equation / D.S. Dzhumabayev // *U.S.S.R. Computational Mathematics and Mathematical Physics*. 1989. – Vol. 29, No. 1. – P. 34-46.
15. Dzhumabaev, D.S. A method for solving the linear boundary value problem for an integro-differential equation / D.S. Dzhumabaev // *Computational Mathematics and Mathematical Physics*. 2010. – Vol. 50, No. 7. – P. 1150-1161.
16. Dzhumabaev, D.S. An algorithm for solving a linear two-point boundary value problem for an integrodifferential equation / D.S. Dzhumabaev // *Computational Mathematics and Mathematical Physics*. 2013. – Vol. 53, No. 6. – P. 736-758.
17. Dzhumabaev, D.S., Bakirova, E.A. Criteria for the unique solvability of a linear two-point boundary value problem for systems of integro-differential equations / D.S. Dzhumabaev, E.A. Bakirova // *Differential Equations*. 2013. – Vol. 49, No. 9. – P. 1087-1102.
18. Dzhumabaev, D.S. Necessary and sufficient conditions for the solvability of linear boundary-value problems for the Fredholm integrodifferential equations / D.S. Dzhumabaev // *Ukrainian Mathematical Journal*. 2015. – Vol. 66, No. 8. – P. 1200-1219.
19. Dzhumabaev, D.S. Solvability of a linear boundary value problem for a Fredholm integro-differential equation with impulsive inputs / D.S. Dzhumabaev // *Differential Equations*. 2015. – Vol. 51, No. 9. – P. 1180-1196.
20. Минглибаева, Б.Б. Коэффициентные признаки однозначной разрешимости линейных двухточечных краевых задач с параметром / Б.Б. Минглибаева // *Математический журнал*. 2003. - Т. 3, No. 2. - С. 55-62.
21. Джумабаев, Д.С., Минглибаева, Б.Б. Корректная разрешимость линейной двухточечной краевой задачи с параметром / Д.С. Джумабаев, Б.Б. Минглибаева // *Математический журнал*. 2004. - Т. 4, No. 1. - С. 41-51.

22. Джумабаев, Д.С., Бакирова, Э.А. Решение линейной краевой задачи для интегро-дифференциального уравнения Фредгольма с параметром / Д.С. Джумабаев, Э.А. Бакирова // Известия МКТУ им. Х.А.Ясави. Серия математика, физика, информатика. Спец. выпуск по материалам конференции «Актуальные проблемы математики», 27-28 апреля 2018г. Туркестан. 2018. – Т. I, No. 1(4). – С. 35-39.

23. Dzhumabaev, D.S. On one approach to solve the linear boundary value problems for Fredholm integro-differential equations / D.S. Dzhumabaev // Journal of Computational and Applied Mathematics. 2016. – Vol. 294, - P. 342-357.

УДК 517.9; 532.5

## ПРОБЛЕМЫ ПРОБЛЕМ УСРЕДНЕНИЯ НЕРАВНОВЕСНОГО ТЕЧЕНИЯ В ПОРИСТЫХ СРЕДАХ

Байшемиров Ж.Д.<sup>1,2</sup>, Жанбырбаев А.Б.<sup>1,2</sup>, Бердышева Д.<sup>1</sup>

<sup>1</sup>Институт информационных и вычислительных технологий КН МОН РК

<sup>2</sup>Казахский национальный педагогический университет имени Абая

**Аннотация.** В данной работе мы даем аналитический обзор проблем усреднения неравновесного течения в пористых средах и методы их развития. Приведены модели: модель релаксации самораспространения фаз в пространстве; модель релаксации активных и пассивных фаз; модель релаксации с переносом интерфейсов и модель релаксации потока в гетерогенной среде.

**Ключевые слова:** математическая модель, пористая среда, капиллярное давление, проницаемость, насыщенность.

### Введение

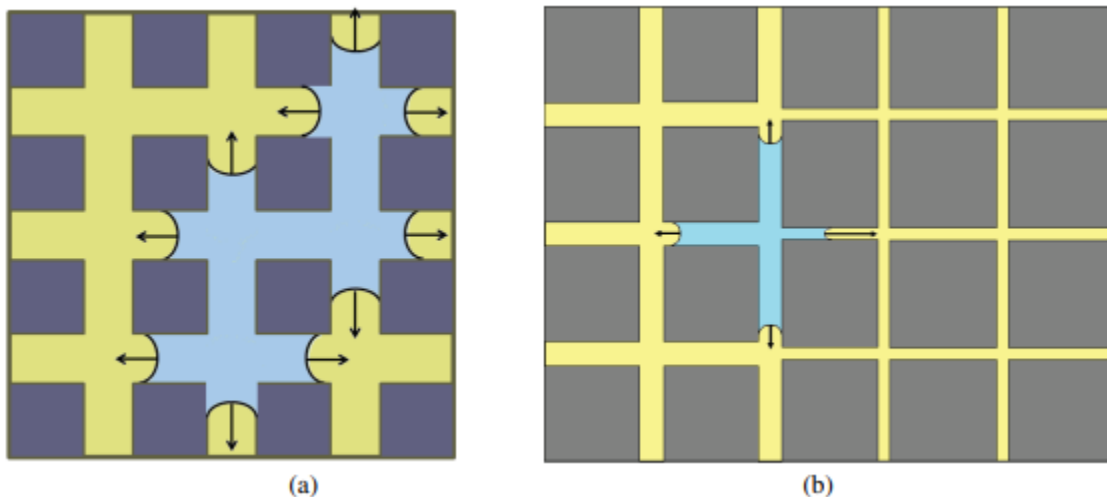
Известно, что классическая модель двухфазного потока в пористых средах, основанная на концепции относительной проницаемости (ОП) и эффективного капиллярного давления (ПК), недостаточна для описания сложных явлений, таких как поток в гетерогенных резервуарах, поток с гетерогенной смачиваемостью и другие. Основные критики касаются того факта, что капиллярные процессы в тонких порах очень медленные, и любое резкое нарушение двухфазной системы вызывает фазовое перераспределение в пространстве, которое, как ожидается, будет медленным, тогда как теория считает его мгновенным. Вот почему концепция капиллярного неравновесия играет одну из самых важных ролей в новых подходах, пытавшихся улучшить классическую теорию, которые сейчас находятся в быстром развитии. Такое резкое и сильное возмущение, которое обязательно провоцирует пространственное фазовое перераспределение, представляет собой сильную неоднородность среды, например чередование плотных пористых пород («блоков») и высокопроницаемых каналов («перелом»). Таким образом, ожидается, что сильная неоднородность среды спровоцирует капиллярную неравновесность. Истинные геологические пористые среды имеют еще более сложную структуру, которая является многомасштабной (чередование нанопор, микропор, трещин и трещины и перелом), так что каждый

уровень гетерогенности может вызывать неравновесность. Взаимодействие неравновесных эффектов в различных масштабах является открытой научной проблемой.

Концепция капиллярного неравновесности была впервые введена Баренблаттом [1] в рамках феноменологического подхода и в дальнейшем была развита рядом авторов. Как уже упоминалось, она тесно связана с концепцией капиллярного движения, самораспространения фаз в пространстве и средней гетерогенностью. Феноменологические подходы изучают только слабую неравновесность, которая может быть описана некоторыми релаксационными уравнениями для капиллярного давления. Однако даже в этом случае феноменологическая теория неспособна ответить на вопрос, как нелинейность двухфазного потока влияет на процесс релаксации и на его параметры. Случай с высокой неравновесностью следует анализировать по-другому.

Для двухфазной жидкости интерфейсы между фазами генерируют поле поверхностных сил, называемое капиллярными силами, которые действуют на обе жидкости. Повышение свободной поверхности воды в тонкой вертикальной стеклянной трубке, погруженной в бассейн, является классическим примером, иллюстрирующим эти эффекты, вызванные капиллярными силами. Эта сила пропорциональна кривизне интерфейса.

В тонкой поре две жидкие фазы создают особый тип интерфейсов, показанных на рис.1, и называются менисками. Это сегменты сферы, кривизна которой обратно пропорциональна радиусу трубы. Таким образом, чем тоньше поры, тем выше капиллярная сила, приложенная к мениску. Равновесная форма фазового кластера определяется равновесием между капиллярными силами, применяемыми ко всем менискам. В однородной среде, показанной на рис.1а (две фазы - синие и желтые), любая воображаемая форма фазового кластера находится в равновесии (как видно, результирующие горизонтальные и вертикальные силы равны нулю).



**Рисунок 1: (а) Капиллярное равновесие в однородной пористой среде и (б) неравновесность в гетерогенной среде (две жидкости - синие и желтые)**

Нарушение может вытеснять жидкости, но не нарушает равновесие капиллярных сил. Следовательно, все возможные геометрические формы синего или желтого кластера механически идентичны. Поэтому макроскопические гидродинамические

параметры жидкостей зависят только от относительного фазового объема (фазового насыщения), но не от формы кластера. Следовательно, классические характеристические функции двухфазного потока (ОП и ПК) независимы от времени и скорости.

Однако гетерогенная пористая среда вызывает капиллярную неравновесность. Среда, показанная на рис.1б, состоит из больших и узких труб. Как видно, синяя жидкость подвергается результирующей силе, которая толкает ее вправо, пока не достигнет состояния равновесия. Поэтому релаксация в равновесное состояние в пористой среде представляет собой движение обеих жидкостей, что провоцирует их пространственное самораспределение. В ходе такого процесса жидкие кластеры постоянно изменяют свою геометрию и их переделку между узкими и большими порами. Это, безусловно, оказывает влияние на макроскопические гидродинамические параметры обеих жидкостей, которые также должны постоянно меняться во времени. Уже в 1950-х годах было известно, что геометрия фазового кластера может изменять характеристики потока при фиксированной фазовой насыщенности [2]. Таким образом, макроскопические характеристики двухфазного потока в гетерогенной среде должны зависеть явно от времени, а не только от насыщения.

Это было подтверждено рядом экспериментальных и численных данных, полученных учеными и инженерами-нефтяниками. Было показано, что в ряде работ, опубликованных много лет назад, основные характерные функции двухфазного потока, такие как ОП и ПК, изменяются во времени с очень медленной стабилизацией: [3], [4], [5]. Эти авторы рассчитали численно ОП и ПК, используя сетевые модели пористых сред. Экспериментально был получен одинаковый результат Топп и другими [6], которые измерили кривые ОП (ПК (S)) для квазистатических, стационарных и динамических условий. Было показано, что капиллярное давление зависит от фазового насыщения  $S$  и времени.

Этот вывод находится в явном противоречии с классической двухфазной теорией, основанной на предположении локального капиллярного равновесия.

Таким образом, капиллярная неравновесность вызывает самораспределение обеих фаз в пространстве. Время капиллярной релаксации в равновесное состояние является характерным временем самораспространения жидкости в пористой сети.

## 2 Модель релаксации самораспространения фаз в пространстве

Первая серия работ, которые разрабатывали зависящие от времени неравновесные капиллярные соотношения, была сделана Баренблаттом [1] с модификациями в [7] и [8]. Авторы согласились с тем, что ОП и ПК зависят от кажущегося насыщения  $\sigma$ , что соответствует равновесному значению, которое может быть достигнуто системой за время  $\tau^*$  после перераспределения фаз в пространстве. Время  $\tau^*$  - это время перераспределения фазы, которое может быть зависимым от насыщения. Разница между истинным насыщением  $S$  и  $\sigma$  определяется только локальной скоростью изменения насыщения  $\partial S / \partial t$ , и эта зависимость считалась линейной:

$$S - \sigma = -\tau^* \frac{\partial S}{\partial t} \quad (1.1)$$

Предполагая, что релаксация  $\tau^*$  достаточно мала (слабая неравновесность), можно получить для ПК или любой другой функции насыщения:

$$p_c(S) = p_c(\sigma) - \tau^* p'_c(\sigma) \frac{\partial S}{\partial t} + \dots$$

или

$$p_c^{dyn} - p_c^{eq} = \tau(S) \frac{\partial S}{\partial t}, \quad (1.2)$$

где  $p_c^{dyn}$  и  $p_c^{eq}$  - неравновесное и равновесное капиллярное давление, и  $\tau$  это новое кинетическое (или релаксационное) время.

Эта фундаментальная связь показывает, что капиллярное давление становится зависящим от скорости или «динамическим».

В работе [8] авторы определили время перераспределения как характерное время перегруппировки фазовых кластеров и последующую модификацию их свойств потока. Значение времени перераспределения зависит от насыщения. Действительно, в пористых средах подвижность фаз зависит непосредственно от кластерной связи. Фаза становится неподвижной, если она теряет связь и трансформируется в систему капель. Тогда при малой насыщенности перераспределение становится невозможным, а время релаксации стремится к бесконечности.

Наконец, по мнению Баренблатта и Виниченко, время релаксации между насыщениями конечных точек примерно обратно пропорционально производной капиллярного давления, близкому к постоянному значению.

Теоретическое обоснование эвристической взаимосвязи (1.2) было получено в нескольких работах с использованием подхода неравновесной термодинамики: сначала Николаевским [9], а затем Марлом [10]. Точнее, связь для производства энтропии была сформулирована как билинейная форма между термодинамическими силами и потоками и предполагала линейные соотношения Онзагера между силами и потоками той же тензорной размерности 3. Развитие уравнения для производства энтропии основано на законах сохранения и нескольких дополнительных предположениях о микроскопическом поведении системы, которые различны для разных авторов. В частности, [11] использовал тот же подход, что и [10], но получил немного другое определение времени релаксации, а также статического капиллярного давления.

Попытки определить время релаксации  $\tau$  были выполнены в [12], [13] и [14], но полученные данные касались одной пористой среды. Используя экспериментальные данные, Стауффер [15] придумал соотношение (1.2). Более того, он предложил формулу для коэффициента пропорциональности, которая считалась постоянной, тогда как  $\tau$  в (1.2) была зависимой от насыщения. В статье [16] исследованы экспериментальные аспекты зависимости капиллярного давления от скорости потока для двухфазного (водно-нефтяного) течения в пористых средах. Он предложил соотношение, аналогичное (1.2). Динамический термин  $\tau \partial_t S$  интерпретировался как вклад вязких сил в падения давления, так как параметр  $\tau$  имеет точно размер вязкости.

Общие принципы построения неравновесных моделей, которые помнят их историю, были изучены Энтовым [17]. Автор выдвинул идеи Баренблатта [1] о несинтезном фазовом перераспределении в пространстве, предполагая, что этот процесс зависит от направления перемещения жидкости. В результате процесс должен зависеть от всей его истории и должен описываться интегральными операторами. Однако в этой статье не была предложена какая-либо конкретная математическая модель. Заметим, что зависимость потока от его истории была объявлена Доддом и

Килем [18], где эффекты длинной памяти были учтены при численном моделировании течения в капиллярных сетях.

### **3 Модель релаксации активных и пассивных фаз**

В ряде работ были использованы другие физические предположения для разработки неравновесной модели. Для каждой фазы предполагалось наличие активной и пассивной (захваченной в поры) компонентов с обменом между ними (т.е. вероятной иммобилизацией активной фазы и наоборот).

Баренблатт и Энтов [19] предположили, что этот процесс обмена неинтересен и его скорость пропорциональна разности между активным и пассивным насыщениями. После математических разработок было показано, что этот подход дает то же формальное математическое соотношение, что и (1.2).

Таким образом, различные физические механизмы приводят к одной и той же математической модели. Это дает более общий смысл параметру  $\tau$ , который отвечает за различные релаксационные процессы фазового перераспределения в пространстве или для обмена между различными псевдофазами.

Экстремальный вариант этой модели был предложен Зарубиным [20], который полностью исключил концепцию ОП и ПК и рассмотрел обменные термины между пассивными и активными компонентами в качестве альтернативы им.

Те же идеи были представлены Булыгиным [21] в терминах динамической эффективной пористости по отношению к активному компоненту. Эта динамическая пористость была экспериментально исследована Плешинским [22], но полученные данные были недостаточными для оценки времени  $\tau$  обмена. Основным недостатком этой модели состоит в том, что предполагается одно и то же время  $\tau$  для мобилизации и захвата фаз. Однако эти два процесса несимметричны.

Другой подход к представлению длинной памяти в макроскопических моделях был разработан Бувичем и Мамбетовым [23] как расширение обменной модели. Вместо пассивных и активных фазовых компонентов каждая фаза представлялась как система нескольких кластеров разной шкалы. Между кластерами существует процесс обмена. Сохранялся только обмен между соседним кластером. Каждый обмен был описан кинетическим законом. Ожидалось, что макромасштабная модель даст интегро-дифференциальные уравнения. Однако была представлена только общая методология.

Более общая модель того же класса была предложена Хильфером в [24], [25], [26], который ввел связанные и несвязанные элементы той же фазы и обмена между ними. Эта концепция более общая, чем модель активных и пассивных фаз. Это позволило автору связать макроскопическую двухфазную теорию течения с теорией перколяции.

### **4 Модель релаксации с переносом интерфейсов**

Как упоминалось выше, одно только насыщение недостаточно для описания геометрии фазового перераспределения в пространстве. Вторым параметром может быть конкретная область интерфейса между фазами,  $A$ . В работе [27] была рассмотрена система с рядом интерфейсов в рамках феноменологического подхода. Была предложена новая идея изучить высокодисперсный интерфейс как индивидуальный континуум и сформулировать для него уравнения баланса массы и импульса. Чтобы гомогенизировать такую систему с множественными разрывами, автор применил технику распределений. В частном случае автор получил известную модель с



поперечной относительной проницаемостью, которая доказала справедливость этого нового подхода.

В работах [28], [29] и [30] авторы ввели два теоретических понятия одновременно: дополнительный континуум Марля, представляющий ансамбль интерфейсов, и концепцию Баренблатта о капиллярной неравновесии. Применяя технику неравновесной термодинамики, авторы получили новую модель двухфазного течения, которая содержала аналогичное соотношение (1.2), но в котором время релаксации и капиллярное давление зависят от насыщения  $S$  и конкретной области  $a$ :

$$P_c^{dyn} - P_c^{stat} = -\tau(S, a) \frac{\partial S}{\partial t}, \quad (1.3)$$

Для конкретной области  $a$ , авторы получили новое уравнение баланса импульса. Эта функция  $\tau$  также изучалась на основе моделирования капиллярной сети. Было показано, что эта функция связана с насыщением, но не уникальным образом. Авторы дали оценки для  $\tau$  систем воздух-вода, а также для нефте-водных систем, интерпретируя собственные данные и данные, полученные несколькими авторами за период с 1967 по 1993 год. Авторы показали, что их модель была без противоречий для других публикаций, основанных на модели релаксации (1.2).

Более развитая форма модели с транспортом интерфейсов была предложена в [31], где система интерфейсов рассматривалась как новый континуум (M-континуум), имеющий конкретные физические свойства и генерирующий векторное поле капиллярных сил, связанных с каждым индивидуальным мениском. Поэтому концепция капиллярности изменила его топологический размер. Такая модель качественно отличается от подхода, связанного с капиллярной неравновесностью.

## 5 Модель релаксации потока в гетерогенной среде

Самораспространение жидкостей в пространстве или обмен между пассивными и активными компонентами жидкости имеют одно и то же происхождение - среднюю гетерогенность. В частности, достаточно рассмотреть среду, состоящую из двух компонентов: высокопроницаемых каналов («трещины») и низкопроницаемых включений («блоки»). Чем выше разность проницаемости блока и трещин, тем выше будет капиллярная неравновесность. Если такая среда имеет периодическую структуру с небольшим периодом (по сравнению с размером домена), ожидается, что будет показано среднее поведение, которое должно включать неравновесные эффекты. Следовательно, общая неравновесная модель может быть получена путем гомогенизации такого типа течения.

Такая модель была разработана Панфиловым [32] и Бурджатом и Панфиловым [33] асимптотической гомогенизацией двухфазного потока в описанной ниже среде. Блоки формировали периодическое поле с периодом  $\varepsilon$ , который был намного меньше, чем макроскопический размер области (среда с двойной пористостью). Степень гетерогенности описывалась двумя параметрами: отношением низкой и высокой проницаемости,  $\omega$  и отношением среднего значения капиллярного давления в блоках по сравнению с величиной в трещинах,  $\delta$ . Случай высокой гетерогенности  $\omega \sim \varepsilon^2$  характеризуется значительным неравномерным неравновесностью, что приводит к появлению длинной памяти. Чтобы избежать проблемы связи между нелинейностью уравнений двухфазного течения и длинной памятью, авторы рассмотрели случай

низкого неравновесного, промежуточного между точками  $\omega \sim 1$  и  $\omega \sim \varepsilon^2$ , более точно, следующим классом сред:

$$\omega \leq 1, \quad \aleph \leq \delta \leq 1, \quad \aleph \equiv \frac{\varepsilon^2}{\omega} \quad (1.4)$$

Для этого случая они вывели полностью гомогенизированную модель формальными асимптотическими разложениями. В частности, было получено следующее неравновесное соотношение между капиллярными давлениями в трещинах и блоках:

$$\frac{1}{\delta} p_c^I(S^I) - p_c^{II}(S^{II}) = -\tau_*(S^I) \frac{\partial S}{\partial t}, \quad (1.5)$$

где  $S^I$  и  $S^{II}$  являются макроскопическая водонасыщенность в блоках и трещинах соответственно, капиллярные давления безразмерны (делятся на среднее значение капиллярного давления в блоках и трещинах соответственно) и порядка 1.

Принимая во внимание тот факт, что процесс в блоках находится в неравновесном отношении по отношению к трещинам, уравнение (1.5), очевидно, идентична (1.2).

Капиллярный релаксационный термин  $\tau^*$  зависит от  $S^I$  и пропорционален к  $\aleph = \varepsilon^2/\omega$ . Он определяется через клеточную проблему, которая имеет явное аналитическое решение.

*Работа выполнена при финансовой поддержке гранта КН МОН РК №AP05132680.*

### Список литературы

1. Баренблатт Г.И. Фильтрация двух несмешивающихся жидкостей в однородной пористой среде // Изв. АН СССР. МЖГ. 1971. - №5. - С.144-151.
2. А. Е. Scheidegger, The Physics of Flow Through Porous Media, (University of Toronto Press, Toronto, 1957).
3. M.J. Blunt, and P. King, Transport in Porous Media, 6, 107 (1991).
4. Ентов В.М., Чен-Син Э. Микромеханика двухфазного течения в пористых средах // Численные методы решения задач фильтрации многофазной несжимаемой жидкости. Новосибирск: ИТПМ, 1987. -С.120-129.
5. A.K. Singhal, and W.H. Somerton, Rev. Inst. Fr. Petrol., 32, 897 (1977).
6. G.C. Topp, A. Klute, and D. B. Peters, Soil. Sci. Am. Proc., 31 312 (1967).
7. Винниченко А. П. К теории нестационарной фильтрации несмешивающихся жидкостей в пористой среде // Изв. АН СССР. МЖГ.- 1978.-№3.-С.57-59.
8. Баренблатт Г.И., Винниченко А.П., УМН. Наука, 3, 35 (1980).
9. Николаевский В.Н., Бондарев Э.А., Миркин М.И. и др. Движение углеводородных смесей в пористой среде. М.: Недра, 1968.
10. С. М Marle, From the pore scale to the macroscale: Equations governing multiphase fluid flow through porous media, in Procs. Euromech143, edited by A. Verruijt and F.D.J. Barends (Balkema, Rotterdam, 1981).
11. F. Kalaydjian, Transport in Porous Media, 2, 537 (1987).
12. Бочаров О. Б., Витовский О. В., Кузнецов В. В. Структура скачка насыщенности при неравновесном вытеснении в пористых средах // МЖГ.- 1990.-№5. - С. 97-104.

13. Бочаров О.Б. Влияние неравновесности на двухфазное течение в пористой среде // Гидродинамика и тепломассообмен в неподвижных зернистых слоях. Новосибирск: Институт теплофизики, 1991. - С.89-95.
14. Кузнецов В.В., Димов С.В. Микроструктура двухфазного течения в пористой среде // Гидродинамика и тепломассоперенос в пористых средах. Новосибирск: институт теплофизики, 1991. - С.71-88.
15. F. Stauffer, Time dependence of the relations between capillary pressure, water content and conductivity during drainage of porous media, in IAHN Symp. on Scale Effects in Porous Media, (Thessaloniki, Greece, 1978).
16. F. Kalaydjian, Society of Petroleum Engineers, 24813, 491 (1992).
17. Ентов В.М. К теории неравновесных эффектов при фильтрации неоднородных жидкостей // Изв. АН СССР. МЖГ. 1980. - №3. - С.52-58.
18. С.G. Dodd, and O.G. Kiel, J. Phys. Chem., 63, 1646 (1959).
19. Баренблатт Г.И., Ентов В.М. Периодическая издания Академия Наук СССР, 33 (1972).
20. Зарубин У.А. . Обменная модель смещения в пористых средах. Известия АН СССР, Механика Жидкости и Газы, (1), 1993.-С. 91-97
21. Булыгин В. В. Гидромеханика нефтяного месторождения. М.: Недра, 1974.
22. Плешинский Б. И., Назаровский Г.А., Гимадиев Г.А., Гимадиев Р.М. Записки Казанского университета, 130, 120 (1970).
23. Бувевич А., Мамбетов У.М., Инженерно-физический журнал, 60, 98 (1991)
24. R. Hilfer, Physical Review E, 58, 2090 (1998).
25. R. Hilfer, and H. Besserer, Old problems and new solutions for multiphase flow in porous media. in: Porous Media: Physics, Models, Simulation, edited by A. Dmitrievsky, and M. Panfilov, (World Scientific, Singapore, 2000).
26. R. Hilfer, and H. Besserer, Physica B, 279, 125 (2000).
27. C. M. Marle, Int. J. Eng. Sci., 20, 643 (1982).
28. S.M. Hassanizadeh, and W.G. Gray, Adv. Water Resources, 13, 169 (1990).
29. S.M. Hassanizadeh, and W.G. Gray, Water Resour. Res., 29, 3389 (1993).
30. M. Panfilov, and I. Panfilova, Transport in porous media, 58, 87 (2005).
31. S.M. Hassanizadeh, and W.G. Gray, Water Resources Research, 29, 3389 (1993).
32. M. Panfilov, ComptesRendus Acad. Sciences, Paris, Ser.IIb, 322, 195 (1996).
33. A. Bourgeat, and M. Panfilov, Computational Geosciences, 2, 191 (1998).

## **КОНФЛИКТНЫЕ СИТУАЦИИ В СИСТЕМЕ ЗАЩИЩЕННОГО ТРАНСГРАНИЧНОГО ОБМЕНА**

**Бегимбаева Е.Е.**

*Институт информационных и вычислительных технологий КН МОН РК,  
Казахстан*

e-mail: *enlik\_89@mail.ru*

***Аннотация.** В статье исследуются возможные конфликтные ситуации в системе защищенного трансграничного обмена. Приведена структурная схема компонентов системы защищенного трансграничного обмена.*

В декабре 2017 года была утверждена государственная программа «Цифровой Казахстан». Одной из задач программы «Цифровой Казахстан» [1] является обеспечение информационной безопасности в сфере ИКТ. В концепции кибербезопасности «Киберцит Казахстан» [2] указано что, одной из ключевых проблем является транснациональный и трансграничный характер многих продуктов ИКТ и международная связанность сетей телекоммуникаций общего пользования, которые используются преступностью в целях совершения противоправных действий в отношении пользователей и операторов ИКТ-услуг и владельцев Интернет-ресурсов, размещенных в национальном сегменте, а также информационных систем, взаимодействующих с Интернетом.

Широкое внедрение цифровых технологий на основе инфокоммуникационных систем привело к необходимости создания защищенных систем обмена электронной информацией. Одна из основных целей организации информационного обмена в электронном виде - гарантированное и своевременное ознакомление с содержанием документов тех лиц, которым они адресованы. Комплекс программно-аппаратных и организационно-правовых решений по разработке системы защищенного информационного обмена должен гарантировать [3]:

- конфиденциальность хранимых, обрабатываемых и передаваемых электронных данных (информации);
- целостность пересылаемой информации;
- наличие и безусловность признания авторства электронного документа;
- надёжность и отказоустойчивость системы информационного обмена.

Решение первых трех требований обеспечивается при помощи средств криптографической защиты информации. Обеспечение конфиденциальности электронных документов обеспечивается при помощи средств шифрования, а целостность и признание авторства обеспечивается при помощи электронной цифровой подписи (ЭЦП) [3], т.е. применение ЭЦП подтверждает авторство (никто, кроме обладателя секретного ключа, не сможет подписать) и целостность (любое изменение документа с подписью приведёт к недействительности ЭЦП).

Именно электронный документ с ЭЦП является доказательной базой при разрешении конфликтной ситуации.



Рисунок - 1 Компоненты системы защищенного трансграничного обмена

Таким образом для обеспечения информационной безопасности казахстанского сегмента защищенного трансграничного информационного взаимодействия необходимо разработать автоматизированную систему защищенного трансграничного обмена электронными документами. Для создания защищенной системы необходим комплексный подход информационной безопасности. На рисунке 1 приведены компоненты, составляющие систему защищенного трансграничного информационного обмена (ТИО):

- системы шифрования и электронной цифровой подписи (ЭЦП);
- система генерации и хранения ключей;
- система разграничения доступа;
- система разрешения конфликтных ситуаций;
- система управления базами данными (СУБД).

Системы шифрования и ЭЦП будут разработаны на основе криптографических алгоритмов шифрования и формирования и проверки ЭЦП соответственно. Эти алгоритмы будут использоваться при шифровании, хранении и передаче электронной информации. В связи с этим необходимы также алгоритмы для генерации секретных ключей для систем шифрования и формирования и проверки ЭЦП.

Объем и поток электронной информации (данных), которая постоянно поступает, накапливается и хранится в рамках одного государственного сегмента и за ее пределами является колоссальным. Для продуктивного использования всего объема информации, обеспечения целостности и доступности, а также контроля качества нужен инструмент управления и хранения. Таким инструментом является СУБД - система управления базами данных.

При отсутствии подобной системы рабочие процессы, задействованные в информационном потоке, значительно усложняются. Вследствие чего могут возникнуть некоторые проблемы, конфликтные ситуации.

При информационном обмене могут по каким либо техническим причинам передаваться искаженные исходные данные. В дальнейшем это часто приводит к негативным последствиям, поскольку при передаче документа далее другому сегменту поступят недостоверные сведения. При этом восстановить корректные исходные данные бывает довольно затруднительно или вовсе невозможно.

В случае использования СУБД такие ошибки возможно отсеять, поскольку предполагается, что вся информация о процессах и данных будет храниться в базах данных.

Поскольку пересылаемые данные хранятся на рабочих компьютерах абонентов в виде набора различных файлов, при передаче абоненты используют более или менее «свежую» или, «свою» версию данных, то, интерпретация этих данных может получиться также различной и неполной.

Также при хранении электронных файлов в базах данных доступ к ним должен быть ранжированным в зависимости от степени доступа субъекта системы к базам данных. Предполагается создать систему, которая позволит настроить права доступа чтения, записи и удаления той или иной информации и разграничить доступ субъектов к объектам сегмента. В Институте информационных и вычислительных технологий имеется необходимый задел для разработки систем разграничения доступа [4,5].

В трансграничном информационном обмене между объектами системы могут возникать различные конфликтные ситуации независимо от формы представления электронной информации. В связи с этим должны быть определены, исследованы и программно реализованы возможные виды конфликтных ситуаций и способы их разрешения [6].

Каждая конфликтная ситуация при трансграничном информационном обмене электронными документами является уникальной. В ТИО можно столкнуться со следующими ситуациями:

### **1. Конфликтные ситуации, связанные с ЭЦП**

Пока нет конфликтов, одни подписывают, другие проверяют и считают подписи действительными. Конфликт возникает в следующих случаях:

- неподтверждение подлинности защищенных электронных документов (ЭД) средствами проверки ЭЦП получателя;
- оспаривание факта идентификации владельца ЭЦП, подписавшего ЭД;
- оспаривание целостности ЭД;
- оспаривание факта отправления и/или получения защищенного ЭД;
- оспаривания времени отправления и/или получения защищенного ЭД;
- оспаривание содержания пересылаемого ЭД;
- оспаривание идентичности экземпляров ЭД и (или) подлинника и копии электронного документа на бумажном носителе;

Разбор конфликтов возможен: а) внутри организации или системы б) с привлечением внешних специалистов в) в судебном порядке.

### **2. Конфликтные ситуации, связанные с СУБД**

Современные СУБД являются многопользовательскими системами, т.е. допускают параллельную одновременную работу большого количества пользователей. При этом пользователи не должны мешать друг другу. Т.к. логической единицей работы для пользователя является транзакция, то работа СУБД должна быть организована так, чтобы у пользователя складывалось впечатление, что их транзакции выполняются независимо от транзакций других пользователей [6].

Простейший и очевидный способ обеспечить такую иллюзию у пользователя состоит в том, чтобы все поступающие транзакции выстраивать в единую очередь и выполнять строго по очереди. Такой способ не годится по очевидным причинам - теряется преимущество параллельной работы. Таким образом, транзакции необходимо выполнять одновременно, но так, чтобы результат был бы такой же, как если бы транзакции выполнялись по очереди. Трудность состоит в том, что если не предпринимать никаких специальных мер, то данные измененных одним пользователем могут быть изменены транзакцией другого пользователя раньше, чем закончится транзакция первого пользователя. В результате, в конце транзакции первый пользователь увидит не результаты своей работы, а неизвестно что.

Часто такая проблема присутствует в казахстанских системах обмена и обработки данных. Например, при покупке онлайн билетов «Қазақстан темир жолы», в случае плохой зоны покрытия интернета при параллельном сеансе двух пользователей, система выдает результаты чужого пользователя.

Итак, анализ проблем параллелизма показывает, что если не предпринимать специальных мер, то при работе в системе нарушается свойство (И) транзакций - изолированность. Транзакции реально мешают друг другу получать правильные результаты.

Однако не всякие транзакции мешают друг другу. Очевидно, что транзакции не мешают друг другу, если они обращаются к разным данным или выполняются в разное время.

В результате конкурирующих транзакции, которые пересекаются по времени и обращаются к одним и тем же данным, между транзакциями возникают конфликты доступа к данным:

– первая транзакция изменила объект и не закончилась. Вторая транзакция пытается изменить этот объект. Результат - потеря обновления.

– первая транзакция прочитала объект и не закончилась. Вторая транзакция пытается изменить этот объект. Результат - несовместимый анализ (неповторяемое считывание).

– первая транзакция изменила объект и не закончилась. Вторая транзакция пытается прочитать этот объект. Результат - чтение "неправильных" данных.

### **3. Конфликтные ситуации, связанные с технической составляющей и программными средствами [8]:**

– недоверие к технологии обмена электронными документами и защиты информации при обмене электронными документами;

– недоверие к используемым структурам и форматам электронного документа;

– недоверие к программно-техническим средствам обеспечения обмена электронными документами и защиты информации при обмене электронными документами.

Указанные ситуации разрешаются путем использования исключительно сертифицированных отечественных систем криптографической защиты информации, в противном случае использование несертифицированных может привести, например, к признанию недействительности ЭЦП и к отмене операции. Аналогично для шифрования - человек, зашифровавший документ, может требовать компенсации за якобы имевший место факт утечки информации, утверждая, что несертифицированная реализация алгоритма шифрования недостаточно надёжна.

Особое место среди конфликтных ситуаций занимают такие ситуации, в которых конфликтующая система может целенаправленно подвергаться атакам, преследующим цель дестабилизировать работу находящейся в конфликте системы путем уменьшения ее надежности. Для повышения надежности и отказоустойчивости работы системы необходимо введение в эту систему и активное использование в процессе конфликта различного вида избыточности, это аппаратную, логическую, информационную, временную. Следовательно, в реальных условиях для успешного участия в конфликтной ситуации защищаемая система трансграничного информационного обмена должна обладать определенным ресурсом для собственной защиты. При этом успешность действий в конфликтной ситуации защищаемой стороны будет определяться выбранной стратегией использования имеющегося ресурса для защиты.

### **Список литературы**

1. Концепция кибербезопасности «Киберщит Казахстан»  
[https://tengrinews.kz/zakon/pravitelstvo\\_respubliki\\_kazahstan\\_premier\\_ministr\\_rk/hozyaystvennaya\\_deyatelnost/id-P1700000407/](https://tengrinews.kz/zakon/pravitelstvo_respubliki_kazahstan_premier_ministr_rk/hozyaystvennaya_deyatelnost/id-P1700000407/)

2. Государственная программа «Цифровой Казахстан»  
[https://primeminister.kz/ru/page/view/gosudarstvennaya\\_programma\\_digital\\_kazahstan](https://primeminister.kz/ru/page/view/gosudarstvennaya_programma_digital_kazahstan)

3. Фатьянов А. А. Правовое регулирование электронного документооборота: учебно-практическое пособие. - М.: Российская газета, 2005. - 200 с.

4. Бияшев Р.Г., Калимолдаев М.Н., Рог О.А. Представление ограничений моделей атрибутного разграничения доступа // Известия НАН РК. Серия физико-математическая. – 2016. - № 1. - С. 58-65.

5. Бияшев Р.Г., Калимолдаев М.Н., Рог О.А. Моделирование семантики типизированного атрибутного разграничения доступа // Проблемы информатики. - 2017, № 1. - С. 25-37.

6. Шокишалов Ж.М., Траншекаралық ақпарат алмасуды ұйымдастыру // Матер. междунар. науч. конф. ИИВТ МОН РК «Современные проблемы информатики и вычислительных технологий». - Алматы, 2016. - С.44-47.

7. Транзакции и параллелизм <http://citforum.ru/database/dblearn/dblearn10.shtml>

8. Разработка и исследование моделей, методов и алгоритмов защищенного трансграничного обмена информации и сближения нормативно-правовой базы для создания и развития информационного пространства стран СНГ: отчет о НИР (промежуточный) / НЦ НТИ: рук. Бияшев Р.Г.; исполн.: Нысанбаева С.Е. - Алматы, 2016. - 138 с. - Инв. № 0216РК00668.

УДК 519.633- 519.688

## ПРИМЕНЕНИЕ ПРЕОБРАЗОВАНИЯ ЛАПЛАСА ДЛЯ СВЕДЕНИЯ УРАВНЕНИЙ ПОРОУПРУГОСТИ К ЭЛИПТИЧЕСКОЙ СИСТЕМЕ УРАВНЕНИЙ

**Бердышев А.С., Блиева Д.Н.**

*Институт информационных и вычислительных технологий КН МОН РК,  
Казахстан*

*e-mail: [berdyshev@mail.ru](mailto:berdyshev@mail.ru), [dana.bliyeva@gmail.com](mailto:dana.bliyeva@gmail.com)*

**Аннотация.** В данной работе рассматривается возможность приведения двумерной динамической задачи для пористых сред, насыщенных жидкостью, к эллиптической системе уравнений с помощью применения преобразования Лапласа. За основу взята линеаризованная модель Доровского В.Н., в которой среда описывается тремя упругими параметрами.

**Ключевые слова:** пористая среда, уравнения пороупругости, интегральное преобразование Лапласа, акустические волны.

### **Введение**

В работе приводится краткое описание и постановка задачи движения сейсмических волн во флюидонасыщенной пористой среде. Именно такая модель является реалистичным представлением пористой среды, состоящей из заполненной вязкой жидкостью упругодеформируемой матрицы. Данная модель позволяет объяснить эффекты, наблюдаемые в процессе сейсмических исследований свойств горных пород при наличии поровой жидкости. В силу необходимости решения практических задач в различных областях геофизики, биомеханики и нефтяной разработки в последние десятилетия получило развитие численное моделирование процесса распространения сейсмических волн во флюидонасыщенных пористых



средах. Для численного моделирования таких процессов обычно используется модель Френкеля-Био [1, 2]. Основным свойством упруго-пористой насыщенной среды в теории Био является наличие дополнительной второй продольной волны, наряду с двумя типами звуковых колебаний: поперечного и продольного. Согласно нелинейной математической модели для пористых сред, построенной в 1981 году В.Н. Доровским [4] на основе общих первых физических принципов, в ней также есть три типа звуковых колебаний: два типа продольных и поперечный. Но в отличие от моделей типа Френкеля-Био линейризованная модель характеризует среду с помощью трех упругих параметров [4, 5]. Эти три упругих параметра взаимнооднозначно определяются с помощью трех скоростей упругих колебаний, что имеет важное значение для численного моделирования распространения упругих волн в пористых средах, при известных значениях пористости, скоростей акустических волн, физических плотностей матрицы и насыщающей жидкости.

Основной целью данной работы является исследование условий существования аналитического решения системы уравнений пороупругости, описывающих движение волн в пористой среде для компонентов скорости, напряжения и порового давления.

В отличие от предыдущих публикаций, где рассматривалась задача в полуплоскости  $x_2 > 0$  будет рассматриваться задача на двух полосах с различными параметрами пористой среды и насыщающей жидкости. Интегральное преобразование Лапласа для производной по времени применяется для решения задач, зависящих от времени, таких как сейсмическое и электромагнитное моделирование, задачи теплопроводности, и т.д.

### Постановка задачи

Перейдем к математической формулировке модели для двумерной динамической задачи. Пусть две полосы находятся в полуплоскости  $x_2 > 0$ , заполненной пористыми средами и насыщенными жидкостями, имеющими различные параметры, характеризующие каждую из полос. Тогда распространение сейсмических волн в данных средах при отсутствии потери энергии описывается следующими начально-краевыми задачами в терминах скорости насыщающей жидкости, скоростей матрицы, давления жидкости и тензора напряжений [5, 7]:

$$\frac{\partial u_i}{\partial t} + \frac{1}{\rho_{0,s}} \frac{\partial \sigma_{ik}}{\partial x_k} + \frac{1}{\rho_0} \frac{\partial p}{\partial x_i} = F_i \quad (1)$$

$$\frac{\partial v_i}{\partial t} + \frac{1}{\rho_0} \frac{\partial p}{\partial x_i} = F_i \quad (2)$$

$$\frac{\partial \sigma_{ik}}{\partial t} + \mu \left( \frac{\partial u_k}{\partial x_i} + \frac{\partial u_i}{\partial x_k} \right) + \left( \frac{\rho_{0,l}}{\rho_0} K - \frac{2}{3} \mu \right) \delta_{ik} \operatorname{div} \vec{u} - \frac{\rho_{0,s}}{\rho_0} K \delta_{ik} \operatorname{div} \vec{w} = 0 \quad (3)$$

$$\frac{\partial p}{\partial t} - (K - \alpha \beta \rho_{0,s}) \operatorname{div} \vec{u} + \alpha \beta \rho_{0,l} \operatorname{div} \vec{w} = 0, \quad (4)$$

$$u_i|_{t=0} = v_i|_{t=0} = \sigma_{ik}|_{t=0} = p|_{t=0} = 0, \quad (5)$$

$$\sigma_{22} + p|_{x_2=0} = \sigma_{12}|_{x_2=0} = \frac{\rho_{0,l}}{\rho_0} p|_{x_2=0} = 0 \quad (6)$$

где индексы  $i, k$  принимают значения  $i=1,2, k=1,2$ , а  $\vec{u}=(u_1, u_2)$  и  $\vec{v}=(v_1, v_2)$ - векторы скорости упругого пористого тела с парциальной плотностью  $\rho_{0,s}$  и жидкости с парциальной плотностью  $\rho_{0,l}$ ,  $p$  - поровое давление,  $\sigma_{ik}$ - компонент тензора напряжений,  $\delta_{i,k}$ - символ Кронекера,  $d_0$  – пористость.

### Получение эллиптической системы уравнений

Для двумерного случая система (1)-(4) состоит из 8 уравнений с 8 неизвестными функциями  $u_1, u_2, v_1, v_2, \sigma_{11}, \sigma_{12}, \sigma_{22}, p$ . Для приведения поставленной задачи к системе эллиптических уравнений применим к системе (1)-(4) интегральное преобразование Лапласа по переменной времени  $t$  вида:

$$\hat{f}(s) = \int_0^{\infty} f(t)e^{-st} dt \quad (7)$$

Используя начальные условия (5), получим эллиптическую систему следующего вида:

$$\frac{1}{\rho_{0,s}} \left( \frac{\partial \hat{\sigma}_{11}}{\partial x_1} + \frac{\partial \hat{\sigma}_{12}}{\partial x_2} \right) + \frac{1}{\rho_0} \frac{\partial \hat{p}}{\partial x_1} + s\hat{u}_1(s, x_1, x_2) = \hat{F}_1 \quad (8)$$

$$\frac{1}{\rho_{0,s}} \left( \frac{\partial \hat{\sigma}_{21}}{\partial x_1} + \frac{\partial \hat{\sigma}_{22}}{\partial x_2} \right) + \frac{1}{\rho_0} \frac{\partial \hat{p}}{\partial x_2} + s\hat{u}_2(s, x_1, x_2) = \hat{F}_2 \quad (9)$$

$$\frac{1}{\rho_0} \frac{\partial \hat{p}}{\partial x_1} + s\hat{v}_1(s, x_1, x_2) = \hat{F}_1 \quad (10)$$

$$\frac{1}{\rho_0} \frac{\partial \hat{p}}{\partial x_2} + s\hat{v}_2(s, x_1, x_2) = \hat{F}_2 \quad (11)$$

$$\left( \frac{4}{3}\mu + \frac{\rho_{0,l}}{\rho_0} K \right) \frac{\partial \hat{u}_1}{\partial x_1} + \left( \frac{\rho_{0,l}}{\rho_0} K - \frac{2}{3}\mu \right) \frac{\partial \hat{u}_2}{\partial x_2} - \frac{\rho_{0,s}}{\rho_0} K \left( \frac{\partial \hat{v}_1}{\partial x_1} + \frac{\partial \hat{v}_2}{\partial x_2} \right) + s\hat{\sigma}_{11} = 0 \quad (12)$$

$$\mu \left( \frac{\partial \hat{u}_2}{\partial x_1} + \frac{\partial \hat{u}_1}{\partial x_2} \right) + \left( \frac{\rho_{0,l}}{\rho_0} K - \frac{2}{3}\mu \right) \left( \frac{\partial \hat{u}_1}{\partial x_1} + \frac{\partial \hat{u}_2}{\partial x_2} \right) - \frac{\rho_{0,s}}{\rho_0} K \left( \frac{\partial \hat{v}_1}{\partial x_1} + \frac{\partial \hat{v}_2}{\partial x_2} \right) + s\hat{\sigma}_{12} = 0 \quad (13)$$

$$\left( \frac{4}{3}\mu + \frac{\rho_{0,l}}{\rho_0} K \right) \frac{\partial \hat{u}_2}{\partial x_2} + \left( \frac{\rho_{0,l}}{\rho_0} K - \frac{2}{3}\mu \right) \frac{\partial \hat{u}_1}{\partial x_1} - \frac{\rho_{0,s}}{\rho_0} K \left( \frac{\partial \hat{v}_1}{\partial x_1} + \frac{\partial \hat{v}_2}{\partial x_2} \right) + s\hat{\sigma}_{22} = 0 \quad (14)$$

$$\left( K - \alpha \frac{\rho_{0,s}}{\rho_0} \left( \frac{\partial \hat{u}_1}{\partial x_1} + \frac{\partial \hat{u}_2}{\partial x_2} \right) + \alpha \frac{\rho_{0,s}}{\rho_0} \left( \frac{\partial \hat{v}_1}{\partial x_1} + \frac{\partial \hat{v}_2}{\partial x_2} \right) - s\hat{p} = 0 \quad (15)$$

$$\hat{\sigma}_{22} + \hat{p}|_{x_2=0} = \hat{\sigma}_{12}|_{x_2=0} = \frac{\rho_{0,l}}{\rho_0} \hat{p}|_{x_2=0} = 0 \quad (16)$$

В результате данного преобразования исходная задача (1) - (6) сводится к двумерной пространственной дифференциальной задаче в спектральной области. Условия разрешимости полученной эллиптической системы будут изучены в ходе дальнейших исследований по данной теме.

*Работа выполнена в рамках проекта грантового финансирования КН МОН РК ПЦФ на 2018-2020 годы по теме: "AP05131026 – Математическое моделирование динамики упруго-*

деформируемых пористых сред с учетом частной зависимости коэффициента трения (с памятью)", рук. проекта д.ф.-м.н., профессор Бердышев А.С.

#### **Список литературы:**

1. Френкель Я.И. К теории сейсмических и сейсмоэлектрических явлений во влажной почве // Изв.АН СССР. Сер.геогр. и геофиз. – 1944.-т.8, No.4 –С133-150
2. Biot М.А. Theory of propagation of elastic waves in a fluid-saturated porous solids // J. Acoustic. Soc. Amer., - 1956, V. 28. P. 168 – 186.
3. Имомназаров Х. Х. , Михайлов А. А. Использование спектрального метода Лагерра для решения линейной двумерной динамической задачи для пористых сред // Сиб. журн. индустр. матем., -2008, том 11, номер 3, С86–95
4. Сорокин К.Э., Имомназаров Х. Х. Численное решение линейной двумерной динамической задачи для пористых сред //Journal of Siberian Federal University, Mathematics&Physics, -2010, 3(2), 256–261 с.
5. Доровский В. Н., Перепечко Ю. В., Роменский Е. И., Волновые процессы в насыщенных пористых упругодеформируемых средах //Физика горения и взрыва,- 1993, No 1, 100–111
6. Blokhin A. M., Dorovsky V. N. Mathematical Modelling in the Theory of Multivelocity Continuum //Nova Sci., New York, -1995, MathSciNet
7. Dorovsky V. N. Imomnazarov Kh. Kh. A Mathematical model for the movement of a conducting liquid through a conducting porous medium //Mathl. Comput. Modelling, Printed in Great Britain, - 1994, Vol. 20, No. 7, pp. 91-97.
8. Abdumauvlen Berdyshev, Kholmatzhon Imomnazarov, Jian-Gang Tang, A. Mikhailov //The Laguerre spectral method as applied to numerical solution of a two-dimensional linear dynamic seismic problem for porous media. //Open Comput.Sci. – 2016, 6, p208-212

## **ИССЛЕДОВАНИЕ ПРОЦЕССОВ ТЕПЛОМАССОПЕРЕНОСА ПРИ ЭЛЕКТРОННО-ЛУЧЕВОЙ СВАРКЕ**

**<sup>1</sup>Блейхер Г.А., <sup>1</sup>Кривобоков В.П., <sup>2</sup>Байдикова Н.В.**

<sup>1</sup> *Национальный исследовательский Томский политехнический университет,  
Россия*

<sup>2</sup> *Карагандинский государственный индустриальный университет, Казахстан  
e-mail: bnatalya@mail.ru*

**Аннотация.** *Исследованы процессы теплопереноса при воздействии на металл остросфокусированных электронных пучков. Сварка с использованием концентрированных пучков электронов получила широкое распространение в автомобильной, авиационной, космической, энергомашиностроительной и других отраслях промышленности. Преимущество данной технологии заключается в том, что имеется возможность получения глубоких и узких сварных швов с минимальной зоной термического влияния, что позволяет сваривать большие толщины за один проход. Выполненные расчеты позволяют определить время существования жидкой*

*фазы, глубину и ширину проплавления, скорости нагрева и охлаждения, рассчитать испарившуюся массу материала изделия при сварке электронным лучом*

В настоящее время существует потребность в высокоэффективных способах сварки, позволяющих получать качественные прецизионные соединения и изделия точной механики. К таким технологиям относится обработка электронным пучком. Сварка с использованием концентрированных пучков электронов получила широкое распространение в автомобильной, авиационной, космической, энергомашиностроительной и других отраслях промышленности.

Преимущество данной технологии заключается в том, что имеется возможность получения глубоких и узких сварных швов с минимальной зоной термического влияния, что позволяет сваривать большие толщины за один проход. Процесс ведется в вакууме, который обеспечивает защиту зону сварки от воздействия различных газов.

С развитием новых систем управления электронным лучом и программирования режимов сварки, значительно расширились возможности совершенствования технологических приемов электронно-лучевой сварки, обеспечивающих разнообразие режимов воздействия луча на поверхность обрабатываемой детали.

Чтобы понимать, как управлять качеством процесса, необходимо знать механизмы, которые при этом происходят. Совокупность таких фундаментальных процессов, как тепло- и массоперенос лежат в основе электронно-лучевой сварки и определяют в итоге конечный результат – фазовые состояния, прочность, твердость, пластичность.

Все расчёты и эксперименты выполнены на примере высокотехнологичного алюминиевого сплава марки 1370 (система Al-Mg-Si-Cu), [1]. Химический состав сплава представлен в таблице 1.

**Таблица 1. Химический состав сплава марки 1370 (система Al-Mg-Si-Cu)**

Элементы сплава	Al	Mg	Si	Mn	Cu	Zn
Содержание, %	93,1	1,74	1,11	0,95	1,9	1,19

В последнее время он получил широкое распространение в подобных изделиях благодаря тому, что обладает хорошей технологической пластичностью при горячей и холодной деформации, значительной прочностью, высоким сопротивлением по отношению к многоцикловой усталости и другими полезными характеристиками. Эти свойства позволяют производить из него путём сварки корпуса с толщиной стенки всего 0,5 мм. Изделия из него заметно легче, чем изготовленные из традиционных алюминиевых сплавов.

Многообразие процессов, протекающих при высоких скоростях и температурах, таких как плавление, испарение, распыление затрудняет их экспериментальное изучение и выработку рациональных режимов сварки. В этом случае для изучения параметров процесса можно использовать моделирование и численный эксперимент. Большое значение имеет возможность прогнозирования таких важных параметров, как распределение температурных полей в свариваемом изделии, формы и размеры сварного шва и зоны термического влияния, температурно-временные параметры процесса сварки для различных участков сварного соединения, скорости нагрева и охлаждения.

Расчёт пространственного распределения поля температур в зоне расплава. Промоделируем распределение тепла в объеме алюминиевого образца при его

облучении электронным пучком. Для расчетов положим источник тепла поверхностным. В качестве начальных условий примем, что температура образца во всем объеме – комнатная.

Основными допущениями модели являются: отсутствие теплообмена с окружающей средой (процесс проводится в вакууме).

Постановка задачи для расчета пространственно-временного поля температур и скорости фронта испарения в цилиндрических координатах имеет следующий вид:

$$\frac{\partial E(z,r,t)}{\partial t} - v_{исн}(r,t) \frac{\partial E(z,r,t)}{\partial z} = \lambda \left( \frac{1}{r} \frac{\partial T(z,r,t)}{\partial r} + \frac{\partial^2 T(z,r,t)}{\partial r^2} + \frac{\partial^2 T(z,r,t)}{\partial z^2} \right)$$

где  $\partial E(z,r,t)$ ,  $T(z,r,t)$  - распределение тепловой составляющей внутренней энергии и температуры в облучаемом образце, R – радиус пучка; r- ширина образца; z- глубина.

Температура плавления и кристаллизации находится из следующих соотношений:

$$T(z,r,t) = \begin{cases} E/c\rho, & E \leq E_m \\ T_m, & E_m < E \leq E_m + q_m \cdot \rho \\ T_m + (E - (E_m + q_m \cdot \rho)) / c\rho, & E > E_m + q_m \cdot \rho \end{cases}$$

1. Начальные условия:

$$T(z,r) = T_0 \quad (T_0 = 300 \text{ K})$$

$$\text{На оси OZ: } \lambda \frac{\partial T}{\partial r} = 0$$

Граничные условия:

1. На облучаемой поверхности:

$$\lambda \frac{\partial T(r,t)}{\partial z} = -q_{всчм}(r,t) + \varepsilon_p T^4 + q_{исн}$$

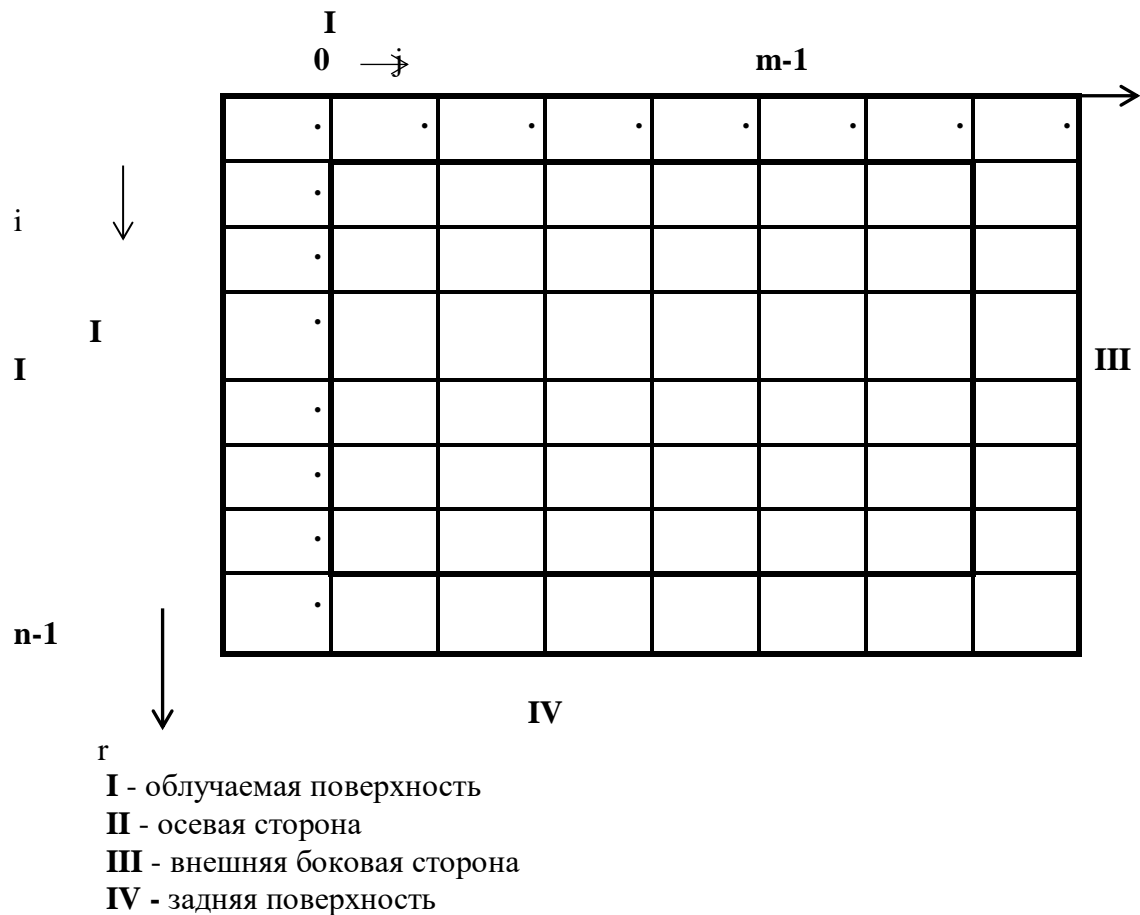
$$\lambda \frac{\partial T(r,t)}{\partial z} \Big|_{z=h} = q_{rad}$$

$$\lambda \frac{\partial T(z,t)}{\partial r} \Big|_{r=R} = q_{rad}$$

Решение задачи выполнено методом конечных разностей по явной схеме для уравнений параболического типа. Вычисления производим на слоях по времени с шагом  $\tau$  и по пространственным координатам  $h_z = \frac{H}{n}$  и  $h_r = \frac{R}{m}$ , где n и m – количество расчетных ячеек по глубине z и радиусу r. Расчет производим в центре ячеек. Время наблюдения (количество слоев) задаем по необходимости. Чтобы численное решение сходилось к решению дифференциальной задачи, достаточно выполнения условия устойчивости разностной схемы:

$$\Delta \tau \leq \frac{\rho \cdot c}{2 \cdot \lambda \cdot \left( \frac{1}{h_r^2} + \frac{1}{h_z^2} \right)}. \quad \text{Вид равномерной}$$

пространственной сетки представлен на рисунке 1.



**Рисунок 1. Равномерная пространственная сетка**

Рассмотрим алгоритм решения:

Тепловая составляющая на новом шаге по времени: для внутренней области

$$E_{i,j}^N = E_{i,j} + \tau \cdot \left[ \begin{aligned} & v_j \cdot \frac{(E_{i+1,j} - E_{i-1,j})}{2 \cdot h_z} + \lambda \cdot \left( \frac{1}{r_j} \cdot \frac{T_{i,j+1} - T_{i,j-1}}{2 \cdot h_r} \right) + \left( \frac{T_{i,j+1} - 2 \cdot T_{i,j} + T_{i,j-1}}{h_r} \right) \cdot \frac{1}{h_r} + \\ & + \left( \frac{T_{i+1,j} - 2 \cdot T_{i,j} + T_{i-1,j}}{h_z} \right) \cdot \frac{1}{h_z} \end{aligned} \right]$$

$1 < j < m-2, 1 < i < n-2$

$E^N$  и  $E$  – тепловые составляющие внутренней энергии на старом и новом слоях по времени соответственно.

Для учета потока на облучаемой поверхности вводим виртуальную ячейку  $E_{-1}$ .

$$\lambda \frac{\partial T(r,t)}{\partial z} = -q_{всвм}(r,t) + \varepsilon_p T^4 + q_{луч} = F$$

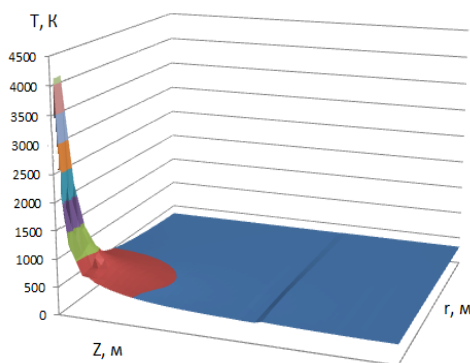
$$\lambda \frac{T_0 - T_{-1}}{h_z} = F; \quad \frac{E_0 - E_{-1}}{h_z} = \frac{c \cdot \rho \cdot F}{\lambda}; \quad E_{-1} = E_0 - \frac{c \cdot \rho \cdot F}{\lambda} \cdot h_z$$

На поверхности  $z = 0$  реализуются граничные условия II рода. Для области при  $i=0$ ,  $1 < j < m-1$  запишем:

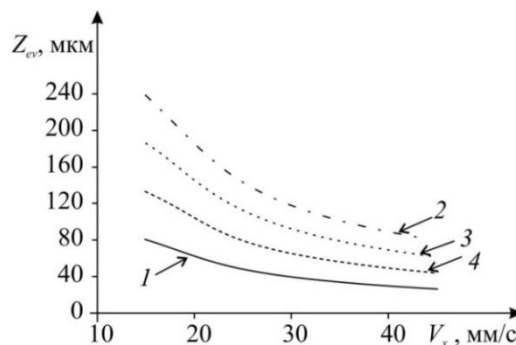
$$E_{0,j}^N = E_{0,j} + \tau \cdot \left[ v_j \cdot \frac{(E_{1,j} - E_{0,j} + c \cdot \rho \cdot F_j^s \cdot h_z)}{2 \cdot h_z} + \lambda \cdot \left( \frac{1}{r_j} \cdot \frac{T_{0,j+1} - T_{0,j-1}}{2 \cdot h_r} \right) + \left( \frac{T_{0,j+1} - 2 \cdot T_{0,j} + T_{0,j-1}}{h_r^2} \right) + \left( \frac{T_{1,j} - T_{0,j} - F_j^s}{h_z} \right) \cdot \frac{1}{h_z} \right]$$

При  $i=0$ ,  $j=m-1$  записываем предыдущее уравнение, но с учетом излучения на смежной границе III. Рассуждая аналогичным образом, решаем уравнения при других краевых условиях. Расчет скорости испарения производили по уравнению Герца-Кнудсена. При этом принимали во внимание величину давления паров вещества мишени на поверхность, то есть задавали соответствующие газодинамические условия.

Произведен расчет температурного поля, получены зависимости толщины испарившегося слоя для алюминиевого сплава 1370 от плотности тока пучка и скорости сканирования с ускоряющим напряжением  $E_0 = 28$  кэВ.



**Рисунок 2 – Температурное поле образца:**  
при  $d_{\text{beam}}=250$  мкм,  $U=28$  кВ,  $I=45$  мА,  
 $V=10$  мм/с



**Рисунок 3 – Зависимость толщины испарившегося слоя от скорости сварки**

Обращает на себя внимание то, что область повышенных значений температуры сосредоточена в тонком приповерхностном слое и слабо расширяется с течением времени облучения, в то время как максимальная температура заметно растет. Полученные в расчетах размеры расплавленной области находятся в удовлетворительном согласии с экспериментом.

В основе всех явлений, вызываемых в твердом теле воздействием мощных пучков заряженных частиц, лежат процессы диссипации энергии, выделенной в приповерхностном слое вещества. Основными каналами диссипации излучения является теплопроводность, фазовые переходы и эрозия вещества

В результате проделанной работы были исследованы процессы тепломассопереноса при воздействии на металл остросфокусированных электронных

пучков. Выполненные расчеты позволяют определить время существования жидкой фазы, глубину и ширину проплавления, скорости нагрева и охлаждения, рассчитать испарившуюся массу материала изделия при сварке электронным лучом.

Для проверки расчетных данных были проведены эксперименты на электронной пушке с плазменным катодом. Подобраны режимы сварки корпуса гироскопа из алюминиевого сплава марки 1370: скорость сварки 25 мм/с, ускоряющее напряжение 28 кВ, плотность тока сварки 255 А/см<sup>2</sup>, длина сварного соединения 973 мм. При данных параметрах масса корпуса гироскопа уменьшится на 170 мг.

### **Литература**

1. ТУ 1-804-398-2002 Листы из алюминиевого сплава марки 1370.
2. Блейхер Г.А., Кривобоков В.П. Эрозия поверхности твердого тела под действием мощных пучков заряженных частиц. / Г.А. Блейхер, В.П. Кривобоков. – Новосибирск: Наука, 2014. - 248 с.
3. Блейхер Г.А., Кривобоков В.П. Модель расчета производительности технологий осаждения покрытий с помощью мощных импульсных пучков заряженных частиц. / Г.А. Блейхер, В.П. Кривобоков. // Известия вузов. Физика. – 2011. Т. 54, №11/2. - С. 142-147.
4. Аброян И.А., Андронов А.Н., Титов А.И. Физические основы электронной и ионной технологии. / И.А. Аброян, А.Н. Андронов, А.И. Титов. – М.: Высшая школа, 1984. - 320 с.
5. Грацианов Ю.А., Путимцев Б.Н., Молотиллов В.В. Металлургия прецизионных сплавов. / Ю.А. Грацианов, Б.Н. Путимцев, В.В. Молотиллов - М.: Металлургия, 1975. - 448 с.

## **ФОРМИРОВАНИЕ СИСТЕМЫ РАБОЧИХ ОСНОВАНИЙ, ДЛЯ АЛГОРИТМА ШИФРОВАНИЯ НА БАЗЕ НЕПОЗИЦИОННЫХ ПОЛИНОМИАЛЬНЫХ СИСТЕМ СЧИСЛЕНИЯ**

**Варенников А.В.**

*Институт информационных и вычислительных технологий КН МОН РК,  
Республика Казахстан  
e-mail: avarennikov@gmail.com*

***Аннотация.** Представлен способ формирования системы рабочих оснований для алгоритма шифрования на базе непозиционных полиномиальных систем счисления с использованием методов проверки полиномов на неприводимость Бен-Ора и Рабина.*

### **Введение**

Суть алгоритма шифрования на базе непозиционных полиномиальных систем счисления (НПСС) заключается в следующем. Исходное открытое сообщение представляется в виде последовательности блоков заданной длины  $N$  (в битах). Каждый блок интерпретируется как последовательность остатков



$\alpha_1(x), \alpha_2(x), \dots, \alpha_n(x)$  от деления некоторого многочлена  $F(x)$  степени не выше  $N$  с двоичными коэффициентами на неприводимые полиномы (рабочие основания)  $p_1(x), p_2(x), \dots, p_n(x)$ :

$$F(x) = (\alpha_1(x), \alpha_2(x), \dots, \alpha_n(x)), \quad (1)$$

где  $F(x) \equiv \alpha_i(x) \pmod{p_i(x)}$ ,  $i = \overline{1, n}$ ,  $n$  – число выбранных рабочих оснований. В соответствии с китайской теоремой об остатках представление (1) является единственным при условии, что все основания НПСС  $p_i(x)$  будут различными.

Рабочие основания выбираются из числа всех неприводимых полиномов с двоичными коэффициентами и степенями  $m_j$ , из условия выполнения уравнения:

$$k_1 m_1 + k_2 m_2 + \dots + k_s m_s = N, \quad (2)$$

где  $1 \leq m_j \leq N$ ,  $k_1 + k_2 + \dots + k_s = n$ .

Для получения криптограммы требуется выполнить наложение секретного ключа (или псевдослучайной последовательности – ПСП)  $G(x)$  на каждый блок исходного сообщения  $F(x)$ . ПСП также интерпретируется как последовательность остатков от деления некоторого многочлена  $G(x)$  степени не выше  $N$  на соответствующие рабочие основания:

$$G(x) = (\beta_1(x), \beta_2(x), \dots, \beta_n(x)), \quad G(x) \equiv \beta_i(x) \pmod{p_i(x)}, \quad i = \overline{1, n}.$$

Тогда криптограмма получается в результате выполнения некоторой функции  $H(F(x), G(x))$ :

$$H(x) = (\omega_1(x), \omega_2(x), \dots, \omega_n(x)), \quad H(x) \equiv \omega_i(x) \pmod{p_i(x)}, \quad i = \overline{1, n}.$$

Применение конкретной функции  $H(F(x), G(x))$  задаёт метод шифрования на базе НПСС. В разработанной криптосистеме в качестве функции  $H(F(x), G(x))$  используется операция умножения модулю:

$$\alpha_i(x) \beta_i(x) \equiv \gamma_i(x) \pmod{p_i(x)}, \quad i = \overline{1, n}. \quad (3)$$

При расшифровании криптограммы  $H(x)$  по известному ключу  $G(x)$  для каждого значения  $\beta_i(x)$  производится вычисление, как следует из (3), обратного (инверсного) многочлена  $\beta_i^{-1}(x)$  из условия выполнения следующего сравнения:

$$\beta_i(x) \beta_i^{-1}(x) \equiv 1 \pmod{p_i(x)}, \quad i = \overline{1, n}. \quad (4)$$

В результате получается многочлен

$$G^{-1}(x) = (\beta_1^{-1}(x), \beta_2^{-1}(x), \dots, \beta_n^{-1}(x)),$$

инверсный к многочлену  $G(x)$ . Тогда элементы последовательности вычетов (1) в соответствии с (4) и (5) восстанавливаются по сравнению:

$$\alpha_i(x) \equiv \beta_i^{-1}(x) \omega_i(x) \pmod{p_i(x)}, \quad i = \overline{1, n}.$$

Таким образом, в рассмотренной модели системы шифрования электронного сообщения заданной длины  $N$  бит в НПСС полным ключом является выбранная система полиномиальных рабочих оснований  $p_1(x), p_2(x), \dots, p_n(x)$  и секретный ключ  $G(x) = (\beta_1(x), \beta_2(x), \dots, \beta_n(x))$ .

Более подробную информацию по алгоритму шифрования на базе НПСС можно найти в работах [1-6].

### 1. Обзор алгоритмов проверки полиномов на неприводимость

Проверка на неприводимость полиномов использует тот факт, что для каждого конечного поля  $\mathbb{F}_q$  и каждого  $n \in \mathbb{N}$  произведение всех нормированных неприводимых полиномов над  $\mathbb{F}_q$ , степень которых делит  $n$ , равно  $x^{q^n} - x$  ([7], теорема 3.20). Поскольку алгоритм шифрования на базе НПСС использует неприводимые полиномы с двоичными коэффициентами, дальнейшие рассуждения приводятся для  $q = 2$ . В частности, все полиномы, принадлежащие полю  $\mathbb{F}_2$ , являются нормированными, а также имеет место равенство  $x^{2^n} - x = x^{2^n} + x$ .

Проверка на неприводимость Бен-Ора [8].

**Алгоритм:** is\_irreducible\_b (Проверка на неприводимость Бен-Ора).

**Вход:** нормированный полином  $f(x) \in \mathbb{F}_2[x]$  степени  $n$ .

**Выход:** либо "Истина" ( $f(x)$  является неприводимым),  
либо "Ложь" ( $f(x)$  является приводимым).

для  $i := 1$  по  $n/2$  цикл

$$g(x) := \text{НОД}(f(x), x^{2^i} + x \bmod f(x));$$

если  $g(x) \neq 1$ , тогда вернуть ("Истина");

**конец цикла;**

возвратить ("Ложь");

Если полином  $f(x)$  является приводимым, тогда должны существовать неприводимые полиномы  $f_1(x), f_2(x), \dots, f_k(x)$  степеней  $n_1, n_2, \dots, n_k$ ,  $\sum_{j=1}^k n_j = n$  такие, их произведение равно  $f(x)$ . Тогда имеем следующее:  $\text{НОД}(x^{2^{n_j}} + x, f_j(x)) = f_j(x)$ , т.е.  $\text{НОД}(x^{2^{n_j}} + x, f_j(x)) \neq 1$ . Отсюда следует, что полином  $f(x)$  степени  $n$  является неприводимым, если значение  $\text{НОД}(x^{2^i} + x, f(x)) = 1$  для всех  $i$ ,  $0 < i < n$ . Согласно теореме Евклида  $\text{НОД}(a, b) = \text{НОД}(b, r)$ , где  $r$  – остаток от деления  $a$  на  $b$ , поэтому вычисляется значение  $\text{НОД}(f(x), (x^{2^i} + x) \bmod f(x))$ . Выполняются не все проверки, а только первая их часть для  $i = 1, \dots, \lfloor \frac{n}{2} \rfloor$ , поскольку, если существует множитель, степень которого больше  $\frac{n}{2}$ , то должны существовать сомножители меньших степеней, которые были бы обнаружены на предыдущих шагах.

Проверка на неприводимость Рабина [9].

**Алгоритм:** is\_irreducible\_r (Проверка на неприводимость Рабина).

**Вход:** нормированный полином  $f(x) \in \mathbb{F}_2[x]$  степени  $n$ ,  
 $p_1, \dots, p_k$  – все простые делители значения  $n$ .

**Выход:** либо "Истина" ( $f(x)$  является неприводимым),  
либо "Ложь" ( $f(x)$  является приводимым).

для  $i := 1$  по  $k$  цикл

$g(x) := \text{НОД}(f(x), x^{2^{n/p_i}} - x \bmod f(x));$   
если  $g \neq 1$ , тогда вернуть ("Истина");

**конец цикла;**

$g(x) := x^{2^n} + x \bmod f(x);$

если  $g(x) = 0$ , тогда вернуть ("Истина");

иначе вернуть ("Ложь");

Проверка на неприводимость Рабина использует тот факт, что если  $p_1, \dots, p_k$  являются всеми простыми делителями значения  $n$ , то нормированный полином  $f(x) \in \mathbb{F}_2[x]$  степени  $n$  является неприводимым тогда и только тогда, когда  $\text{НОД}(f(x), x^{2^{n/p_i}} + x \bmod f(x)) = 1$  для  $1 \leq i \leq k$ , и  $f(x)$  является делителем  $x^{2^n} + x$  (см.[9], лемма 1).

Сравнительный анализ этих двух алгоритмов можно найти в [10]. Там же приводятся теоретические и практические результаты, которые показывают, что алгоритм Бен-Ора является более эффективным, поскольку полиномы больших степеней, в случае, если они являются приводимыми, практически всегда разлагаются на неприводимые полиномы, среди которых найдется полином малой степени, что позволяет быстро обнаружить приводимые полиномы на начальных шагах этого алгоритма.

## 2. Алгоритм нахождения неприводимых полиномов

Основная идея алгоритма нахождения неприводимого полинома заключается в выборе случайным образом полинома заданной степени  $n$  и проверке его на неприводимость с использованием алгоритма Бен-Ора и/или Рабина. Учитывая тот факт, что количество нормированных неприводимых полиномов  $I_n$  степени  $n$  равно  $I_n = \frac{1}{n} \sum_{d|n} \mu(d) 2^{n/d} \approx \frac{1}{n} 2^n$ , где  $\mu(d)$  – функция Мёбиуса,  $d$  – делители числа  $n$ , вероятность того, что выбранный случайным образом полином степени  $n$  окажется неприводимым, равна  $1/n$ . [11]

**Алгоритм:** random\_irreducible (Поиск неприводимого полинома).

**Вход:**  $n$  – значение степени искомого неприводимого полинома.

**Выход:** неприводимый полином  $f(x) \in \mathbb{F}_2[x]$  степени  $n$ .

**цикл**

$f(x) := \text{ГПСЧ}(2^n, 2^{n+1} - 1);$   
если is\_irreducible\_b( $f(x)$ ) = "Истина" и  
is\_irreducible\_r( $f(x)$ ) = "Истина"  
тогда вернуть ( $f(x)$ );

**конец цикла;**

Здесь используется функция-генератор псевдослучайных чисел ГПСЧ( $a, b$ ), порождающая последовательность чисел из диапазона  $[a, b]$ , элементы которой подчиняются равномерному распределению.

### 3. Алгоритм формирования системы рабочих оснований

Как уже было сказано выше, система рабочих оснований представляет собой набор неприводимых полиномов  $p_1(x), p_2(x), \dots, p_n(x)$ , используемых алгоритмом шифрования на базе НПСС в качестве модулей при выполнении операций зашифрования (3) и расшифрования (4). Рабочие основания выбираются из числа всех неприводимых полиномов с двоичными коэффициентами и степенями  $m_j$ , из условия выполнения уравнения (2).

**Алгоритм:** workbase\_system (Формирование системы рабочих оснований).

**Вход:**  $N$  – размер блока/ключа (в битах).

**Выход:**  $p_1(x), p_2(x), \dots, p_n(x)$  – неприводимые полиномы над полем  $\mathbb{F}_2$ , образующие систему рабочих оснований.

$n := N; i = 1;$

**пока** ( $n > 0$ ) **цикл**

**если** ( $n < d_{min}$ ) **тогда**  $m_i := n$  **иначе**  $m_i := \text{ГПСЧ}(d_{min}, n);$

$p_i(x) := \text{random\_irreducible}(m_i);$

**если** ( $\forall j \neq i$ )  $p_j(x) \neq p_i(x)$  **тогда**  $n := n - m_i; i := i + 1;$

**конец цикла;**

возвратить  $(p_1(x), p_2(x), \dots, p_n(x))$

Данный алгоритм учитывает, что количество неприводимых полиномов малых степеней ограничено, что может привести к ошибке из-за невозможности выполнить условие  $(\forall j \neq i) p_j(x) \neq p_i(x)$ , поэтому в качестве значений степеней рабочих оснований (за исключением последнего) рекомендуется использовать такие, чтобы  $m_i \times I_{m_i} \geq N$ , где  $I_{m_i}$  – количество неприводимых полиномов степени  $m_i$ . В таблице 1 показаны наименьшие допустимые значения степеней рабочих оснований  $d_{min}$  для различных значений  $N$ .

**Таблица 1. Наименьшие допустимые значения степеней рабочих оснований**

$N$	$d_{min}$	$I_{d_{min}}$
32	6	9
64	7	18
128	8	30
256	9	56

### Заключение

Описанные алгоритмы нашли применение в программе шифрования файлов на базе непозиционных полиномиальных систем счисления «NPSS Crypto 1.1», разработанной в Республиканском государственном предприятии «Институт информационных и вычислительных технологий» Комитета науки Министерства образования и науки Республики Казахстан (авторское свидетельство № 0201 от 16 января 2018 г.).

### **Список использованных источников**

1. Нысанбаева С.Е. Разработка и исследование криптографических систем на базе непозиционных полиномиальных систем счисления: дисс. докт. техн. наук. – Алматы, 2009. – 240 с.
2. Бияшев Р.Г., Нысанбаева С.Е., Капалова Н.А. Разработка систем криптографической защиты с заданными характеристиками. // Проблемы информатики. 2013. №2. – с. 30-36.
3. Разработка и исследование моделей национального алгоритма шифрования на базе модулярной арифметики // Отчет о научно-исследовательской работе по проекту программно-целевого финансирования О.0678, № гос. регистрации 0115РК01304. – Алматы, 2017. – 171 с.
4. Нысанбаева С. Е., Капалова Н.А., Дюсенбаев Д.С., Дузбаев Т.Т. Алгоритм шифрования на базе китайской теоремы об остатках // Материалы научной конференции Института информационных и вычислительных технологий МОН РК «Современные проблемы информатики и вычислительных технологий». Алматы, 2017.– с. 191-196.
5. Бияшев Р.Г., Варенников А.В., Нысанбаева С.Е. Алгоритм формирования полных ключей для системы шифрования на базе непозиционных полиномиальных систем счисления // Материалы научной конференции Института информационных и вычислительных технологий МОН РК «Современные проблемы информатики и вычислительных технологий». Алматы, 2017.– с. 44-48.
6. Бияшев Р.Г., Варенников А.В., Нысанбаева С.Е. Программная реализация алгоритмов шифрования на базе непозиционных полиномиальных систем счисления // Материалы II Международной научной конференции «Информатика и прикладная математика». – Алматы, 2017. - часть II, С. 172-178.
7. Лидл Р., Нидеррайтер Г. Конечные поля (в 2-х томах), т. 1. – М.: Мир, 1988. – 430 с.
8. Ben-Or, M. Probabilistic algorithms in finite fields. // In Proc. 22th IEEE Annual Symposium on Foundations of Computer Science, 1981. p. 394-398
9. Rabin, M. Probabilistic algorithms in finite fields. // Technical Report MIT/LCS/TR-213. – Massachusetts Institute of Technology, 1979. – 26 p.
10. Shuhong Gao, Daniel Panario. Tests and Constructions of Irreducible polynomials over Finite Fields // Foundations of Computational Mathematics. – Springer, 1997. – p. 346-361.
11. Jörg Arndt. Matters Computational Ideas, Algorithms, Source Code. – Springer, 2011. – 966 p.

## **СПЕКТРАЛЬНЫЕ ХАРАКТЕРИСТИКИ ВОЛОКОННЫХ РЕШЕТОК БРЭГГА**

**Вуйцик В.<sup>1,2</sup>, Калижанова А.У.<sup>1,3</sup>, Кисала П.<sup>1,2</sup>, Кашаганова Г.Б.<sup>1,4</sup>,  
Амиргалиева С.Н.<sup>1,2</sup>, Картбаев Т.С.<sup>1,3</sup>, Оразбеков Ж.<sup>1</sup>**

<sup>1</sup>*Институт информационных и вычислительных технологий КН МОН РК,  
Казахстан*

<sup>2</sup>*Люблинский технический университет, Польша*

<sup>3</sup>Алматинский университет энергетики и связи, Казахстан

<sup>4</sup>Казахстанско-Американский университет, Казахстан  
e-mail: kalizhanova\_aliya@mail.ru

**Аннотация.** В статье рассмотрены спектральные характеристики волоконных решеток Брэгга при различных значениях показателя преломления  $n$ .

Волоконные решетки Брэгга (ВРБ) широко используются в различных областях современной техники и технологии. Каждое применение предъявляет особые требования к спектральным характеристикам ВРБ, которые задаются на стадии изготовления решеток. Наиболее важными параметрами ВРБ являются их спектральные характеристики: спектр пропускания и спектр отражения. В большинстве случаев измерение спектров решеток осуществляется с использованием широкополосного источника оптического излучения и анализатора спектра.

Для регистрации спектров с высоким спектральным разрешением ( $<0.1$  нм) и большим динамическим диапазоном ( $>20$  дБ) требуется широкополосный источник света, обладающий высокой спектральной яркостью. В качестве таких источников, как правило, используются полупроводниковые люминесцентные диоды и волоконные люминесцентные источники. Недостатками таких источников является относительно небольшая спектральная ширина излучения ( $\sim 30 - 40$  нм) и, в ряде случаев, невысокая временная стабильность спектра излучения, осложняющая проведение нормированных абсолютных измерений спектров решеток. В связи с указанными недостатками, часто, при проведении измерений в качестве источника оптического излучения используют вольфрамную галогенную лампу.

В настоящее время, в большинстве случаев для регистрации спектров используются стандартные анализаторы спектра производства ряда иностранных фирм, которые, как правило, имеют спектральный диапазон измерений 600–1700 нм, спектральное разрешение 0.1 нм и лучше, предел обнаружения оптического сигнала на уровне 1 пВт ( $-80 \div -90$  дБм). Время регистрации спектра, как правило, зависит от уровня входного сигнала и ширины измеряемого спектрального диапазона и обычно находится в пределах интервала  $1 \div 100$  сек. Современные спектроанализаторы имеют встроенный интерфейс обмена данными с персональным компьютером, поэтому результаты измерений могут синхронно заноситься в память компьютера, что особенно важно при проведении измерений динамики записи или отжига решеток, изучения их спектрального отклика на внешние динамические воздействия и т.д.

Спектры пропускания регистрируются непосредственно при прохождении излучения источника через световод, с записанной в нем решеткой. При измерении спектров отражения отраженное излучение отводится на спектроанализатор с помощью волоконно-оптического ответвителя или циркулятора с известными спектральными характеристиками. В такой схеме регистрации отраженного сигнала оптические потери, вносимые циркулятором, составляют  $\sim 1$  дБ, в то время как, использование ответвителя приводит к дополнительным потерям не менее 6 дБ. Вместе с тем циркулятор имеет спектрально ограниченную область использования и является дорогостоящим оптическим элементом.

Таким образом, совокупность всех вышеперечисленных факторов определяет выбор схемы регистрации спектров отражения и пропускания волоконных решеток ПП в каждом конкретном случае.

Следует отметить, что при измерении спектров отражения/пропускания, как правило, не учитываются фазовые характеристики волоконных решеток. Однако существуют приложения, где дисперсионные свойства решеток, которые определяются их фазовыми характеристиками, играют основную роль. К таким приложениям относятся, например, компенсация дисперсии [1] и изменение формы коротких оптических импульсов [2].

В большинстве применений волоконных решеток требуется, чтобы рассчитанные спектральные свойства были реализованы максимально точно.

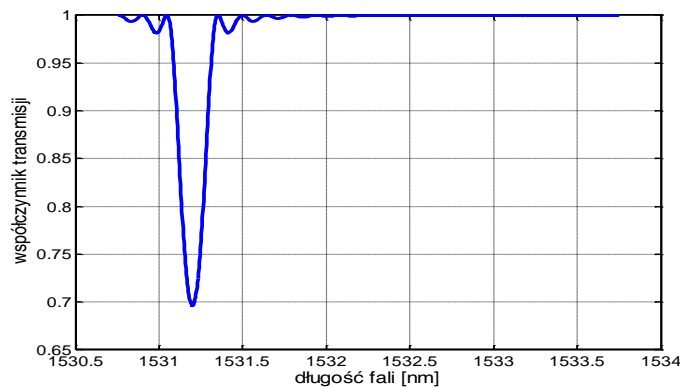
Для математического моделирования спектральных характеристик ВРБ применяются различные методы:

- метод матриц переноса;
- метод связанных мод;
- метод Рунге-Кутты четвертого порядка;
- метод Рунда.

Получение наиболее точных спектральных характеристик возможно при использовании метода связанных мод. Наибольшее приближение к реальным характеристикам ВРБ, подверженных изгибному натяжению, получают с помощью метода матриц переноса для отрезков до и после изгиба.

Метод матриц переноса является более точным и эффективным методом численного моделирования. Характеристики ВРБ описываются произведением матриц 2x2 каждого однородного элемента решетки, что позволяет в итоге получить полную матрицу переноса ВРБ. ВРБ разбивается на N однородных секций, для которых определяются коэффициенты отражения и пропускания.

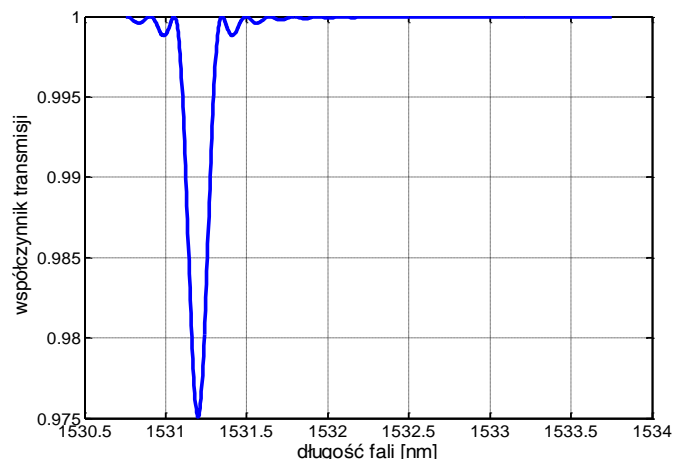
Используя различные пакеты прикладных программ, можно наглядно увидеть результаты математического моделирования спектральных характеристик ВРБ.



**Рисунок 1 – Спектральная характеристика решетки при значении  $\delta n = 0,0008$ ,  $n_{eff} = 1,44688$**

Для исследования спектральных характеристик ВРБ использовали решетку  $L = 1$  мм с величиной модуляции показателя преломления  $\delta n = 0,0008$ , эффективный показатель преломления  $n_{eff} = 1,44688$  и период сетки  $\Lambda = 168,43$  нм.

Из графика видно, что спектр пропускания достигает максимального значения около 0,7 при длине волны Брэгга  $\lambda_{БВ} = 1531,2$  нм. Для улучшения спектра пропускания решили рассмотреть влияние модуляции показателя преломления. Теперь для примера возьмем значение модуляции показателя преломления  $\delta n = 0,0002$ . Результат показано на рисунке 2.

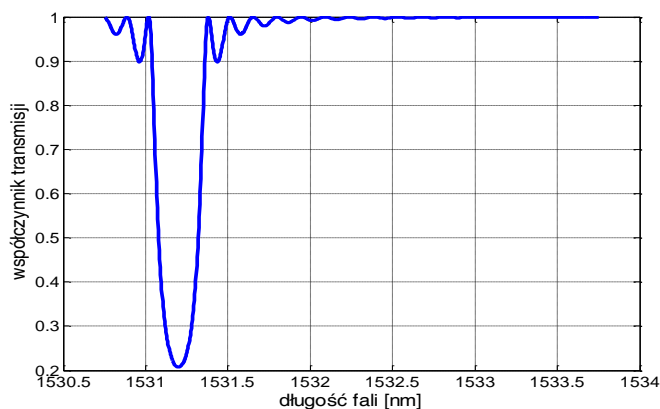


**Рисунок 2 – Спектральная характеристика решетки при значении  $\delta n = 0,0002$ ,  $n_{\text{eff}} = 1,44688$**

Из графика видно, ухудшение оптических параметров решетки. Коэффициент пропускание упал до значения 0,975 для длины волны Брэгга.

Если изготовить решетку с такими параметрами, то ее эффективность отражения длин волн, аналогичная длине Брэгга, была бы настолько мала, что такие решетки нельзя использовать в качестве оптических волоконных датчиков.

Таким образом, моделируемая решетка Брэгга с длиной волны Брэгга будет характеризоваться коэффициентом пропускания, ограниченным 30%.

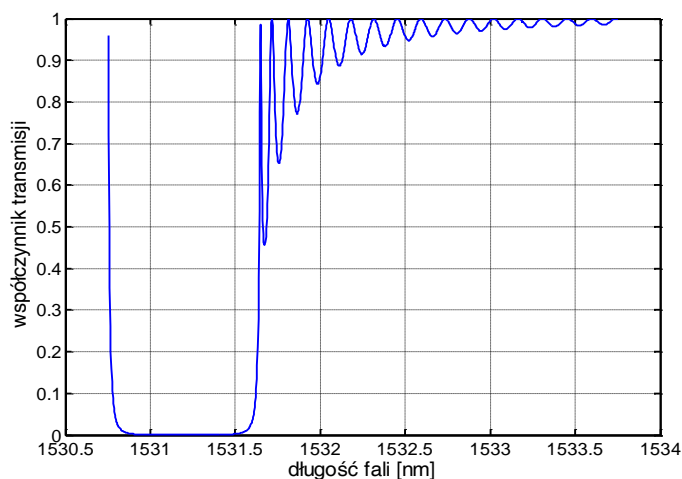


**Рисунок 3 – Спектральная характеристика решетки при значении  $\delta n = 0,0002$**

При модуляции показателя преломления  $\delta n = 0,00002$ , была получена характеристика пропускания со значительно улучшенными оптическими параметрами. Коэффициент передачи упал до 0,2 для длины волны Брэгга. Результаты численных расчетов, полученных для этого случая, показаны на рисунке 3.

Далее рассматриваем возможности дальнейшего уменьшения коэффициента пропускания света в диапазоне длин волн, соответствующими брэгговскому отражению. При значении модуляции показателя преломления  $\delta n = 0,004$ , было достигнуто значительное улучшение его оптических параметров. Было показано, что отражение такой структуры для длин волн в диапазоне отражения Брэгга может быть уменьшено до 0,02.





**Рисунок 4 – Спектральная характеристика решетки с модуляцией показателя преломления равны  $\delta n = 0,0008$ . Длина сетки  $L = 1$  мм**

На рисунке 4 показаны теоретические характеристики решетки Брэгга для случая, когда изменение модуля показателя преломления составляет  $\delta n = 0,0008$ . Оказывается, что в этом случае коэффициент отражения достигает теоретического значения, равного 100% для длины волны в диапазоне Брэггов. Кроме того, это значение поддерживается для диапазона длин волн около 0,5 нм. На основании полученных результатов было показано, что в этом случае изменение длины решетки от 5 до 10 мм существенно не влияет на коэффициент отражения и форму характеристики пропускания.

В связи с вышеизложенным, можно сделать вывод, что при моделировании датчиков значение модуляции показателя преломления не влияет на длину волны Брэгга. Ни одна из спектральных характеристик не показывает сдвиг в поле длин волн из-за изменения модуляции показателя преломления. Это означает, что брэгговская волна не может быть предопределена в структуре системы записи, а изменение значения модуляции  $\delta n$  не изменит ее.

### Литература

1. R. Kashyap, H-G. Froehlich, A. Swanton, DJ. Armes, "1.3 m long super-step-chirped fibre Bragg grating with a continuous delay of 13.5 ns and bandwidth 10 nm for broadband dispersion compensation", *Electronics Letters*, Vol. 32, No. 19, 1807-1809, 1996.
2. K.O. Hill, S. Theriault, B. Malo, F. Belodeau, T. Kitagawa, D.C. Johnson, J. Albert, T. Takiguchi, T. Takaoka, K. Hagimoto, "Chirped in-fibre grating dispersion compensators: Linearisation of dispersion characteristic and demonstration of dispersion compensation in 100 km, 10 Gbit/s optical fibre link", *Electronics Letters*, Vol. 30, No. 21, 1755-1756, 1994.
3. Кашаганова Г.Б. Вуйцик В., Калижанова А.У., Козбакова А.Х. Волоконно-оптические датчики температуры. Вестник №6 (124) КазНИТУ имени К.И. Сатпаева 2017.

## **К ВОПРОСУ МОДЕЛИРОВАНИЯ БУДУЩЕГО КЛИМАТА ЦЕНТРАЛЬНОЙ АЗИИ В РАМКАХ МЕЖДУНАРОДНОГО ПРОЕКТА CORDEX (COORDINATED REGIONAL CLIMATE DOWNSCALING EXPERIMENT)**

**Есеркепова И.Б., Юничева Н.Р., Нурсейтов Д.Б., Бостанбеков К.,  
Гаврук С.В.**

*Институт информационных и вычислительных технологий КН МОН РК  
naduni@mail.ru, k.bostanbekov@norlist.kz*

***Аннотация.** Прогнозы будущего климата являются основным источником научно обоснованной информации для решения вопросов, касающихся политики Казахстана в области изменения климата, проблем адаптации, стратегий планирования деятельности по смягчению воздействий, управлению рисками, для планирования развития таких отраслей экономики, как сельское хозяйство, энергетика, промышленность, водные ресурсы, биоразнообразие.*

*Вопросы оценки характеристик будущего климата на основе региональных гидродинамических моделей являются актуальными в современных научных исследованиях.*

*В работе предполагается определить параметры моделей регионального климата наилучшим образом соответствующие данным наблюдений за климатическими характеристиками, исследовать чувствительность моделей к повышению концентрации парниковых газов за базовый период, а также получить проекции будущего климата по различным сценариям стабилизации парниковых газов.*

Большая часть региональной информации об изменении климата основана на использовании моделей общей циркуляции сопряженной системы атмосфера–океан (МОЦАО), применяемых в течение последних 30 лет в рамках Всемирной программы исследований климата (ВПИК) [1-14]. Однако горизонтальное разрешение большинства современных МОЦАО пока что составляет порядка нескольких сотен километров [15]. Это не позволяет им фиксировать эффекты локальных воздействий (например, характеристики сложной топографии и землепользования), которые вносят изменения в климатический сигнал в мелких масштабах. Чтобы устранить эту проблему, разрабатываются различные методы «регионализации» или «даунскейлинга» (уменьшения масштаба) для пространственного уточнения климатической информации, получаемой с помощью МОЦАО [16].

Приемы регионального даунскейлинга климата (РДК) все более широко используются при рассмотрении самых разных проблем изменения климата, и в настоящее время они являются важным методом, применяемым в исследованиях изменения климата [17]. Одной из причин недостаточно активного использования продукции на основе РДК является отсутствие скоординированной структуры для оценки методов, основанных на РДК, и получения качественных ансамблевых перспективных оценок, позволяющих охарактеризовать неопределенности, лежащие в основе региональных перспективных оценок изменения климата. Большинство исследований на основе РДК обособлены и связаны с конкретными интересами, поэтому в настоящее время нет всестороннего представления о региональных

перспективных оценках изменения климата, основанных на экспериментах с использованием РДК. В связи с этим в рамках Всемирной программы исследований климата (ВПИК) создана Целевая группа по региональному даунскейлингу климата.



Рисунок 1 – Домен - Центральная Азия

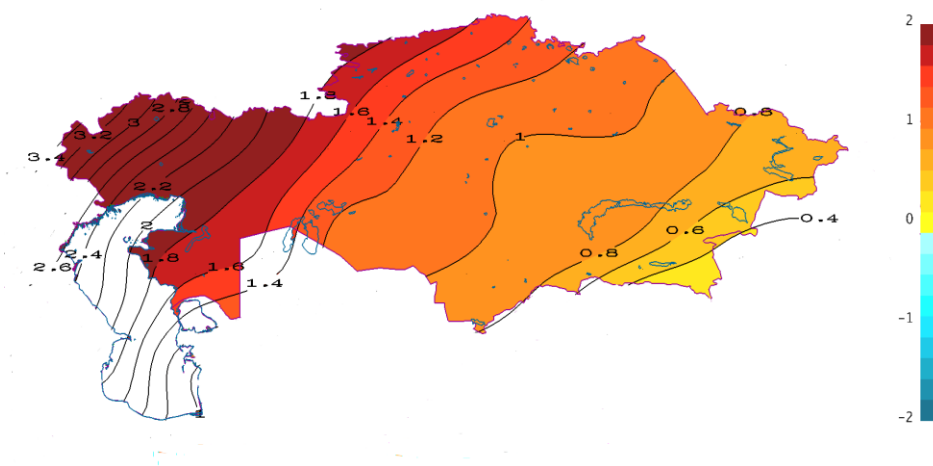


Рисунок 2 – Разница приземных температур воздуха средних за период (2031-2040 гг.) и базовый период (1989 – 2008гг.) по сценарию с уровнем стабилизации  $8.5 \text{ Вт/м}^2$

Первым результатом деятельности этой группы стало создание структуры под названием «Эксперимент по скоординированному региональному даунскейлингу климата (КОРДЭКС)» [18,19]. КОРДЭКС, по существу, имеет двойственную цель: (1) оценить и сравнить рабочие характеристики моделей (структура оценки моделей) и (2) организовать ряд экспериментов для получения перспективных оценок климата, которые будут использоваться в исследованиях воздействия и адаптации (база данных перспективных оценок климата). Цель КОРДЭКС – обеспечить структуру, доступную для широкого научного сообщества при максимальном использовании результатов.

Поэтому области КОРДЭКС охватывают большую часть земной поверхности земного шара, включая регион Центральной Азии.

У нас также имеется опыт по моделированию будущего регионального климата приобретенный в ходе выполнения проекта по подготовке III-VI Национального сообщения Республики Казахстан по изменению климата. Выполнен ряд численных экспериментов по оценке таких характеристик будущего климата, как приземная температура и количество осадков.

Национальные и международные исследования об оценке климата играют важную роль в обеспечении научной основы для понимания и оценки воздействий изменчивости и изменения климата на секторы экономики, такие, как сельское хозяйство и производство продовольствия, водное хозяйство, энергетика и транспорт. Основным элементом этих работ об оценке являются проекции будущего климата, основанные на использовании моделей, которые не только прогнозируют физические показатели будущего климата, но также косвенно предоставляют информацию о социальных последствиях ожидаемых изменений.

Таким образом, проекции будущего климата являются основным источником научно обоснованной информации для решения вопросов, касающихся политики в области климата, проблем адаптации, стратегий планирования деятельности по смягчению воздействий, управлению рисками, для планирования развития таких отраслей экономики, как сельское хозяйство, энергетика, промышленность, водные ресурсы, биоразнообразии. Скоординированный механизм вносит большой вклад в решение задачи по обеспечению многоступенчатого процесса по получению качественно подготовленных проекций климата и преобразования итоговой информации в формат, понятный для конечных пользователей/лиц, принимающих решения.

В данной работе будут определены параметры моделей регионального климата наилучшим образом соответствующие данным наблюдений за климатическими характеристиками.

Исследования будут выполняться на базе расчетов по ансамблю моделей общей циркуляции атмосферы и океана (МОЦАО), участвующие в международном проекте сравнения климатических моделей CMIP5. Модели климата постепенно перерастают в модели земной системы в целом, которые уже не только рассчитывают динамику атмосферы и океана, но и включают в себя детальные подмодели химии атмосферы, растительности, почвы, химии и биологии моря и других процессов и явлений, влияющих на климат. В программу проекта CMIP5 включена группа специальных экспериментов, направленных на установление причин наблюдаемых климатических изменений и оценку сравнительного значения вызывающих их внешних воздействий. Оценка влияния отдельных внешних воздействий на климат основывается на сравнении наблюдаемых изменений с ожидаемой реакцией на предполагаемые воздействия, полученной с помощью климатических моделей.

В отличие от сценариев, используемых в четвертом цикле оценок МГЭИК, которые основывались на сценариях выбросов парниковых газов СДСВ (IPCC, 2000), следующее поколение экспериментов по моделированию сценариев основано на так называемых исходных траекториях концентраций (RCP), т.е. заданных траекториях концентраций парниковых газов на протяжении XXI века, соответствующих различным уровням стабилизации радиационного воздействия, которые будут достигнуты к 2100 г. Были выбраны четыре сценария с уровнями стабилизации 2.9, 4.5, 8.5 и 11.2 Вт/м<sup>2</sup> (которые соответственно называются RCP2.9, RCP4.5, RCP8.5 и RCP11.2). В рамках CMIP5 при моделировании с использованием глобальных моделей

наивысший приоритет отдан уровням RCP4.5 и RCP8.5, которые приблизительно соответствуют сценариям выбросов МГЭИК СДСВ В1 и А1В соответственно. Поэтому предполагается, что одни и те же сценарии будут представлять собой модельные эксперименты КОРДЭКС наивысшего приоритета. В идеальном случае все региональные модельные эксперименты должны охватывать период 1951-2100 гг. для того, чтобы включить последний исторический период плюс весь XXI век. Однако для многих групп расчеты могут оказаться слишком трудоемкими, чтобы выполнять моделирование КОРДЭКС для всего вышеуказанного периода. Поэтому период 1951-2100 гг. разделен на пять 30-летних временных отрезков, и участвующие группы должны моделировать эти отрезки в следующем порядке приоритетности: 1981-2010 гг., 2041-2070 гг., 2011-2040 гг., 2071-2100 гг. и 1951-1980 гг. Первый период (1981-2010 гг.) является исходным для оценки модели и расчета изменения климата.

Второй временной отрезок, охватывающий будущий период, был выбран как компромисс между потребностями сообщества с точки зрения будущего периода времени и потребностями в получении четкого сигнала изменения. Все участвующие группы должны как минимум выполнить эти два временных отрезка, чтобы иметь приемлемый ряд модельных экспериментов для анализа и взаимного сравнения. Для обеспечения широкого участия, для первой фазы моделирования в рамках КОРДЭКС, установлено стандартное горизонтальное разрешение равное ~50 км (или 0,5 градуса).

Для получения оценок рабочих характеристик гидродинамических моделей, для отдельных областей будет выполнен ряд численных экспериментов с использованием данных реанализа Европейского центра среднесрочных прогнозов погоды (ЕЦСПП) ERA-Interim, которые охватывают период 1979-2005 гг.

Ключевым аспектом программы КОРДЭКС является наличие высокопроизводительных вычислительных средств и управление большим количеством требуемых входных данных моделей, а также выходная продукция моделей и взаимное сравнения, которые можно будет получить с помощью этой программы.

*Данная работа выполнена при поддержке грантового финансирования научных проектов КН МОН РК № AP05135848 «Моделирование будущего климата Центральной Азии в рамках международного проекта CORDEX (COordinated Regional climate Downscaling Experiment)».*

### **Литература**

1. Alapaty, K., R. Mathur, J. Pleim, Ch. Hogrefe, S. T. Rao, V. Ramaswamy, S. Galmarini, M. Schaap, R. Vautard, P. Makar, A. Baklanov, G. Kallos, B. Vogel, R. Sokhi (2011) New Directions: Understanding Interactions of Air Quality and Climate Change at Regional Scales. Atmospheric Environment. 49: 419–421.
2. Baklanov A., 2010: Chemical Weather Forecasting: A New Concept of Integrated Modelling. Advances in Science and Research, 4: 23-27.
3. Baklanov, A. and B. Grisogono (Eds.), 2007: Atmospheric Boundary Layers: Nature, Theory and Application to Environmental Modelling and Security. Springer, 248 p., ISBN: 978-0-387-74318-9
4. Baklanov A., U. Korsholm, A. Mahura, C. Petersen, A. Gross, 2008: ENVIRO-HIRLAM: on-line coupled modelling of urban meteorology and air pollution. Advances in Science and Research, 2, 41-46.

5. Baklanov A, J.H. Sørensen, A. Mahura, 2007: Methodology for Probabilistic Atmospheric Studies using Long-Term Dispersion Modelling. *Environ. Model. Assess.*, DOI 10.1007/s10666-007-9124-4
6. Baklanov A., O. Hänninen, L. H. Slørdal, J. Kukkonen, N. Bjergene, B. Fay, S. Finardi, S. C. Hoe, M. Jantunen, A. Karppinen, A. Rasmussen, A. Skouloudis, R. S. Sokhi, J. H. Sørensen, and V. Ødegaard. Integrated Systems for Forecasting Urban Meteorology, Air Pollution and Population Exposure. *Atmos. Chem. Phys.* 7: 855–874
7. Baklanov, A., P. Mestayer, A. Clappier, S. Zilitinkevich, S. Joffre, A. Mahura, N.W. Nielsen, 2008: Towards improving the simulation of meteorological fields in urban areas through updated/advanced surface fluxes description. *Atmospheric Chemistry and Physics*, 8, 523-543.
8. Baklanov, A., S. Grimmond, A. Mahura, M. Athanassiadou, 2009: Meteorological and Air Quality Models for Urban Areas. Springer, 2009, 184 p.
9. Baklanov, A., A. Mahura, R. Sokhi, 2010: Integrated systems of meso-meteorological and chemical transport models, Springer, 190 p., ISBN 978-3-642-13979-6
10. Baklanov A., B. Grisogono, R. Bornstein, L. Mahrt, S. Zilitinkevich, P. Taylor, S.E. Larsen, M.W. Rotach, H. J. S. Fernando (2011): The Nature, Theory, and Modeling of Atmospheric Planetary Boundary Layers. *Bull. Amer. Meteor. Soc.*, 92, 123–128. doi: 10.1175/2010BAMS2797.1
11. Grell, G. and A. Baklanov (2011): Integrated Modeling for Forecasting Weather and Air Quality: A Call for Fully Coupled Approaches. *Atmospheric Environment*, doi:10.1016/j.atmosenv.2011.01.017.
12. Penenko V., A.Baklanov, E.Tsvetora, A.Mahura, 2011: Direct and Inverse Problems in a Variational Concept of Environmental Modelling. *Pure and Applied GEophysics*, Vol.169, No 3/4, DOI 10.1007/s00024-011-0380-5.
13. Zilitinkevich SS, Hunt JCR, Grachev AA, Esau IN, Lalas DP, Akylas E, Tombrou M, Fairall CW, Fernando HJS, Baklanov A, Joffre SM (2006) The influence of large convective eddies on the surface layer turbulence. *Quart J Roy Meteorol Soc* 132: 1423-1456.
14. Zilitinkevich, S.S., I. Mammarella, A.A. Baklanov, and S.M. Joffre, 2008: The effect of stratification on the roughness length and displacement height. *Boundary-Layer Meteorology*. 129: 179-190.
15. *Hydrol. Earth Syst. Sci.*, 17, 2013, p.3661–3677: Comparison of climate change signals in CMIP3 and CMIP5 multi-model ensembles and implications for Central Asian glaciers. A.F. Lutz, W.W. Immerzeel, A. Gobiet, F. Pellicciotti, and M.F.P. Bierkens
16. *Asia-Pac.J. Atmospheric. Science*, 50(1), 2014, p.121-131: A history of mesoscale model development. J. Dudhia
17. *Asia-Pac.J. Atmospheric. Science*, 50(1), 2014, p.83-104: Dynamical downscaling: Fundamental issues from an NWP point of view and recommendations. Hong S.-Y. and M. Kanamitsu
18. *WMO Bulletin* 58 (3), July 2009, p.175-193: Addressing climate information needs at the regional level: the CORDEX framework. Filippo Giorgi, Colin Jones and Ghassem R. Asrar
19. *Climate Research*, Vol. 52, 2012, p.63–76: Simulation of temperature and precipitation climatology for the Central Asia CORDEX domain using RegCM 4.0. Tugba Ozturk, Hamza Altinsoy, Murat Türke, M. Levent Kurnaz

# ИССЛЕДОВАНИЕ РЕКУРРЕНТНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ МОДЕЛИРОВАНИЯ ЕСТЕСТВЕННЫХ ЯЗЫКОВ ФЛЕКТИВНЫХ КЛАССОВ

Жомартова Л.М.<sup>1</sup>, Рахимова Д.Р.<sup>2</sup>

<sup>1</sup>Казахский национальный университет имени аль-Фараби

<sup>2</sup>Институт информационных и вычислительных технологий МОН РК

e-mail: laurazhomartova@gmail.com<sup>1</sup>, di.diva@mail.ru<sup>2</sup>

***Аннотация.** Нейронные сети становятся все более популярными для моделирования сложных интеллектуальных систем. В то время как сети передачи данных используют только непредвиденные обстоятельства. С другой стороны, хорошо известно, что повторяющиеся сети трудно обучить, и вряд ли они будут демонстрировать полный потенциал повторяющихся моделей. Эти проблемы решаются с помощью архитектуры нейронной сети Long Short Term Memory. В этой работе мы анализируем этот тип сети для построения моделей языков флективного типа.*

***Ключевые слова:** нейронная сеть, глубинное обучение, NLP, LSTM, рекуррентные нейронные сети.*

## **Введение**

В последние года нейронные сети активно применяются для различных сложных и интеллектуальных систем. Возможности применения нейронных сетей очень различен, но при всей функциональности есть некие сложности и недостатки. Представьте, что хотите классифицировать, какое событие происходит в каждой точке фильма. Пока неясно, как традиционная нейронная сеть может использовать свои рассуждения о предыдущих событиях в фильме, чтобы сообщить более поздние. Рекуррентные нейронные сети решают эту проблему. Они представляют собой сети с циклом, что позволяет информации сохраняться.

Например, в предложении «облака в небе» попытаемся предсказать последнее слово. В данном случае не нужен никакой дополнительный контекст, это довольно очевидно, что искомое слово будет «небо». Но есть и другие ситуации, например, решая ту же задачу, но уже в тексте: «Я вырос во Франции... Я свободно говорю по-французски». Последняя информация говорит о том, что следующее слово, вероятно, является названием языка. Но для конкретики, какой язык, необходим контекст из прошлого, в данном случае – это Франция. Вполне возможно, что зазор (расстояние) между соответствующей информацией и точкой, где это необходимо, может быть очень большим. К сожалению, с увеличением разрыва рекуррентная нейронная сеть оказывается не в состоянии научиться соединять информацию. Следует отметить, что в пределах одного предложения такая сеть может работать весьма неплохо для предсказания слов, но для больших текстов это становится невозможным. Теоретически рекуррентная нейронная сеть способна обрабатывать такие «долгосрочные зависимости». Человек мог бы тщательно подобрать параметры для сети, чтобы решить эту проблему. Но на практике сеть не способна обучиться этому. Данная проблема была рассмотрена Josef Hochreiter в работе «Untersuchungen zu dynamischen neuronalen Netzen» [15] и Yoshua Bengio в работе «Learning Long-Term Dependencies with Gradient Descent is Difficult» [13] и была названа проблемой

исчезающего градиента. Это проблема не только рекуррентных нейронных сетей, это фундаментальная проблема всех нейронных сетей. Но не все так плохо. В 1997 году Hochreiter & Schmidhuber в работе «LONG SHORT-TERM MEMORY» была введена в пользование сеть под названием Long Short Term Memory networks (LSTM). LSTM – особый вид рекуррентных нейронных сетей, которые способны к обучению долгосрочных зависимостей. Они работают чрезвычайно хорошо на различных типах задач и сейчас широко применяются.

В данной работе будет представлена модель нейронной сети LSTM в области моделирования естественных флективных языков.

### **Проблема статистического моделирования флективного языка и исчезающего градиента**

Известно, что проблема статистического моделирования флективных языков представляет большую сложность, чем английского языка [1]. Основные проблемы возникают вследствие большого количества морфологических форм слов (лемм) и более свободного порядка слов [2]. Обе проблемы в результате усиливают разреженность данных и снижают эффективность n-граммных моделей. В то время как использование n-граммных моделей на первых стадиях распознавания сегодня является стандартной практикой [3], возможности для последующей обработки в рамках алгоритма распознавания, осуществляющего несколько проходов по входным данным, гораздо шире. Например, для переранжирования гипотез, возвращаемых процедурой лучевого поиска Витерби, может быть использована морфологическая, синтаксическая и семантическая информация. В последнем случае значения слов представляются посредством вложения слов в некоторое векторное пространство. К методам, осуществляющим такие вложения, относятся: латентно-семантический анализ [4], вероятностное тематическое моделирование [5] или нейронные сети [6]. В 2010 г. была представлена языковая модель на рекуррентной нейронной сети (RNNLM) [7]. Модель была предложена для английского языка

При наличии словаря существенного объема статистическое моделирование флективных языков составляет дополнительную техническую проблему для нейросетевого подхода. Большое количество различных словоформ приводит к пропорционально большему размеру выходного слоя, сложность алгоритма обучения линейна по объему выходного слоя.

Чтобы обойти эту проблему, можно было бы использовать следующий алгоритм работы:

1. Каждое входное слово предварительно лемматизируется внешним морфологическим анализатором. Леммы используются для предсказания последующих лемм.

2. Далее предсказанной леммы запускается линейный классификатор (например, логистическая регрессия), предсказывающий словоформу по лемме и морфологическим признакам контекста.

Данный подход позволяет миновать проблему разрастания словаря. Другой подход мог бы состоять в том, чтобы разделить выходной слой на два вектора – словарный (леммы) и морфологический (морфологические признаки). Ошибка предсказания в данном случае получалась бы суммированием ошибок на двух векторах.

Пусть имеется языковая модель, с помощью которой необходимо предсказать следующее слово, используя предыдущие. Когда разрыв между важной информацией и отрывком, где она необходима. RNN можно обучить использовать информацию,



полученную ранее. Но если необходим контекст, как в примере “Я выросла во Франции. Я свободно говорю по-французски”, то разрыв между существенной информацией и местом вставки становится шире. К сожалению, по мере роста разрыва, RNN невозможно обучить связывать информацию. Одной из проблем RNN является то, что её ранние модели очень сложно обучать из-за неустойчивого градиентного спуска. Обучение в предыдущих слоях происходит очень медленно, так как градиент становится всё меньше и меньше при обратном распространении. То есть если сеть работает довольно долго, то градиент может стать крайне неустойчивым.

$$\frac{\partial E}{\partial W} = \sum_{t=1}^T \frac{\partial E_t}{\partial W}$$

$$\frac{\partial E_t}{\partial W} = \sum_{k=1}^t \frac{\partial E_t}{\partial y_t} \frac{\partial y_t}{\partial h_t} \frac{\partial h_t}{\partial h_k} \frac{\partial h_k}{\partial W}$$

$$\frac{\partial h_t}{\partial h_k} = \prod_{j=k+1}^t \frac{\partial h_j}{\partial h_{j-1}} = \prod_{j=k+1}^t W^T * \text{diag} [f'(h_{j-1})]$$

$$\frac{\partial E}{\partial W} = \sum_{i=1}^T \sum_{k=1}^t \frac{\partial E_t}{\partial y_t} \frac{\partial y_t}{\partial h_t} \left( \prod_{j=k+1}^i \frac{\partial h_j}{\partial h_{j-1}} \right) \frac{\partial h_k}{\partial W}$$

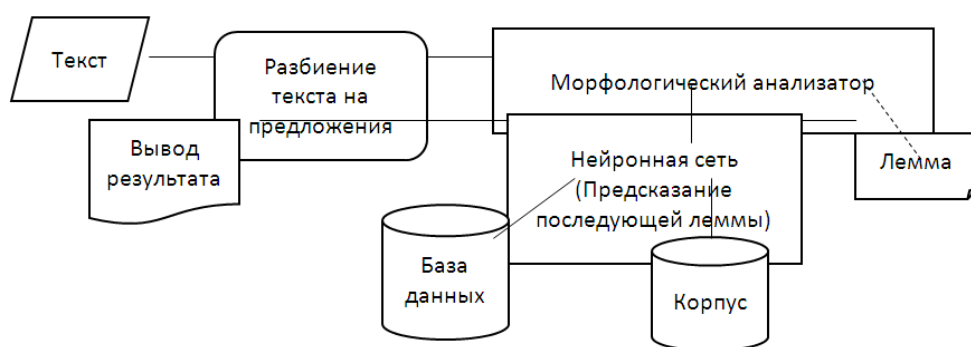
Итоговое выражение

$$\left\| \frac{\partial h_t}{\partial h_k} \right\| = \left\| \prod_{j=k+1}^t \frac{\partial h_j}{\partial h_{j-1}} \right\| \leq (\beta_w \beta_n)^{t-k}$$

Когда значение градиента становится равным нулю, оно может оставаться необнаруженным, однако качество обучения модели для отдаленных слов в корпусе резко уменьшается. Это вызвано проблемой исчезающего градиента (vanishing gradient problem) [9]. В теории RNN может обрабатывать подобные долговременные зависимости, однако на практике внимательный подбор параметров для решения игрушечных проблем не срабатывает. Эта проблема разрешается добавлением модулей долгой краткосрочной памяти (long short-term memory) в рекуррентную сеть [17].

### Архитектура сети для моделирования языка

Первым этапом будет подготовка корпуса для обучения нейронной сети. Система предполагает использование внешнего морфологического анализатора. И реализация линейного классификатора. Таким образом архитектура приложения будет выглядеть следующим образом:



**Рисунок 1 – архитектура модели**

### **Выводы**

В статье была рассмотрена нейронная сеть для проверки их применимости для моделирования казахского языка. В ходе исследования были изучены модели неронных сетей для различного типа языков, в том числе русского и чешского языка, так как эти языки так же как казахский являются флективными языками. Следующим этапом работы будет реализация подобранной модели нейронной сети с использованием языка python на базе TensorFlow технологии и их применение для обработки текста с использованием корпуса казахского языка.

### **Список литературы**

1. Oparin I. Language Models for Automatic Speech Recognition of Inflectional Languages. PhD thesis. University of West Bohemia, Pilsen. 2008. P. 125.
2. E.W.D. Whittaker. Statistical Language Modeling for Automatic Speech Recognition of Russian and English. PhD Thesis. Cambridge University. 2000. P. 141.
3. Deoras A., Mikolov T., Kombrik S. Approximate inference: A sampling based modeling technique to capture complex dependencies in a language model, Speech Communication, Speech Communication, 2013.Vol. 55. № 1. P. 162.
4. Bellegarda J. Exploiting latent semantic information in statistical language modeling, Proceedings of the IEEE, August, 2000. № 88. P. 1279.
5. Gildea D., Hofmann T. Topic-based language models using EM. History. Proceedings of the 6th European Conference on Speech Communication and Technology, 1999. № 6. P. 2167.
6. Andrieu C. et al. An introduction to MCMC for machine learning. Machine learning, 2003. Vol. 50. N 1-2. P. 5.
7. Mikolov T. et al. Recurrent neural network based language model. INTERSPEECH. 2010. Vol. 2. P. 3. Mikhail S. Kudinov. On Applicability of Recurrent Neural Networks to Language Modelling for Inflective Languages
8. Mikolov T. Statistical Language Models based on Neural Networks. PhD thesis. Brno: University of Technology. 2012. P. 133.
9. Vazhenina D., Markov K., Zelezny M. et al. Evaluation of Advanced Language Modeling Techniques for Russian LVCSR, SPECOM 2013, LNAI 8113. 2013. P. 124.
10. Mikolov T. et al. Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781, 2013.

11. Mikolov T., Sutskever I., Chen K., Corrado G. and Dean J. Distributed Representations of Words and Phrases and their Compositionality, Proceedings of NIPS, URL: <http://arxiv.org/pdf/1301.3781v3.pdf>, 2013. [12] Elman J. Finding Structure in Time, Cognitive Science, 1990. № 14. P. 179.
12. Bengio Y., Simard P., Frasconi P. Learning long-term dependencies with gradient descent is difficult, Neural Networks, IEEE Transactions on, 1994. Vol. 5. № 2. P. 157.
13. Pascanu R., Mikolov T., Bengio Y. On the difficulty of training recurrent neural networks, arXiv preprint arXiv:1211.5063, 2012.
14. Hochreiter S., Schmidhuber J. Bridging long time lags by weight guessing and “Long ShortTerm Memory”, Spatiotemporal models in biological and artificial systems, 1996. Vol. 37. P. 65.
15. Muzychka S., Romanenko A., Piontkovskaja I. Conditional Random Field for morphological disambiguation in Russian, Conference Dialog, 2014. P. 11.
16. Mikolov T., Kombrik S. RNNLM – Recurrent Neural Network Modeling Toolkit. ASRU, 2011. P. 456.

## **РЕАЛИЗАЦИЯ КАРКАСНОЙ ТЕХНОЛОГИИ В AUTODESK INVENTOR**

**Искакова М.Т.<sup>1,2</sup>, Калижанова А.У.<sup>1,2</sup>, Айткулов Ж.С.<sup>3</sup>,  
Тогжанова Г.О.<sup>3</sup>**

<sup>1</sup>*Казахский национальный университет им. аль-Фараби*

<sup>2</sup>*Институт информационных и вычислительных технологий КН МОН РК*

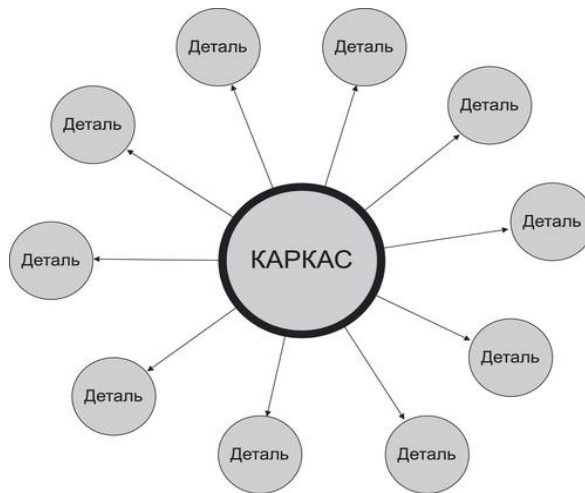
<sup>3</sup>*Алматинский университет энергетики и связи*

e-mail: [iskakovatolya@gmail.com](mailto:iskakovatolya@gmail.com)

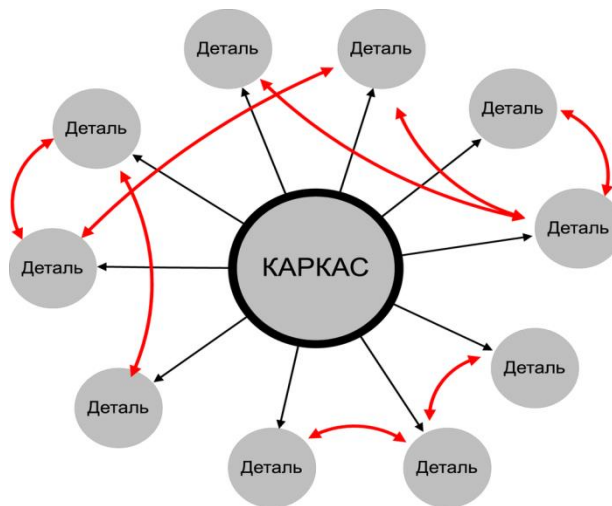
***Аннотация.** Одной из реализаций методологии проектирования и разработки для централизованного управления изменениями трудных сборок при модификации внешних параметров проекта является каркасное моделирование. Каркасная технология предполагает наличие у каждого элемента сборки только одного общего для всех источника данных о геометрии, ориентации и независимых параметрах задачи. Этим источником и является каркас.*

С точки зрения Autodesk Inventor, каркас представляет собой обычный файл детали, однако эта деталь играет в сборке совершенно особую роль. Все прочие элементы сборки создаются как производные детали от каркаса. Чтобы это было возможным, каркас должен содержать в себе все эскизы, параметры, рабочую геометрию, поверхности и даже, твердые тела, необходимые для создания основных конструктивных элементов всех будущих компонентов сборки. Полученные таким образом элементы имеют «от рождения» общую с каркасом систему координат, что позволяет легко зафиксировать их положение в сборке относительно ее главных осей. Таким образом, обеспечивается ассоциативная связь компонентов сборки с опорной геометрией в виде каркаса. Такая система связей «каркас – деталь» устойчива к операциям реструктуризации сборки верхнего уровня, что выгодно отличает ее от ассоциативности на основе адаптивных сборочных зависимостей. [1]

В англоязычной литературе термину «каркасное моделирование» соответствуют синонимы «Skeleton Modeling» и «Master Part Modeling». Строгого определения каркасного моделирования не существует. В зависимости от специфики проектируемых изделий применяются самые разнообразные варианты каркасной технологии, но в ее основе всегда лежит стройная и четкая система прямых зависимостей всех основных компонентов сборки от геометрии и параметров каркаса. Следует отметить две особенности данной технологии. В каркасном подходе последовательно реализуется принцип единственности источника данных о геометрии сборки и ее составляющих.



**Рис. 1– Классическая одноуровневая схема связей компонентов сборки с каркасом**



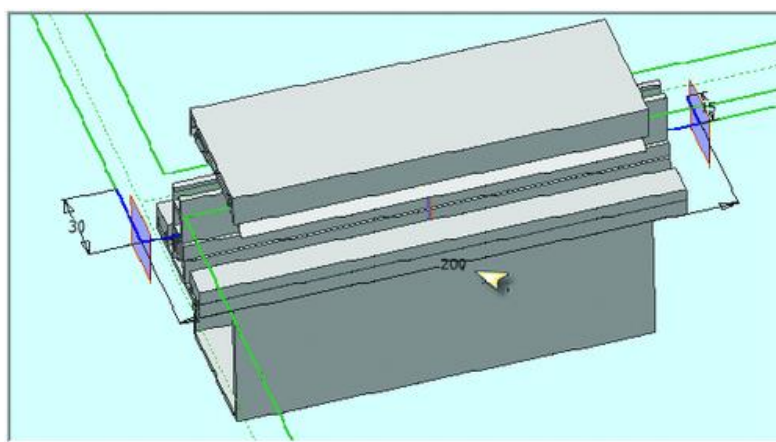
**Рис. 2– Схема взаимных связей компонентов сборки при наличии адаптивных зависимостей (отмечены красными дугowymi стрелками)**

В сборке, построенной на основе каркаса, каждая деталь в идеале имеет связь лишь с родительским каркасом. Система зависимостей имеет топологию типа «звезда» Рис. 1. Такая структура связей устойчива к операциям реструктуризации сборки

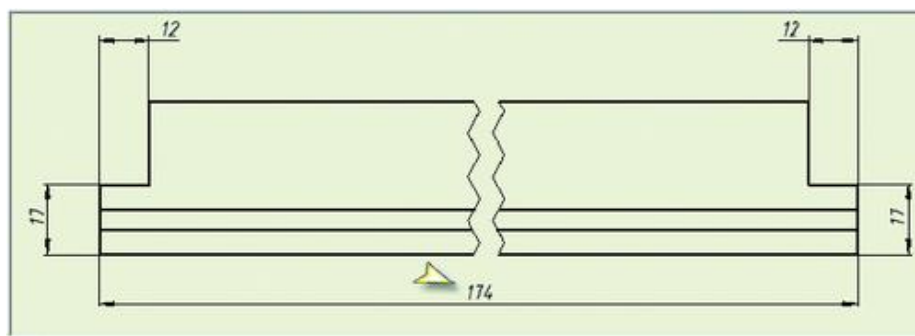
верхнего уровня. Это важный элемент системы обеспечения живучести больших, сложных сборок и проекта в целом. Минимизация сложных сборочных зависимостей «деталь ↔ деталь» и отсутствие адаптивных связей Рис. 2, помогают сделать процесс обновления сборки менее ресурсоемким и более надежным.

Каждая деталь типовой сборки создается как производный компонент от локального каркаса своей сборки и связана только с ним. Такое построение гарантирует ассоциативную связь геометрии компонентов сборки с геометрией локального каркаса. В локальном каркасе предусматриваются управляемые извне параметры и геометрия, множество которых определяется спецификой области проектирования.

Поскольку типовая параметризованная сборка создается заранее, мы в состоянии снабдить ее всеми необходимыми чертежами и спецификациями, подготовленными вручную с использованием штатного пользовательского инструментария ArhiCad Рис.3 и Рис. 4.



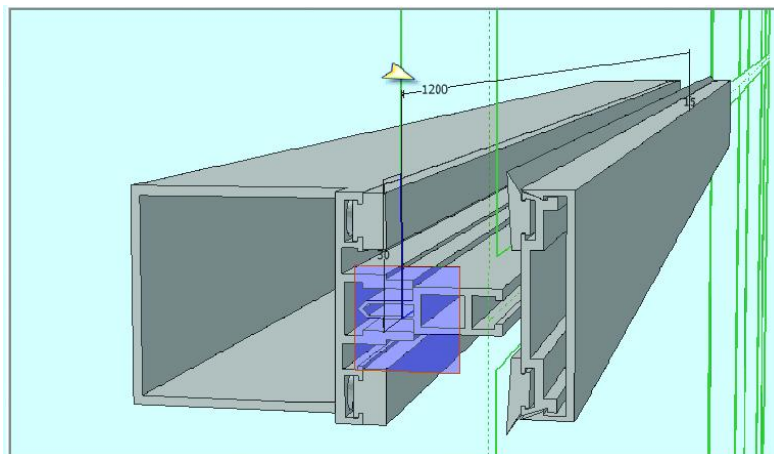
**Рис. 3 – Типовая ригельная сборка до установления связи с геометрией главного каркаса**



**Рис. 4 – Ассоциативный чертеж ригеля в библиотечной сборке отражает длину детали до установления связи локального каркаса сборки с главным каркасом**

На третьем этапе на локальный каркас накладываются все требуемые геометрические и размерные зависимости, после чего он окончательно принимает

геометрию, предписанную глобальным каркасом. В силу врожденных «генетических» связей обновятся все компоненты под сборки – клона, а также ассоциативно связанные с ними чертежи. На последнем, четвертом этапе обновленная под сборка – клон вставляется в главную сборку и фиксируется в ее системе координат. Благодаря установленной связи двух каркасов дальнейшее обновление компонентов под сборки-клона при изменениях геометрии и параметров глобального каркаса обрабатывается уже средствами ArhiCad Inventor автоматически Рис. 5.



**Рис. 5 – Длина деталей сборки, определяемая глобальным каркасом**

Представленная технология позволяет одновременно решить две задачи:

- обеспечивается ассоциативная связь каждого компонента модели с геометрией глобального каркаса через локальный каркас в качестве посредника;
- типовые под сборки могут содержать полные комплекты заранее подготовленных и оформленных рабочих чертежей, ассоциативно связанных с моделями. Немаловажно и отсутствие каких-либо ограничений на количество уровней в каркасной схеме («вторая производная» — не предел).

Каркасное моделирование является мощной и достаточно гибкой технологией создания сложных сборок в среде ArhiCad Inventor. К преимуществам многоуровневых схем «каркас в каркасе» следует отнести высокую производительность создания больших сборок, возможность получения типовой чертежной документации и существенное – на порядок – упрощение модели главного каркаса.

За плюсы приходится расплачиваться необходимостью тщательно планировать свои действия и анализировать удаленные последствия решений, принимаемых на этапе создания каркаса. Это требует от пользователя определенной квалификации. Но затраченные усилия окупаются строгой внутренней логикой получаемых моделей, их живучестью к изменениям внешних параметров, что в конечном итоге приводит к существенному снижению издержек на внесение в проект неизбежных исправлений. По сути дела, при каркасном моделировании мы неявным образом программируем ArhiCad Inventor на автоматическое отслеживание разнообразных зависимостей геометрии и состава сложных сборок от внешних параметров. Это создает надежный фундамент для быстрого реагирования на уточнения заказчиком исходной постановки задачи, что в рыночных условиях вполне конвертируется в денежные знаки.[2]

## Литература

1. Средства реализации каркасной технологии в Autodesk Inventor - [http://www.cadmater.ru/magazin/articles/cm\\_45\\_inventor.html](http://www.cadmater.ru/magazin/articles/cm_45_inventor.html) (05.06.2018г.)
2. Каркасное моделирование, или О пользе вторых производных - [http://studbooks.net/2322021/informatika/sredstva\\_realizatsii\\_karkasnoy\\_tehnologii\\_autod esk\\_inventor](http://studbooks.net/2322021/informatika/sredstva_realizatsii_karkasnoy_tehnologii_autodesk_inventor) (05.06.2018г.)
3. Беляков, С. И. Перспективы развития производственного потенциала строительных предприятий в современных условиях / С. И. Беляков // Недвижимость. Экономика. Управление. - 2009. - № 1. - С. 54-57.
4. Благовещенский Ф. А., Архитектурные конструкции. [учебник по специальности "Архитектура"] - 2007
5. Ильичев В. Ю. Основы проектирования экобиозащитных систем: Учебное пособие /В. Ю. Ильичев, А. С. Гринин; Под ред. А. С. Гринина. -М.: ЮНИТИ-ДАНА, 2002. -207 с.
6. Исаев М. И. Теория коррозионных процессов. Учебник. - М.: Метал-лургия, 1997. - 344 с.

## МНОГОУРОВНЕВОЕ ОПРЕДЕЛЕНИЕ МОДЕЛИ ТИПИЗИРОВАННОГО АТРИБУТНОГО РАЗГРАНИЧЕНИЯ ДОСТУПА

**Калимолдаев М.Н., Бияшев Р.Г., Рог О.А.**

*Институт информационных и вычислительных технологий КН МОН РК,  
Казахстан*

*e-mail: mnk@ipic.kz , brg@ipic.kz , olga@ipic.kz*

***Аннотация.** Сформулированы требования к методам разграничения доступа, обеспечивающие безопасное совместное использование информационных ресурсов, как в локальных, так и глобальных вычислительных средах.*

*Приводится многоуровневое определение модели разрабатываемого типизированного атрибутного разграничения доступа (ТАРД).*

*Показано, что модель ТАРД, определяемая на следующих уровнях: мета, объектном и уровне матрицы доступа с помощью языков логической и алгебраической спецификаций, отвечает поставленным требованиям универсальности, гибкости, удобства администрирования, способствующим обеспечению безопасности разграничения доступа вне зависимости от типа операционной среды.*

### **1. Введение**

В результате развития новых технологий вычислительные среды эволюционировали от централизованных систем до распределенных гетерогенных систем, осуществляющих децентрализованное хранение и обработку данных.

Методы предоставления доступа к информации становятся сервисно-ориентированными, что позволяет эксплуатировать одни и те же ресурсы в режиме совместного использования и требуют обеспечения их защиты путем организации

разграничения доступа на основе общих правил, оформляемых в виде политик разграничения доступа.

На сегодняшний день распределенные вычислительные среды предоставляют коллективный доступ к ресурсам различным видам пользователей. При этом субъекты получают доступ к объектам системы в соответствии с политикой авторизации, определяющей допуск к ресурсам определенных видов согласно предоставленным полномочиям с одновременным запретом различных видов несанкционированного доступа.

подавляющее большинство приложений снабжаются средствами контроля доступа в той или иной форме. Системы разграничения доступа, будучи важнейшими компонентами систем защиты, наиболее подвержены рискам из-за возможных ошибок в конфигурации политик разграничения доступа.

Актуальность проблемы защиты, а, следовательно, требования к политикам разграничения доступа растут по мере увеличения объемов хранимых данных и роста сложности программного обеспечения для их обработки.

## **2. Недостатки существующих моделей разграничения доступа и пути их преодоления**

Необходимо признать, что применяемые в настоящее время модели не обеспечивают безопасность защищаемых ими информационных ресурсов.

С начала 1970-х годов было разработано много моделей, основными из которых являются модели дискреционного разграничения доступа DAC (Discretionary Access Control), мандатного разграничения доступа MAC (Mandatory Access Control) и ролевого разграничения доступа RBAC (Role Based Access Control) [1, 2, 3].

Задачи идентификации сущностей в этих моделях выполняются путем присвоения субъектам и объектам уникальных имен, ввиду чего модели называют идентификаторными. Доступ субъекта к объекту осуществляется на основе проверки имен или приписанных им ролей.

Основным недостатком данных моделей является то, что они не учитывают дополнительных параметров разграничения доступа, что ведет к «грубому» разграничению доступа («coarse-grained access control») и служит причиной появления «избыточности прав доступа» у пользователей, нарушающей безопасность.

Также, идентификаторные модели не содержат средств администрирования полномочий. В широкомасштабных системах управление правами большого числа пользователей и машин становится слишком сложным и подверженным ошибкам.

В работе [4] были проанализированы недостатки существующих моделей и сформулированы проблемы, требующие решения при организации разграничения доступа.

Прежде всего, необходимо решить вопросы идентификации с целью избавления от избыточных прав путем обеспечения точного наименования сущностей.

Следующим недостатком является недостаток гибкости. Единственной политики безопасности, обеспечиваемой традиционными моделями DAC, MAC, RBAC в автономных приложениях, недостаточно для защиты данных в сложных системах, требующих интегрированных политик разграничения доступа для одновременного выполнения различных критериев защиты.

Отмечается необходимость введения новой парадигмы – «множественной политики», означающей, что в системе должна быть предусмотрена возможность применения разных политик авторизации в зависимости от требований безопасности среды, в которой функционирует система.



Для этого в системе должны присутствовать средства конструирования различных политик без реконфигурации самой системы. Как правило, политики конструируются на основе единого объекта - политики высокого уровня, или метаполитики, реализуемой обычно в виде фреймворков. Это позволяет создавать и применять политики для разграничения доступа по различным признакам, делая защиту многокритериальной.

В последнее время наблюдается значительный рост числа крупномасштабных распределенных открытых систем, состоящих из независимых автономных доменов. В таких системах необходимо идентифицировать ресурсы и пользователей наборами характеристик, т.к. решения о предоставлении доступа должны приниматься с учетом оценки целых наборов атрибутов субъектов и объектов.

В работе [10] были приведены итоговые требования к методам разграничения доступа, обеспечивающие безопасное совместное использование информационных ресурсов.

Было установлено, что модели разграничения доступа должны обладать универсальностью, обеспечивать удобство администрирования и гибкость управления безопасностью.

Для выполнения данных требований модели должны обладать следующими характеристиками:

- иметь формальное определение,
- содержать средства администрирования,
- обеспечивать возможность применения множественных политик и включать способы их конструирования,
- язык описания модели должен быть высокоуровневым, т.е. не зависеть от операционной среды.

Были перечислены основные задачи, требующие решения при определении политик разграничения доступа и конструировании соответствующих им моделей, а именно задача точной идентификации сущностей, задача обеспечения возможностей динамического конструирования политик и задача использования множественных политик в рамках одной системы.

### **3. Методы атрибутивного разграничения доступа**

Для решения проблем был предложен атрибутивный метод разграничения доступа (Attribute Based Access Control, ABAC). Его основу составляет безидентификаторный подход, который заключается в обозначении субъектов и объектов совокупностями атрибутов, и позволяющий принимать решение по управлению доступом без предварительного знания субъектов или их отношения к поставщику услуг.

Наиболее общим определением ABAC является следующее. Атрибутивное разграничение доступа – это метод, посредством которого запросы субъекта на выполнение определенных операций над объектом удовлетворяются или отвергаются на основе приписанных им атрибутов, условий среды выполнения и набора политик, сформулированных с учетом этих условий и атрибутов [5, 6].

Преимущество ABAC состоит в том, что оно позволяет создавать политики доступа на основе атрибутов пользователей и объектов, а не назначать роли, права собственности или метки безопасности вручную системным администратором. Это упрощает администрирование в сложных системах с большим числом пользователей, устраняя необходимость ручного вмешательства при авторизации пользователей для определенных ролей или уровней безопасности, а также создавая возможность

автоматизации решения по управлению доступом для удаленных пользователей из других доменов.

Система именования сущностей атрибутами обеспечивает точность идентификации и, следовательно, «точное» разграничение доступа («fine-grained access control») в процессе организации защищенного использования ресурсов, не допуская избыточных прав доступа у пользователей.

Языки спецификации моделей АВАС дают гибкость и выразительную мощь описаниям политик безопасности. При этом многие из них разрешают моделировать традиционные методы разграничения доступа – DAC, MAC, RBAC.

Модели атрибутного разграничения доступа находят применение в самых разных областях современных вычислений для защиты приложений, баз данных, файловых серверов, в облачных средах и больших данных.

За последнее время было разработано множество АВАС-моделей, как базовых, так и специализированных [7, 8, 9].

Их объединяет то, что они могут рассматриваться в качестве основополагающих моделей нового направления защиты, способных решать задачи разграничения доступа, поставленные в [4, 10].

Типичная АВАС-модель содержит следующие компоненты:

- атрибуты пользователей;
- атрибуты объектов;
- атрибуты контекста или вычислительной среды;
- политики авторизации, основанные на этих атрибутах.

Атрибуты определяются в виде функций, аргументами которых служат субъекты или объекты, и которые возвращают результат в виде значений их атрибутов.

Политика авторизации предоставляет группам пользователей определенные виды доступа (такие, например, как чтение и запись) к заданным объектам на основе оценки значений их атрибутов.

Преимущество данного подхода для представления атрибутных политик разграничения доступа заключается в его простоте и легкости использования. Создание новых правил авторизации не представляет трудностей, так как не включает дополнительных расходов, требуемых, например, для инжиниринга ролей в RBAC. Эти правила способны гибко и в сжатой форме описывать даже сложные политики. Не существует ограничений на количество используемых в них атрибутов и сложность языка описания правил [5].

С другой стороны, создание выразительных вычислительных языков для спецификации атрибутных правил разграничения доступа делает задачи вычисления значений разнородных атрибутов в процессе конструирования и выполнения политик NP-полными или даже неразрешимыми, что служит, вместе с отсутствием формальных определений моделей и сложностью администрирования, главным препятствием широкому применению метода АВАС.

#### **4. Типизированное атрибутное разграничение доступа**

В работах [11-20] приводятся описание и формальное представление разрабатываемого метода и модели типизированного атрибутного разграничения доступа (ТАРД).

Модель ТАРД, будучи определенной на нескольких уровнях с помощью языков логической и алгебраической спецификации, обеспечивает решение основных задач политиками разграничения доступа, создавая таким образом необходимые условия для безопасного доступа к защищаемым информационным ресурсам.

Модель типизированного атрибутного разграничения доступа принадлежит классу моделей АВАС, но, в отличие от АВАС, атрибутам безопасности сущностей ТАРД приписаны определенные типы. Решение о возможности доступа принимается на основе обработки однотипных атрибутов пары субъект-объект.

Описываемая модель ТАРД может быть отнесена к разряду моделей разграничения доступа общего назначения, основанной на логических формулах. В ее состав входят:

- средства определения возможности доступа субъектов к объектам в соответствии с их полномочиями;

- средства двухступенчатого администрирования – для конструирования политик авторизации и для управления идентификацией сущностей в процессе разграничения доступа.

В соответствии с определением типа Т, модель ТАРД имеет многоуровневое представление, которое позволяет конструировать типы, представляющие собой различные политики безопасности, включая традиционные DAC, MAC, RBAC, реализуя таким образом парадигму «множественной политики».

Идентификация сущностей ТАРД осуществляется путем их типизации. Предоставляется возможность множественной типизации сущностей, каждая из которых может иметь один атрибут определенного типа или ряд атрибутов различных типов.

Тип Т определяется как математический объект, содержащий домен типа - конечное множество всевозможных значений атрибутов в виде полного частично упорядоченного множества, структурированного отношением предшествования, и определенных на нем операций типизации Type, представленных непрерывными монотонными функциями, а также операции доступа Acc [11, 13, 17, 18].

Операция типизации типа Т присваивает сущности e метку безопасности T(e):

$$T(e) = \text{Type}(e).$$

Множественная метка безопасности сущности e в виде кортежа  $T_1(e), \dots, T_k(e)$  присваивается ей в результате применения операций типизации типов  $T_1, \dots, T_k$ .

Предикат Acc осуществляет сравнение меток безопасности типа Т субъекта и объекта, разрешая/отвергая возможность доступа. Истинностное значение функции означает разрешение на доступ субъекту s к объекту o:

$$\text{Acc}(T(s), T(o)) = \text{true/false}.$$

Набор операций типа Т образует механизм разграничения доступа, а также служит средством реализации политики типизированного атрибутного разграничения доступа P(T), определяемой типом Т [13].

Спецификация политики безопасности P(T) задается конкретным видом структуры домена типа и соответствующими этой структуре операциями.

Тип Т служит ограничением на значения атрибутов и круг операций с атрибутами данного типа. Данное обстоятельство лежит в основе принципа безопасности моделей ТАРД [16].

Метод ТАРД представляется на следующих уровнях: метауровень МЕТА, объектный уровень OBJ, уровень матрицы доступа AM. В соответствии с этим дается многоуровневое определение типа Т.

На уровне МЕТА тип Т представлен метатипом  $T_{МЕТА}$ , являющимся обобщенным представлением политики  $P_{МЕТА}(T)$ . Он служит для порождения политик типизированного атрибутивного разграничения доступа объектного уровня  $P_{ОВJ}(T)$ .

Тип Т объектного уровня – это конкретные политики разграничения доступа  $P_{ОВJ}(T)$ , получаемые из метаполитики  $P_{МЕТА}(T)$  путем моделирования ее семантики. Тип  $T_{ОВJ}$  является интенциональным представлением типа Т.

Тип  $T_{AM}$  уровня АМ – это реализация политики типизированного атрибутивного разграничения доступа  $P_{AM}(T)$  в виде множества типизированных переменных, образующих матрицу доступа и являющихся экстенциональным представлением типа Т.

Построена модель  $M(T)$ , которая является формальным представлением типа Т. Она предназначена для конструирования и представления предметной области ТАРД путем моделирования ее семантики.

Модель, также как и тип, имеет многоуровневое определение.

**Метамодель**  $M_{МЕТА}(T)$  представляет метатип  $T_{МЕТА}$ . Играет роль фреймворка или инструмента для создания конкретных моделей ТАРД.

**Объектная модель**  $M_{ОВJ}(T)$  представляет объектный тип  $T_{ОВJ}$ . Является моделью конкретной политики безопасности типа Т, полученной из метатипа  $T_{МЕТА}$ .

**Матрица доступа**  $M_{AM}(T)$  – сложноструктурированная область типизированных значений атрибутов в виде меток безопасности сущностей. Содержит результаты выполнения политики безопасности  $P(T)$ , получаемые в процессе функционирования модели  $M_{ОВJ}(T)$ .

Результаты работы модели на разных уровнях определения зависят от выполняемых ею операций.

На метауровне МЕТА осуществляется конструирование типа  $T_{ОВJ}$  специальной операцией интерпретации I. Конструирование типа заключается в определении вида структуры домена типа  $T_{ОВJ}$  как подструктуры домена типа  $T_{МЕТА}$ , и присвоении значений его элементам. Результатом является создание различных моделей объектного уровня  $M_{ОВJ}^i(T)$ , предназначенных для формирования матрицы доступа и предоставления доступа субъектов к объектам:

$$M_{МЕТА}(T, I) \rightarrow M_{ОВJ}^i(T).$$

Формирование матрицы доступа осуществляется операциями  $T_{ур}$  сконструированных моделей  $M_{ОВJ}^i(T)$ , которое заключается в присвоении полномочий сущностям в виде их меток безопасности:

$$M_{ОВJ}^i(T, Type(e)) \rightarrow M_{AM}^i(T).$$

На уровне АМ производится обработка создаваемой матрицы доступа. При этом выдается разрешение на доступ согласно критерию, задаваемому типом  $T_{ОВJ}$ , в результате выполнения операции доступа Acc модели  $M_{ОВJ}^i(T)$ :

$$M_{ОВJ}^i(T, Acc(T(s), T(o))) \rightarrow \{true, false\}.$$

Одновременное использование полученных моделей в рамках одной системы обеспечивает разграничение доступа по ряду критериев, определяемых политиками  $T_1, \dots, T_K$ . Критерием при этом служит результат оценки значения предиката  $MM_{AM}$

$(T_1, \dots, T_k)$ , который предоставляет доступ субъекту  $s$  к объекту  $o$  при условии одновременного выполнения критериев, задаваемых типами  $T_1, \dots, T_k$ :

$$MM_{AM}(T_1, \dots, T_k) = M_{AM}(T_1, Acc_1) \wedge \dots \wedge M_{AM}(T_k, Acc_k). \quad (1)$$

Построено формальное представление модели ТАРД, определяемое на уровнях МЕТА, OBJ и АМ в виде языковых спецификаций  $T^{AS}$  и  $T^{LS}$ , взаимосвязь которых определяет ее функционирование [17, 18, 19].

Модель ТАРД представима в виде пары

$$M = (T^{AS}, T^{LS}),$$

где  $T^{AS} = (D, \sigma)$  - многоуровневое определение многосортной алгебраической системы, состоящей из домена  $D$  и сигнатуры  $\sigma$  операций типа  $T$ , которая служит для представления состояния каждого уровня модели:

$$T^{AS} = (T_{META}^{AS}, T_{OBJ}^{AS}, T_{AM}^{AS}).$$

$T^{LS}$  – многосортная логическая система, включающая язык  $L$  с алфавитом, представленным доменом  $D$  и правилами грамматики – операциями типа  $T$ , а также аксиомами  $Ax$  и правилами вывода  $Inf$ :

$$T^{LS} = (L, Ax, Inf).$$

Функционирование многоуровневой модели ТАРД заключается в моделировании семантики следующего уровня путем интерпретации модели, представляющей предыдущий уровень. При этом  $T^{LS}$ , представленная в виде тройки  $(T_{META}^{LS}, T_{OBJ}^{LS}, T_{AM}^{LS})$ , образует систему семантического моделирования.

Ввиду того, что определение  $T^{LS}$  включает функции интерпретации, схему взаимодействия спецификаций модели ТАРД можно представить следующим образом:

Метауровень МЕТА:

$$T_{META}^{LS}(T_{META}^{AS}) \rightarrow T_{OBJ}^{AS};$$

Объектный уровень OBJ:

$$T_{OBJ}^{LS}(T_{OBJ}^{AS}) \rightarrow T_{AM}^{AS}.$$

Уровень матрицы доступа АМ:

$$T_{AM}^{LS}(T_{AM}^{AS}) \rightarrow \{true, false\}.$$

В целях реализации модели ТАРД для каждого ее уровня дается формальное определение системы типизированного атрибутивного разграничения доступа как программы (или программной системы)  $S(T)$  на языке  $T$ , основанной на модели  $M(T)$ . Она обеспечивает разграничение доступа в соответствии с политикой, представленной типом  $T$  [19].

Программа  $S$  является общим представлением функции семантического моделирования состояний предметной области ТАРД.

Многокритериальное разграничение доступа в соответствии с критериями политик, представленных типами  $T_1, \dots, T_k$ , обеспечивается программой, построенной на основе нескольких моделей  $M(T_1), \dots, M(T_k)$  согласно критерию (1).

Архитектура системы ТАРД состоит из модуля настройки и модуля выполнения, соответствующих мета- и объектным уровням определения модели ТАРД.

Модуль настройки (Администратор 1) осуществляет конфигурирование системы путем задания нескольких типов атрибутов  $T_1, \dots, T_k$  (которое заключается в определении их структуры и задании значений элементов доменов), служащих критериями разграничения доступа. Набор этих критериев должен отвечать требованиям защиты данной информационной системы.

Модуль выполнения осуществляет две функции - функцию администрирования (Администратор 2), присваивая сущностям (субъектам и объектам) кортежи значений типов  $(T_1(e), \dots, T_k(e))$  в качестве их меток безопасности, и функцию разграничения, которая обеспечивает доступ субъектов к объектам на основе сравнения их меток безопасности одинаковых типов.

Необходимо отметить, что выразительность метода ТАРД несколько ограничена по сравнению с общепринятыми АВАС методами.

Описанная модель ТАРД обладает следующими характеристиками:

- принцип обработки атрибутов одинаковых типов является предпосылкой обеспечения скорости вычислений их значений;
- имеет возможность формально доказывать правильность решений о предоставлении доступа, используя дедуктивный аппарат логической спецификации модели;
- обеспечивает наглядность и контроль процесса администрирования;
- способна динамически конструировать новые модели разграничения доступа вместе с возможностью моделировать традиционные DAC, MAC, RBAC;
- модель непосредственно реализуема на языках функционального и логического программирования с использованием аппарата программирования в ограничениях.

Перечисленные особенности позволяют использовать системы типизированного атрибутного разграничения доступа в качестве центров управления политиками безопасности, создаваемых в локальных и глобальных вычислительных средах.

На основе этих центров, рассматриваемых в качестве виртуальных организаций, могут создаваться домены администрирования, поддающиеся локализации и обеспечивающие в силу этого полную защиту информации на контролируемых ими участках распределенных гетерогенных вычислительных сред [20].

### **Заключение**

Рассмотрены основные методы разграничения доступа в их развитии – от традиционных моделей DAC, MAC и RBAC до последних разработок - моделей, основанных на методе атрибутного разграничения доступа АВАС.

Приведен перечень недостатков традиционных моделей и пути их преодоления. Сформулированы требования, обеспечивающие гибкость управления безопасностью, универсальность и удобство администрирования, которым должны отвечать модели разграничения доступа, функционирующие как в локальных, так и распределенных гетерогенных средах.

Описана разрабатываемая в настоящее время модель типизированного атрибутного разграничения доступа, показано, что она отвечает основным

требованиям построения моделей, обеспечивающим безопасное пользование разделяемыми ресурсами.

### Литература

1. Sandhu R.S., Samarati P. Access control: principle and practice. *Communications Magazine*, IEEE, 32(9):40–48, 1994.
2. Девянин П.Н. Модели безопасности компьютерных систем: Учеб. пособие для студ. высш. учеб. заведений. – М.: Издательский центр «Академия», 2005. – 144 с.
3. Гайдамакин Н.А. Разграничение доступа к информации в компьютерных системах / – Екатеринбург: Изд-во Урал. ун-та, 2003 . – 328 с.
4. Hosmer H. 1993. The multipolicy paradigm for trusted systems. In *Proceedings on the 1992-1993 workshop on New security paradigms (NSPW '92-93)*, J. Bret Michael, Victoria Ashby, and Catherine Meadows (Eds.). ACM, New York, NY, USA, 19-32. DOI=<http://dx.doi.org/10.1145/283751.283768>
5. Hu V. C., Ferraiolo D., Kuhn R., Schnitzer A., Sandlin K., Miller R., and Scarfone K. Guide to attribute based access control (ABAC) definition and considerations. NIST Special Publication, 800:162, 2014. <http://dx.doi.org/10.6028/NIST.SP.800-162>
6. Servos D., Osborn S. Current Research and Open Problems in Attribute-Based Access Control. *ACM Computing Surveys (CSUR)* 49, 4, Article 65 (January 2017), 45 pages. DOI: 10.1145/3007204
7. Jin X., Krishnan R., Sandhu R.S. A unified attribute-based access control model covering DAC, MAC and RBAC. *DBSec*, 12:41–55, 2012.
8. Servos D. and Osborn S. L. HGABAC: Towards a formal model of hierarchical attribute-based access control. In *Foundations and Practice of Security*, pages 187–204. Springer, 2014.
9. Yuan E. and Tong J. Attributed based access control (ABAC) for web services. In *Proceedings. 2005 IEEE International Conference on Web Service*. IEEE, 2005.
10. Калимолдаев М.Н., Бияшев Р.Г., Рог О.А. Анализ атрибутивных методов разграничения доступа // Принята к печати в журнале ПДМ (Прикладная дискретная математика) 23.04.2018.
11. Калимолдаев М.Н., Бияшев Р.Г., Рог О.А. Формальное представление функциональной модели многокритериальной системы разграничения и контроля доступа к информационным ресурсам // *Проблемы информатики*. – 2014. – № 1(22). – С. 43-55.
12. Rog O.A. Polymorphic typing of entities in the multi-criteria system of access control and a task of constructing types // *Information Technologies, Management and Society*. The 12 th International Scientific Conference Information Technologies and Management. 2014 April 16 – 17. Riga, 2014. - с. 66.
13. Бияшев Р.Г., Калимолдаев М.Н., Рог О.А. Полиморфная типизация сущностей и задача конструирования механизма многокритериального разграничения доступа. // *Известия НАН РК. Серия физико-математическая*. – 2014. – № 5. – С. 33-41.
14. Бияшев Р.Г., Калимолдаев М.Н., Рог О.А. Конструирование систем многокритериального атрибутивного разграничения доступа в облачных структурах. // *Одиннадцатая Международная Азиатская Школа-семинар "Проблемы оптимизации сложных систем"* Кыргызская Республика, Иссык-Кульская область, г. Чолпон-Ата, 27 июля - 7 августа 2015 г. – С. 148-152.
15. Бияшев Р.Г., Калимолдаев М.Н., Рог О.А. Логический подход к организации многокритериального атрибутивного разграничения доступа. // *Abstracts of the International Conference "Computational and Informational Technologies in Science,*

Engineering and Education” (September 24-27, 2015). – Almaty: Қазақ университеті, 2015. – Р.86.

16. Бияшев Р.Г., Калимолдаев М.Н., Рог О.А. Представление ограничений моделей атрибутного разграничения доступа // Известия НАН РК. Серия физико-математическая. – 2016. – № 1. – С. 58-65.

17. Бияшев Р.Г., Калимолдаев М.Н., Рог О.А. Моделирование семантики типизированного атрибутного разграничения доступа // журнал Проблемы информатики, 2017, № 1. С. 25-37.

18. Калимолдаев М.Н., Бияшев Р.Г., Рог О.А. Применение логики для построения моделей разграничения доступа к информации // «Доклады Национальной Академии Наук Республики Казахстан» 2017, №3. С. 48-54.

19. Калимолдаев М.Н., Бияшев Р.Г., Рог О.А. Основы архитектуры программных систем для осуществления типизированного атрибутного разграничения доступа // Современные проблемы информатики и вычислительных технологий: Мат. науч. конф. (29-30 июня 2017 г). – Алматы, 2017, – С. 88-95.

20. Калимолдаев М.Н., Бияшев Р.Г., Рог О.А. О применении типизированного атрибутного разграничения доступа в глобальных вычислительных средах // Известия Научно-Технического Общества «КАХАК» Алматы, 2017 г., № 3 (58), – С. 30-36.

УДК 004.891

## ИНСТРУМЕНТЫ ДЛЯ ОБРАБОТКИ ПОТОКОВЫХ ДАННЫХ

**Калимолдаев М.Н., Магзом М.М., Косынбай Е.Б.**

*Институт информационных и вычислительных технологий*

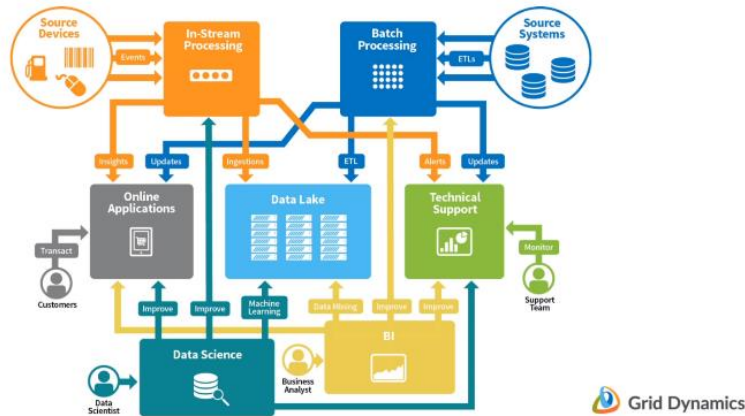
**Аннотация.** Факт, прогрессивного роста количества и качества источников данных, определяет одно из ключевых требований к проектированию архитектур на основе распределенных систем, способных обрабатывать потоковые данные с миллиардов устройств одновременно и производить анализ миллионов событий одновременно. В данной статье описываются принципы реализации архитектуры для обработки потоковых данных, необходимый набор элементов системы и их реализация.

**Ключевые слова:** инструменты обработки потоковых данных, распределенная архитектура, большие данные.

### **Введение**

Проблемы технологического характера обработки потоковых больших данных существуют по сей день, решением данных проблем может быть разработка комплексной модульной системы для отслеживания отклонений и событий данных, автономной или свободно интегрируемой в различные информационные системы, основанной на архитектуре потоковой обработки больших данных на основе распределенных вычислений.





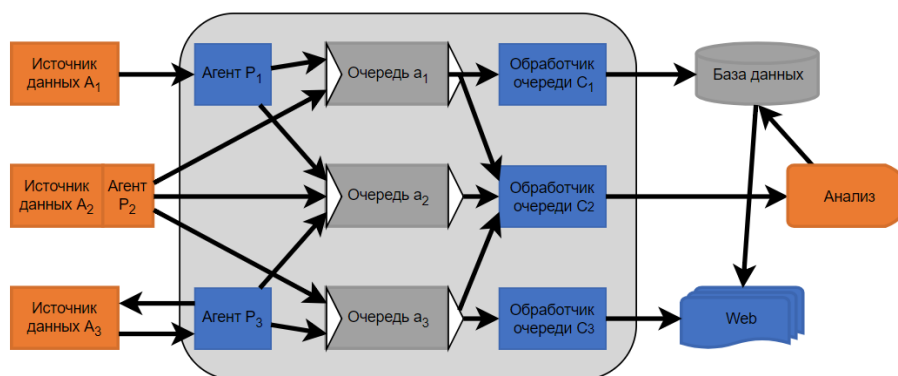
**Рис. 1. Эталонную архитектуру сервиса для потоковой обработки больших данных [1]**

К данной задаче, ранее, уже были предложены решения, одним из самых популярных предложений является концепция компании Grid Dynamics, на рисунке 1 [1]. В данной статье предлагается один из множества вариантов реализации архитектуры с подобной концепцией, предложения.

Важно упомянуть что абсолютно все программные элементы, отмеченные в архитектуре, могут быть реализованы за счет компонентов с открытым исходным кодом, готовых для образования масштабируемого облачного сервиса, что позволит в свою очередь значительно сократить бюджет ИТ инфраструктуры.

### Сбор информации

Данный модуль отвечает за непосредственный сбор данных от источников данных вне зависимости push или pull методов поддерживаемых источниками данных, а также проведение первоначальной обработки данных, приведение данных к стандартному виду внутренней экосистемы. Алгоритм его работы проиллюстрирован на рисунке 2.



А – Источники данных; Р – Агенты для постановки данных в очередь;  
а – очереди сообщений; С – обработчики очередей

**Рис. 2. Алгоритм функционирования модуля сбора данных**

Система сбора информации представляет собой совокупность распределенных агентов. Входные агенты, принимающие поток данных или забирающие порционно от источников данных, далее очереди данных, выходные агенты обработки очереди и передающие данные внутри системы соответствующим модулям. Наличие множества агентов решает проблемы различных форматов данных, т.к. позволяет проводить первоначальную обработку данных, снижая вычислительную нагрузку на блок анализа.

Apache Kafka - распределённая стриминговая платформа, а также очень быстрый распределенный брокер сообщений, позволяющий реализовать требуемый функционал, разработана компанией LinkedIn, впоследствии передана Apache Foundation [2]. Агенты Kafka (producer и consumer) могут быть реализованы разных современных языках программирования, таких как Python, Java, Scala. В основу архитектуры Kafka легла паттерн концепция Издатель-подписчик. Сравним функционал Kafka с аналогами.

С одной стороны, функционал Kafka ограничен только одним паттерном, тем самым уступая по функционалу многим другим платформам, к примеру платформе RabbitMQ [3]. Однако иные платформы удаляют сообщения после успешной доставки, Kafka же хранит сообщения любой необходимый период, по умолчанию - неделя, тем самым повышая гарантии доставки сообщений, в случае сбоя последующих зависимых модулей. Также требования к пропускной способности сети и ресурсам оборудования, на котором будет запущена Kafka ниже, чем у популярных брокеров сообщений ActiveMQ и RabbitMQ, так цифры пропускной способности: 20 кБит/с сообщений в секунду у Kafka против 100 кБит/с у RabbitMQ. С точки зрения кластеризации Kafka имеет преимущество перед аналогами простотой конфигурирования и масштабированием. С точки зрения кластеризации, один Kafka-брокер - это кластер, чтобы добавить второй экземпляр Kafka, достаточно указать первому о наличии второго. Kafka сохраняет сообщения на диск, что позволяет между узлами кластера синхронизировать конечные разметки, вместо репликации данных, при этом сохраняется логика обработки топиков. Также на выходе пропускная способность вырастает экспоненциально, т.к. Один топик может обрабатываться несколькими брокерами. [4]

Непосредственно LinkedIn применял Kafka для сбора и агрегации пользовательской активности на сайте: идентификация манипуляций, поиска, посещение страниц; отправляя данные на топики, а уже оттуда забирали другие сервисы: аналитики, обработки, постоянного хранения, и т.п. Также Kafka превосходно используется для агрегации логов, для примера вместо того, чтобы манипулировать файлами, можно сохранять историю действий для повторения потока данных-событий на иных копиях файлов, который можно перенаправлять, анализировать и хранить.

### **Хранение данных**

Модуль предназначен для длительного хранения данных системы, как исходных данных, так и полученных на основе исходных данных путем вычислений. Данный модуль, также, как и предыдущий должен быть масштабируемым, поддерживать кластеризацию, распределение обработки данных и высокую отказоустойчивость.

Apache Cassandra - распределённая система управления базами данных, относящаяся к классу NoSQL-систем и рассчитанная на создание высокомасштабируемых и надёжных хранилищ огромных массивов данных, представленных в виде хэша. Изначально Cassandra была разработана в Facebook, в

2009 году передана Apache Software Foundation, где продолжается развитие проекта. Промышленные решения на базе Cassandra развёрнуты для обеспечения сервисов таких компаний, как Cisco, IBM, Cloudkick, Reddit, Digg, Rackspace, Apple и Twitter. К 2011 году крупнейший кластер серверов, обслуживающий единую базу данных под управлением Cassandra, насчитывал более 400 машин и содержал данные размером более 300 ТБ. [5]

Концепция Cassandra основана на понятии - пространства ключей (keyspace), равнозначна понятию (database schema) в реляционной модели. Далее в пространство ключей входят колоночные семейства (column family), что соответствует понятию реляционной таблицы. Далее колоночные семейства включают в себя колонки (column), которые объединяются при помощи ключа (row key) в записи (row). Колонка состоит из трех частей: имени (column name), метки времени (timestamp) и значения (value). Колонки в пределах записи упорядочены. В отличие от реляционной БД, никаких ограничений на то, чтобы записи (строки) содержали колонки с такими же именами, как и в других записях - нет. Также в последних версиях Cassandra появилась возможность выполнять запросы определения и изменения данных (DDL, DML) при помощи языка CQL - язык запросов, предлагает модель, близкую к SQL в том смысле, что данные помещаются в таблицы, содержащей строку из столбцов, а также создавать вторичные индексы (secondary indices). Структура Cassandra приведена на рисунке 3. [6]

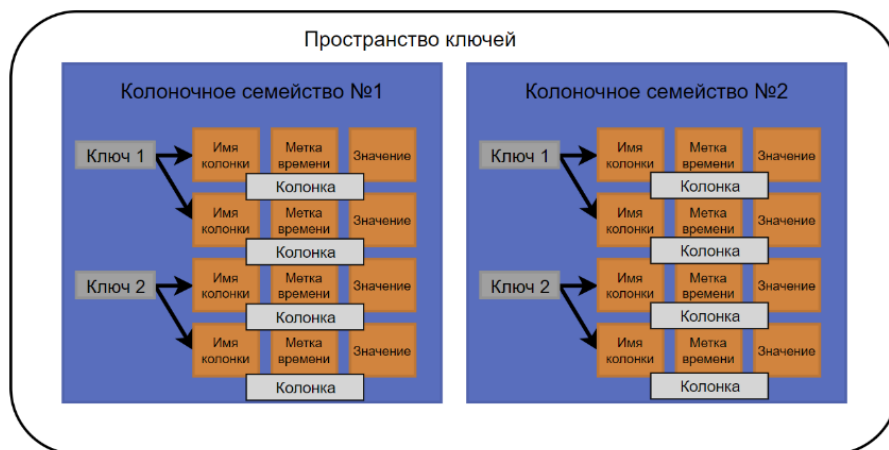


Рис. 3. Структура хранения данных Cassandra

Конкретное значение, хранимое в Cassandra, идентифицируется: пространством ключей - это привязка к приложению (предметной области). Позволяет на одном кластере размещать данные разных приложений; колоночным семейством - это привязка к запросу; ключом - это привязка к узлу кластера. От ключа зависит, на какие узлы попадут сохранённые колонки; именем колонки - это привязка к атрибуту в записи. Позволяет в одной записи хранить несколько значений.

### Анализ данных

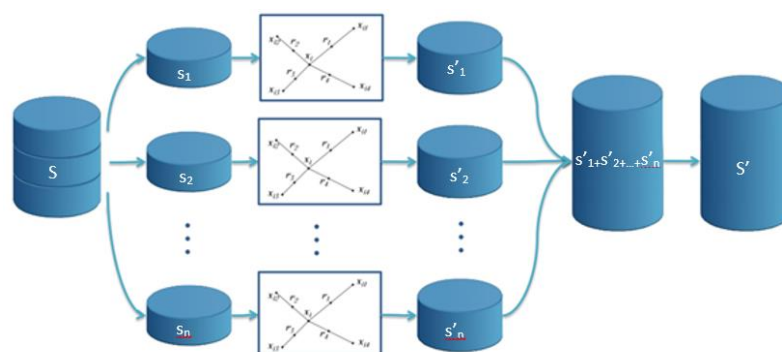
Модуль анализа функционально реализует модель классификации и выделения ключевых признаков в данных, а также алгоритмы сравнения и анализа. В задачи блока анализа также входит взаимодействие с системами хранения данных, для сохранения результатов.

Выполняет роль реализует алгоритмы классификации. Также, в пределах данного блока производится хранение информации, получаемой от каждого блока сбора данных для дальнейшего анализа, в том числе хранения операционных данных и взаимодействие с файловым хранилищем.

Функциональные задачи, стоящие перед блоком анализа: прием сообщений от агентов сбора данных; анализ поступивших данных; организация хранения данных, полученных от агентов сбора.

Модуль реализует саму модель анализа, алгоритмы формирования признаков из поступающих данных. В основу данного блока вошла модель распределённых вычислений MapReduce, являющаяся частью экосистемы проектов Hadoop. [7]

Работа MapReduce состоит из двух шагов: Map и Reduce, названных так по аналогии с одноименными функциями высшего порядка, map и reduce. На Map-шаге происходит предварительная обработка входных данных. Для этого один из вычислительных узлов (называемый главным узлом - master node) получает входные данные задачи, разделяет их на части и передает другим компьютерам (рабочим узлам - worker node) для предварительной обработки. На Reduce-шаге происходит свёртка предварительно обработанных данных. Главный узел получает ответы от рабочих узлов и на их основе формирует результат - решение задачи, которая изначально формулировалась. Преимущество MapReduce заключается в том, что он позволяет распределенно производить операции предварительной обработки и свертки. Операции предварительной обработки работают независимо друг от друга и могут производиться параллельно [7]. Принцип действия MapReduce проиллюстрирована на рисунке 4.

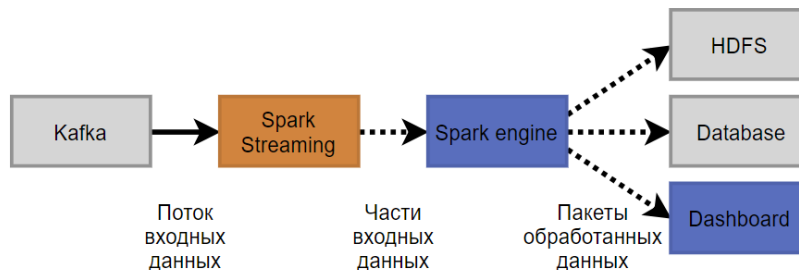


S – Исходный набор данных,  $S_i$  – Части исходного набора данных,  
 $S'_i$  - Результат операции Map,  $S'_1+S'_n$  – Результат операции Reduce,  
 $S'$  - Выходной набор данных.

**Рис. 4. Схема работы MapReduce**

Для реализации данного модуля было определено использовать Apache Spark с библиотекой Spark Streaming API. Данное совмещение позволяет использовать эффективность и простоту функционирования Hadoop, Spark поддерживает Scala, Python, Java, R и другие популярные языки разработки, и имеет большую базу открытого кода. В частности, рассматриваемые в диссертации подходы и технологии, а также методы, существуют в виде реализации на данных языках, имеют наработки программного кода в открытых репозиториях, что делает доступными для модификации. На рисунке 5 представлено, как можно применять Apache Spark и Apache Spark API с другими модулями проектируемой архитектуры

Также перечисленные языки программирования поддерживают широкий спектр статистических и численных методов и обладают хорошей расширяемостью с помощью большого количества многофункциональных пакетов. Пакеты представляют собой библиотеки для работы специфических функций или специальных областей применения.



**Рис. 5. Применение Apache Spark и Apache Spark API с Apache Kafka и другими модулями проектируемой архитектуры**

Несмотря на сходство с Hadoop, Spark представляет собой новую кластерную вычислительную среду, обладающую полезными особенностями. Например, Spark предназначен для решения в вычислительном кластере задач определенного типа, в которых рабочий набор данных многократно используется в параллельных операциях (например, как в задачах машинного обучения). Для оптимизации задач этого типа в Spark вводится понятие кластерных вычислений в памяти, когда наборы данных могут временно храниться в оперативной памяти для уменьшения времени доступа. Так, Spark бравитурет понятием устойчивого распределенного набора данных (resilient distributed datasets RDD). RDD - коллекция неизменяемых объектов, распределенных по множеству узлов. Эти коллекции устойчивы, потому что в случае потери части набора данных они могут восстанавливаться. Процесс восстановления части набора данных опирается на механизм отказоустойчивости, поддерживающий родословную (или информацию, которая позволяет восстанавливать часть набора данных с помощью процесса, в результате которого эти данные были получены). RDD представляет собой объект Scala и может создаваться из файла; в виде параллельного среза (распределенного по узлам); как преобразование другой RDD; и, наконец, путем изменения персистенции, существующей RDD, такой как запрос на кэширование в памяти. [8]

Использование таких технологий совместно с кластерными базами данных позволяет существенно уменьшить задержки доступа к данным, что критично для анализа больших объемов информации. С другой стороны, данный подход позволяет манипулировать данными без записи ее на дисковое хранилище, однако, учитывая факт ограниченности оперативной памяти по объему, стоит учесть хранение в файловой системе, поддерживающая единый доступ среди всех вычислительных кластеров.

Как и Hadoop, Spark поддерживает одноузловые и много узловые кластеры. При много узловой работе Spark опирается на менеджера кластера Mesos. Mesos обеспечивает эффективную платформу распределения ресурсов и изоляции распределенных приложений. [9]

Рассмотрим подробнее популярные языки поддерживаемые Spark. Scala - мультипарадигмальный язык, гладко и удобно поддерживает языковые функции, характерные для императивных, функциональных и объектно-ориентированных

языков. С точки зрения объектно-ориентированного программирования каждое значение представляет собой объект. Аналогично, с точки зрения функционального программирования каждая функция - значение. В том числе, Scala статически типизируется с помощью выразительной и безопасной типовой системы. Scala функционирует в различных Web частях приложений Twitter, LinkedIn и Foursquare, а также является объектом интересов для EDF Trading и др. финансовых учреждений.

Scala представляет собой язык виртуальной машины (VM) и работает непосредственно на Java Virtual Machine (JVM) с использованием Java Runtime Environment версии 2 посредством байт-кодов, генерируемых компилятором Scala. Этот подход позволяет Scala выполнять почти все, что работает на JVM.

### **Визуализация данных**

Зачастую люди плохо воспринимают данные, представленные в виде текста или таблицы, что в свою очередь усложняет понимание событий и создает сложности в понимании ценностей представленных чисел.

Модуль визуализации должен удовлетворять современным стандартам UX, максимально возможно понятно визуализировать данные, но также соответствовать техническим требованиям, таких, как кроссплатформенность, доступность и безопасность, а также перспективу расширения клиентского приложения.

Учитывая определённые требования, была определена необходимость разработки Web клиента для визуализации данных мониторинга в виде графиков и диаграмм. Архитектура клиента основана на популярном клиент-серверном концепте - backend/frontend. Таким образом, разделяя общую экосистему системы мониторинга с частью визуализации - backend'ом, что в свою очередь будет удовлетворять необходимым гарантиям безопасности.

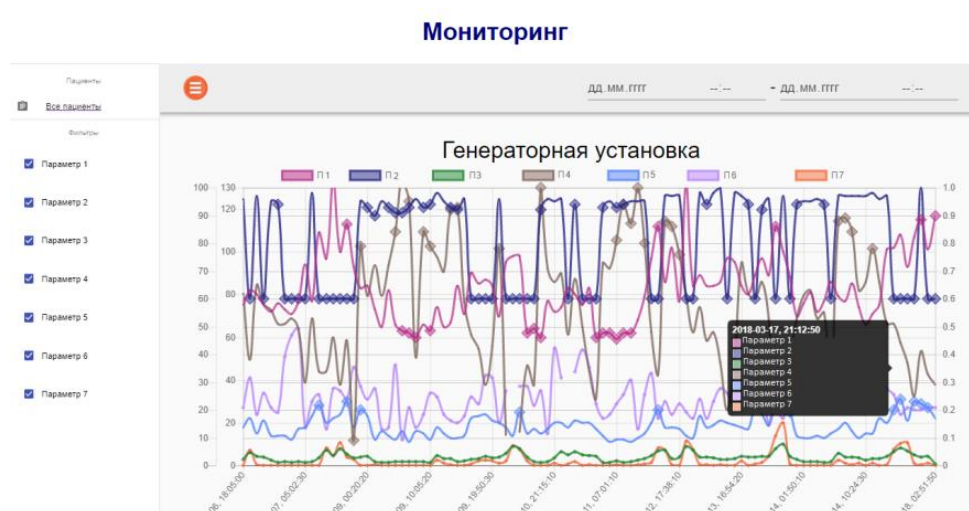
Web приложение является кроссплатформенным интерфейсом, не требующем установки, однако требовательным к подключению к сети, функционирующем через программы браузеры на почти всех популярных платформах, в том числе Android OS, iOS, Windows Phone. Backend - серверная часть, выполняющая подключение к элементам внутренней экосистемы, взаимодействующая с базой данных Cassandra, способного разграничивать права доступа пользователей, для взаимодействия с клиентами обычно применяется защищенный функционал API. [10] Frontend - клиентская часть приложения, предоставляющая интерфейс взаимодействия между человеком-пользователем системы и всей системой через backend. На рисунке 6 проиллюстрирован пример пользовательского интерфейса Web приложения.

Backend - функционирует на стороне сервера (не в браузере), т.е. специально выделенной под определённые задачи вычислительной машине и только, что в свою очередь предоставляет широкий выбор инструментом, доступных для разработки, это означает, что вы можете использовать любой универсальный язык программирования: Ruby, PHP, Python, Java, JavaScript / Node, bash. Это также означает, что вы можете использовать системы управления базами данных, такие как MySQL, PostgreSQL, MongoDB, Redis, Memcached.

Однако Backend - был определен Django, основанный на языке python, который поддерживается Kafka, Cassandra и другими модулями, позволит построить гибкую и перспективную систему. [11]

Frontend –должен быть основан на технологиях поддерживаемых браузерами, для реализации тонкого клиента в Web посредством HTML, CSS и JavaScript. HTML (HyperText Markup Language) сообщает браузеру, каково содержание страницы, например, «заголовок», «параграф», «список», «элемент списка». CSS (Cascading Style

Sheets) передает браузеру, как отображать элементы. JavaScript задает динамику элементам представления и управляет параметрами браузера при взаимодействии.



**Рис. 6. Пример пользовательского интерфейса (Front End), разработанного с применением React.js и Chart.js**

Большинство сайтов не применяют большой объем JavaScript кода, однако его применение позволяет манипулировать содержимым страницы без визуальных мерцаний экрана для пользователя и технически без пересоздания страницы в целом. В связи с чем, для Frontend был определен React.js - одна из самых популярных библиотек для создания сложных Frontend-приложений. Одна из главных особенностей React.js - свобода действий, существует огромное количество подходов к построению приложений с его помощью (redux, mobx и другие). [12]

### **Заключение**

Согласно представленному описанию, решение проблемы технологического характера обработки потоковых больших данных реализуемы в рамках интеграции современных программных инструментов. Каждый элемент системы может быть отдельным модулем, в свою очередь, автономным и свободно интегрируемым с другими системами в различные конфигурации, тем самым модифицируя различные архитектуры потоковой обработки больших данных на основе распределенных вычислений.

### **Список литературы**

1. Овчиников А. Эталонная архитектура сервиса из компонентов потоковой обработки больших данных // Сборник докладов конференции «Разработка ПО»: сайт - Центр Digital October, 28-30 октября 2016. Режим доступа: <http://goo.gl/CAF1G7> (Дата обращения: 27.02.2017).
2. Apache Kafka // Wikipedia свободная энциклопедия: сайт. Режим доступа: [https://ru.wikipedia.org/wiki/Apache\\_Kafka](https://ru.wikipedia.org/wiki/Apache_Kafka) (Дата обращения 13.10.2017).
3. RabbitMQ vs. Kafka vs. ActiveMQ // Техноблог StackShare.io. Режим доступа: <https://stackshare.io/stackups/activemq-vs-kafka-vs-rabbitmq> (Дата обращения: 13.10.2017).

4. Humphrey P. Understanding When to use RabbitMQ or Apache Kafka // Сайт. Режим доступа: <https://content.pivotal.io/blog/understanding-when-to-use-rabbitmq-or-apache-kafka> (Дата обращения 15.11.2017)
5. Apache Cassandra // Wkipedia свободная энциклопедия: сайт. Режим доступа: [https://ru.wikipedia.org/wiki/Аpache\\_Cassandra](https://ru.wikipedia.org/wiki/Аpache_Cassandra) (Дата обращения 13.10.2017).
6. Apache Cassandra Documentation // Официальный портал Apache Cassandra: сайт. Режим доступа: <http://cassandra.apache.org/doc/latest/> (Дата обращения: 19.02.2018).
7. Apache Hadoop Documentation // Официальный портал Apache Hadoop: сайт. Режим доступа: <http://hadoop.apache.org/docs/current/> (Дата обращения: 13.11.2017).
8. М.Джонс. Spark, альтернатива для быстрого анализа данных // IBM developerWorks: сайт. Режим доступа: <https://www.ibm.com/developerworks/ru/library/os-spark/#fig1> (Дата обращения: 17.12.2017).
9. Native App vs. Web App: What's the Difference? // UpWork: сайт. Режим доступа: <https://www.upwork.com/hiring/mobile/native-app-vs-web-app-for-mobile/> (Дата обращения: 19.03.2018).
10. Django Documentation // Django Project: сайт. Режим доступа: <https://www.djangoproject.com/en/2.0/> (Дата обращения: 21.01.2018).
11. React JS Documentation // React JS: сайт. Режим доступа: <https://reactjs.org/docs/try-react.html> (Дата обращения: 21.01.2018).
12. Redux JS Documentation // Redux JS: сайт. Режим доступа: <https://redux.js.org/> (Дата обращения: 21.01.2018).

## **О ПРЕДВАРИТЕЛЬНОМ РАСЧЕТЕ ПАРАМЕТРОВ НЕПОЗИЦИОННОЙ СИСТЕМЫ СЧИСЛЕНИЯ**

**Калимолдаев М.Н., Тынымбаев С.Т., Магзом М.М.**

*Институт информационных и вычислительных технологий КН МОН РК,  
Казахстан  
e-mail: [magzomxzn@gmail.com](mailto:magzomxzn@gmail.com)*

***Аннотация.** В данном обзоре приведены промежуточные результаты выполнения задачи проекта грантового финансирования КН МОН РК «Разработка программно-аппаратных средств для криптосистем на базе непозиционной системы счисления». Идея данного проекта заключается в разработке различных вариантов умножителей полиномов по модулю неприводимых полиномов и их реализация на базе программируемых логических интегральных схем (ПЛИС) с целью создания программно-аппаратной реализации криптосистем на базе непозиционной полиномиальной системы счисления, и создание программных средств для выполнения предварительных вычислений параметров непозиционной системы счисления.*

### **Введение**

Основной целью данного проекта является разработка программно-аппаратных средств, выполняющих различные вычисления в непозиционной системе счисления



для эффективной реализации алгоритмов криптографической защиты информации. Для достижения поставленной цели проекта первой задачей ставится разработка программ для выполнения предварительных вычислений для подготовки данных к зашифрованию и расшифрованию: формирование системы оснований непозиционной полиномиальной системы счисления (НПСС) исходя из размера шифруемого блока, интерпретации входного блока данных и ключевой последовательности в виде остатков.

Для начала опишем необходимые для выполнения преобразования в НПСС. После рассмотрим программную и программно-аппаратную реализацию системы предварительных вычислений непозиционной системы счисления.

### Криптографическое преобразование на базе НПСС

Сегодня разработка криптографических систем, использующих нетрадиционный алгебраический подход при создании надёжных, и в то же время, производительных шифров, представляет значительный интерес. Этим и обусловлены исследования по разработке криптографических систем на базе систем остаточных классов (СОК) и непозиционных систем счисления [1, 2].

Аналогично китайской теореме об остатках [3], в полиномиальных системах счисления в остаточных классах любой полином может быть представлено своими остатками (вычетами) от деления на систему оснований, состоящей из неприводимых многочленов над полем  $GF(2)$  [4,5].

Для формирования НПСС при шифровании блока длиной  $N$  бит из множества всех неприводимых многочленов степени не выше значения  $N$  выбираются рабочие основания

$$p_1(x), p_2(x), \dots, p_s(x). \quad (1)$$

Все выбираемые основания должны отличаться друг от друга, даже если они являются неприводимыми полиномами одной степени. Тогда в этой системе любой многочлен степени меньше суммы степеней всех рабочих оснований (1) имеет единственное представление в виде последовательности остатков (вычетов) от деления его на данные основания. Таким образом, блок открытого текста и ключевая последовательность длиной  $N$  бит могут быть представлены в виде последовательностей вычетов  $F(x) = (\alpha_1(x), \alpha_2(x), \dots, \alpha_s(x))$  и  $G(x) = (\beta_1(x), \beta_2(x), \dots, \beta_s(x))$  соответственно, полученных в результате деления по рабочим основаниям системы. Покажем пример простого преобразования открытого текста в шифр.

Шифрованное сообщение формируется в результате умножения многочленов  $F(x)$  и  $G(x)$ :

$$F(x)G(x) \equiv H(x) \pmod{P(x)}, \quad (2)$$

т.е. представлено в виде остатков от деления произведений  $\alpha_i(x)\beta_i(x)$  на соответствующие основания  $p_i(x)$

Для расшифрования необходимо вычислить обратный многочлен  $G^{-1}(x) = (\beta_1^{-1}(x), \beta_2^{-1}(x), \dots, \beta_s^{-1}(x))$ . Тогда исходное сообщение восстанавливается по сравнению:

$$F(x) \equiv G^{-1}(x)H(x) \pmod{P(x)}. \quad (3)$$

Секретный ключ, используемый для шифрования, называется полным. Полный ключ состоит из ключевой псевдослучайной последовательности и выбранной системы полиномиальных оснований, которая также держится в секрете. С ростом порядка неприводимых многочленов с двоичными коэффициентами их количество стремительно растёт, в связи с этим очевиден широкий выбор полиномиальных оснований. Количество вариантов выбора НПСС существенно возрастает с увеличением длины шифруемого блока [4].

Предварительные расчёты параметров НПСС включают в себя:

- создание базы данных неприводимых полиномов с коэффициентами над  $GF(2)$ ;
- определение необходимой длины блока и ключа шифрования;
- формирование в соответствии с длиной блока системы оснований;
- подбор и проверка содержания системы оснований;
- предварительная интерпретация входного блока данных в виде остатков от деления на выбранную систему оснований.

#### **Программная реализация системы предварительного расчёта**

*Предварительный расчёт параметров НПСС выполняется в программном пакете, на основе разработанной ранее библиотеки для работы с полиномами с коэффициентами над  $GF(2)$ . Для обеспечения доступа к средствам расчёта был разработан веб-интерфейс на основе фреймворка Spring.*

Фреймворк Spring MVC реализовывает архитектуру паттерна Model — View — Controller при помощи слабо связанных готовых компонентов. Паттерн MVC разделяет аспекты приложения (логику ввода, бизнес-логику и логику UI), обеспечивая при этом свободную связь между ними.

- Model (Модель) инкапсулирует (объединяет) данные приложения, в целом они состоят из математической-компьютерных моделей для работы в НПСС.
- View (Отображение, Вид) отвечает за отображение данных Модели, — как правило, генерируя HTML, которые мы видим в своём браузере.
- Controller (Контроллер) обрабатывает запрос пользователя, создаёт соответствующую Модель и передаёт её для отображения в Вид.

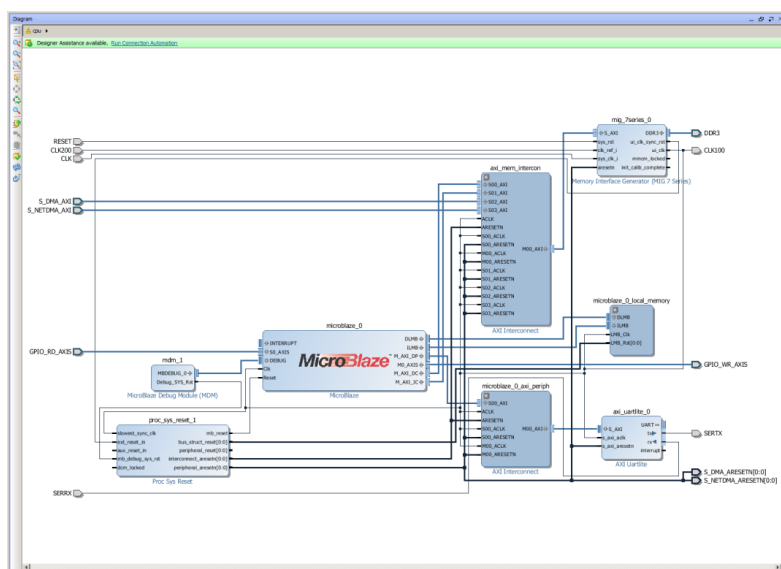
После окончательного выбора компонентов и создания была получена архитектура web-приложения с готовыми для дальнейшего наполнения директориями. Фрагменты для взаимодействия с визуальной частью, графические стили CSS, стандартные страницы HTML и функционал JavaScript располагаются в “resources”, а модели для расчёта и логика управления, соответственно, размещается в “java”.

#### **Программно-аппаратная реализация системы предварительного расчёта**

Введётся работа по разработке и реализации программно-аппаратной системы для предварительного расчёта параметров непозиционной системы счисления. В такой реализации основная трудоёмкая операция – деление полинома по модулю неприводимого полинома – выполняется аппаратно на устройстве умножения, схема которого была представлена в [6].

В качестве аппаратной платформы используется FPGA Artix-7 (XC7A35TICSG324-1L) на основе отладочной платы Artu A7 от компании Digilent [7].

В программной части данной работы используется микропроцессорное ядро Microblaze, реализуемое на основе ПЛИС компании Xilinx. MicroBlaze является составной частью пакета Embedded Development Kit (EDK), предлагаемого фирмой Xilinx в качестве основного инструмента разработки и отладки встраиваемых микропроцессорных систем на основе ПЛИС серий FPGA. Элементы семейства MicroBlaze представляют собой встраиваемые микропроцессорные ядра с RISC-архитектурой, которые предназначены для применения в системах, выполняемых соответственно на основе ПЛИС. Схема процессора и периферийных устройств показана на рисунке 1.

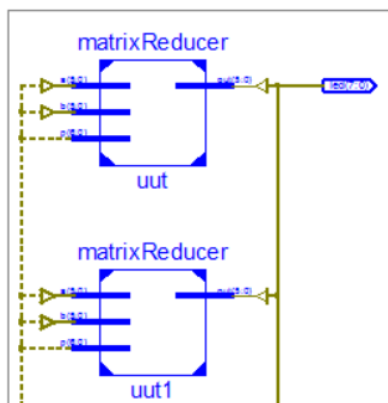


**Рис 1. Микропроцессор Microblaze в среде разработки Vivado**

Данные по полиномам для умножения подготавливаются в программном микропроцессоре Microblaze, после чего для их умножения полинома по модулю неприводимого полинома эти данные передаются в устройство умножителя.

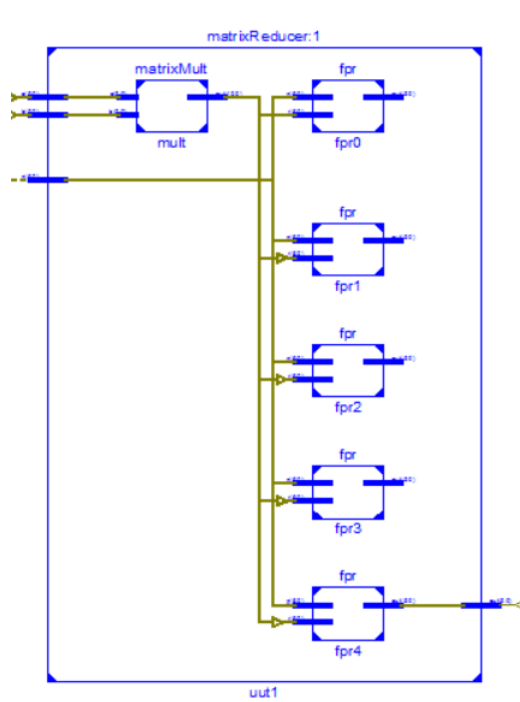
Для описания разработанного умножителя была использована среда проектирования Xilinx Vivado. Описание схемы состоит из следующих компонентов:

- Модуль верхнего уровня. В данном модуле определяются умножители matrixReducer, их входы и выходы. Пример данного модуля с двумя умножителями показан в рисунке 2.



**Рис. 2. Фрагмент модуля верхнего уровня схемы**

– Каждый модуль умножителя matrixReducer состоит из модуля конъюкторов строк матрицы и дерева сумматоров, и из модулей ФЧО. Пример 6-ти разрядного умножителя показан на рисунке 3.



**Рис. 3. Структура модуля одного умножителя по модулю неприводимых Полиномов**

- Модуль дерева сумматоров и конъюкторов строк матрицы.
- Модуль ФЧО.

### **Заключение**

Были рассмотрены промежуточные результаты выполнения задачи «Разработка программ для выполнения предварительных вычислений параметров НПС» проекта по разработке программно-аппаратных средств, выполняющих различные вычисления в непозиционной системе счисления. На данный момент получены программы для

вычисления параметров непозиционной системы счисления, с целью последующего применения данных параметров при шифровании в криптосистеме на основе НПСС. Кроме этого, рассматриваются вопросы выполнения упомянутых предварительных вычислений в программно-аппаратном виде на плате с ПЛИС и программным микропроцессором.

Данные средства и инструменты будут использованы в дальнейших исследованиях по разработке криптосистем на основе НПСС.

### **Литература**

- 1 Svoboda A. Valach M. Operatorove obvody // Stroje Na Zpracovani Informaci – 1955. – Vol 3. – P. 247-295.
- 2 Акушский И.Я., Юдицкий Д.И. Машинная арифметика в остаточных классах. – М.: Советское радио, 1968.
- 3 Калимолдаев М.Н., Бияшев Р.Г., Нысанбаева С.Е., Бегимбаева Е.Е., Магзом М.М. История развития системы остаточных классов в компьютерной технике // Вестник КазНУ. - Алматы, 2017. – №2. – С. 416-419.
- 4 Бияшев Р.Г., Нысанбаева С.Е. Алгоритм формирования электронной цифровой подписи с возможностью обнаружения и исправления ошибки // Кибернетика и системный анализ. – 2012. – Т. 48, № 4. – С. 14-23.
- 5 Амербаев В.М., Бияшев, Р.Г., Нысанбаева С.Е. Применение непозиционных систем счисления при криптографической защите // Известия Национальной академии наук Республики Казахстан. Серия физико-математическая. – Алматы: Гылым, 2005. – № 3. – С. 84-89.
- 6 Тынымбаев С.Т., Капалова Н.А., Магзом М.М. Разработка и реализация аппаратного умножителя в непозиционной системе счисления // Материалы научной конференции ИИВТ МОН РК «Современные проблемы информатики и вычислительных технологий». – Алматы, 2017. – С. 263-270.
- 7 Arty A7: Artix-7 FPGA Development Board for Makers and Hobbyists . – <https://store.digilentinc.com/artix-a7-artix-7-fpga-development-board-for-makers-and-hobbyists/> (май 2018)

## **АНАЛИЗ МЕТОДОВ АНГЛО-КАЗАХСКОГО ВЫРАВНИВАНИЯ СЛОВ**

**Картбаев А.Ж.**

*Казахский национальный университет имени аль-Фараби  
a.kartbayev@gmail.com*

***Аннотация.** Эта работа посвящена проблеме о наиболее широко используемых подходах к выравниванию слов для машинного перевода, где мы хотим вкратце описать, как системы машинного перевода эффективно используют выравнивание слов. Системы машинного перевода, как правило, подразделяются на два типа моделей: модель перевода, которая является моделью перевода между исходными и целевыми языками, а также языковой моделью, которая является моделью сформированных предложений на целевом языке. Чтобы перевести новое исходное*

*предложение, мы рассчитываем на максимальную вероятность этих двух моделей, а также мы будем искать целевую строку, которая является одновременно вероятным сопоставлением исходной и целевой строки, а также является правильным целевым предложением. Модель перевода оценивается с использованием выравнивания слов параллельного корпуса. Качество языковой модели оценивается на языке оригинала.*

В этой работе мы хотели бы обосновать утверждение о том, что улучшение согласованности выравнивания слов приведет к улучшению систем машинного перевода. Предположим, что существуют выравнивания фиксированного параллельного текста, которые значительно лучше, чем выравнивания, создаваемые нашей базовой системой. Мы разработаем усовершенствованные выравнивания с помощью системы, которая сообщает нашей системе выравнивания о том, как улучшить работу, и она знает, насколько это эффективно. Мы измерим производительность статистического машинного перевода при использовании новой системы выравнивания и сравниваем результат с использованием базового варианта. Мы делаем это с помощью BLEU метрики, как показали исследования[1], она хорошо соотносится с человеческими суждениями качества. Улучшенное выравнивание из базовой системы обычно приводит к высокому BLEU баллу, чем доказывает, что улучшение выравнивания может повысить производительность машинного перевода. Определение предсказания вероятностей для этой задачи является ее основной проблемой. На данный момент наша система не находится у верхней границы производительности. С учетом имеющихся вычислительных ресурсов мы смогли найти лучшую систему, которая позволила бы повысить производительность. Далее в работе мы покажем, как получить улучшенные выравнивания слов с этой системой.

Существует множество результатов, в которых система выравнивания успешно повысила оценку в соответствии со встроенными метриками качества выравнивания, многие из этих подходов приводят к значительному увеличению производительности перевода. Мы покажем, что это объясняется тем, что две встроенных метрики качества выравнивания, которые обычно используются, хорошо измеряют полезность выравнивания для перевода. Эти метрики являются хорошо известными сбалансированными F-мерами[2] и показателем ошибок выравнивания (AER). Мы рассчитали корреляцию между этими метриками и BLEU метрикой и показали, что эта корреляция хорошая. Кратким математическим описанием корреляции является коэффициент определения, который является квадратом коэффициента корреляции Пирсона. Для задачи выравнивания с использованием широко изученного набора данных, показатель ошибок выравнивания показывает необходимую корреляцию с BLEU. Для одной и той же задачи сбалансированная мера F-мера также показывает корреляцию с BLEU.

Вероятностный подход к решению подобных задач исходит из предположения, что система рассматриваемых величин обладает определенным совместным распределением вероятностей. Основные свойства коэффициента корреляции:

- 1)  $\rho(x,y) \leq 1$ ;
- 2) если  $X, Y$  независимы, то  $\rho(X,Y)=0$ ;
- 3) если  $X, Y$  связаны между собой линейной зависимостью, т.е.  $Y=aX+b$ , то  $|\rho(X,Y)|=1$ . При этом чем ближе  $|\rho|$  к 1, тем лучше линейная зависимость между  $X$  и  $Y$ .

Коэффициент корреляции Пирсона применяется в случае, если изучаемые случайные величины предположительно распределены по нормальному закону. Он

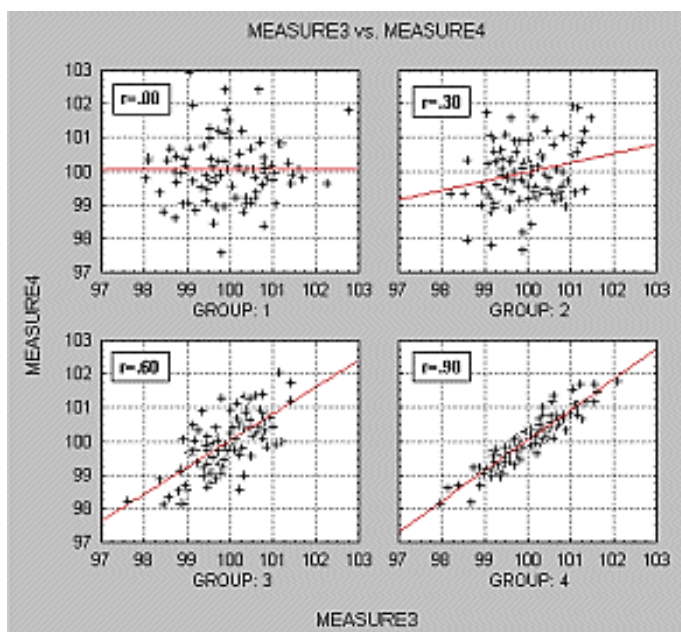
обозначается  $\rho(X, Y)$  - для двух случайных величин  $X$  и  $Y$ , - и рассчитывается с помощью соотношения:

$$\rho(X, Y) = \frac{M((X - M(X)) \cdot (Y - M(Y)))}{\sigma(X)\sigma(Y)}. \quad (1)$$

Здесь  $M$  и  $\sigma$  обозначают математическое ожидание и среднеквадратическое отклонение случайной величины. Для результата  $n$  опытов коэффициент корреляции рассчитывается по формуле:

$$r = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2} \sqrt{\sum (Y_i - \bar{Y})^2}} \quad (2)$$

Наша цель состоит в том, чтобы автоматически производить выравнивания, которые смогут использоваться для создания высококачественных систем машинного перевода.



**Рис. 1** Диаграммы корреляции между разными группами выравнивания

Выравнивания с аннотациями для человека часто содержат выравнивания M-N, где несколько исходных слов выровнены по нескольким целевым словам, а результирующая единица может быть разложена еще больше. Исходные или целевые слова в одной единице иногда не являются последовательными. Одним из важных недостатков многих систем является неспособность моделей правильно моделировать непрерывные выравнивания “многие ко многим”. Мы исправляем эту проблему. Наша новая генеративная модель способна моделировать выравнивания, состоящие из “многих ко многим” непоследовательных минимальных корреспондентий без использования эвристических правил.

Мы обсуждаем различные предположения в отношении структуры выравниваний, которые были сделаны в ходе предыдущей работы. Использование

допущения от 1 до N широко распространено, вероятно, благодаря успеху моделей выравнивания слов IBM[3]. Если у нас есть доступ к еще более аннотированным данным, мы можем получить дополнительные параметры непосредственно, но это всего лишь часть всех параметров, которые нам нужно найти по крупным корпусам.

Мы разрабатываем новую модель, которая обучается под частичным контролем, эта модель использует богатые возможности, полученные из нашей новой генеративной модели, но также позволяет легко интегрировать новые источники знаний, которые могли бы стать проблемой для добавления в генеративную историю. Это приводит к значительному увеличению точности выравнивания и точности перевода относительно исходной отметки. Одна из проблем с гипотетическим выравниванием заключается в том, что если мы смогли бы легко внедрить источник знаний, который использовал бы строку сходства в модели выравнивания, мы могли бы легко решить эту задачу.

### **Литература**

1. Brown P., Della Pietra V., Mercer R., Della Pietra S., and Lai J. An estimate of an upper bound for the entropy of English // Computational Linguistics. – 1992. – vol. 18, no. 3. – p.31-40.
2. Melamed D. Models of translational equivalence among words // Computational Linguistics. – 2000. – vol. 26, no. 2. – pp. 221–249.
3. Brown P.F., Pietra V.J.D., Pietra S.A.D., Mercer R.L. The mathematics of statistical machine translation: parameter estimation // Computational Linguistics – vol.19, no.2. – 1993. – pp. 263-311.

## **ОНТОЛОГИЧЕСКИЙ ПОДХОД К СЕМАНТИЧЕСКОМУ МОДЕЛИРОВАНИЮ ОБРАЗОВАТЕЛЬНЫХ ПРОГРАММ В ВЫСШЕМ ОБРАЗОВАНИИ**

**Кубеков Б.С., Утегенова А.У., Науменко В.В., Аленова Р.А.**

*Институт информационных и вычислительных технологий КН МОН РК  
e-mail: b.kubekov@mail.ru*

***Аннотация.** С развитием технологий в области IT, появились инновационные возможности работы с веб-технологиями в электронном обучении. Наибольшую популяризацию в электронном обучении получил онтологический инжиниринг. Наши исследования сосредоточились на обучении с использованием концепций онтологического подхода по представлению многоуровневых структур знаний: учебные программы, учебный план, учебный предмет и учебный контент. В статье освещены основные этапы разработки методики проектирования учебных планов на основе метода проектного обучения и концепций всемирной инициативы CDIO, с использованием Smart-технологий. Исследование ориентировано на индивидуализацию траекторий обучения с точки зрения предпочтений и индивидуальных возможностей обучающихся, подготовку обучающихся к успешной профессиональной деятельности в условиях цифрового общества и умной экономики.*



## **Введение**

Образовательные технологии базируются на современных достижениях информационных технологий, на привлечении инженерии знаний и, в частности, онтологического инжиниринга. Несмотря на гибкие и разнообразные методы и средства, потребуются существенные усилия для создания методик и методов представления и организации знаний, а также для проектирования образовательных ресурсов, которые позволяли бы в полной мере использовать множество педагогических, образовательных и информационных парадигм в рамках единой образовательной парадигмы.

Известная онтологическая триада Аристотеля «концепт-атрибут-отношение» стала основой всех современных достижений онтологов, начиная от языка OWL и онтологического редактора Protégé до мультиагентных технологий и прикладных информационных систем.

Настоящий проект актуализирован процессами научных исследований в отношении обновления, накопления, передачи и повторного использования знаний. Анализ международных научных журналов «Прикладные онтологии», «Онтология проектирования», и трудов ассоциаций, лабораторий, компаний, форумов и других объединений, свидетельствуют о росте интереса к онтологии, как в научном, так и в прикладном плане.

Актуальность исследования заключается в:

- поддержке процедур правительства Республики Казахстан по реализации Государственной программы по развитию науки и образования на 2016- ., и программы по цифролизации отраслей экономики;
- предоставлении новой парадигмы информационной базы обучения, отвечающей потребностям цифровой экономики, с акцентом на развитие навыков в анализе информации и креативности мышления;
- диверсификации методик электронной педагогики и ассортимента образовательных услуг за счет освоения новых образовательных, в том числе Smart-технологий, с целью повышения эффективности и качества обучения;
- в модернизации учебных программ и планов подготовки инженерных кадров, с целью развития эпистемологической функции компетентностного подхода;
- в необходимости применения методов инженерии знаний и онтологического инжиниринга, а также внедрения информационных технологий и педагогических средств интенсификации учебно-познавательной деятельности обучающихся.

Исследуемое направление отражает исследования структуры формального представления данных и разработке синтаксиса языка спецификации знаний для глубокой модернизации образовательных программ в области техники и технологий, в которых объединены четыре подхода: всемирная инициатива CDIO, метод форсайтных исследований, проектный метод обучения и онтологический инжиниринг.

Цель статьи заключается в разработке интеллектуального сервиса и модели представления и организации знаний с использованием компетентностного и онтологического подходов; разработке методики проектирования дисциплин и учебных планов многоуровневой системы, основанных на проектном методе обучения и концепциях Всемирной инициативы CDIO.

## **1 Проектный метод обучения**

Современная IT индустрия предъявляет высокие требования к специалистам, которых готовят высшие учебные заведения. Образование должно стать ресурсом

подготовки специалистов, способных воспринимать любые новые веяния в науке, технологии и бизнесе[1].

Исследовательская работа – это в основном кропотливая индивидуальная работа, которая подразумевает получение собственного теоретического или экспериментального материала, на основании которого необходимо провести анализ и сделать выводы об исследуемой области. Принципиальной особенностью таких работ является неопределенность получаемого результата в результате проведенных исследований, по отношению к планируемому[1].

Тогда как выполнение проекта – это всегда коллективная, творческая практическая работа, предназначенная для получения определенного продукта или научно-технического результата. Такая работа подразумевает четкое, однозначное формирование поставленной задачи, определение сроков выполнения намеченного, определение требований к разрабатываемому объекту[1].

Исходя из отмеченного, можно сделать выводы, что студентам будет несколько тяжело выполнять исследовательскую работу, так как для нее необходим опыт исследователя, а выполнение проекта будет подстраиваться под те компетенции, которыми владеет учащийся на каждом этапе проекта. Участники проекта могут перераспределять роли и мотивировать остальных участников на приобретение новых компетенций.

Новой формой мотивации студентов к обучению является привлечение их к работам над реальными задачами, путем организации соответствующих тематических проектов, формирования проектных групп для их выполнения и введения в учебный план отдельной дисциплины «Групповое проектное обучение». Под основной задачей проектного обучения будем понимать усиление когнитивной способности студентов, которая будет оцениваться качеством решения задач проекта, степенью его реализации[1].

Основные требования, предъявляемые к проектному обучению:

1. наличие значимой в теоретическом, исследовательском, техническом плане задачи;
2. исследование существующих решений, для достижений оригинальности и технической новизны проекта;
3. командный подход для выполнения проекта;
4. наличие результатов выполнения проекта в виде, разработанного решения, публикаций, активное участие в мероприятиях НИРС университета, стартапах;
5. привлечение инвестиций для последующей коммерциализации.

Представленные требования помогут организовать у учащихся специальные компетенции в соответствии с постоянно возрастающими требованиями рынка IT отрасли.

Применение проектного метода обучения раскроет приобретение навыков как в программировании и разработке, так и в дальнейшей реализации продаж и поддержке программного продукта.

Внедрение проектного обучения в образовательные учреждения, планируемые в перспективе, будут служить площадкой для организации бизнеса с нуля. Данная методика неоднократно зарекомендовала себя как наиболее гибкий инструмент для формирования быстрого старта обучающихся, так как предполагается возможность полного или частичного участия образовательных структур в содействии и консультации по вопросам управления знаниями и по созданию инновационных решений. Предполагается развитие программных продуктов open source, и только после этого продажа обновлений и дополнительных функций, т.е. создание

минимальных рисков для инвесторов и заинтересованных сторон. Будущие предприниматели могут самостоятельно тестировать рынок IT решений в различных отраслях, тем самым делать попытки по улучшению микроклимата страны.

Промежуточные опросы студентов, преподавателей и инвесторов, выявили основные проблемы:

1. Необходимость структурирования знаний и переход к новой образовательной парадигме представления и организации знаний с привлечением онтологического инжиниринга, в учебных заведениях.

2. Исследование методик повышения эффективности педагогических и информационных технологий, по развитию эпистемологической функции компетентностного подхода.

3. Исследование эмпирических моделей формализации и отображения знаний образовательной области, моделей спецификации знаний на основе онтологического подхода, технологий и инструментальных средств Smart-образования.

4. Исследование методики применения онтологического подхода для проектирования онтологий учебного контента дисциплин и знаниевых компонент планируемого обучения образовательной среды.

5. Организация и представление моделей знаний и формирование знаниевых компонент планируемого обучения CDIO Syllabus,.

6. Разработка и обоснование проектных и инфраструктурных решений образовательной среды.

## **2 Концепция CDIO и проектно-ориентированная образовательная среда**

Всемирная инициатива CDIO - это проектно-ориентированная технология обучения, которая ориентирована на студента и интегрирована с проблемами и опытом реального производства[2].

Проектный подход и этапы всемирной инициативы CDIO следует рассматривать как контекст образовательной среды, который будет способствовать приобретению знаний, умений и практических навыков профессиональной инженерной деятельности обучающегося. В рамках такой образовательной среды должны происходить формирование комплексного подхода по определению образовательных потребностей студента и разработки последовательности учебных мероприятий, направленных на их удовлетворение.

Таким образом, ориентированная на проектный подход образовательная среда, определенная в рамках этапов *Conceiving, Designing, Implementing* и *Operating* инициативы CDIO, направлена на усвоение и глубокое понимание теоретических основ инжиниринга, применение которых будет способствовать приобретению знаний, практических навыков и умений, или, другими словами, на приобретение профессиональных компетентностей в будущей деятельности выпускника[2].

Для решения данной задачи большое значение должно отводиться применению современных педагогических подходов и инновационных методик представления и организации знаний, с целью:

- повышения когнитивной способности студентов;
- успешного усвоения необходимых знаний для решения практических учебно-производственных задач через призму этапов инициативы CDIO;
- усвоения обучающимися проектного опыта и активного применения полученных знаний в учебной и профессиональной деятельности.

Чтобы гарантировать достижение поставленных целей необходимо рассматривать инициативу CDIO как технологический процесс, в котором технология

определения образовательных потребностей учитывала бы требуемый уровень компетентностей при выполнении практических учебно-производственных задач, и имела возможность моделирования и конструирования образовательных (знаниевых) компонент планируемого обучения для последующего проектирования дисциплин учебного плана[3].

Сформулированные таким образом требуемые результаты обучения служат основой для определения целей обучения, моделирования и последующего проектирования образовательных программ, процессов обучения и оценивания достижений студентов.

### **3 Компетентностные модели и этапы CDIO**

Совершенствование качества подготовки специалистов должно основываться на компетентностном подходе, рассматриваемом с точки зрения профессиональных компетенций, из которых выделим базовые, как основа профессии и дополнительные, принадлежащие другим типам деятельности, но значимые для определенного рабочего места или профиля выпускника[3].

Для каждого из этапов инициативы CDIO определим уровни профессиональных компетенций, в рамках каждого из которых определяются базовые и дополнительные компетенции. При этом:

- профессиональные компетенции, для любого этапа инициативы CDIO, являются обязательными по определению;
- базовые и дополнительные компетенции, для любого этапа инициативы CDIO, могут быть обязательными, необязательными и альтернативными компетенциями;
- профессиональные, базовые и дополнительные компетенции описываются с помощью онтологий опорных понятий и соответствующими выражениями знания, где опорные понятия представляют собой семантические знания предметной области;
- идентификация опорных понятий профессиональных, базовых и дополнительных компетенций происходит с помощью конкретизирующих понятий.

Реализация компетентностных моделей и этапов CDIO не ограничивается участием в образовательном процессе только выпускающих кафедр, а предполагает сотрудничество различных кафедр и факультетов, бизнес-инкубатора, предприятий, заинтересованных в выполнении разрабатываемых проектов и профессиональной подготовке будущих выпускников. Это поможет более глубокому погружению в профессиональные компетенции[4].

Обязательная работа и формирование группового стиля мышления и исполнительности на всех этапах выполнения работ по проекту, поможет формированию общего представления, начиная от идеи и до реализации готового программного продукта. При этом распределяются роли, формируются индивидуальные задачи, назначаются ответственные исполнители из числа участников проекта.

Проектный подход предполагает выполнение всевозможных типов проектов (научных, учебных, социально-экономических), целью которых, как правило, является реализация идеи в бизнесе. Результаты выполнения проекта имеют реальную перспективу дальнейшего продвижения и коммерциализации через бизнес-инкубатор, технико-внедренческую зону или самостоятельно, с целью выхода на реальный рынок и являются начальным этапом в цепочке от идеи до создания малых предприятий.

#### **4 Онтологический инжиниринг как инструмент формирования компетентностной модели**

Онтология определяет концептуализацию, которая лежит в основе формализма представления семантических знаний. Моделирование характеристик является основной методикой выявления и фиксации общности и изменчивости на понятиях в онтологии и на свойствах характеристик, что дает возможность разрабатывать повторно используемые образовательные компоненты и использовать их как при проектировании дисциплин учебного плана специальности, так и для адаптации образовательных программ к требуемым профессиональным компетенциям выпускника[2].

Образовательные системы должны быть гибкими и адаптивными за счет анализа и реализации абстракций общности и изменчивости образовательных средств и ресурсов, что накладывает определенную обязанность по подбору таких абстракций, которые со временем будут оставаться неизменными и адаптируемыми к новым условиям и требованиям. Такая стабильность наиболее характерна для абстракций общеобразовательных дисциплин учебного плана, частично для базовых дисциплин, и в большей степени изменчивость характерна для абстракций цикла профилирующих дисциплин специальности. Поэтому при планировании учебного процесса следует основываться, прежде всего, на профилирующих дисциплинах и на знании о том, каким изменениям будут подвергаться требования к профессиональным компетенциям выпускника в будущем.

Применение онтологии и характеристических моделей для обучения требует: во-первых, проведение анализа структуры и организации контента учебного материала дисциплин; во-вторых, создание визуально-мыслительных образов, с помощью ассоциативного связывания понятий и структурных элементов содержания учебного материала для формирования целостной системы знаний в рамках отдельной дисциплины, курса и специальности; и в-третьих, включение механизмов взаимодействия визуально-мыслительных образов учебного материала, способствующих приобретению обучающимися профессиональных навыков и практических умений[2].

Процесс построения онтологии начинается с определения базовых абстракций предметной области, так называемых, опорных понятий, составляющих семантические знания предметной области. Каждое опорное понятие, в свою очередь, может быть идентифицировано набором конкретизирующих понятий. Каждое такое представление, таким образом, есть онтология опорного понятия, представляемая в виде понятийного графа.

Если онтология используется для представления и организации знаний предметной области, то с помощью характеристической модели можно описывать свойства общности и изменчивости понятий онтологии.

Моделирование характеристик - самая значимая заслуга инженерии знаний. Проектирование систем обучения невозможно без образовательных компонент повторного использования, что, в свою очередь, невозможно без анализа и представления общности и изменчивости понятий дисциплин специальности и соответствующих образовательных средств и учебных ресурсов. Образовательные средства и учебные ресурсы многократного применения должны отличаться от традиционных учебных ресурсов значительно более существенной изменчивостью, а основной методикой выявления и фиксации изменчивости является именно моделирование характеристик. Характеристики совершенно незаменимы при формулировании кратких описаний учебных ресурсов, между экземплярами которых

существуют значительные различия, а также в дифференциации экземпляров этих ресурсов.

Таким образом, характеристические модели дают возможность применения дополнительных формализмов для моделирования и представления семантического содержания образовательных средств и учебных ресурсов.

Если онтология используется для представления и организации знаний предметной области, то с помощью характеристической модели описываются свойства общности и изменчивости понятий онтологии.

Процесс построения онтологии начинается с определения базовых абстракций предметной области - опорных понятий, составляющих семантические знания предметной области, каждое из которых идентифицируется набором конкретизирующих понятий. Визуально онтология опорного понятия представляется понятийным графом.

Основное назначение онтологии заключается в определении формальной семантики знания в сочетании с удобными формами хранения и представления для восприятия человеком[1].

Потребность в разработке онтологии возникает в связи с необходимостью:

1. совместно использовать общее понимание структуры информации;
2. повторно использовать знания в предметной области;
3. отделять знания в предметной области от оперативных знаний;
4. проводить анализ знаний в предметной области.

По мере развития базы знаний содержания учебных дисциплин следует разработать общий подход по онтологическому анализу учебного контента и автоматизации процессов представления, организации и визуализации знаний на базе повторно используемых образовательных компонентов и, как следствие, созданию интегрированных учебных курсов, совместных программ обучения в рамках реализации академической мобильности и возможности гибкого планирования учебного процесса[5].

### **Заключение**

Таким образом, технология проектного обучения, как новая парадигма профессиональной подготовки в вузе, будет способствовать повышению уровня обучения будущих специалистов, наработке профессиональных компетенций, связанных не только с получением аналитических знаний и умений, но и с приобретением опыта в области предпринимательства и бизнеса.

Получаемые студентами в ходе группового проектного обучения навыки позволяют им правильно выстраивать приоритеты не только в плане реализации проекта, но и проводить форсайтные исследования по дальнейшей перспективе развития рассматриваемой области.

Практически все бизнес идеи подразумевают коллективное участие разработчиков, а существующая система организации учебного процесса не всегда способствует в решении этого вопроса, что сказывается на успешности их адаптации и трудоустройстве в ИТ-рынке.

Таким образом, для реализации целей исследования были сформулированы следующие задачи.

1) Исследованы эмпирические модели формализации знаний предметной области, моделей спецификации знаний на основе онтологического подхода, технологий и инструментальных средств.

2) Исследованы модели формальных систем и процедур вывода знаний образовательной парадигмы, развиваемой в рамках компетентностной модели и метода проектного обучения Всемирной инициативы CDIO.

3) Проведен анализ и адаптация моделей и знаниевых компонент информационно-образовательной среды.

4) Проведен анализ и обоснование проектных решений по аппаратно-программному обеспечению информационно-образовательной среды.

Данные исследования выполняются в рамках грантового проекта: "Исследование и разработка моделей и методики представления и организации знаний с применением онтологического подхода и инструментальных средств Smart-технологии, при реализации образовательных программ и процессов", руководитель Кубеков Б.С.

#### **Литература**

1. Боков Л.А., Катаев М.Ю., Поздеева А.Ф. Технология группового проектного обучения в вузе как составляющая методики подготовки инновационно-активных специалистов // Современные проблемы науки и образования. – 2013. – № 6.

2. Кубеков Б.С., Утегенова А.У., Зыкин С.Л. Концепция парадигмы в задаче представления и организации знаний. Сборник трудов Международной научно-практической конференции "Информатизация образования: тенденции, перспективы, инновации" (ИТО-КФО 2015), 27-30 апреля 2015г. Алушта, Республика Крым, 2015, с.25-30. ISBN 978-5-905813-05-4

3. Кубеков Б.С., Утегенова А.У., Антонов А.В. Онтологический подход в планируемом обучении. Пятнадцатая открытая Всероссийская конференция "Преподавание информационных технологий в Российской Федерации", 11-13.05.2017г., г.Архангельск, Северный (Арктический) Федеральный университет. с. 103-105

4. Bulat Kubekov, Beyer Ditur, Anar Utegenova Natalia Zhaksybaeva, Innovative paradigm of education of knowledge –competency form based on ontology, Journal of Theoretical and Applied Information Technology, E-ISSN 1817-3195, ISSN 1992-8645, Vol 95. No 21-2017.

5. Антонов А.В., Кубеков Б.С., Утегенова А.У., Проектный подход и планируемое обучение на основе онтологии. Двенадцатая Международная Азиатская школа-семинар «Проблемы оптимизации сложных систем». г. Новосибирск Российской Федерации, 12-16 декабря 2016 года.

## **К ВОПРОСУ О КЛАССИФИКАЦИИ ТИПОВ СВИДЕТЕЛЬСТВ В БАЙЕСОВСКИХ СЕТЯХ**

**Литвиненко Н.Г., Литвиненко А.Г., Шаяхметова А.С.,  
Мамырбаев О.Ж.**

*Институт информационных и вычислительных технологий КН МОН РК  
e-mail: n.litvinenko@inbox.ru, lag@list.ru, a\_sh\_s83\_83@mail.ru,  
morkenj@mail.ru*

***Аннотация.** Статья посвящена вопросам использования свидетельств при исследовании различных проблем с использованием аппарата байесовских сетей.*

*Первоначально в моделях с байесовскими сетями использовались лишь «жесткие» свидетельства. Почти сразу стало понятно, что «жестких» свидетельств недостаточно для построения байесовских моделей, достаточно полно и адекватно отражающих исследуемые проблемы. Возникло понятие «мягких» свидетельств. В настоящее время, например, в разработках алгоритмов для пакета **BayesiaLab** используется четыре типа свидетельств. Вопросам классификации свидетельств, адекватного понимания роли свидетельств в различных моделях с использованием байесовских сетей, как это видят авторы, посвящена данная статья. В статье, для построения примеров, использовались учебные версии пакетов **BayesiaLab** и **AgenaRisk**. Для первоначального понимания основ байесовских сетей читатель может воспользоваться [1], [2], [3], [4]. Для понимания работы пакета **BayesiaLab** отсылаем к технической документации [5]. Чтобы понять, как работает **AgenaRisk** можно посмотреть техническую документацию [6], [7].*

**Введение.** Искусственный интеллект прочно вошел в жизнь исследователей проблем в различных областях науки, экономики, общественной жизни и производства. Наиболее популярным направлением использования искусственного интеллекта с начала 2000 годов стало использование в исследованиях аппарата байесовских сетей. Был разработан хороший математический аппарат по байесовским сетям. Открылись широкие возможности для исследования проблем в различных областях науки. Однако наличие больших объемов статистических данных по исследуемой проблеме, которые требовалось обрабатывать в русле требований аппарата байесовских сетей, еще сильнее усложнили задачу рядовым исследователям. Требовался хороший программный продукт, который избавил бы рядовых исследователей от объемных рутинных вычислений, от необходимости вручную обрабатывать большие объемы данных. Появилось много интересных пакетов для решения задач, связанных с байесовскими сетями. Наиболее популярные на начало 2018 года пакеты: **BayesiaLab**, **AgenaRisk**, **Bayes Server**, **Netica**, **Hugin Expert**, **BayesFusion** и др.

В данной статье мы пытаемся озвучить некоторые разногласия в использовании свидетельств для байесовских сетей. Авторы статьи, изучив наиболее востребованные на рынке программные пакеты для работы с байесовскими сетями, пытались понять тенденции развития понятия «мягких» свидетельств и найти перспективные пути разработки алгоритмов и программ для расчетов с байесовскими сетями.

**Основные понятия.** Кратко напомним основные понятия, связанные с байесовскими сетями.

Байесовской сетью называется ациклический ориентированный граф, каждой вершине которого соответствует случайная переменная, а дуги графа отражают отношения условной независимости между этими переменными. Вершины графа часто называют узлами [8], [9], [10], [11].

Если информация в каком-либо узле становится известной, говорят, что вершина получила свидетельство. В настоящее время свидетельства в основном делятся на два класса: «жесткие» и «мягкие». Под «жесткими» свидетельствами понимают безусловную информацию, которую принимает переменная в рассматриваемом узле. Если информация для переменной не до конца определена, носит некоторый вероятностный характер, то говорят, что узел получил «мягкое» свидетельство.

Примеры свидетельств.



1. Среди населения исследуемого города по статистике 10% больных бронхитом. Это переменная, присвоенная некоторому узлу байесовской сети.

2. После медицинского обследования конкретного человека выяснилось, что он болен бронхитом. Это «жесткое» свидетельство, присвоенное рассматриваемому узлу.

3. В результате медицинского обследования конкретного человека не удалось точно выяснить, болен или нет бронхитом данный человек. Можно лишь утверждать, что этот человек с вероятностью 70% болен бронхитом. Это «мягкое» свидетельство, присвоенное рассматриваемому узлу.

Ясно, что во всех трех случаях алгоритм расчета байесовской сети, т.е. расчет состояния других узлов должен быть разным. Интуитивно понятно, что наибольшие затруднения будут при расчете байесовской сети в случае наличия в ней «мягких» свидетельств. В частности, это связано с причинами, по которым данное «мягкое» свидетельство не может быть «жестким».

### **Пример байесовской сети**

Рассмотрим распространенный пример байесовской сети, часто используемый во многих статьях. На этом примере рассмотрим некоторые проблемы, связанные с использованием свидетельств в байесовских моделях. Настройки байесовской сети взяты произвольно, к действительности не имеют никакого отношения.

На рисунке 1 изображена байесовская сеть, отражающая причинно-следственные связи между заболеванием раком и туберкулезом в зависимости от некоторых сопутствующих факторов.

Байесовская сеть уже настроена. Априорно считаем, что:

1. 10% жителей посещало Азию, 90% не посещало Азию.
2. Среди жителей 25% - юноши, 40% - среднего возраста, 35% - преклонного возраста.
3. Статистика курящих показывает, что:
  - a. Среди молодежи курящих - 70%, некурящих – 30%.
  - b. Среди людей среднего возраста курящих - 50%, некурящих – 50%.
  - c. Среди людей старшего возраста курящих - 25%, некурящих – 75%.
4. Статистика показывает, что среди курящих болен бронхитом в среднем – 60%, не болен – 40%, среди некурящих болен бронхитом в среднем – 30%, не болен – 70%.
5. Статистика показывает:
  - a. Молодые и курящие болеют раком – 3%, не болеют – 97%.
  - b. Молодые и некурящие болеют раком – 1%, не болеют – 99%.
  - c. Люди среднего возраста и курящие болеют раком – 5%, не болеют – 95%.
  - d. Люди среднего возраста и некурящие болеют раком – 2%, не болеют – 98%.
  - e. Люди старшего возраста и курящие болеют раком – 20%, не болеют – 80%.
  - f. Люди старшего возраста и некурящие болеют раком – 10%, не болеют – 90%.

Исходя из приведенных первичных настроек байесовской сети и проведенных расчетов для случайного респондента, о котором ничего не известно, видим следующее:

- Вероятность того, что респондент курит равна 53.75% (не курит = 46.25%).
- Вероятность того, что респондент не болен бронхитом равна 56.12% (болеет = 43.88%).
- Вероятность того, что респондент не болен туберкулезом равна 98.10% (болеет = 1.9%).

- Вероятность того, что респондент не болен раком равна 93.63% (болеет = 6.37%).
- Вероятность того, что респондент не болен туберкулезом или раком равна 91.85% (болеет = 8.15%).

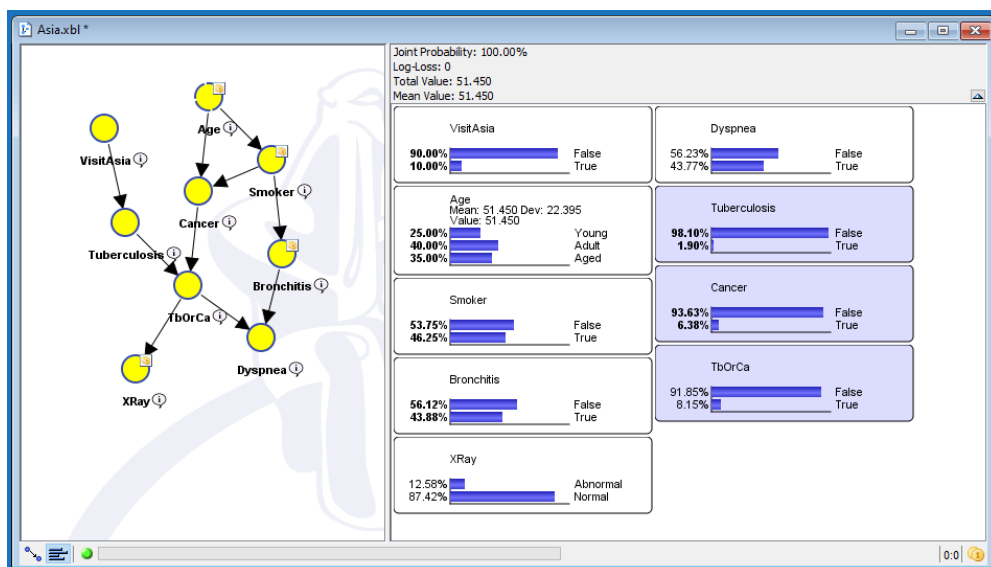


Рис. 1 Байесовская сеть «Азия»

### Жесткие свидетельства

Предположим, что требуется определить вероятность заболевания раком или туберкулезом конкретного человека, о котором есть некоторая информация. Воспользуемся при этом построенной выше моделью.

Допустим об исследуемом человеке известно следующее - этот человек среднего возраста; этот человек курит; человек болеет бронхитом. Таким образом мы имеем три жестких свидетельства.

На рисунке 2 показана байесовская сеть с указанными свидетельствами.

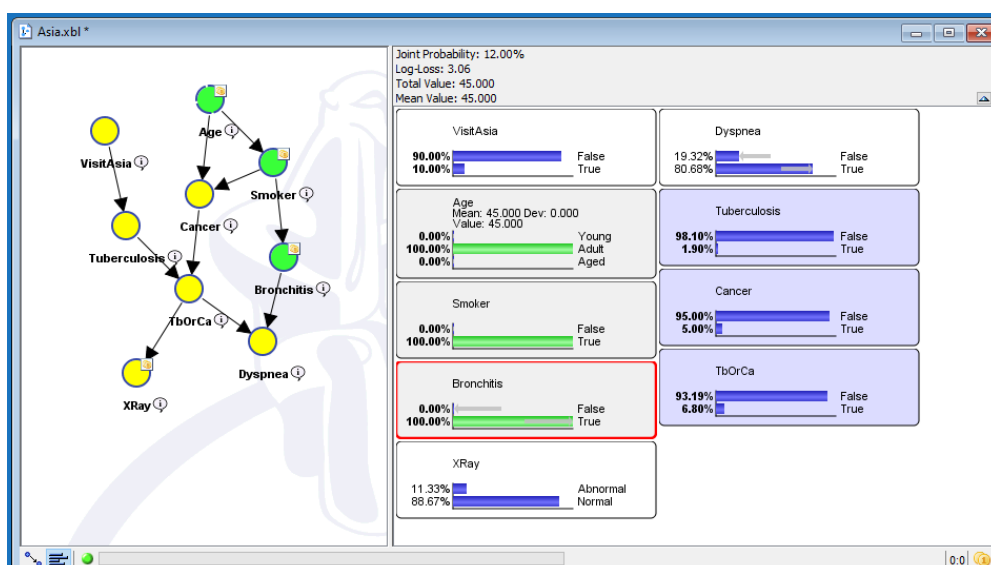


Рис. 2 Байесовская сеть с жесткими свидетельствами

Исходя из приведенных первичных настроек байесовской сети, обозначенных свидетельствах и проведенных расчетов видим следующее:

- Вероятность того, что респондент курит равна 100% (Свидетельство).
- Вероятность того, что респондент среднего возраста равна 100% (Свидетельство).
- Вероятность того, что респондент не болен туберкулезом равна 98.10% (болеет = 1.9%).
- Вероятность того, что респондент не болен раком равна 95.00% (болеет = 5.00%).
- Вероятность того, что респондент не болен туберкулезом или раком равна 93.19% (болеет = 6.81%).

Заметим, что жесткие свидетельства не изменяются ни при каких расчетах. Вероятности других факторов изменяются при назначении жестких свидетельств.

### Мягкие свидетельства

Как и в предыдущем примере, предположим, что требуется с помощью описанной выше модели определить вероятность заболевания раком или туберкулезом конкретного человека. Однако исходные данные о данном респонденте несколько изменим - этот человек либо молодой, либо среднего возраста (с равной вероятностью), но не преклонного возраста; этот человек курит; с вероятностью 70% человек болен бронхитом.

Данные свидетельства несколько изменены по сравнению с предыдущими. Второе свидетельство осталось жестким, как и в предыдущем варианте. Первое свидетельство носит мягкий характер. Мы визуально определили, что это человек не преклонного возраста, но определить, юноша это или человек среднего возраста, сложно. Есть серьезные подозрения, что человек болен бронхитом, но утверждать это мы не можем.

На рисунке 3 показана байесовская сеть с указанными свидетельствами.

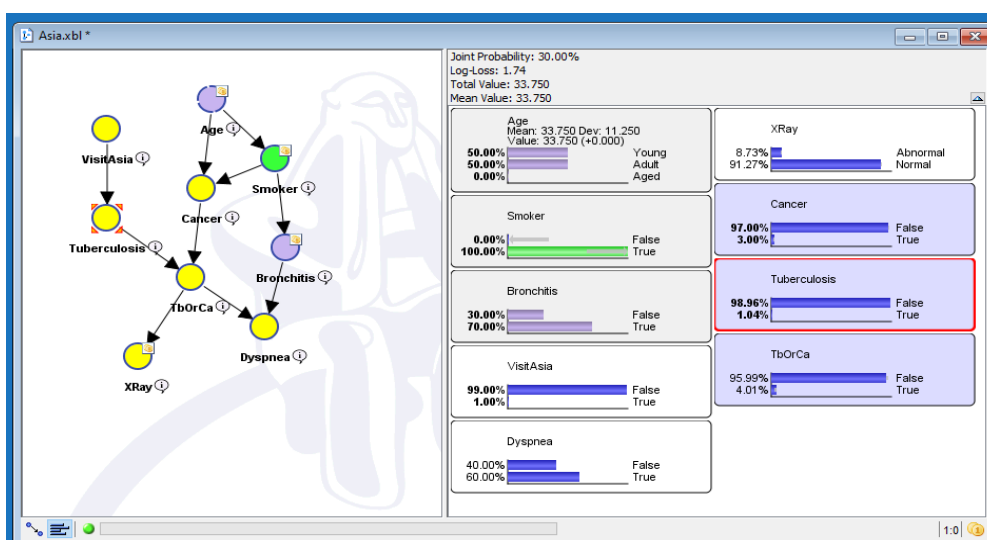
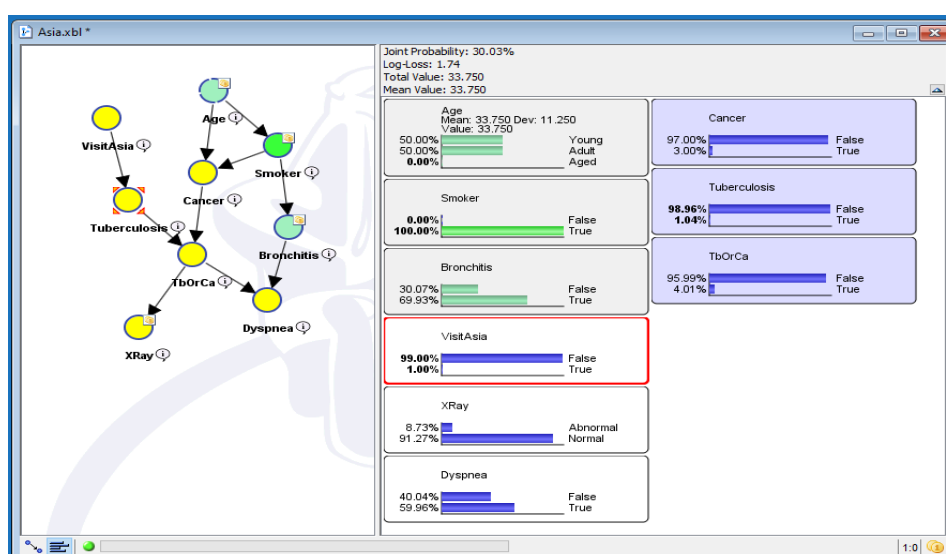


Рис. 3 Байесовская сеть с мягкими свидетельствами

Исходя из приведенных первичных настроек байесовской сети, обозначенных свидетельствах и проведенных расчетов видим следующие новые расчетные данные:

- Вероятность того, что респондент курит равна 100% (Жесткое свидетельство).
- Респондент не преклонного возраста (Мягкое свидетельство).
- Респондент с вероятностью 70% человек болеет бронхитом (Мягкое свидетельство).
- Вероятность того, что респондент не болеет туберкулезом равна 98.96% (болеет = 1.04%).
- Вероятность того, что респондент не болеет раком равна 97.00% (болеет = 3.00%).
- Вероятность того, что респондент не болеет туберкулезом или раком равна 95.99% (болеет = 4.01%).

Использование мягких свидетельств ставит перед разработчиками алгоритмов и перед исследователями ряд вопросов. Каким наиболее разумным способом использовать мягкие свидетельства? Например, должны ли меняться сами мягкие свидетельства в процессе расчета байесовской сети, или они должны оставаться неизменными. На рисунке 3 мягкие свидетельства остаются неизменными. Вводя те же самые свидетельства, но, разрешив изменяться мягким свидетельствам, получим новое решение. На рисунке 4 приведены те же мягкие свидетельства, что и на рисунке 3, но разрешено изменение свидетельств в процессе пересчета сети.



**Рис. 4 Байесовская сеть, позволяющая изменение свидетельств в процессе пересчета**

Пока решение полностью совпадает с решением на рисунке 3. Попробуем установить жесткое свидетельство – респондент болеет раком. Мы видим (рисунок 5), что наши мягкие свидетельства сильно изменились. Выяснилось, наиболее вероятно, что человек среднего возраста (83.33%).

Из приведенных примеров видно, что из-за достаточно сложных взаимосвязей в байесовской сети результаты часто становятся неожиданными. Необходимо время чтобы осмыслить и понять полученные результаты. А если сеть достаточно сложная, что-либо проверить вручную невозможно. При этом интуитивно понятно, что многие непонятные моменты вызваны именно использованием «мягких» свидетельств.

Желание более ясно понимать расчеты в байесовских сетях с «мягкими» свидетельствами заставляет нас внимательно разобраться с проблемами, возникающими при использовании «мягких» свидетельств.

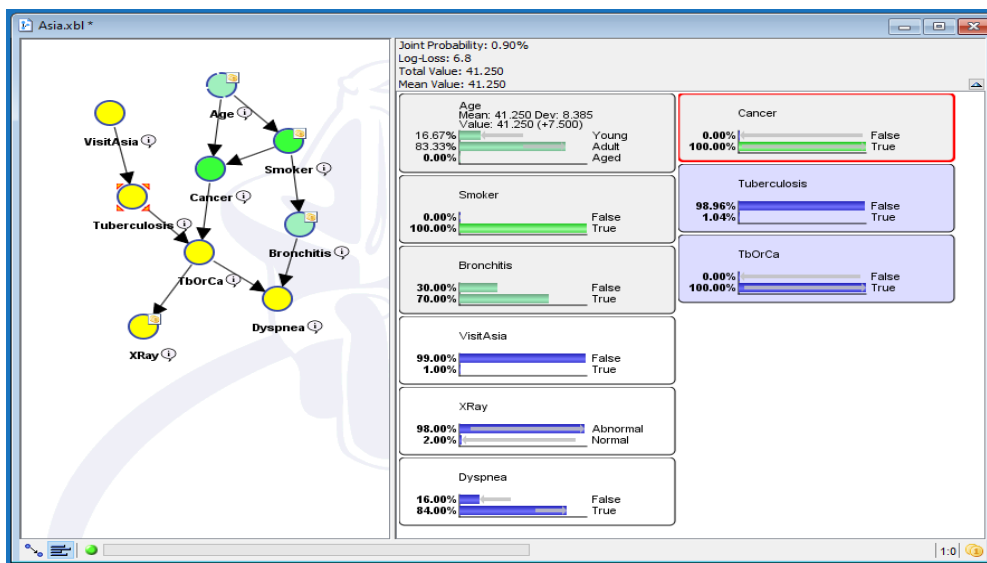


Рис. 5 Изменение свидетельств в Байесовской сети

### Проблемы

В данной работе мы пытаемся систематизировать проблемы, связанные с использованием «мягких» свидетельств, а также предложить варианты решения этих проблем. Проблемы, связанные с «мягкими» доказательствами, можно разделить на три типа:

1. **Возникновение и определение «мягких» свидетельств.** Случаи, связанные с возникновением и определением «мягких» свидетельств можно разделить на следующие:

а. Вероятностное определение «мягкого» свидетельства. Например, согласно медицинскому обследованию вероятность заболевания конкретным человеком бронхитом равна 70%.

Данная возможность реализована, например, в **BayesiaLab**. Однако необходимо определить, что в данном случае мы понимаем под термином «свидетельство». С одной стороны, сам термин говорит о том, что это что-то неизменное и не должно изменяться в процессе дальнейших пересчетов сети, с другой стороны «вероятность 70%» позволяет думать, что «мягкое» свидетельство может изменяться в каких-то разумных пределах. В **BayesiaLab** существует две возможности: зафиксировать «мягкое» свидетельство или позволить ему изменяться в любых пределах. Однако, в реальной ситуации часто требуется, чтобы «мягкое» свидетельство могло изменяться в процессе пересчета сети, но лишь в **заданных пользователем пределах**. Например, вероятность заболевания бронхитом должна находиться в пределах 60% - 80%. Неправильно, если в процессе пересчета сети вероятность окажется равной 50% или 90%. Наверное, некоторыми искусственными методами в **BayesiaLab** этого можно добиться, но нам кажется, что такие случаи достаточно типичны и должен быть простой механизм, позволяющий пользователю указывать нужный диапазон вероятностей. Данный вариант «мягкого» свидетельства рассматривается в пункте (b.). Но также должен существовать механизм настройки изменений, который бы

позволял бы «мягкому» свидетельству изменяться в некоторых заранее оговоренных пределах в случае, если пользователь собственные пределы не задавал.

b. Вероятностное ограничение «мягкого» свидетельства. Например, согласно медицинскому обследованию вероятность заболевания конкретным человеком бронхитом не ниже 70% (т.е. от 70% до 100%). Разумно предположить, что, задав таким образом «мягкое» свидетельство, пользователь считает, что в процессе пересчета байесовской сети свидетельство может изменяться, но только в указанных пределах. Возможности устанавливать жестко значение данного свидетельства программа не должна предусматривать.

c. Вероятностное определение «мягкого» свидетельства по совокупности свойств. Допустим, переменная, соответствующая узлу «Age» может принимать с какой-то вероятностью три значения – «молодой», «средний возраст», «преклонный возраст». Известно, что исследуемый человек с вероятностью 95% либо среднего, либо преклонного возраста. Должна существовать возможность определить «мягкое» свидетельство таким образом, чтобы во время пересчета вероятности среднего и преклонного возраста изменялись произвольно, но сумма вероятностей всегда была бы равна 95%.

2. **Возможность изменения узлов и правила изменения узлов.** Можно выделить следующие варианты решения задания «мягких» свидетельств:

a. «Мягкое» свидетельство не должно изменяться во время других корректировок сети и пересчета. Этот вариант реализован в **BayesiaLab**. Вероятность свидетельства не изменяется во время пересчета.

b. Во время корректировок и пересчета сети «мягкое» свидетельство может изменяться всегда и произвольно, в зависимости алгоритма пересчета. Этот вариант реализован в **BayesiaLab**. Произвольность изменения вероятности «мягкого» свидетельства не всегда устраивает исследователя.

c. Во время корректировок и пересчета сети, если возникает конфликт свидетельств, может быть три варианта:

i. Разрешить «мягкому» свидетельству (или нескольким свидетельствам) изменяться вне указанных диапазонов для предотвращения возможных конфликтов.

ii. Завершить аварийно пересчет сети с указанием проблемного свидетельства (или свидетельств).

iii. Позволить исследователю принять решение.

d. Во время корректировок и пересчета сети «мягкое» свидетельство может изменяться в особых случаях, при выполнении каких-либо условий.

e. Во время корректировок и пересчета сети «мягкое» свидетельство может изменяться в ограниченных пределах.

f. Во время корректировок и пересчета сети «мягкое» свидетельство может изменяться в особых случаях, при выполнении каких-либо условий, но в ограниченных пределах.

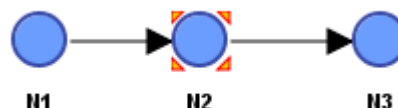


Рис. 6 Причинно-вероятностная связь между узлами

g. Если узел **N2** принимает «мягкое» свидетельство, причинно-вероятностная связь между узлами **N1** и **N3** нарушается. Расчеты узлов происходят по-другому.

h. Если узел **N2** принимает «мягкое» свидетельство, причинно-вероятностная связь между узлами **N1** и **N3** не нарушается, расчет узлов **N1, N2, N3** проводится по стандартным правилам.

i. Если узел **N2** принимает «мягкое» свидетельство, причинно-вероятностная связь между узлами **N1** и **N3** не нарушается, но расчет узлов **N1, N2, N3** проводится по другим правилам.

3. **Выбор алгоритмов расчёта сети.** После ввода «мягких» свидетельств, естественно, должны корректироваться и алгоритмы расчета сети. Новые алгоритмы должны учитывать тип введенных «мягких» свидетельств, а также их порядок обработки. Выделим, на наш взгляд, основные моменты, которые нужно учитывать при разработке алгоритма для случая использования «жестких» и «мягких» свидетельств:

a. До ввода свидетельств, как жестких, так и мягких, порядок расчетов узлов сети удобно расположить от родителей к дочерним элементам. Удобно, если расчет будет происходить по поколениям.

b. В случае, если возникают конфликты по вине «жестких» свидетельств, необходимо аварийно прервать расчет и выдать поясняющее сообщение.

c. В случае, если возникают конфликты по вине «мягких» свидетельств, которые нельзя корректировать, или разрешено корректировать в указанных пределах, можно поступить тремя способами:

i. Разрешить «мягкому» свидетельству (или нескольким свидетельствам) изменяться вне указанных диапазонов для предотвращения возможных конфликтов.

ii. Завершить аварийно пересчет сети с указанием проблемного свидетельства (или свидетельств).

iii. Позволить исследователю принять решение.

d. В случае ввода «жестких» свидетельств, а также «мягких» свидетельств, не подлежащих корректировке необходимо определить оптимальный порядок расчета узлов.

e. В случае ввода всех свидетельств необходимо определить порядок корректировки «мягких» свидетельств.

Изложение материала в данном пункте - это наше понимание какими должны быть алгоритмы для расчетов в байесовской сети, основанное на анализе работы существующих пакетов для работы с байесовскими сетями, а также понимание пути построения алгоритмов и разработки программного обеспечения для грантового проекта «Разработка и программная реализация пакета прикладных программ для решения прикладных задач по байесовским сетям».

В первом варианте разработки программного продукта мы попытаемся частично реализовать наше понимание «мягких» свидетельств.

**Заключение.** Понятие «мягких» свидетельств постоянно развивается. Это находит отражение, например, в разработках **BayesiaLab**. На наш взгляд развитие понятия «мягкие свидетельства» - достаточно перспективное направление в развитии байесовских сетей. Мы пытаемся понять тенденции развития данного понятия и его перспективы для разработки программного продукта по грантовому проекту.

## Литература

1. Barber, D. Bayesian Reasoning and Machine Learning. - <http://web4.cs.ucl.ac.uk/staff/D.Barber/textbook/020217.pdf> (19.05.2018)
2. Barber, D. Machine Learning. A probabilistic Approach. - <https://pdfs.semanticscholar.org/7bc7/54bc548f32b9ac53df67e3171e8e4df66d15.pdf> (26.05.2018)
3. Jensen, F.V., Nielsen T.D. Bayesian Networks and Decision Graphs. – Springer, 2007. – 447 p.
4. Murphy K.P. An introduction to graphical models. – <http://www.stat.yale.edu/~jtc5/BioinformaticsCourse2001/MurphyBayesNetIntro.pdf> (20.05.2018)
5. Conrady S., Jouffe L. Bayesian Networks & BayesiaLab. A Practical Introduction for Researches. ISBN: 978-0-9965333-0-0
6. AgenaRisk 7.0 User Manual. – <http://www.agenarisk.com> (26.05.2018)
7. Getting Started with AgenaRisk. – <http://www.agenarisk.com> (20.05.2018)
8. Murphy K.P. Machine Learning. A Probabilistic Perspective. - The MIT Press. – 1067 p.
9. Neapolitan R.E. Learning Bayesian Networks. – [http://www.cs.technion.ac.il/~dang/books/Learning%20Bayesian%20Networks\(Neapolitan,%20Richard\).pdf](http://www.cs.technion.ac.il/~dang/books/Learning%20Bayesian%20Networks(Neapolitan,%20Richard).pdf) (15.05.2018)
10. Бидюк П.И., Терентьев А.Н. Построение и методы обучения Байесовских сетей //Таврический вестник информатики и математики, 2004. – №2. - с.139-154
11. Гончаров М. Байесовские сети. – <http://www.businessdataanalytics.ru/download/BayesianNetworks.pdf> (26.05.2018)

## **БИОМЕТРИЧЕСКАЯ ИДЕНТИФИКАЦИЯ ЧЕЛОВЕКА ПО ЛИЦУ**

**Мазаков Т.Ж., Айпанов Ш.А., Тусупова С.А., Байрбекова Г.С.,  
Зиятбекова Г.З., Мазакова А.Т.**

*Институт информационных и вычислительных технологий КН МОН РК  
e-mail: [tmazakov@mail.ru](mailto:tmazakov@mail.ru)*

***Аннотация.** Статья посвящена биометрической идентификации человека по лицу. В качестве информативных признаков предложены двумерные и трехмерные характеристики лица человека, учитывающие площадь и объем.*

*Для учета таких явлений, как сдвиг портрета, разный масштаб фотографий и наклон идентифицируемого лица, разработан сложный алгоритм идентификации.*

***Ключевые слова:** защита информации, двухмерное и трехмерное изображение, идентификация.*

**Введение.** Проблема защиты информации и информационной безопасности является одним из важнейших аспектов развития современного общества. В настоящее время решение этой проблемы в области разработки и эксплуатации информационных систем различного назначения связано с разработкой всевозможных требований к обеспечению их безопасности и созданием программно-аппаратных средств от несанкционированного доступа.



Автоматическое распознавание лица для установления личности имеет большое количество приложений в различных областях. Проблемы общественной безопасности, потребность в удаленной аутентификации, развитие человеко-машинных интерфейсов вызывает повышенный интерес к данной технологии.

Уже несколько десятилетий ведется разработка методов распознавания лиц, тем не менее, эта проблема по-прежнему далека от завершения. Из-за переменчивых условий визуализации лиц, связанных с освещением, положением головы по отношению к камере, старением, мимикой и другими факторами, автоматическое распознавание лиц является непростой задачей. Накладывая жесткие ограничения на процесс фотосъемки лиц, при проектировании систем стараются избежать негативного воздействия указанных факторов. Однако наибольший практический интерес представляет задача распознавания лиц на изображениях, полученных в неконтролируемых условиях. Благодаря развитию методов машинного обучения и появлению больших баз фотографий для тренировки систем, в последние годы наблюдается значительный прогресс в данной области [1, 2].

В Указе Президента Республики Казахстан от 10 октября 2006 года N 199 «О Концепции информационной безопасности Республики Казахстан» отмечено: «Анализ современного состояния информационной безопасности в Казахстане показывает, что ее уровень в настоящее время не соответствует потребностям человека, общества и государства» и в качестве основной цели обеспечения информационной безопасности указано: «создание и укрепление национальной системы защиты информации, в том числе в государственных информационных ресурсах».

31 января 2017 г. Президент Республики Казахстан Нурсултан Абишевич Назарбаев обратился к казахстанцам с посланием «Третья модернизация Казахстана: глобальная конкурентоспособность». В этом обращении отмечена необходимость разработки и принятия программы «Цифровой Казахстан». В связи с этим по поручению Н.А. Назарбаева в целях обеспечения информационной безопасности общества и государства в сфере информатизации и связи, а также защиты неприкосновенности частной жизни граждан при использовании ими информационно-коммуникационных технологий разработана Концепция «Киберцит Казахстана». В ней отмечено, что особого внимания требуют вопросы подготовки кадров в вузах Казахстана по информационной безопасности и разработки отечественных средств защиты информации.

В связи с развитием аппаратно-программного обеспечения в последние годы многие разработчики перешли от исследования двумерных изображений лица к трехмерным, что позволило перейти от обработки характерных точек лица к объемным характеристикам [3].

**Основная часть.** Обработка видеок кадров разрабатываемой системой должна включать модули, представленные на рисунке 1.

Первый этап – построение модели распознаваемого объекта. Нами выбраны 25 параметров учитывающие такие характеристики лица человека как площадь и объем.

Второй этап – выделение информативных признаков. Ниже представлены алгоритмы обработки информационных параметров для идентификации человека по лицу.

Третий этап – проверка подлинности распознаваемого объекта.

Четвертый этап – сопоставление информативных признаков объекта с БД. На данном этапе необходимо учитывать такие факторы как масштаб, положение лица и т.д.

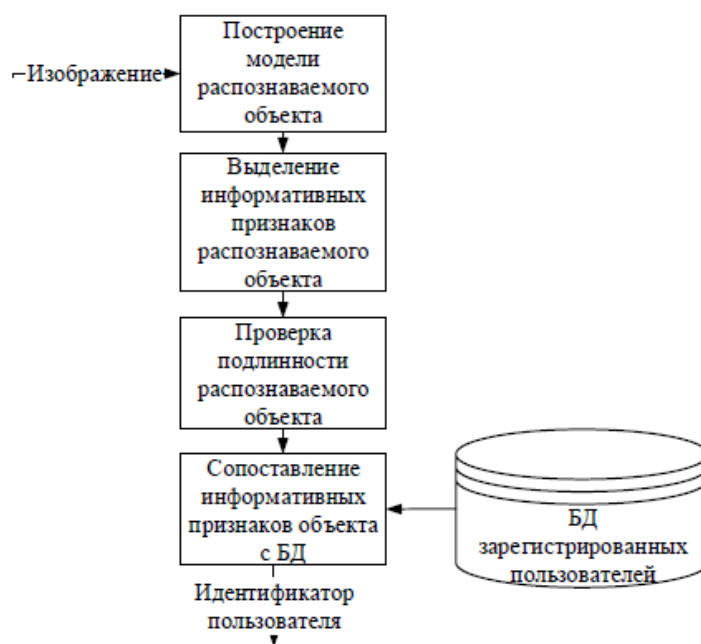


Рисунок 1. Структура системы распознавания изображений лиц

Для представления геометрии облака точек, выбрана модель отображения данных в формате PLY. PLY — формат файлов описания геометрии, известный также как Polygon File Format и Stanford Triangle Format. Он был разработан, главным образом, для хранения трёхмерных данных 3D сканеров. Формат поддерживает относительно простое описание объекта как списка плоских полигонов. PLY может хранить множество свойств объекта, включающее: цвет и прозрачность, нормали к поверхности, текстурные координаты и т.д. Формат позволяет иметь различные свойства передней и задней грани полигона. Существует 2 версии формата PLY: ASCII и в виде бинарного файла.

Файлы устроены в виде заголовка, в котором определяются элементы полисеток и их типы, и следующего за ним списка самих элементов. Элементы — это обычно вершины и грани, но могут включаться другие сущности, такие как рёбра или полосы треугольников (triangle strips).

Алгоритм построения матрицы высот базируется на методе интерполяции поверхностей. В нем неравномерно распределенные точки в трехмерном пространстве интерполируются непрерывной функцией двух независимых переменных. Рассмотрим создание регулярной матрицы высот. Для построения матрицы высот выполняются следующие этапы: формирование опорных узлов, вычисление матрицы ближайших точек и матрицы расстояний, интерполяция узлов, корректировка матрицы высот.

Ввиду того, что данные в формате PLY представлены поточечно координатами  $(x, y, z)$  разработана программа линейной интерполяции, которая строит регулярную матрицу высот. При работе программы запрашивает количество строк и столбцов результирующей матрицы, определяет диапазон представления данных  $(Xmin, Xmax)$ ,  $(Ymin, Ymax)$  и  $(Zmin, Zmax)$ . В результирующий текстовый файл в справочный блок

записываются данные о количестве строк и столбцов, матрицы и диапазон представления данных ( $X_{min}$ ,  $X_{max}$ ), ( $Y_{min}$ ,  $Y_{max}$ ) и ( $Z_{min}$ ,  $Z_{max}$ ).

Значения элементов матрицы вычисляются с помощью билинейной функции, также вычислить приближенное значение функции внутри этого треугольника можно с помощью билинейной функции.

$f(x, y) \approx F(x, y) = ax + by + c$  находя коэффициенты  $a$ ,  $b$ ,  $c$  из условий

$$ax_A + by_A + c = f_A,$$

$$ax_B + by_B + c = f_B,$$

$$ax_D + by_D + c = f_D,$$

где  $\{x_A, y_A, f_A\}$ ,  $\{x_B, y_B, f_B\}$ ,  $\{x_D, y_D, f_D\}$  – координаты вершин  $A$ ,  $B$ ,  $D$  некоторого треугольника на треугольной расчетной сетке. Погрешность такой интерполяции для функции  $f(x, y)$  с непрерывными вторыми производными будет  $O(h^2)$ , где  $h$  — длина наибольшей стороны треугольника  $ABD$ . Причем треугольники  $A$ ,  $B$ ,  $D$  выбираются исходя из условия минимума расстояния между вершинами и точкой  $\{x, y\}$ , в которой нужно вычислить интерполяционное значение. Текст программы приведен в приложении А.

Далее в результативный файл записывается построчно элементы вычисленной матрицы высот.

В случае недостаточного количества строк и столбцов или не гладкости поверхности разработана программа кубической интерполяции, которая из исходной регулярной матрицы высот рассчитывает новую регулярную матрицу высот с большими размерами матрицы [1x3].

На базе СУБД Visual FoxPro (реализована интерфейсная часть), включающая следующие режимы: 1) биологические характеристики, 2) параметры характеристик, 3) исходные базы данных, 4) настройка базы данных, 5) классификация, 6) простая идентификация, 7) сложная идентификация.

На данный момент в качестве биологической характеристики включен «видеообраз лица». В дальнейшем планируется работа со следующими биологическими характеристиками: «термограмма лица», «отпечаток пальца», «геометрия руки», «голос» и др.

Для параметров характеристик введены следующие типы:

1 - координата точки; 2 - расстояние (число); 3 - площадь; 4 - объем.

В дальнейшем по мере введения новых биометрических характеристик будут вводиться новые типы.

Для характеристики «видеообраз лица» определены ряд параметров, представляющих собой:

1) точку - координаты зрачков глаз, переносицы, кончика носа,

2) расстояние – между глазами, между переносицей и кончиком носа, основанием носа,

3) периметр - треугольника (зрачки глаз и кончик носа), треугольника (переносица и основание носа),

4) площадь – изолинии глазниц, изолинии носа,

5) объем – глазниц, носа.

В качестве исходных данных для «видеообраза лица» используется объемная 3d-модель, представленная как регулярная матрица высот [4].

Алгоритмы обработки информационных параметров для идентификации человека по лицу:

1. Координаты центра зрачка левого глаза -  $(P1x, P1y)$ : определяются из графического файла с фотографией человека;
2. Координаты центра зрачка правого глаза -  $(P2x, P2y)$ : определяются из графического файла с фотографией человека;
3. Расстояние между зрачками -  $P3$ : вычисляются через данные  $(P1x, P1y)$  и  $(P2x, P2y)$ ;
4. Глубина левой глазницы -  $P4$ : определяются из файла с 3D-данными лица человека;
5. Площадь первой изолинии левой глазницы -  $P5$ : вычисляются файла с 3d-данными лица человека путем аппроксимации изолинии эллипсом;
6. Площадь второй изолинии левой глазницы -  $P6$ : вычисляются файла с 3d-данными лица человека путем аппроксимации изолинии эллипсом;
7. Площадь третьей изолинии левой глазницы -  $P7$ : вычисляются файла с 3d-данными лица человека путем аппроксимации изолинии эллипсом;
8. Объем левой глазницы -  $P8$ : вычисляются через данные  $P5, P6, P7$  и расстояние (шаг) между изолиниями;
9. Глубина правой глазницы -  $P9$ : вычисляется аналогично параметру  $P4$ ;
10. Площадь первой изолинии правой глазницы – $P10$ : вычисляется аналогично параметру  $P5$ ;
11. Площадь второй изолинии правой глазницы –  $P11$ : вычисляется аналогично параметру  $P6$ ;
12. Площадь третьей изолинии правой глазницы –  $P12$ : вычисляется аналогично параметру  $P7$ ;
13. Объем правой глазницы –  $P13$ : вычисляется аналогично параметру  $P8$ ;
14. Координаты кончика носа –  $(P14x, P14y)$ : определяются из графического файла с фотографией человека;
15. Координаты левого основания носа –  $(P15x, P15y)$ : определяются из графического файла с фотографией человека;
16. Координаты правого основания носа –  $(P16x, P16y)$ : определяются из графического файла с фотографией человека;
17. Высота кончика носа –  $P17$ : определяются из файла с 3d-данными лица человека;
18. Высота переносицы –  $P18$ : определяются из файла с 3d-данными лица человека;
19. Площадь первой изолинии левой носа –  $P19$ : вычисляются файла с 3d-данными лица человека путем аппроксимации изолинии трапецией;
20. Площадь второй изолинии левой носа– $P20$ : вычисляются файла с 3d-данными лица человека путем аппроксимации изолинии трапецией;
21. Площадь третьей изолинии левой носа – $P21$ : вычисляются файла с 3d-данными лица человека путем аппроксимации изолинии трапецией;
22. Объем носа –  $P22$ : вычисляются через данные  $P19, P20, P21$  и расстояние (шаг) между изолиниями;
23. Периметр треугольника «нос-глаза» – $P23$ : вычисляются через данные  $(P1x, P1y), (P2x, P2y)$  и  $(P14x, P14y)$ ;
24. Площадь треугольника «нос-глаза» – $P24$ : вычисляются через данные  $(P1x, P1y), (P2x, P2y)$  и  $(P14x, P14y)$  по формуле Герона;
25. Отношение «Расстояние между зрачками»/«Расстояние между переносицей и основанием носа» –  $P25$ : вычисляются через данные  $(P1x, P1y), (P2x, P2y)$  и  $(P14x, P14y)$ .

*Примечание.* Площади фигуры, ограниченной изолинией вычисляется по формуле Гаусса:

$$A = \frac{1}{2} \left| \sum_{i=1}^{n-1} x_i y_{i+1} + x_n y_1 - \sum_{i=1}^{n-1} x_{i+1} y_i - x_1 y_n \right|$$

$$= \frac{1}{2} |x_1 y_2 + x_2 y_3 + \dots + x_{n-1} y_n + x_n y_1 - x_2 y_1 - x_3 y_2 - \dots - x_n y_{n-1} - x_1 y_n|$$

где  $A$  — площадь многоугольника,  $n$  — количество сторон многоугольника,  $(x_i, y_i)$ ,  $i = 1, 2, \dots, n$  — координаты вершин многоугольника.

Режим — «классификация». В программе реализована классификация по одному или нескольким параметрам. В данном случае задача классификации состоит в упрощении матрицы данных, слишком обширной для непосредственного анализа человеком. В программе реализованы различные алгоритмы классификации, так называемые стратегии объединения (агломеративные системы): гибкая стратегия, стратегия ближайшего соседа, стратегия дальнего соседа, стратегия группового среднего, центроидная стратегия, стратегия на сумме квадратов.

Режим — «простая идентификация». В данном режиме для вводимого изображения лица, которое необходимо идентифицировать, вычисляется степень соответствия его с каждым из изображений, внесенных в таблицу IsxDan.dbf. При этом для исключения сдвига координаты всех точек сдвигаются относительно правого глаза.

Режим — «сложная идентификация». В данном режиме для вводимого изображения лица, которое необходимо идентифицировать, вычисляется степень соответствия его с каждым из изображений, внесенных в таблицу IsxDan.dbf. При этом реализованы следующие режимы;

1) для исключения сдвига координаты всех точек смещаются относительно правого глаза,

2) для исключения наклона лица за точку все координаты точек поворачиваются на угол, соответствующий разности ординат левого и правого глаз в соответствии со следующими шагами:

- пусть  $(x_1, y_1)$  — координаты правого глаза,  $(x_2, y_2)$  — координаты левого глаза,  $(x_3, y_3)$  — координаты любой точки, которую надо повернуть в соответствии с алгоритмом,  $(x'_2, y'_2)$  — координаты нового положения левого глаза,  $(x'_3, y'_3)$  — координаты нового положения точки  $(x_3, y_3)$ .

Перенос начала координат в точку  $(x_1, y_1)$

$$2: (x_2 - x_1, y_2 - y_1), \quad 3: (x_3 - x_1, y_3 - y_1),$$

$$2': (x'_2 - x_1, y'_2 - y_1), \quad 3': (x'_3 - x_1, y'_3 - y_1),$$

Вычисляется матрица поворота на угол  $\alpha$ :  $A = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}$

$$\text{Дано: } 2 \rightarrow 2' \begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \end{pmatrix} = A \begin{pmatrix} x'_2 - x_1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos \alpha (x'_2 - x_1) \\ -\sin \alpha (x'_2 - x_1) \end{pmatrix}$$

откуда

$$\cos \alpha = \frac{x_2 - x_1}{x'_2 - x_1}, \sin \alpha = \frac{y_1 - y_2}{x'_2 - x_1}.$$

Приведенная формула корректна, т.к.

$$|x'_2 - x_1| = \sqrt{(x'_2 - x_1)^2 + (y'_2 - y_1)^2}$$

Из формулы

$$\begin{pmatrix} x'_3 - x_1 \\ y'_3 - y_1 \end{pmatrix} = A^T \begin{pmatrix} x_3 - x_1 \\ y_3 - y_1 \end{pmatrix} = \frac{1}{x'_2 - x_1} \begin{pmatrix} x_2 - x_1 & y_2 - y_1 \\ y_1 - y_2 & x_2 - x_1 \end{pmatrix} \begin{pmatrix} x_3 - x_1 \\ y_3 - y_1 \end{pmatrix}$$

находим координаты нового положения точки 3

$$x'_3 = \frac{(x_2 - x_1)(x_3 - x_1) + (y_2 - y_1)(y_3 - y_1)}{x'_2 - x_1} + x_1$$

$$y'_3 = \frac{(y_1 - y_2)(x_3 - x_1) + (x_2 - x_1)(y_3 - y_1)}{x'_2 - x_1} + y_1$$

3) для исключения возможности рисунков разного масштаба, все данные преобразуются пропорционально расстояниям между глазами.

### **Заключение**

Разработана АРМ «Биометрическая система защиты информации». Определены структуры таблиц базы данных и их взаимосвязь. Разработаны классы, соответствующие спроектированным модулям и реализующие алгоритм работы приложения. Приведены подробные описания их атрибутов и методов. Разработан графический интерфейс пользователя приложения. Описаны различные рабочие области интерфейса. Выявлены основные проблемы, возникшие в ходе разработки биометрических систем аутентификации и идентификации.

Для учета таких явлений, как сдвиг портрета, разный масштаб фотографий и наклон идентифицируемого лица, разработан сложный алгоритм идентификации.

*Работа выполнена за счет средств грантового финансирования научных исследований на 2018-2020 годы по проекту AP05131027 «Разработка биометрических методов и средств защиты информации».*

### **Литература**

1. Колесниченко Н.М., Черняева Н.Н. Инженерная и компьютерная графика. – М.: ИНФРА-Инженерия, 2018. – 236с.
2. Компьютерная геометрия /Голованов Н.Н., Ильютко Д.П., Носовский Г.В., Фоменко А.Т. – М.: Изд. центр «Академия», 2006. – 512 с.
3. Кухарев Г.А., Каменская Е.И., Матвеев Ю.Н., Щеголева Н.Л. Методы обработки и распознавания изображений лиц в задачах биометрии. – М.: Политехника, 2013. – 416 с.
4. Джомартова Ш.А., Мазакон Т.Ж., Мазакон А.Т. Автоматизированная система поиска кольцевых структур //Вестник национальной инженерной академии Республики Казахстан, 2016, № 1 (59), с.59-64.

## ОБРАБОТКА СИГНАЛОВ ЭКГ НА МИКРОПРОЦЕССОРАХ

**Мазаков Т.Ж., Исимов Н.Т., Зиятбекова Г.З., Жолмагаметова Б.Р.,  
Джомартова Д.Т., Ыдырышбаева М.Б.**

*Институт информационных и вычислительных технологий КН МОН РК  
e-mail: tmazakov@mail.ru*

**Аннотация.** Разработан экспериментальный вариант системы психологического тестирования с фиксированием физиологических параметров тестируемого в реальном времени. В качестве источников физиологических данных определены данные электрокардиограммы (ЭКГ). Программно-аппаратный комплекс психофизиологического тестирования позволяет при ответе на каждый вопрос теста фиксировать и оценивать психофизиологическое состояние тестируемого, что дает дополнительную информацию для психолога.

**Ключевые слова:** электрокардиограмма, микропроцессор, обработка сигналов, психологические тесты.

**Введение.** В эпоху научно-технического прогресса с его напряженными ритмами, новыми специфическими условиями деятельности человека, значительно возрастают требования к его интеллектуальным, эмоциональным и волевым ресурсам. В этой связи особенно остро возникает потребность со стороны кадровых подразделений организаций в объективном психофизиологическом портрете личности. Основным аппаратом психологов являются психологические тесты. Однако, как показывает практика, в связи общей доступностью к тестам в последнее время увеличится эффект субъективизма.

Бурное развитие компьютерной техники способствовали автоматизации проведения и обработки психологического тестирования [1] и применения новых методов математической обработки биомедицинских данных [2]. Современные возможности по разработке различных датчиков [3] и удешевление микропроцессоров также открыли широкую возможность по внедрению аппаратно-программных средств оценки психофизиологического портрета личности [4 -6].

Вышеизложенные обстоятельства диктуют необходимость в создании объективной системы профессионального отбора личности.

### **Основная часть**

Для системы профессионального отбора определен ряд тестов (рекомендованных психологическими службами кадровых подразделений). Разрабатывается система психологического тестирования на казахском и русском языках с фиксированием физиологических параметров тестируемого в реальном времени. В качестве источников физиологических данных определены данные ЭКГ. На платформе Arduino [7] разработаны система приема и обработки данных с датчиков ЭКГ.

Для подключения датчиков ЭКГ использована микросхема AD8232 (продукт компании Analog Devices), который представляет собой интегрированный блок обработки сигнала для ЭКГ и других биопотенциальных задач [8]. Микросхема

предназначена для получения, усиления и фильтрации слабых биопотенциальных сигналов в условиях сильных помех.

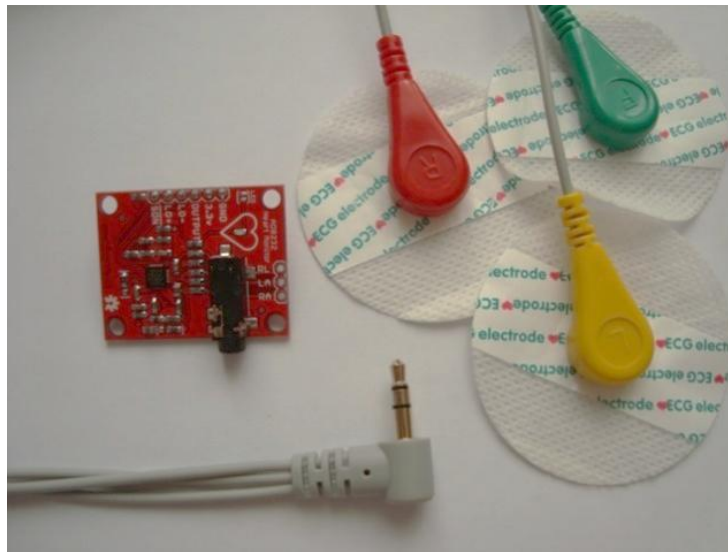


Рис.1 Общий вид микросхемы AD8232 и датчиков ЭКГ

Основные характеристики AD8232 (рисунок 1):

- Низкое потребление тока: 170 мкА,
- Напряжение питания: однополярное от 2 до 3,5 В,
- Rail to Rail выходной сигнал,
- Количество электродов: 2 или 3,
- Количество отведений ЭКГ: 1,
- Встроенный фильтр ВЧ помех,
- 2-полюсный фильтр высоких частот,
- 3-полюсный фильтр низких частот,
- Коэффициент ослабления синфазного сигнала: 80 дБ,
- Детектор контакта электродов,
- Выходной сигнал: аналоговый [8].

Для получения кардиограммы электроды прикрепляются на грудь и конечности, с которых снимаются сигналы электрической активности сердца.

Отличительной особенностью данного модуля является его компактность, и внешнее подключение к компьютерам, что позволяет создавать мобильные системы диагностического оборудования. Устройство соединяется с компьютером через USB-разъем.

Программно-аппаратный комплекс психофизиологического тестирования позволяет при ответе на каждый вопрос теста фиксировать и оценивать психофизиологическое состояние тестируемого, что дает дополнительную информацию для психолога.

При обработке физиологических данных вычисляются следующие параметры ЭКГ, необходимые для математической модели оценки состояния испытуемого.

Данные от датчиков ЭКГ поступают в виде числа, характеризующего амплитуду сигнала ЭКГ, с интенсивностью (частотой) 160 отсчетов за 1 сек. Обозначим через  $x_i^{(k)}$  –*i*-ый сигнал ЭКГ относящийся к *k*-му воздействию (пробе). На рисунке 2 представлен



график изменения ЭКГ по времени. На рисунке 3 представлен общий вид одного периода сигнала ЭКГ.

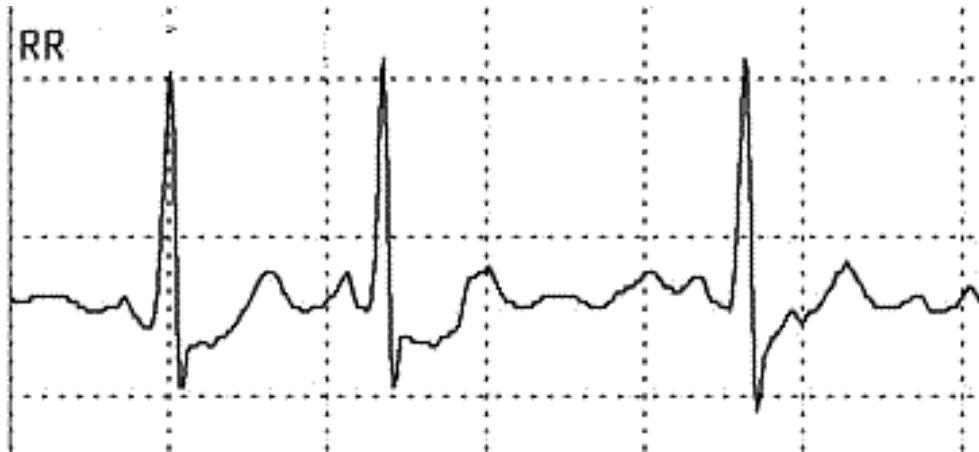


Рис.2 График изменения сигнала ЭКГ

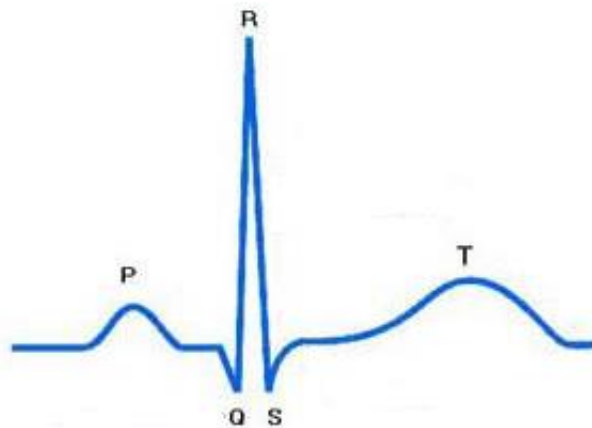


Рис.3 Общий вид одного периода сигнала ЭКГ

В качестве измеряемых параметров используются следующие данные:

- минимальная амплитуда ЭКГ;
- максимальная амплитуда ЭКГ;
- среднее значение амплитуды;
- среднеквадратичное отклонение;
- минимальное значение RR-интервала
- максимальное значение RR-интервала
- минимальное значение T-пика амплитуды
- максимальное значение T-пика амплитуды
- минимальное смещение T-пика
- максимальное смещение T-пика
- минимальное значение интеграла RR-интервала
- максимальное значение интеграла RR-интервала
- минимальное значение сдвиговой функции
- максимальное значение сдвиговой функции.

Первые четыре параметра вычисляются по формулам:

$$\begin{aligned}
 x_{cp}^{(k)} &= \frac{1}{n_k} \sum_{i=1}^{n_k} x_i^{(k)}, \\
 x_{min}^{(k)} &= \min_{i=1, n_k} x_i^{(k)}, \\
 x_{max}^{(k)} &= \max_{i=1, n_k} x_i^{(k)}, \\
 \sigma^{(k)} &= \frac{1}{n_k} \sqrt{\sum_{i=1}^{n_k} (x_i^{(k)} - x_{cp}^{(k)})^2}, \\
 z^{(k)} &= \max\{x_{max}^{(k)} - x_{min}^{(k)}, \sigma^{(k)}\},
 \end{aligned} \tag{1}$$

Для вычисления остальных характеристик применяются следующие процедуры. В силу периодического характера сигнала, выделяются каждый отдельно выбранный RR-интервал. Во время обработки фоновых данных, формируется усредненная форма сигнала ЭКГ, свойственная конкретному тестируемому в спокойной обстановке. Для фонового усредненного RR-интервала обозначим через  $Y_i, i = 1, 100$ . Таким образом, вектор  $Y$  характеризует форму индивидуального фонового RR-интервала. В процессе обработки данных ЭКГ, поступающих во время последующих воздействий (вопросов-ответов) выделяются соответственно RR-интервалы. Обозначим через  $Z_i, i = 1, L_r$ . Здесь через  $L_r$  обозначена длина очередного RR-интервала. При обработке вектор  $Z$  выделяется - зубец, который характеризуется смещением  $L_t$  относительно начала RR-интервала и амплитудой. Вычисляется площадь RR-интервала –  $S$ . Вычисляется значение сдвиговой функции  $F$ :

$$\begin{aligned}
 S &= \int z(t) dt = \sum_{i=1}^{L_r} z_i, \\
 F &= \sum_{i=1}^{100} (y_i - z_i)^2,
 \end{aligned} \tag{2}$$

При изменении психофизиологического состояния исследуемого (например, при стрессе) учащается или становится реже дыхание (что фиксируется изменением размаха RR-интервалов), уменьшается амплитуда - пика (сигнал “размазывается”), изменяется амплитуда и положение -пика. Все перечисленные признаки могут быть определены программно и использованы при диагностике исследуемого.

В качестве дополнительного параметра анализируется также время ответа на каждый вопрос теста.

## Заключение

Разработан аппаратно-программный комплекс психофизиологического тестирования, позволяющий фиксировать и оценивать психофизиологическое состояние тестируемого при ответе на каждый вопрос теста. Разработан графический интерфейс пользователя приложения. Отличительной особенностью данного модуля является его компактность, и внешнее подключение к компьютерам, что позволяет создавать мобильные системы диагностического оборудования. В дальнейшем планируется создание интерфейса на казахском и русском языках, что расширит контингент пользователей комплекса.

Ожидается использование аппаратно-программного комплекса для получения психофизиологического портрета личности при приеме на работу в государственные и частные организации, а также на службу в правоохранительные органы.

Работа выполнена за счет средств грантового финансирования научных исследований на 2018-2020 годы по проекту АР05132044 «Разработка аппаратно-медицинского комплекса оценки психофизиологических параметров человека».

### **Литература**

1. Дюк В.А. Компьютерная психодиагностика. – Санкт-Петербург: Братство, 1994. -364с.
2. Донцов В.И., Крутько В.Н., Кудашов А.А. Виртуальные приборы в биологии и медицине. М.:Ленанд 2009. – 216 с.
3. Шарапов В.М. и др. Датчики. - М.: Техносфера, 2012. -624 с.
4. Кулачев А.П. Компьютерная электрофизиология и функциональная диагностика. – М.: Форум, ИНРФА-М, 2010. – 640 с.
5. Новые методы электрокардиографии //Под ред. Грачева С.В., Иванова Г.Г., Сыркина А.Л. – М.: Техносфера, 2007.- 552 с.
6. Дмитриева Н.В. Системная электрофизиология. Системный анализ электрофизиологических процессов. – М.: Сайнс-пресс, 2008. – 256 с.
7. Петин В.А. Проекты с использованием контроллера Arduino. – Санкт-Петербург: БХВ-Петербург, 2014. – 400с.
8. Оптические датчики сердечного ритма. Простой кардиомонитор своими руками. - <http://schem.net/medic/medic37.php> (05 июня 2018 года).

## **МЕТРИКА ФИШЕРА – РАО НА СФЕРЕ ДЛЯ ТОПОЛОГИЧЕСКОГО РАСПОЗНАВАНИЯ ТЕКСТУР ЦИФРОВЫХ ИЗОБРАЖЕНИЙ**

**Макаренко Н.Г.<sup>1,2</sup>, Рыбинцев А.С.<sup>1</sup>, Князева И.С.<sup>1,2</sup>, Волобуев Д.М.<sup>1,2</sup>,  
Каримова Л.М.<sup>2</sup>, Пак И.Т.<sup>2</sup>.**

<sup>1</sup>Главная астрономическая обсерватория РАН, Санкт-Петербург, Россия

<sup>2</sup>Институт информационных и вычислительных технологий КН МОН РК,  
Казахстан

e-mail: [ng-makar@mail.ru](mailto:ng-makar@mail.ru)

**Аннотация.** В статье обсуждается распознавание текстур на цифровых изображениях методами вычислительной топологии и римановой геометрии. Топологические свойства паттернов представлены отрезками (баркодами), полученными при фильтрации по уровню фотометрической меры. Начало баркода кодирует уровень на котором появляется топологическое свойство (компонента связности и/или «дыра»), а его конец – уровень на котором свойство исчезает. Баркоды удобно параметризовать координатами их концов в прямоугольной системе координат «рождение» и «смерть» топологического свойства. Такое представление в форме облака точек на плоскости, называют диаграммой персистентности (ДП). Существуют разные способы векторизации этого (мульти)множества. Наиболее простым является получение диффузного образа ДП с помощью свертки облака точек с гауссовским ядром. Интегрируя этот образ можно получить оценку плотности распределения вероятностей – pdf в конечном окне. Эти объекты принадлежат подмножеству Банахова пространства, выделенного условиями положительности и нормировки. Плотности не образуют векторного пространства. Если использовать полуплотности, т.е. квадратные корни из pdf то мы получим единичную сферу. Метрика Фишера-Рао в касательном расслоении к ней имеет структуру гильбертова скалярного произведения. Тогда, расстояние между двумя pdf измеряется углом на сфере. Дополнительным преимуществом является возможность усреднять плотности по Карчеру-Фреше. Мы применяем этот подход для описания диаграмм персистентности, полученных по топологии природных текстур. Мы показываем, что результаты распознавания классов текстур существенно лучше, по сравнению с другими способами векторизации (ДП).

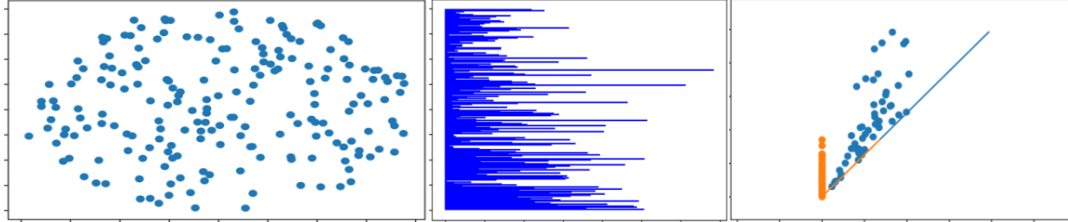
**Введение.** Для описания паттернов цифровых изображений мы используем TDA-Topological Data Analysis [1,2]. TDA не требует каких-либо априорных предположений о природе источника данных и позволяет извлечь новые знания из изменения формы окрестностей точек в пространстве признаков. Для того чтобы продемонстрировать основные идеи для изображений рассмотрим следующий пример [3].

Пусть  $V$  метрическое пространство и  $f: V \rightarrow \mathbb{R}$  гладкая функция. Обозначим через  $V_u[f]$  множество выбросов функции выше уровня  $u$ , т.е.

$$V_u[f] \equiv V_u = \{t \in V | f(t) \in [u, \infty)\} \equiv f^{-1}([u, \infty)).$$

Здесь  $f^{-1}$  обозначает обратную к  $f$  функцию, т.е. прообраз. Для графика хорошей функции в 1D, прообразы выбросов  $f^{-1}$  принадлежат непустому множеству конечных интервалов. Очевидно, что для  $u \geq s$ , справедливо вложение  $V_u \subseteq V_s$ . Тогда, для упорядоченного набора уровней, выбросы вверх (супер-уровни) образуют называемую фильтрацию, как последовательность вложенных друг в друга стратифицированных пространств:  $V_1 \supseteq V_2 \supseteq V_3 \supseteq \dots$  Одномерный интервал, т.е. отрезок, ограничен двумя крайними точками, образующими его нульмерную границу. В случае 2D мы получим для  $V_u[f]$  двухмерные области с одномерными и возможно, нульмерными границами. В общем случае, конструкцию, состоящую из объединения конечных непересекающихся множеств разной размерности называется стратифицированным пространством. Примером может служить прямоугольный фрагмент, состоящий из 4-х нульмерных страт, 4 –х одномерных страт и одного Если

размерность  $\dim V = d$ , то для гладкой функции, типично  $\dim V_u[f] = d$ . Время жизни топологического свойства, измеренного в функции высоты, называют персистентностью. Длину отрезка – баркод-можно (см. Рис.1) параметризовать на плоскости в координатах рождение – смерть.



**Рис.1 Фильтрация 200 точек выбранных случайно из эллиптической области. Баркоды (в центре) и диаграмма персистентности, справа**

Ранее, мы уже рассматривали возможности ТДА для описания текстур [4,5]. Основной проблемой является преобразование мультимножества –ДП в векторное пространство. Для этого существует несколько подходов[6]. Но все они векторизуют ДП в плотность вероятности,  $pdf$ , которые расположены в пространстве Банаха:

$$\mathcal{P} = \left\{ g: [0,1] \rightarrow \mathbb{R}_{\geq 0} \mid \int_0^1 g(t) dt = 1 \right\} \quad (1)$$

Приведем эвристические соображения о способе введения римановой метрики на множестве плотностей распределения вероятностей  $p(x|\theta)$ . Они основаны на метрике Фишера-Рао [7], в которой метрический тензор определяется для функции правдоподобия выражением:

$$g_{ik} = \int p_i(\theta) \frac{\partial \log p_i(\theta)}{\partial \theta_j} \frac{\partial \log p_i(\theta)}{\partial \theta_k} d\theta_j d\theta_k, \quad (2)$$

Здесь базисные вектора задаются касательными вдоль криволинейных параметров функции правдоподобия. С использованием взвешенного скалярного произведения. В касательном пространстве,  $T_g(\mathcal{P})$  при переходе к новым переменным  $y_i = \sqrt{p_i}$  квадрат расстояния между бесконечно близкими точками определяется формулой:

$$ds^2 = \sum_i (dy_i)^2 = \sum_i d\sqrt{p_i} d\sqrt{p_i} = \frac{1}{4} \sum_i \frac{\|\nabla p\|^2}{p_i} \quad (3)$$

Эта метрика не является Гильбертовой: числитель последнего члена в (2) образован векторами касательного пространства, но в знаменателе стоит точка самого многообразия.

Рассмотрим метрику Фишера Рао на сужении строго-положительных функций  $\mathcal{P}_0 = \{g \in \mathcal{P} \mid g > 0\}$ . Для  $v_1, v_2 \in T_g(\mathcal{P})$ , непараметрическая метрика Фишера-Рао определяется тогда как непрерывный аналог (3):

$$\left( (v_1, v_2) \right)_g = \int_0^1 v_1(t)v_2(t) \frac{1}{g(t)} dt \quad (4)$$

При замене  $q(t) = \sqrt{g(t)}$  pdf на полуплотность

$$\int_0^1 g(t) dt = 1 \rightarrow \int_0^1 q^2(t) dt = 1 \quad (5)$$

Таким образом,  $q(t)$  является элементом положительного ортанта  $S_\infty^+$ , бесконечномерной сферы  $S_\infty = \{f \in \mathbb{L}^2 \mid \|f\| = 1\}$ ,  $S_\infty \subset \mathbb{L}^2$ . Можно показать [7], что метрика (4) в  $\mathbb{L}^2$  метрику с точностью до константы при описанной замене переменной:

$$\begin{aligned} \left( (v_1, v_2) \right)_g &= \int_0^1 v_1(t)v_2(t) \frac{1}{g(t)} dt = 4 \int_0^1 \sqrt{g(t)} w_1(t) \sqrt{g(t)} w_2(t) \frac{1}{g(t)} dt \propto \\ &\propto \int_0^1 w_1(t) w_2(t) dt = \langle w_1 w_2 \rangle_Q, \quad v(t) = 2\sqrt{g(t)} w(t), w \in T_q(Q) \end{aligned} \quad (6)$$

Это обычная гильбертова метрика в  $T_q(Q)$  и мы можем теперь вычислять расстояние между двумя pdf как:

$$\left( w_1, w_2 \right) = \|w_1\| \|w_2\| \cos \varphi, \quad \varphi = \arccos \frac{\langle w_1, w_2 \rangle}{\|w_1\| \|w_2\|} \quad (7)$$

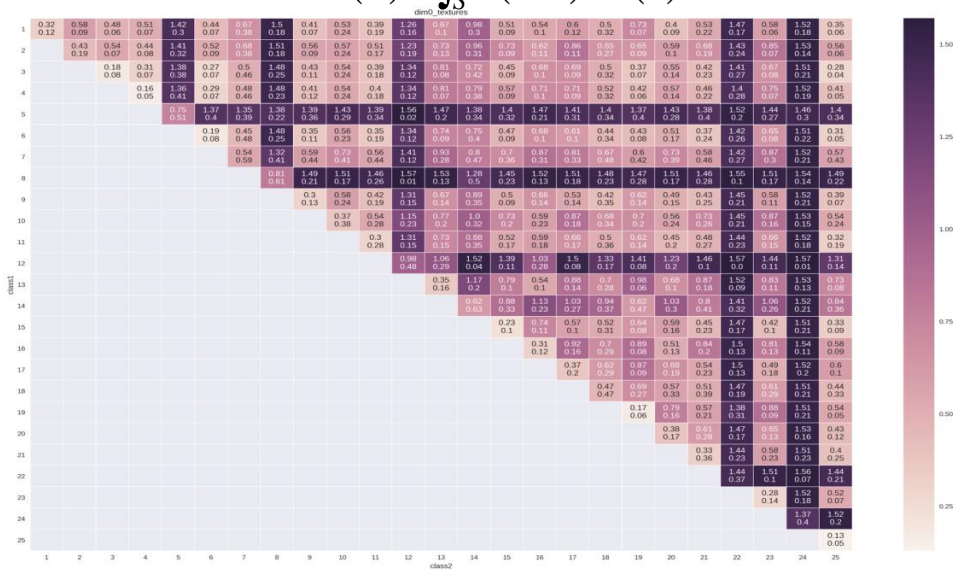
Заметим, что существуют определенные геометрические ограничения на оценку (7). Они связаны с условием локального диффеоморфизма некоторой окрестности нуля в касательном расслоении, на некоторую окрестность базы [7]. Так, максимальное расстояние, для которого справедливо экспоненциальное отображение реализующее диффеоморфизм соответствует углу  $\varphi = \pi/2 \approx 1.57$ .

**Численные эксперименты.** Мы вычислили диаграммы персистентности первых двух классов гомологий для всех 25 классов текстов стандартной базы HR-изображений [4,5]. Все диаграммы персистентности были векторизованы с помощью персистентных изображений [6] и пересчитаны как pdf на окно 100x100. Используя собственное программное обеспечение в среде Python, мы вычислили все попарные расстояния между pdf по формулам (7) для каждого из чисел Бетти. Результаты для ДП числа Бетти -0 приведены на рисунке 2.

Примером максимального расстояния является  $d(12;24) = d(8;12) = d(12;22) = 1.57$ . Минимальное расстояние между текстурами демонстрирует класс №25:  $d(25,25) = 0.13$ .

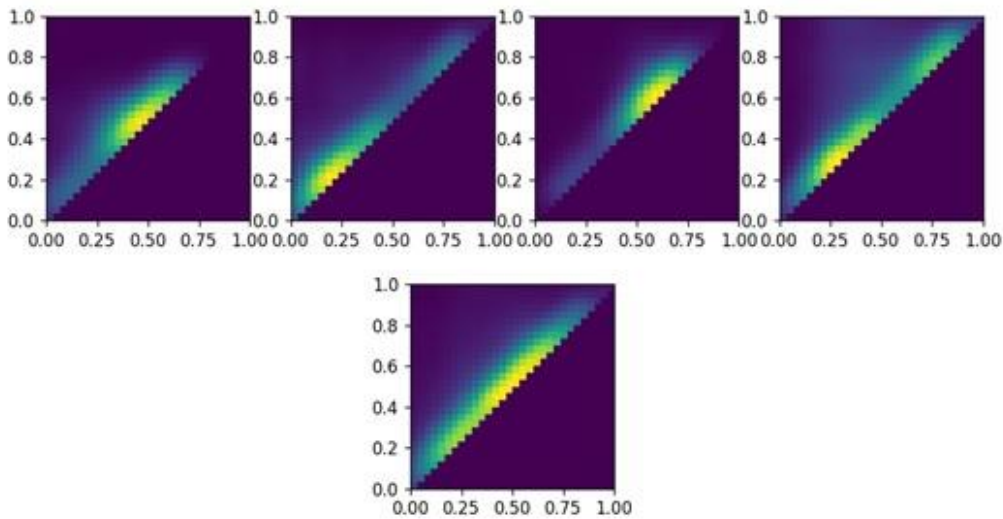
Для усреднения pdf удобно использовать среднее Карчера [7], которое является вариантом среднего по Фреше. Определим дисперсию Карчера  $\rho: S_N \rightarrow \mathbb{R}_{\geq 0}$  относительно точки  $P$  и плотности распределения вероятности  $f(q)$  точек  $q$  образующих референц меру:

$$\rho(p) = \int_S d(p, q)^2 f(q) dq \quad (8)$$



**Рисунок 2 - Среднее (арифметическое) значение (верхняя цифра в ячейке) и дисперсия (нижняя) Риманова расстояния между всеми парами диаграмм персистентности текстур всех классов для Бетти 0. В строках и столбцах представлены номера классов текстур согласно [4]**

Средним по Карчеру называют функцию  $\mu_{kar} = f(q)$ , которая доставляет локальный минимум дисперсии  $\rho(p)$ . На рисунке 3 приведена иллюстрация дисперсии относительно среднего по Карчеру по сравнению с произвольно выбранной точкой.



**Рисунок 3 - Пример усреднения по Карчеру 4-х pdf. Среднее показано на нижней панели**

Используя среднее по Карчеру, были вычислены средние Римановы расстояния от каждой текстуры класса до среднего, по Карчеру для класса. Результаты, которые мы не можем привести здесь из-за недостатка места, показывают, что использование римановой метрики позволяет уверенно различать текстуры.

*Работа выполнялась в рамках Гранта АР 05134227, МОН РК*

### **Литература**

1. Edelsbrunner H., Harer J., *Computational Topology, An Introduction*, Amer. Math. Soc. 2009. 241 p.
2. Ghrist R. W. *Elementary applied topology*. – Seattle : Createspace, 2014
3. Adler R., Taylor J. E. *Topological Complexity of Smooth Random Functions: École D'Été de Probabilités de Saint-Flour XXXIX-2009*. – Springer Science & Business Media, 2011.
4. Макаренко Н.Г. и др. Распознавание текстур на цифровых изображениях методами вычислительной топологии //Современные проблемы дистанционного зондирования Земли из космоса. – 2015. – Т. 12. – №. 1. – С. 131.
5. Makarenko, N., Kalimoldayev, M., Pak, I., & Yessenaliyeva, A. (2016). Texture recognition by the methods of topological data analysis. *Open Engineering*, 6(1).
6. Berry, E., Chen, Y. C., Cisewski-Kehe, J., & Fasy, B. T. (2018). Functional Summaries of Persistence Diagrams. *arXiv preprint arXiv:1804.01618*.
7. Srivastava A., Klassen E. P. *Functional and shape data analysis*. – New York : Springer, 2016.

## **СИСТЕМА УПРАВЛЕНИЯ БИОНИЧЕСКОЙ РУКОЙ С ПОМОЩЬЮ ЭМГ ДАТЧИКОВ**

**Мамырбаев О.Ж., Сундетов Т., Жакупбеков Т.Е.**

*РГП "Институт информационных и вычислительных технологий"  
КН МОН РК*

**Аннотация.** *Одна часть человеческого тела, в которой очень полезная в этой жизни рука. Эта жизнь будет нарушена, если одна или обе руки не будут работать из-за инвалидности или ампутации. Одним из вкладов, которые помогают этому условию, является бионическая рука. В этом исследовании была разработана бионическая рука, состоящая из шести сервомоторов для движения пальцев. Датчик ЭМГ используется в этой бионической руке, он расположен в предплечье. Он использовался для записи сигналов мышц. Выделение мышечного сигнала было сформулировано с использованием среднеквадратичного значения среднеквадратичного значения. Вид бионного движения руки был найден путем идентификации картины сигнала движения с использованием обратного распространения через искусственную нейронную сеть. Выходные данные, полученные от обратного распространения, обрабатывались микроконтроллером и передавались сервомоторам для получения движения бионической руки. Исходя из*



экспериментальных данных, в этой системе было проверено пять видов движения, среднее значение распознавания для каждого движения.

**Ключевые слова:** бионическая рука, датчик ЭМГ, микроконтроллер, сервомотор.

## Введение

Наряду с развитием робототехнических технологий в области роботов, много исследований сделали, чтобы помочь людям в повседневной жизни. Одно из этих исследований в настоящее время разрабатывается как бионический робот. Бионический робот - это робот, который имплантируется людям, чтобы заменить функцию и роль отсутствующих частей человеческого тела, таких как руки или ноги. Части тела человека очень важны для людей, чтобы выполнять повседневную деятельность, одна из которых - рука. Потеря руки из-за несчастного случая или паралича вызовет у людей трудности с выполнением их повседневной деятельности. Например, человеку, потерявшему руку, будет сложно выполнять задачи, связанные с функцией человеческой руки. Исследователи пытаются разработать бионический робот для решения вышеупомянутых проблем. Бионическая рука будет контролироваться через миоэлектрические сигналы, генерируемые мышцами человека. В принципе, у человеческого тела есть электричество. Хотя мышцы имеют электрические сигналы, но сила результирующего напряжения очень мала. Небольшой электрический ток может быть подключен с помощью набора электронных устройств (датчиков). В этом исследовании датчик, ограничен только датчиком ЭМГ (электромиограмма), использовался для получения сигналов ЭМГ, генерируемых движением мышц на руке человека. Данные считываются ЭМГ, а затем отправляются через Bluetooth на персональный компьютер (ПК) для обработки, а затем отправляются протезу, который оснащен специальным микроконтроллером. Полученные данные будут обработаны, а затем применены в виде команд для управления шестью сервоприводами, встроенными в механику роботоманипуляторов. Все вышеперечисленные процессы создадут новую функцию, похожую на реальные человеческие руки.

## 1. Методика

Проектирование всей системы для этого исследования показано в Рисунок 1.

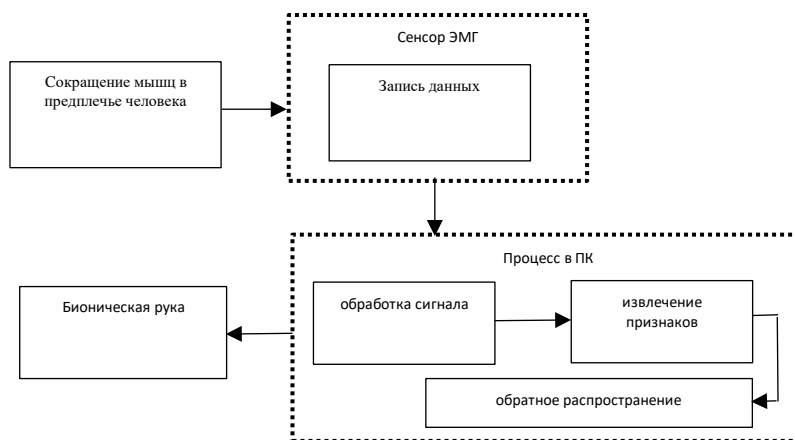


Рис. 1. Схема проектирования системы

В этом исследовании первое, что нужно сделать, это получить сигналы от мышечные сокращения в руках человека с использованием датчика ЭМГ.

Размещение датчика не должно быть произвольным, поскольку оно имеет восемь электродов, которые будут использоваться для записи активности нескольких мышц, расположенных на предплечье человека. Эти датчики записывают данные сокращения мышц, а затем данные обрабатываются на ПК для извлечения функций, которые будут установлены как шаблоны и классифицируются методом обратного распространения. Из этой классификации, он будет генерировать тип движения, которое уже распознано из входного сигнала, затем принимает решение и отправит его роботу-биону, чтобы выполнить то же движение руками пользователя.

## 2. Обработка сигналов

В этом исследовании протестирован входной сигнал в систему по-прежнему за короткое время (50 мс). Сигналы все еще остаются во временном домене, потому что, если использовать предварительную обработку, такую как преобразование Фурье, чтобы создать лучшую функцию сигнала, оказалось, что медленный отклик на производительность системы должен быть близок к реальному времени.

Хотя движения нет, человеческие мышцы все еще излучают сигналы, регистрируемые датчиком ЭМГ [1], это видно из сохранения величины амплитуды электрической мышцы, хотя и очень малой. В этом исследовании сигналы называются шумами, поэтому он должен различать активный сигнал и шум. Активный сигнал получается из среднего количества незанятых данных, которые затем вычисляются, чтобы найти значение стандартного отклонения для получения максимального значения незанятых данных. Если данные больше, чем максимальные данные бездействия, то данные являются активным сигналом [2]. Значение стандартного отклонения и максимальных холостых данных может быть рассчитано с использованием уравнений 1 и 2.

$$SD = \sqrt{\frac{\sum_{i=0}^N (X_i - \mu)^2}{N}} \quad (1)$$

где:

SD = стандартное отклонение

N = данные выборки

X = необработанные данные простоя

$\mu$  = среднее значение простаивающих данных

$$\text{Max}_{\text{idle data}} = X_{\text{idle data}} + SD \quad (2)$$

где:

Max = максимальное значение незадействованных данных

X = среднее количество незанятых данных

SD = стандартное отклонение

## 3. Отвлечение внимания

Необходимо выполнить выделение функций, чтобы получить характеристики входного сигнала для каждого движения. В этом исследовании исследователи внедряют СК-метод для получения характеристик сигнала в каждом движении. СК (средний квадрат) является «квадратным, а затем усредненным [3], а затем квадратным корнем». СК выполняется для определения характерных признаков каждого

движения. В этом исследовании характеристики, используемые для идентификации каждого движения, представляют собой среднеквадратичное значение каждого ЭМГ [4]. Значение СК каждого канала может быть рассчитано с использованием уравнения 3.

$$СК = \sqrt{\frac{\sum_0^n Xn^2}{N}} \quad (3)$$

где:

СК = значение СК (среднеквадратическое значение)

Xn = n-ые данные

N = количество сигналов

Если СК-значение каждого канала найдено, то следующим шагом будет поиск минимального и максимального значений СК-значения [5]. Минимальные и максимальные значения используются для нормализации данных ко всем каналам с диапазоном значений между 0 и 1. Нормализованные данные могут быть рассчитаны с использованием уравнения 4.

$$Dn = \frac{X - \text{dataMin}}{2\text{dataMax} - \text{dataMin}} \quad (4)$$

где :

Dn = нормализованные данные

X = значение СК

dataMin = минимальные данные

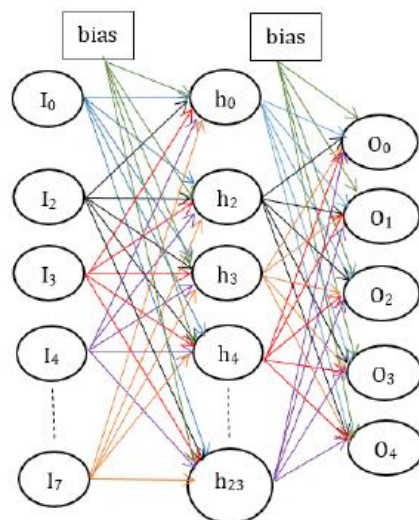
dataMax = максимальные данные

#### 4. Обратное распространение

Метод, используемый в этом исследовании для распознавания шаблонов сигналов, - это обратное распространение. Обратное распространение является одной из искусственных нейронных сетей с контролируемым обучением. Обратное распространение имеет три уровня: слой ввода, скрытый слой и выходной слой. В этом исследовании метод обратного распространения выбирается потому, что он имеет прямое и обратное обучение, поэтому более точно распознавать шаблоны сигналов и классифицировать их как тип движения. Ввод от обратного распространения по этому сигналу является нормированным СК-данным каждого канала [6]. Таким образом, сетевая архитектура состоит из восьми входных узлов, 24 скрытых узлов и 5 выходных узлов. Полученный результат отправляется на микроконтроллер на бионическую руку. Ожидается создание сети, которая может быть использована для классификации 5 движений [7]. Архитектурный дизайн для обратного распространения этой системы показан на рисунке 2.

В этом исследовании данные обучения, используемые для обучения по методу обратной пропаганды, все еще выполняются на этапе испытаний только одним пользователем, принимая 10-кратные данные каждого движения. Поскольку только один пользователь, это исследование не провело проверку системы, которая построена, например, с помощью набора данных. Затем данные, используемые для обучения системы, составляли до 50 данных для каждого движения, каждый раз в 10 раз превышал данные.

Чтобы получить самый прекрасный архитектурный дизайн, исследователи проверили значения параметра, изменив количество параметров скрытого слоя, которые необходимо приравнять, а именно скорость обучения и ошибки.



**Рис. 2. Структура обратной связи структуры**

В этом исследовании, отмеченном изменением параметров скрытого слоя, с частотой обучения на 0,6 и ошибкой равна 0,00001. Во время учебного процесса в нескольких скрытых слоях 8 количество итераций мало, но есть достаточно высокая погрешность. В то время учебный процесс по ряду скрытых 16 обнаружил число итераций для больше, но все же есть ошибка, почти равная числу скрытого 8. Между тем, во время учебного процесса на количество скрытого слоя 24 получило количество итераций довольно много по сравнению скрытых 8 и 16, но результирующая ошибка значение мало. Поэтому, основываясь на тестировании, которое было выполнено по числу спрятанного использования этой системы требует высокой степени точности так что исследователи решили использовать номер 24, потому что это имеет высокую скрытую небольшую ошибку.

## 5. Результат

Фаза тестирования движения бионического руки - конечный результат вышеупомянутых шагов. На этом тесте протез уже интегрирован с ПК, который используется для обработки сигналов, которые были распознаны путем обратного распространения. На этом этапе робот-манипулятор получает данные, которые были обработаны на ПК и распознаны обратным распространением в качестве желаемого шаблона движения. Из результатов СК каждого движения, соответственно, получается совершенно разные модели. Эти шаблоны различают каждое движение и будут использоваться как вход для распространения и классификации назад. Результаты испытаний показали, что точность распознавания образов сжимания составляет около 65%, распознавание рисунка составляет около 75%, распознавание образов волны в около 100%, распознавание образов волны составляет около 75%, и распознавание рисунка распространения составляет около 100%. Общий процент полученной ошибки составляет около 17%.

## Список литературы

1. Setiawan, Cahyo, "Tangan Robot Bionik dengan Kendali Otot Berbasis EMG" // Computer Engineering, Electronic Engineering Polytechnic Institute of Surabaya, Surabaya, Indonesia, 2015.

2. Malik, Iqbal Maulana, "Sistem Pengenalan Pola Gerak Tangan Untuk Persentase Interaktif Dengan Sensor EMG (Electromyogram)" // Computer Engineering, Electronic Engineering Polytechnic Institute of Surabaya, Surabaya, Indonesia, 2015.
3. Suresh.M, et al, "Electromyography Analysis for Person Identification" // Kalpataru Institute of Technology, India, 2011.
4. Nomiyaari, "Perancangan dan Pembuatan Modul ECG dan EMG dalam satu unit PC" // Electronic Engineering, Electronic Engineering Polytechnic Institute of Surabaya, Surabaya, Indonesia, 2011.
5. B. Peerdeman, D. Boere, H. Witteveen, R.H. in t Veld, H. Hermens, S. Stramigioli, et al., "Myoelectric forearm prostheses: State of the art from a user-centered perspective" // JRRD, vol. 48, no. 6, pp. 719-738, 2011.
6. A.H. Al-Timemy, G. Bugmann, J. Escudero and N. Outram, "Classification of finger movements for the dexterous hand prosthesis control with surface electromyography" // IEEE J. Biomed. Health Informatics, vol. 17, no. 3, pp. 608-618, 2013.
7. C. Miller, Real-Time Feature Extraction and Classification of Prehensile // EMG Signals, 2008

## **АНАЛИЗ ЗАЩИТЫ ДОСТУПА СВЯЗИ РОБОТОТЕХНИЧЕСКИХ СИСТЕМ**

**Марденов Е.М.**

*Институт информационных и вычислительных технологий КН МОН РК  
e-mail: [ivideooperator@mail.ru](mailto:ivideooperator@mail.ru)*

*В современной робототехнике одним из важнейших стало дистанционное управление робототехническими системами оператором или запрограммированными командами с сервера управления. Дистанционное управление является одним из самых уязвимых элементов в робототехнической системе, к которым относятся перехват, просмотр, подмена передаваемой управляющей команды. Целью данной статьи является анализ и обзор публикаций, охватывающих тематику «повышение надежности защиты беспроводных систем безопасности».*

В настоящее время, как в Казахстане, так и за рубежом ведутся активные работы по созданию мобильных робототехнических комплексов (МРК).

Одним из главных являются задачи, в ходе которых мобильный робот действует в условиях, опасных и труднодоступных для нахождения человека. Перспективным является использование автоматизированных робототехнических комплексов в боевых условиях, когда имеется прямая угроза жизни оператора.

Построение автономных роботов является достаточно сложной задачей. Дистанционное управление осуществляется за счет беспроводной системы связи, которая является одним из наиболее уязвимых элементов робототехнического комплекса, подверженных различным угрозам. В качестве преднамеренных дестабилизирующих факторов для радиоканала в целом можно выделить следующие виды: перехват, просмотр, подмена, радиоэлектронное подавление. Следует заметить, что данные дестабилизирующие факторы могут применяться не только поодиночке,

но и в комплексе. Комплексные угрозы, воздействующие на беспроводную систему безопасности одновременно, могут серьезно дестабилизировать ее работу.

Проанализировав материалы публикации, защиты доступа управления связи робототехнических систем можно разделить на следующие типы

- криптографические;
- широкополосные сигналы;
- изменение параметров радиосигнала (частота, мощность);
- смешанные методы.

Проведем анализ отмеченных методов с точки зрения целесообразности их использования при управлении робототехническими комплексами специального назначения.

Рассмотрим криптографические методы защиты информации.

Общепризнанное преимущество симметричных систем заключается в более высокой скорости шифрования, в меньших размерах используемого ключа, обеспечивающего стойкость, и в более высокой степени научной обоснованности криптографической стойкости. Однако асимметричные криптосистемы позволяют реализовать протоколы, заметно расширяющие область применения криптографической защиты информации. Поэтому нередко используются гибридные криптосистемы, сочетающие принципиальные преимущества как симметричных, так и асимметричных криптосистем.

Наиболее выгодным представляется протокол секретной связи с использованием гибридной криптосистемы, в котором асимметричный алгоритм используется для засекречивания и распределения ключей связи, а алгоритм с секретным ключом связи используется для защиты данных. Кроме того, такой протокол допускает уничтожение секретного сеансового ключа сразу после завершения сеанса, что существенно снижает опасность его компрометации.

Для компенсации недостатков симметричного шифрования в настоящее время широко применяется комбинированная (гибридная) криптографическая схема, где с помощью асимметричного шифрования передается сеансовый ключ, используемый сторонами для обмена данными с помощью симметричного шифрования. Важным свойством симметричных шифров является невозможность их использования для подтверждения авторства, так как ключ известен каждой стороне.

В работе [3] робототехнический комплекс, предназначенный для дистанционной работы в условиях боевых действий, а также в труднодоступных и опасных для присутствия человека местах. Технический результат - сбор и передача разведывательной информации, охрана или патрулирование гражданских и военных объектов, проведение антитеррористических операций в городских и полевых условиях, ведение стрельбы по различным видам целей в дневных и ночных условиях. Представлена информационно-вычислительно система связана со всеми системами и подсистемами РТП и обеспечивает решение информационных и расчетных задач, обработку информации от ПДУ, телеметрической информации и видеоизображения, формирование информационных команд, управление аппаратными средствами РТП, дистанционное и полуавтономное управление движением РТП, запоминание маршрута движения, возвращение РТП в исходную точку по пройденному маршруту, сжатие информации и ее криптографическую защиту при передаче по каналу связи, решение топогеодезических задач, управление датчиками и приводами. БК 25 работает с ПДУ 8, БК 26 - с ПДУ 8 или АОПР 9. Распределение информационных потоков в основном происходит в ММП 27.

В работе [4], управление осуществляется через беспроводную сеть, сетевые помехи или помехи извне и предотвращение запуска робота (фальсификация командных команд, чтение команд команды). Препятствие обеспечивается сквозным шифрованием (E2EE). Особенностью переходного шифрования являются пароли обмена на стороне клиента и сервера, обмен информацией и риск кражи информации, такие как «люди в окружающей среде», потому что расшифрованная информация может занять много времени, не зная общего ключевого слова возможно.

Рассмотрим использование широкополосных сигналов (ШПС). В работе [5] исследуется задача удаленного управления наземными роботами в рамках ФЦП «Исследования и разработки по приоритетным направлениям развития научно-технического комплекса России на 2007–2013 гг.». В результате исследований было установлено, что использование универсальных радиомодемов УКВ-диапазона для передачи данных с использованием сигнально-кодовых конструкций на основе OFDM-модуляции, работающих в диапазоне частот 100- 500 МГц, обеспечивает надежную защищенную передачу информации в экстремальных условиях (многолучевое распространение сигнала, эффект Доплера, замирание сигнала) с возможностью увеличения зоны покрытия за счет ретрансляторов сигнала. Однако, влияние преднамеренных помех и НСД к радиоканалу пока исследовано не было. Известен мобильный робот-разведчик «Инженер», который обладает стереозрением и может поднимать видеокамеру на уровень 130 см. Данный робот использует псевдослучайную перестройку частоты (ППРЧ) радиоканала для защиты сжатых видеоданных от перехвата и подавления преднамеренными помехами. В настоящее время данный робот проходит полевые испытания.

В работе [6] группа изобретений относится к технике предотвращения недозволенного использования или кражи транспортных средств. Предполагается наличие системы дистанционного управления, включающей носимое передающее устройство и приемное устройство, установленное на объекте дистанционного управления. Во время передачи радиосигнала носимого передающего устройства, который состоит из нескольких последовательных отдельных кодовых посылок, с помощью приемопередающего устройства, дополнительно установленного на объекте дистанционного управления, принимают первую кодовую посылку. Во время передачи первой и каждой последующей кодовых посылок носимого передающего устройства излучают помеху с длительностью, меньшей, чем четверть длительности кодовой посылки. При этом, по меньшей мере, во время передачи двух кодовых посылок помехи излучают таким образом, чтобы они перекрывали различные участки кодовых посылок. В течение приема кодовых посылок выявляют и восстанавливают код кодовой посылки носимого устройства. После окончания передачи радиосигнала от носимого передающего устройства с помощью дополнительного приемопередающего устройства передают радиосигнал, содержащий наряду с ложными кодовыми посылками, по меньшей мере, одну кодовую посылку с выявленным и восстановленным кодом с мощностью, меньшей мощности радиосигнала носимого передающего устройства. Группа изобретений позволяет снизить риск перехвата и расшифровки сигнала между носимым передатчиком и приемником, установленном на защищаемом объекте, в частности, на транспортном средстве.

ШПС обладают следующими общими преимуществами: высокой помехоустойчивостью и защищенностью от НСД, возможностью борьбы с многолучевостью и интерференцией сигнала. К общим недостаткам систем передачи информации с ШПС можно отнести сложность приемно-передающей аппаратуры,

необходимость наличия устойчивой синхронизации, малый объем используемых последовательностей, а также наличие внутрисистемных помех.

Рассмотрим использование изменения параметров радиосигнала (частота, мощность) как метода защиты беспроводного канала от НСД.

В работе [7] описывается система управления боевой роботизированной платформой. Указывается, что благодаря данному способу, в том числе, возможно, перестраивать параметры радиоканала (частота, режим канала), мощность излучения и параметры маскировки данных. Перестройка параметров радиоканала и мощности излучения без дополнительных средств защиты не является эффективным методом защиты от НСД и помех.

Рассмотрим смешанные методы защиты. Известен подход для удаленного управления роботом через беспроводное подключение 3G с помощью специально разработанного программного обеспечения, которое позволяет снизить количество передаваемой информации для устойчивости канала связи с роботом. В качестве метода защиты применяются криптография и широкополосные сигналы.

В работе [8] предлагается для повышения защищенности радиоканала робототехнических средств использовать мобильные ретрансляторные центры (МРЦ). МРЦ позволяют создавать защищенные VPN соединения с использованием ресурсов спутниковой связи, систем сотовой связи (GPRS/ GSM, 3G) и систем беспроводной передачи данных IEEE 802.16. В качестве метода защиты применяются VPN, криптография и широкополосные сигналы.

Недостатки для криптографических и VPN и методов применения широкополосных сигналов следует добавить изначальное предназначение сетей IEEE 802.16, GSM, 3G для решения задач информационного обмена широкого круга пользователей, что может вызывать сложности их применения в специальных целях.

Проанализировав можно увидеть, что широкое распространение в качестве каналов беспроводной связи наземных роботов специального назначения получили стандарты передачи информации IEEE 802.15.1, IEEE 802.15.4, IEEE 802.11, IEEE 802.16, GSM, 3G. Встроенные системы защиты информации (криптография, широкополосные сигналы), используемые в упомянутых стандартах, хорошо описаны и исследованы, но в специальных целях их применение неэффективно, поскольку связано с появлением существенных недостатков.

Для повышения защищенности информационного обмена в роботизированных комплексах необходима разработка принципиально новых систем связи оператора (сервера) с роботом, обеспечивающих скрытые и помехозащищенные режимы приема и передачи сигналов.

Применение криптографии, как способа защиты, обеспечивает защиту от просмотра и подмены трафика, однако данный метод неэффективен при перехвате информации и подавлении информационного канала преднамеренными помехами.

Эффективным методом, обеспечивающим одновременно защиту от просмотра, подмены, перехвата и наложения помех, выступают широкополосные сигналы. В источнике описано использование широкополосных сигналов для систем связи наземных роботов стран-участниц блока НАТО. Данные системы связи должны отвечать следующим требованиям:

- широкая полоса пропускания;
- скрытность;
- помехозащищенность;
- противодействие несанкционированному доступу.



Рассмотренные выше методы защиты каналов беспроводной связи на основе применения широкополосных сигналов могут быть применены, однако они обладают следующими недостатками: незначительное количество последовательностей, отсутствие возможности периодической смены используемых сигналов.

По мнению авторов, перспективным направлением в защите беспроводного канала управления от НСД является использование хаотических широкополосных сигналов. Предлагаемый подход основан на использовании перезаписываемых накопителей хаотических последовательностей, позволяющих использовать уникальные наборы хаотических сигналов. Данный подход позволит значительно увеличить защищенность беспроводного канала управления робототехническими комплексами специального назначения, как от несанкционированного доступа, так и от воздействия помех.

### **Выводы**

Проанализировав материалы защиты управления роботов от несанкционированного доступа связи, можно сделать выводы, что:

- распространение в управлении робототехническими системами специального назначения получили каналы беспроводной связи стандартов IEEE 802.15.1, IEEE 802.15.4, IEEE 802.11, IEEE 802.16, GSM, 3G, которые сравнительно легко подавляются средствами радиоэлектронной борьбы;

- ШПС является эффективным методом одновременной защиты как от НСД, так и от помех;

- основными недостатками ШПС являются незначительное количество последовательностей и отсутствие возможности периодической смены используемых сигналов;

- в качестве перспективного направления, позволяющего уменьшить недостатки систем с ШПС, предлагается использование перезаписываемых накопителей хаотических последовательностей, позволяющих использовать уникальные наборы хаотических сигналов.

### **Литература**

1. Левчук В. С. Критерии надежности беспроводных охранно-пожарных систем // Грани безопасности. 2007. № 3 (45). С. 42–43.

2. Успенский, Александр Юрьевич, Защита информации в радиоканалах мобильных робототехнических комплексов

3. Громов Владимир Вячеславович (RU), Липсман Давид Лазорович (RU). роботизированная транспортная платформа публикация патента: 10.02.2014

4. Т. У. Исламгожаев А. И. Елеусинов, Ш. С. Мажитов, С. К. Джолдасбаев, Ж. А. Бурибаев. Мобильді роботтың басқару командаларының қауіпсіздігін қамтамасыз етуге позициялық емес полиномды санау жүйесіне негізделген шифрлеу алгоритімін қолдану.

5. Гугалов К.Г. Оборудование беспроводной передачи информации для наземных робототехнических комплексов // Технологии и средства связи. 2013. № 3. С. 14–15.

6. Буслаев Владимир Михайлович (RU) В60R25/00 Оборудование транспортных средств для предотвращения или обнаружения недозволенного использования или кражи транспортных средств

7. Патент РФ 2544740. Система управления боевой роботизированной платформой / Громов В.В., Липсман Д.Л., Мосалев С.М., Рыбкин И.С., Сеницын Д.И., Фуфаев Д.А., Хитров В.А. Заявл. 26.11.2013. Опубл. 20.03.2015. Бюл. № 8. 10 с.

8. Цариченко С.Г. Экстремальная робототехника в МЧС России – задачи и перспективы // Bezpieczeństwo i Technika Pożarnicza. 2012. № 4(28). С. 97–105.

9. Жук Александр Павлович, Анализ методов защиты от несанкционированного доступа беспроводных каналов связи робототехнических систем

## **СИСТЕМА МОНИТОРИНГА И ПРОВЕРКИ ДОСТОВЕРНОСТИ ПУБЛИКАЦИЙ В СМИ**

**Мусабаев Р.Р., Красовицкий А.М., Уалиева И.М., Касымжанов Б.К., Ибраева В.М., Мейрамбеккызы Ж., Тасболатулы Н., Мукашев А.Ш.**

*Институт информационных и вычислительных технологий КН МОН РК,  
Казахстан  
Лаборатория анализа и моделирования информационных процессов  
e-mail: rustam@ipic.kz*

***Аннотация.** В статье рассматривается методология разработки системы мониторинга и достоверности публикаций в средствах массовой информации (СМИ). Предлагается методика расчета степени достоверности отдельно взятой публикации, на основе которой рассчитывается степень достоверности публикации по автору, конкретному СМИ и степень достоверности СМИ в целом. Показаны информативные признаки, используемые для предварительной классификации и оценки степени недостоверности. Описан прототип системы автоматической параметризации, классификации и оценивания недостоверности. Описаны будущие шаги для разработки информационной системы мониторинга и проверки достоверности публикаций.*

***Ключевые слова:** средства массовой информации, система автоматизированного мониторинга СМИ, достоверность публикации, информативные признаки*

### **Введение**

Проблема отношения и освещения средствами массовой информации работы властных структур и жизни общества в РК имеет особую значимость, так как именно СМИ оказывают существенное влияние на формирование общественного мнения. Фактически СМИ приобрели функцию «четвёртой власти», но в погоне за рейтингами и в угоду читательской аудитории, не все журналисты/редакторы СМИ ответственно используют имеющиеся у них возможности формирования общественного мнения. К типичным злоупотреблениям со стороны СМИ относится публикация недостоверной информации, сознательное манипулирование общественным мнением в угоду определённым политическим или экономическим интересам, направленная дискредитация целевых субъектов и т.д. Задача разрабатываемой системы состоит в автоматическом выявлении недостоверных публикаций. Авторы предлагают общий

подход к выявлению недостоверных публикаций в СМИ, основанный на расчете информативных признаков.

### **Информативные признаки для автоматического выявления недостоверных публикаций**

По умолчанию любая публикация является достоверной, пока не доказано, что она недостоверна. А. А. Морозова в своей статье [1] выделяет основные причины недостоверности информации в социальных сетях, в частности «ВКонтакте», из них можно выделить общие как для официальных средств массовой информации, так и для социальных сетей, критерии оценок:

- преднамеренная гиперболизация информации;
- преднамеренное искажение информации или публикация заведомо ложной информации для привлечения внимания;
- умалчивание фактов, неполнота информации в преследовании идеологических целей;
- случайное искажение информации при ее передаче от первоисточника к транслятору;
- ненамеренная неверная интерпретация фактов и явлений из-за сложности информации или некомпетентности автора.

Также отсутствие аффилиации к первоисточникам или анонимность автора – все эти факторы также снижают степень достоверности публикации.

К достоверным критериям можно отнести:

- источник информации, его авторитетность;
- компетентность и неанонимность автора статьи;
- имеющиеся фактические доказательства информации, переданной автором;
- представленная полнота информации и наличие фактических данных;
- отсутствие излишней эмоциональности, упор на факты автором публикации.

В любом случае, достоверность отдельно взятой публикации становится основой доверия к автору публикации, к изданию, ее опубликовавшему, и в целом к средствам массовой информации нашей страны.

Исходя из причин недостоверности информации и критериев ее достоверности, мы предлагаем набор следующих информативных признаков для автоматического определения достоверности публикации:

- степень позитивной направленности публикации;
- степень негативной направленности публикации;
- степень генерализации;
- степень политизированности;
- степень конкретизации;
- степень субъективности (степень выраженности мнения автора по предмету публикации);
- степень ангажированности (предвзятости, пристрастности);
- степень актуальности (степень актуальности информации на сегодняшний день);
- степень оригинальности (уникальности);
- степень привлекательности публикации (насколько публикация может способна привлечь к себе внимание потенциальных читателей);
- степень казахстанского содержания;

- степень провокационности (манипулятивности);
- степень связанности с другими публикациями по обсуждаемым вопросам;
- степень противоречия другим публикациям;
- степень плагиаризма;
- степень соответствия названия содержанию;
- наличие ссылок на первоисточники;
- наличие имени автора;
- реально ли существует человек, написавший данную публикацию;
- наличие поздних корректировок.

Очевидно, что формирование оптимального набора информативных признаков, обеспечивающих наилучшее качество классификации, требует дальнейших исследований. Предполагается, что будут разработаны алгоритмы ранжирования и валидности информативных признаков, значимых в процессе классификации и имеющих семантическую интерпретацию, на основе экспертных оценок, так и с использованием математического аппарата.

Введем модель численной оценки достоверности отдельно взятой публикации или степень ее достоверности. Степень достоверности отдельно взятой публикации  $P_i$  – это численная оценка, которая показывает то, насколько публикация достоверна или в какой степени можно доверять данной публикации. Степень достоверности  $P_i$  публикации является базовым показателем для последующего расчета степени достоверности публикаций по отдельному журналисту AFF (Author False Factor), степени достоверности конкретного печатного СМИ MFF (Media False Factor) и степени достоверности всех печатных СМИ в целом в рамках страны MMFF (Mass Media False Factor):

$$AFF = \frac{1}{N^a} \sum_{i=1}^{N^a} P_i^a \quad (1)$$

$$MFF = \frac{1}{N^m} \sum_{i=1}^{N^m} P_i^m \quad (2)$$

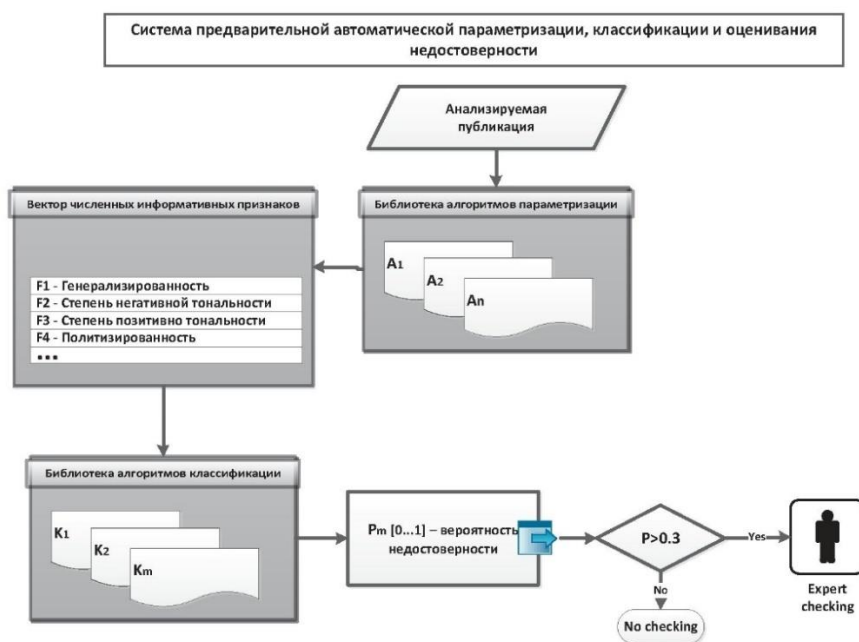
$$MMFF = \frac{1}{M} \sum_{i=1}^M MFF_i \quad (3)$$

где  $P_i^a, P_i^m$  – соответствующие степени достоверности публикаций отдельного автора и отдельного СМИ;  $N^a$  – количество публикаций автора  $a$ ;  $N^m$  – количество публикаций печатного  $m$ -го СМИ;  $M$  – количество печатных СМИ.

### **Описание прототипа системы автоматической классификации публикаций на основе набора информативных признаков**

Рассмотренные выше признаки войдут в библиотеки алгоритмов классификации в системе автоматической параметризации, классификации и оценивания недостоверности, далее система (рис. 1).

Задача системы – предварительная автоматическая классификация отдельных публикаций на основе набора информативных признаков с целью формирования множества подозрительных публикаций, требующих экспертного анализа.



**Рис. 1 Система автоматической параметризации, классификации и оценивания достоверности**

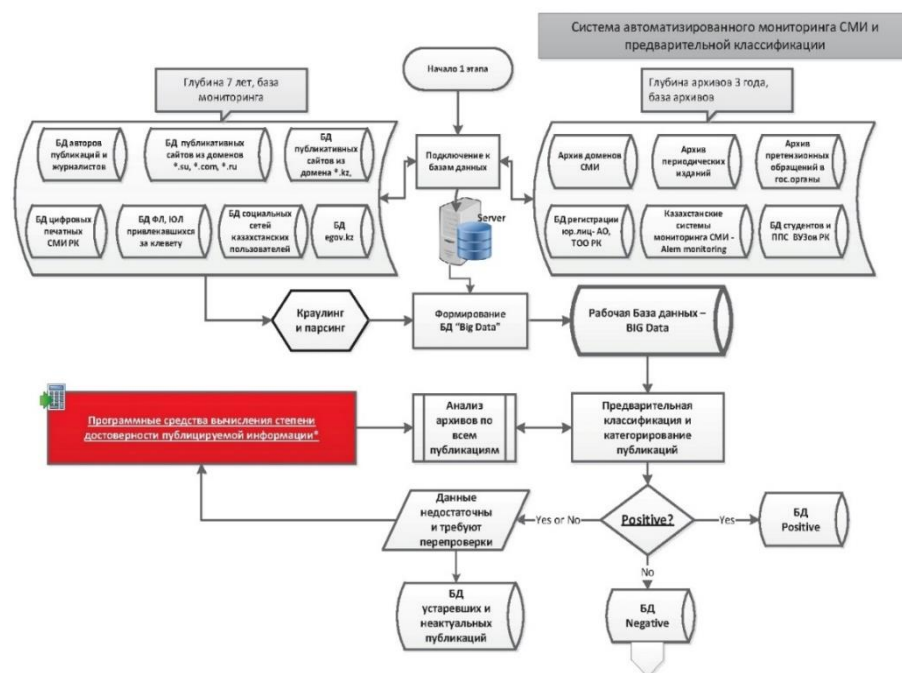
Система включает библиотеки алгоритмов параметризации и алгоритмов классификации. Задача подсистемы «Библиотека алгоритмов параметризации» состоит в обеспечении первоначальной обработки анализируемой публикации: нормализации, параметризации и анализа текстовых данных. Задача подсистемы «Библиотека алгоритмов классификации» состоит в автоматической классификации анализируемых публикаций для дальнейшего экспертного анализа [2]. В случае, если система распознала публикацию как публикацию с низкой степенью достоверности, то публикация подвергается дальнейшей экспертной оценке. Предполагается, что на основе алгоритмов вычисления информативных признаков будет разработана еще одна подсистема – «Библиотека признаков».

### **Общая схема системы автоматизированного мониторинга СМИ в сети интернет, формирования полнотекстовой базы данных и ее классификации**

По нашей программе ПЦФ предлагаемый подход к выявлению недостоверных публикаций в СМИ, основанный на расчете информативных признаков, мы будем апробировать на большом текстовом корпусе объемом не менее 30 млн словоупотреблений, глубиной в семь лет официальных средств массовой информации РК, в котором будут храниться статьи, собранные из открытых источников. Общая система автоматизированного мониторинга СМИ будет включать следующие функциональные подсистемы (рис. 2):

- подсистема сбора, обработки данных, предназначение которой – обзор источников данных (официальные, полуофициальные средства массовой информации, публикации, статьи и т.д.), сбор данных из систем источников глубиной 7 лет, начиная с 2011 года. В базу мониторинга предлагается включить базы данных авторов публикаций и журналистов, базы данных публикативных сайтов, цифровых печатных средств массовой информации РК, базы данных юридических лиц, физических лиц,

привлекавшихся за клевету, базы данных социальных сетей казахстанских пользователей и базы egov.kz, краулинг и парсинг;



**Рис. 2** Схема системы автоматизированного мониторинга СМИ в сети интернет, формирования полнотекстовой базы данных и предварительной классификации тональности текста

- подсистема хранения данных в базах данных Big Data. Основная база данных будет содержать «сырые» данные с частичным форматированием, такие как: ссылка на источник, автор публикации, название статьи, ключевые слова, абстракт, саму статью и т.д. Рабочая база данных будет содержать упорядоченные данные, которые будут необходимы для предварительной оценки информативных признаков достоверности (положительная тональность, негативная тональность, генерализация и т.д.);
- подсистема обработки и анализа средств массовой информации, предназначение которой вычисление степени достоверности публикаций по отдельному журналисту, конкретного печатного СМИ, всех печатных СМИ, основанное на информативных признаках. Подсистема включает программные средства вычисления тональности текста и размещения оценок в базы данных позитивных и негативных публикаций, программные средства расчета степени генерализации, вычисления степени фальсифицированности публикаций, соответствия публикации авторству и построения «родственных связей» - аффилированности между авторами, сайтами, аккаунтами и средствами массовой информации и т.д., согласно списка информативных признаков. На рисунке 3 показаны подсистема сбора, обработки данных и модуль вычисления тональности текста подсистемы обработки и анализа;
- подсистема визуализации данных, задача которой формирование отчетности по каждому из признаков.

## **Заключение**

На данный момент авторами решены следующие задачи:

- Выявлены основные причины недостоверности информации в публикациях;
- Выявлены информативные признаки для автоматического определения достоверности публикации;
  - Разработана методология и алгоритмы оценки степени генерализации, позитивной тональности текста, негативной тональности текста, субъективизма и конкретизации на основе словарного подхода;
  - Разработан прототип информационной системы параметризации, классификации и оценивания недостоверности публикации на основе экспертных оценок и получены предварительные результаты;
  - Проведен сравнительный анализ предварительных результатов и оценок экспертов;
  - Определена общая схема системы автоматизированного мониторинга СМИ в сети интернет, формирования полнотекстовой базы данных и ее классификации.

В дальнейшем предлагается решить следующие задачи:

1. Организовать серверное пространство для рабочей базы данных и настройка СУБД;
2. Разработать систему автоматического мониторинга публикаций в сети Интернет и формирования их полнотекстовой БД;
3. Разработать алгоритмы предварительной нормализации, параметризации и анализа текстовых данных;
4. Построить алгоритмы автоматического определения автора анонимной публикации на основе анализа и сопоставления с мониторинговой БД;
5. Сформировать библиотеку алгоритмов вычисления информативных признаков значимых в процессе классификации и имеющих семантическую интерпретацию;
6. Исследовать оптимальный состав вектора информативных признаков обеспечивающего наилучшее качество классификации анализируемых публикаций;
7. Построить алгоритмы предварительной автоматической классификации отдельных публикаций на основе анализа набора их информативных признаков с целью формирования множества подозрительных публикаций требующих экспертного анализа;
8. Построить алгоритмы экспертного анализа и классификации подозрительных публикаций;
9. Разработать алгоритмы для вычисления PFF(Publication False Factor) - степени достоверности отдельной публикации;
10. Разработать алгоритмы для вычисления AFF(Author False Factor) - степени достоверности публикаций определённого автора;
11. Разработать алгоритмы для вычисления MFF(Media False Factor) - степени достоверности публикаций определённого печатного СМИ;
12. Разработать аналитические инструменты для анализа MFF и прогнозирования её динамики в разрезе различных СМИ.

*Задача мониторинга и проверки достоверности публикаций в средствах массовой информации проводится в рамках ПЦФ «Разработка информационных технологий и систем для стимулирования устойчивого развития личности как одна из основ развития цифрового Казахстана».*

### **Список опубликованных статей в рамках проекта**

1. Бисаринов Б., Мусабаев Р.Р., Бисаринова А. Collective method in solving the Big Data clustering problems// Материалы 16-ой Международной научной конференции "Информационные Технологии и Менеджмент 2018", стр. 75-76, Рига, Латвия, 26-27 апреля 2018 г.
2. Красовицкий А.М., Мусабаев Р.Р. Energy maximization approach for solving clustering problem// Материалы 16-ой Международной научной конференции "Информационные Технологии и Менеджмент 2018", стр. 77-78, Рига, Латвия, 26-27 апреля 2018 г.
3. Мусабаев Р.Р. Development of social trust information system in the education field using the "Big Data" technologies// Материалы 16-ой Международной научной конференции "Информационные Технологии и Менеджмент 2018", стр. 79-80, Рига, Латвия, 26-27 апреля 2018 г.
4. Уалиева И.М., Красовицкий А.М., Мусабаев Р.Р. Generalization rate in open publication materials in Russian Language// Материалы 16-ой Международной научной конференции "Информационные Технологии и Менеджмент 2018", стр. 81-82, Рига, Латвия, 26-27 апреля 2018 г.
5. Красовицкий А.М., Мусабаев Р.Р. Energy-based centroid identification and cluster propagation with noise detection// Материалы 10-ой Международной конференции по вычислительному коллективному интеллекту, Бристоль, Великобритания, 5-7 сентября 2018 г. (принята)
6. Уалиева И.М., Красовицкий А.М., Мейрамбеккызы Ж., Мусабаев Р.Р. Lexicon-based approach in generalization estimation in Russian media// Материалы Конференции по эмпирическим методам в обработке естественного языка, EMNLP 2018, Брюссель, Бельгия, 31 октября – 4 ноября 2018 г. (в процессе принятия)
7. Уалиева И.М., Красовицкий А.М., Мейрамбеккызы Ж., Мусабаев Р.Р. The recognition of the informative features with semi-automatic compilation of dictionaries of lexical forms: text generalization, sentiment analysis// Материалы Конференции по искусственному интеллекту и естественному языку, AINL 2018, Санкт-Петербург, Россия, 17-19 октября 2018 г. (в процессе принятия)

### **Список литературы**

1. Морозова А. А. Достоверность информации в социальных сетях и критерии ее верификации (на примере «ВКонтакте») // Вестник Челябинского государственного университета, no. 6 (402), 2017, pp. 75-83. КиберЛенинка: <https://cyberleninka.ru/article/n/dostovernost-informatsii-v-sotsialnyh-setyah-i-kriterii-ee-verifikatsii-na-primere-vkontakte>
2. Ualiyeva I., Krassovitsky A., Mussabayev R. Generalization rate evaluation in open publication materials in Russian language. The 16th International Conference Information Technologies and Management, April 26-27, 2018, Information Systems Management Institute, Riga, Latvia



# АЛГОРИТМ ИЕРАРХИЧЕСКОЙ КЛАСТЕРИЗАЦИИ НА ОСНОВЕ ЭНЕРГЕТИЧЕСКОГО КРИТЕРИЯ

Мусабаев Р.Р., Красовицкий А.М.

*Институт информационных и вычислительных технологий КН МОН РК  
Лаборатория анализа и моделирования информационных процессов  
e-mail: rustam@ipic.kz*

***Аннотация.** Рассматривается новый алгоритм иерархической кластеризации для разделения наборов объектов на группы в соответствии со схожестью их атрибутов. Нами предложен параметрический алгоритм конкурентного роста кластеров за счет поглощения/слияния граничных кластеров. Объекты поглощаются соседним кластером, если достигается максимум энергии кластеров.*

***Ключевые слова:** кластеризация на основе энергии, дендрограмма, анализ данных*

## **Введение**

В последнее время в литературе предложен ряд различных подходов, использующих энергию в кластеризации. К примеру, в [1] рассматривается функция энергии, построенная на основе энергетической функции Ляпунова. Качество кластеризации измеряется с точки зрения пространственных и временных характеристик для потоковых наборов данных. Идея заключается в том, чтобы оценить временную эволюцию потока входных данных, и затем зафиксировать изменения поведения и тенденции в кластеризации. В [2] вводится метод оптимизации с использованием энергии, который помогает эффективно отобразить произвольный граф на плоскости. Разбитые на участки краевые сегменты входного графа фильтруются и иерархически кластеризуются с помощью функции энергии. Последний дает количественное представление о скоплении вершин графа и используется в качестве критерия для выбора регионов с наибольшей плотностью (перегруженностью) ребер, подлежащих перестроению. В [3] используется модель притяжения/отталкивания для поиска оптимального разложения графа. Введенная функция энергии охватывает как оценку степеней вершин/краев, так и их положений, поэтому в рамках данной модели можно принимать решение об извлечении подграфа в качестве характеризующего кластера. В [4] проблема сегментации изображения рассматривается с точки зрения объединения нескольких известных подходов к решению данной задачи. Чтобы сочетать различные конфликтующие и взаимодополняющие критерии (например, вариация информации на основе региона и  $F$ -меру на основе контура), вводится критерий глобальной согласованности энергии слияния. Как следствие, сегментация изображения обрабатывается алгоритмом  $k$ -средних в разных пространствах, соответствующих цветовым пространствам  $RGB$ , с последующей максимизацией энергии для решения задачи согласованной кластеризации.

В приведенных выше публикациях показано, что с помощью энергии может быть построено и измерено некоторое обобщающее состояние системы во время процесса формирования кластеров. Это открывает возможность решения проблемы кластеризации при помощи оптимизации единого критерия, основанного на нескольких смежных и взаимоисключающих показателей рассматриваемой системы.

В нашей работе используется метод для измерения локальных свойств точек, сгруппированных с помощью специальной функции энергии и отражающей характеристики пространственной кластеризации. Реализованный алгоритм кластеризации, стремится достигнуть максимум энергии для каждого формируемого кластера и полную энергию всей кластеризации. Одно из главных предположений на котором основывается данный подход состоит в том, что каждая точка кластеризации  $s$  имеет некоторый квант энергии, который непрерывно излучается. Мы предполагаем, что излучаемая энергия максимальна для ближайшей окрестности  $s$  и нелинейно затухает с расстоянием.

Мы применяем подход распространения кластеров путем включения точек в один из кластеров таким образом, чтобы оптимизировать энергию этих кластеров. В частности, наш алгоритм сортирует кластеры, после чего находим пару, которая дает максимальное увеличение энергии в соответствии с функцией  $p_q$  и сливает два кластера в один. В нашем алгоритме точки кластеризации не могут переходить из одного кластера в другой. Эта процедура повторяется до тех пор, пока не останется одного единственного кластера.

### Конструкция дендрограмм на основе энергетических функций

Для построения алгоритма кластеризации мы используем подход к измерению энергетической функции  $p_q$ . Идея алгоритмов следующая: сначала выделяем каждую точку как кластер с одним элементом. Затем мы применяем процедуру объединения кластеров для той пары кластеров  $(i', j')$ , которая дает  $E_{i,j} = \sum_{s \in cl_i} \sum_{s' \in cl_j} p_q [dist(s, s')]$  максимальное значение для всех кластеров  $i, j$ .  $E_{i,j}$  выступает критерием кумулятивной энергии между кластерами  $i$  и  $j$ .  $d(s, s')$  – это евклидово расстояние между точками  $s$  и  $s'$ . Нелинейная эвристическая функция энергии  $p_q: [0,1] \rightarrow [0,1]$  определяется действительным параметром  $q \in (-1,1)$ . Были рассмотрены следующие параметрические функции энергии:

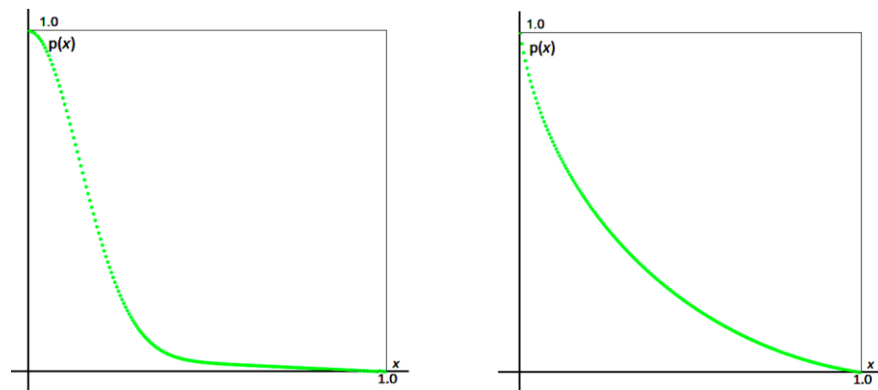


Рис. 1. Нелинейные эвристические функции энергии  $p_q$ .

#### Алгоритм:

**Входные данные:**  $S$  множество точек в нормализованном виде,  $N = |S|$ ;

**Параметры:**  $\delta$  – шаг дискретизации,  $q$  – параметр функции энергии  $p_q$ ,  $n$  – требуемое количество кластеров

#### Выходные данные:

$T = \{ (s, d, clusterID, clusterID') \mid s \in S, d, clusterID, clusterID' > 0 \}$  множество кортежей, которые образуют древовидную структуру для кластеризации, где

$s$  - это точка данных кластера, для которой индекс  $clusterID$  на шаге  $d$  изменяется на  $clusterID'$ .

**for**  $s_i \in S$

    точке  $s_i$  назначается кластер  $Cl_i$

$Cl = \{Cl_i\}$

$d = 0, R_d = \delta, T = \emptyset$

**while**  $|Cl| > n$

**for**  $Cl_i \in Cl$

**for**  $Cl_j \in Cl$

**if**  $dist(s, s') < R_d$

$$E_{i,j} = \sum_{s \in Cl_i} \sum_{s' \in Cl_j} p_q[dist(s, s')]$$

$(i', j') \leftarrow \arg \max_{i,j} E_{i,j}$

**if**  $(i', j')$  не определено

$R_d = R_d + \delta$

**else**

$d = d + 1$

**if**  $|Cl_i| > |Cl_j|$

**Merge**( $j', i', d$ )

**else**

**Merge**( $i', j', d$ )

**end while**

**Function Merge**( $i, j, d$ ):

**for**  $s \in Cl_i$

        replace  $s$  to  $Cl_j$

        add ( $s, d, i, j$ ) to  $T$

Данный алгоритм производит слияние кластеров, для которых значение вычисленной энергии  $E_{i,j}$  максимально. Алгоритм останавливается, как только создано заданное количество кластеров  $n$ .

Данный подход требует определения параметров  $q$ ,  $\delta$  и  $n$ . Данная задача может быть решена путем идентификации параметров итеративного поиска: в качестве задачи оптимизации точности кластеризации. Точность вычисляется путем сравнения кластеров, которые мы вычисляем с помощью известного распределения по кластерам из соответствующего исходного набора данных. Для решения данной задачи алгоритм кластеризации инкапсулируется в «черный ящик», так что его параметры используются как входные данные, при этом мы получаем результирующую точность в качестве оценочного критерия качества. Подобный подход был апробирован авторами на аналогичных задачах с использованием функции энергии [5, 6] на тестовых данных из общедоступной базы данных UCI [7].

### Заключение

Алгоритм позволяет получить дендрограмму  $T$  для набора данных  $S$ . Таким образом исходные точки данных формируют иерархическую структуру представления. Однако этот подход требует дальнейшего изучения, поскольку он требует ручной пошаговой настройки кластеризации и определения терминальных критериев, для качественно выполненной кластеризации. Данный метод является аналогом метода

иерархической кластеризации “полной связи” [8]. Однако в нашем подходе критерий для установления связанности вычисляется с применением эвристической нелинейной функции энергии. Это приводит к тому, что данная реализация не является эффективной с точки зрения времени ее исполнения. Иными словами, мы улучшаем качество кластеризации за счет обработки избыточного объема информации о пространственном распределении входных данных. Подобный подход можно рассматривать как компромисс между эффективностью по времени и качеством кластеризации.

Мы реализовали вычислительно устойчивый алгоритм для иерархической кластеризации и определения его параметров, который может дать приемлемое качество разбиения исходных данных по мере их близости (схожести). В дальнейшем мы планируем усовершенствование этой вычислительной модели для использования в реальных проблемно-ориентированных задачах кластеризации и классификации. Однако ее эффективное использование в конкретных приложениях может потребовать дальнейшего изучения предметной области. Что важно в нашем подходе, так это возможность адаптировать его к целому ряду типичных наборов, данных требующих кластеризации и достичь приемлемой точности.

Заметим, что процедура конкурентного роста кластеров начиная с единичных кластеров может быть реализована и более оптимальными способами. Например, путем включения параллелизма в алгоритм кластеризации.

#### **Список литературы**

1. Albertini, M.K., de Mello, R.F. Energy-based function to evaluate data stream clustering. *Advances in Data Analysis and Classification* 7(4), 435-464, 2013, DOI 10.1007/s11634-013-0145-3. URL <https://doi.org/10.1007/s11634-013-0145-3>
2. Zhou, H., Yuan, X., Cui, W., Qu, H., Chen, B. Energy-based hierarchical edge clustering of graphs. In: *PacificVIS '08*, vol. Visualization Symposium, pp. 55-61, 2008
3. Noack, A.: Energy-based clustering of graphs with nonuniform degrees. In: P. Healy, N.S. Nikolov (eds.) *Graph Drawing*, pp. 309-320, 2006, Springer Berlin Heidelberg, Berlin, Heidelberg
4. Khelifi, L., Mignotte, M. Efa-bmfm: A multi-criteria framework for the fusion of colour image segmentation. *Information Fusion* 38, 104-121, 2017
5. Красовицкий А.М., Мусабаев Р.Р. Energy maximization approach for solving clustering problem// *Материалы 16-ой Международной научной конференции "Информационные Технологии и Менеджмент 2018"*, стр. 77-78, Рига, Латвия, 26-27 апреля 2018 г.
6. Красовицкий А.М., Мусабаев Р.Р. Energy-based centroid identification and cluster propagation with noise detection// *Материалы 10-ой Международной конференции по вычислительному коллективному интеллекту*, Бристоль, Великобритания, 5-7 сентября 2018 г. (accepted)
7. Online clustering data sets UCI. URL <https://archive.ics.uci.edu/ml/datasets.html> (accessed 09.06.2018)
8. Defays D. An efficient algorithm for a complete link method. *The Computer Journal*. British Computer Society. 20 (4): 364–366. doi:10.1093/comjnl/20.4.364, 1977

# ОСОБЕННОСТИ ОБРАБОТКИ ТЕКСТОВ ЕСТЕСТВЕННОГО ЯЗЫКА В РАЗРАБОТКЕ ИНТЕЛЛЕКТУАЛЬНОЙ ПОИСКОВОЙ СИСТЕМЫ

Мусаев М.С.<sup>1</sup>, Рахимова Д.Р.<sup>2</sup>

<sup>1</sup>Казахский национальный университет имени аль-Фараби

<sup>2</sup>Институт информационных и вычислительных технологий

e-mail: magzhanmusaev@gmail.com<sup>1</sup>, di.diva@mail.ru<sup>2</sup>

**Аннотация.** В статье анализируются современные системы автоматической обработки текстов на естественном языке, трудности их использования в лингвистических исследованиях и предлагаются процедуры анализа исходного запроса пользователя на естественном языке

**Ключевые слова:** обработка; естественный язык; лингвистический; процедуры; запрос.

Значительное место в технологиях текстового поиска занимает обработка естественных языков. Под обработкой естественных языков (Natural Language Processing, NLP) понимается решение задач, связанных с пониманием, анализом, выполнением различных операций над текстами, а также их генерацией [6]. Примеры подобных задач: классификация, кластеризация хранимых коллекций документов, глубинный анализ текстов, перевод документов с одного языка на другой и т.д.

До организации процесса поиска программный препроцессор должен «перевести» запрос пользователя на поиск с естественного языка на конкретизированный информационно-поисковый язык. В настоящее время интеллектуальные поисковые системы могут решать следующие достаточно сложные задачи.

Автоматически предопределять язык запроса. Данная функция разрешает ограничить языковой сектор Сети, в котором станет выполняться разведка запрашиваемой данной, собственно, что с учетом число данных в виртуальной среде, позитивно воздействует на процессе поиска.

Вычеркивать неинформативные тексты (стоп-слова), т.е. текст служебных частей речи, не несущих практически никакой смысловой нагрузки, или кое-какие более общеупотребительные текста. Их удаление важно уменьшает размер индекса и наращивает релевантность итогов розыска.

Проводить лингвистический анализ исходного запроса пользователя, а также отысканных текстовых документов. Лингвистический анализ включает следующие процедуры:

- 1) лексический (графематический) анализ;
- 2) морфологический анализ;
- 3) синтаксический анализ;
- 4) семантический анализ.

Лексический анализ заключается в выявлении составляющих структуры слова: параграфов, абзацев, предложений, отдельных текстов и т.д. При данном ведется определение языка слова и присутствия лексических выражений (жаргонных текстов и т.п.). Осуществление такового на подобии анализа не вызывает затруднений. Морфологический анализ заключается в автоматическом распознавании

морфологических данных всякого текста. При проведении предоставленного на подобии анализа почти все находится в зависимости от натурального языка, на котором написан текст документа. К примеру, информационно-поисковая система довольно отлично распознает русский язык по причине его развитой морфологии. В случае с английским (британским) языком имеют все шансы возникновения проблемы по причине его аналитической природы. Целью синтаксического анализа считается нахождение синтаксической зависимости текстов в предложении и возведение его активного дерева. Механически отличаются смысловые составляющие тирады: закономерный тип, закономерный сказуемое, объекты (прямые и косвенные дополнения), атрибуты (определения) и всевозможные облики событий. Синтаксический анализ – довольно сложная задачка, сложность которой впрямую находится в зависимости от числа слов в предложении и использованных в нем правил. Есть некоторое количество способов проведения такового анализа, которые воплощены в программных продуктах на подобии Ergo Linguistic Technologies Parser, Link Parser, Functional Dependency Grammar и др. Семантический анализ ведется с целью определения значения слова. Заключение данной задачки относится к непросто реализуемым процессам по причине надобности существования безукоризненного механизма экспертной оценки свойства данных. В следствие этого на нынешний день системы, способные выполнить семантический тест в полном размере, отсутствуют.

Целью морфологического анализа является определение морфологических характеристик слова и его основной словоформы. Особенности анализа сильно зависят от выбранного естественного языка

Целью синтаксического анализа является определение синтаксической зависимости слов в предложении. Так как казахский язык относится к агглютинативной группе языков, наличием тесной связи между семантикой и синтаксисом, процедура автоматизированного синтаксического анализа текста является трудоемкой. Сложность алгоритма увеличивается экспоненциально при увеличении количества слов в предложении и числа используемых правил. Разработки в области семантического анализа текста связаны с областью искусственного интеллекта, делающей акцент на смысловом понимании текста.

Целью графематического анализа является выделения элементов структуры текста: параграфов, абзацев, предложений, отдельных слов и т. д

В некоторых интеллектуальных информационно-поисковых системах применяется способ «неточного поиска», позволяющий предопределять веб-странички, которые имеют все шансы быть релевантными запросу на поиск даже при их неточном совпадении. Системы могут корректировать ряд опечаток в запросе пользователя на базе входящего в их состав массива омонимов и слов, имеющих альтернативное написание.

Таким образом, главные отличия интеллектуальной поисковой системы от традиционной заключаются в умении системы возделывать требования пользователя, сформулированные в случайной форме на естественном языке, а еще представлять итоги поиска не в форме ранжированных адресов веб-сайтов, а в виде определенного текстового куска, содержащего подходящую пользователю информацию.

Для решения задачи обработки текста можно использовать метод машинного обучения. Чтобы их использовать нам понадобится размеченный набор текстов  $T$  для обучения системы, т.е. каждому учебному тексту присваивается метка определённого класса. Также необходимо выбрать способ формализации этих данных, т.е. некоторым способом построить отображение  $f$  из множества текстов  $T$  в пространство признаков  $X$

$$f: T \rightarrow X \quad (1)$$

Функцию  $f$  называют процедурой извлечения признаков или feature extraction.

После того как мы определим  $f$  и пространство признаков  $X$  будет построено, каждому тексту из  $T$  поставлена в соответствие точка из  $X$ , появится возможность применять математические методы для разделения всех точек  $X$  на подмножества. То есть это и есть задача кластеризации точек из  $X$ , задача сортировки текстов по темам сводится к задаче классификации точек из  $X$ . Формально говоря, нам необходимо построить отображение  $g$  из множества вектор-признаков  $X$  во множество меток  $L$ . Итого наша задача обработки текста разделяется на два основных этапа.

- построение пространства признаков (извлечение признаков)
- разделение пространства признаков на части

Далее метод извлечения признаков из текста под названием частотный анализ (другое название - мешок слов и его модификации. Для формирования вектора признаков текста, частотный анализ считает количество повторов каждого слова в тексте. Если более детально, то схема частотного анализа выглядит следующим образом:

- построить словарь  $V$  из всех слов, используемых в исходных текстах  $T$ ;
- для каждого текста  $t_i \in T$  и для каждого слова из словаря  $v_j \in V$ ;
- посчитать число вхождений  $x_{ij}$  слова  $v_j$  в текст  $t_i$

Таким образом для каждого текста  $t_i \in T$  мы получаем вектор целых неотрицательных чисел  $x_i$ , длина которого равна количеству слов в словаре  $V$ . Это базовый вид метода частотного анализа текстов. Для текстов разного размера величина значений частоты  $x$  может сильно отличаться, т.е. чем больше текст, тем больше может быть повторов слов. Для уменьшения этого эффекта применяется метод нормализованного частотного анализа или TF (term frequency), значения частоты  $x$  делятся на общее число слов в тексте  $t$ .

Формирование словаря - важная часть работы системы обработки текстов. От того какие слова там окажутся, зависит информативность вектор-признаков и соответственно качество работы системы. Словарь формируется из всего набора текстов исходя из задачи.

Например, предлоги (және, немесе, егер) встречаются в любом тексте. Для решения задачи сортировки текстов по темам эти слова не информативны, поэтому их можно удалить из словаря.

Однако, для задачи определения авторства текста частота употребления предлогов может оказаться очень даже полезными признаками. Удалять часто употребляемые слова или нет? Существует компромиссный вариант формирования вектор-признаков. Имя ему TF-IDF, этот метод не выбрасывает часто употребляемые слова из словаря, но уменьшает их вес в вектор-признаке. Для этого для всех слов словаря вычисляется коэффициент обратной частоты IDF (inverse document frequency), выглядит это следующим образом.

$$\text{IDF}(v_i) = \log \frac{\text{size}(T)}{\text{size}(T(v_i))} \quad (2)$$

т.е. логарифм дроби, у которой в числителе общее количество текстов, а в знаменателе количество текстов, содержащих слово  $v_i$ , т.е. чем чаще встречается слово, тем меньше значение его IDF.

Существует ещё одна проблема. В сложных естественных языках одно и то же слово может принимать разные формы (падежи), и в словарь частотного анализа могут попадать все словоформы, отличающиеся предложениями и/или окончаниями. Из-за этого может сильно увеличиваться размер словаря и соответственно размер набора данных для обучения, что может вызывать падение производительности системы и ухудшать обобщающие способности классификатора (переобучение). Существует несколько способов решения этой задачи. Первый это лемматизация - все слова в тексте приводятся к нормальной форме (единственное число, именительного падежа). Лемматизация требует словарей языков, на которых написан текст, она выполняет поиск для всех слов текста их нормальной формы по словарям и это может существенно снижать производительность. Вторым методом это стемминг - выделение основы слов путём отбрасывания приставок и окончаний. Этот способ нормализации текста работает гораздо быстрее чем лемматизация. Он менее качественный, но для частотного анализа его вполне достаточно. Также можно использовать хэширование - необходимо построить функцию отображающую множество слов во множество хэш-кодов фиксированной длины, причём, стремиться к тому, чтобы похожие слова (словоформы одного слова) имели одинаковый хэш. Перед выполнением частотного анализа, все слова в тексте и словаре заменяются хэш-кодами, что позволяет сократить количество ненужных повторов.

Общая схема системы состоит из следующих этапов работы:

1. собрать и разметить учебные тексты  $T$
2. удалить из текстов лишние слова (исходя из задачи)
3. выполнить стемминг (нормализация слов)
4. построить словарь  $V$  из очищенного набора текстов  $T$
5. с помощью словаря  $V$  выполнить частотный анализ текстов
6. построить матрицу учебных примеров  $X$  из векторов частотных характеристик и списка меток текстов  $L$
7. случайным образом перемешать набор  $X$  (переставить строки матрицы)
8. разделить набор  $X$  на две части - учебный набор  $X_1$ , тестовый набор  $X_t$
9. обучить классификатор (или кластеризатор)  $g$  на наборе  $X_1$  и сохранить его параметры  $W$
10. протестировать качество работы системы на наборе  $X_t$

Прежде чем классификатор начнет работать его необходимо обучить на множестве учебных текстов. Эту процедуру можно разбить на четыре этапа:

1. Определение количества классов, с которыми будет работать система. Подбор тематических учебных текстов, в одинаковом объеме для каждого класса. Тексты должны в точности соответствовать своей теме.
2. Составление словаря из множества учебных текстов.
3. Частотный анализ множества учебных текстов, с использованием полученного словаря. Получаем множество учебных векторов для нейронной сети.
4. Обучение нейросетевого классификатора на наборе векторов частотных характеристик учебных текстов. Далее рассматривается решение этой задачи.

Далее мы рассмотрим реализацию описанной выше схемы средствами языка Python. Для этого воспользуемся языком Python и библиотекой Scikit-learn. Последняя представляет собой большой набор реализаций разных методов машинного обучения и связанных с этой функцией. Начнём с данных, Scikit-learn содержит в себе различные наборы данных для тестирования программ, среди них набор 20newsgroups из репозитория UCI.



Второй инструмент, который мы рассмотрим, это Stemmer для выполнения стемминга (выделения основы слов). Stemmer не является частью Scikit-learn и устанавливается отдельно (например, с помощью утилиты pip).

Для выполнения частотного анализа в Scikit-learn есть следующие инструменты.

- CountVectorizer - составляет словарь и заменяет слова в тексте номерами в словаре.

- HashingVectorizer - аналогичная функциональность CountVectorizer но с предварительным хешированием слов в тексте

- TfidfTransformer - частотный анализатор TF-IDF

- TfidfVectorizer - CountVectorizer и TfidfTransformer в одной упаковке

Тесты будут проводиться на наборе из нескольких текстов на казахском языке, разбитых на несколько групп по темам. Тексты будут храниться в формате sqlite. В качестве классификатора используем логистическую регрессию из пакета Scikit-learn. Для каждого выполним тестирование на 10 случайных разбиениях набора (учебный и тестовый). Далее рассмотрим работу кластеризатора, т.е. решение задачи поиска похожих текстов. Поиск похожих текстов выполняется аналогичным классификации образом, только вместо логистической регрессии выполняется кластеризация - KMeans из пакета Scikit-learn.

Описанный метод может недостаточно хорошо работать на больших текстах, где частотная характеристика может размываться большим количеством разных слов. На очень маленьких текстах тоже могут быть не лучшие результаты, потому как мало информации для частотного анализа. Этот метод может давать сбои на задачах где важен порядок слов в тексте. В остальных случаях частотный анализ текстов можно успешно применять для автоматической сортировки массива текстов.

### Список литературы

1. Алгоритмы поисковых систем [Электронный ресурс].— Режим доступа: <http://xbb.uz/SEO/Algoritmy-poiskovy-h-sistem>. — Дата доступа: 12.11.2015.

2. Описание технологии лингвистического анализа [Электронный ресурс]. — Режим доступа: <http://asknet.ru/Technology/techdescr.htm>. — Дата доступа: 23.10.2015.

3. Концептуальные положения создания программного обеспечения поисковой машины, реализующей морфологическую, синтаксическую и семантическую обработку текстовой информации [Электронный ресурс]. — Режим доступа: <http://www.shkolagym.ru/obls/konceptualeniye-polojeniya-sozdaniya-programmnogo-obespecheniya>. — Дата доступа: 21.12.2015.

4. Семантический поиск [Электронный ресурс]. — Режим доступа: <http://www.seonews.ru/analytics/semanti-cheskiy-poisk>. — Дата доступа: 21.12. 2015.

5. Е.С.Борисов Классификатор текстов на естественном языке. - <http://mechanoid.kiev.ua/neural-net-classifier-text.html>

## МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ДИНАМИЧЕСКОЙ СИСТЕМЫ

**Мустафин С.А., Сериков Б.Б.**

*Казахский Национальный Университет им. Аль-Фараби, Алматы*  
e-mail: [sam@ipic.kz](mailto:sam@ipic.kz), [serikov.bagdat@mail.ru](mailto:serikov.bagdat@mail.ru)

***Аннотация:** Математическое определение понятия динамической системы. На примере динамических систем, описываемых обыкновенными дифференциальными уравнениями, иллюстрируются четыре типа решений: состояние равновесия, устойчивое периодическое решение, квазипериодическое и хаотическое решения. Вводится понятие странного аттрактора, обсуждаются основные свойства регулярных и хаотических решений.*

***Ключевые слова:** Динамическая система, дифференциальное уравнение, фазовое пространство, колебательные системы, дискретное время.*

## **1. Введение**

Одной из важных научных проблем естествознания является решение задачи предсказания поведения изучаемого объекта во времени и пространстве на основе определенных знаний о его начальном состоянии. Эта задача сводится к нахождению некоторого закона, который позволяет по имеющейся информации об объекте в начальный момент времени  $t_0$  в точке пространства  $x_0$  определить его будущее в любой момент времени  $t > t_0$ . В зависимости от степени сложности самого объекта этот закон может быть детерминированным или вероятностным, может описывать эволюцию объекта только во времени, только в пространстве, а может описывать пространственно-временную эволюцию.[1]

## **2 Постановка задачи**

Под динамической системой понимают любой объект или процесс, для которого однозначно определено понятие состояния как совокупности некоторых величин в данный момент времени и задан закон, который описывает изменение (эволюцию) начального состояния с течением времени. Этот закон позволяет по начальному состоянию прогнозировать будущее состояние динамической системы, его называют законом эволюции. Динамические системы – это механические, физические, химические и биологические объекты, вычислительные процессы и процессы преобразования информации, совершаемые в соответствии с конкретными алгоритмами. Описания динамических систем для задания закона эволюции также разнообразны: с помощью дифференциальных уравнений, дискретных отображений, теории графов, теории марковских цепей и т.д. Выбор одного из способов описания задает конкретный вид математической модели соответствующей динамической системы.[2] Математическая модель динамической системы считается заданной, если введены параметры (координаты) системы, определяющие однозначно ее состояние, и указан закон эволюции. В зависимости от степени приближения одной и той же системе могут быть поставлены в соответствие различные математические модели.

## **3 Существующие решения**

Рассмотрим динамические системы, моделируемые конечным числом обыкновенных дифференциальных уравнений. Применительно к таким системам сохранились представления и терминология, первоначально возникшие в механике. В рассматриваемом случае для определения динамической системы необходимо указать объект, допускающий описание состояния заданием величин  $x_1, x_2, \dots, x_n$  в некоторый момент времени  $t = t_0$ . Величины  $x_i$  могут принимать произвольные значения, причем двум различным наборам величин  $x_i$  и отвечают два разных состояния. Закон эволюции динамической системы во времени записывается системой обыкновенных дифференциальных уравнений.

$$\frac{dx}{dt} = x_i = f_i(x_1, x_2, \dots, x_n), \quad i=1, 2, \dots, N. \quad (1)$$

Если рассматривать величины  $x_1, x_2, \dots, x_n$  как координаты точки  $x$  в  $N$ -мерном пространстве, то получается наглядное геометрическое представление состояния динамической системы в виде этой точки, которую называют изображающей, а чаще фазовой точкой, а пространство состояний – фазовым пространством динамической системы.[2][3] Изменению состояния системы во времени отвечает движение фазовой точки вдоль некоторой линии, называемой фазовой траекторией. В фазовом пространстве системы уравнениями (1) определяется векторное поле скоростей, сопоставляющее каждой точке  $x$  выходящий из нее вектор скорости  $F(x)$ , компоненты которого даются правыми частями уравнений (1).

#### 4 Использование колебательные системы и их свойства

Важную группу динамических систем представляют системы, в которых возможны колебания. Колебательные системы с точки зрения их математических моделей разделяют на определенные классы. Различают линейные и нелинейные колебательные системы, сосредоточенные и распределенные, консервативные и диссипативные, автономные и неавтономные. Особый класс представляют так называемые автоколебательные системы. Основные свойства указанных систем подробно обсуждаются в работах по теории колебаний.[4]

Колебательная система называется линейной или нелинейной в зависимости от того, линейна или нелинейна описывающая ее система дифференциальных уравнений. Линейные системы являются частным случаем нелинейных. Однако в силу принципиальной важности линейных систем при исследовании вопросов устойчивости колебаний, а также возможности использования принципа суперпозиции решений такая классификация оправданна.

Динамические системы, моделируемые конечным числом обыкновенных дифференциальных уравнений, называют сосредоточенными или точечными системами. Они описываются с помощью конечномерного фазового пространства и характеризуются конечным числом степеней свободы. Одна и та же система в различных условиях может рассматриваться либо как сосредоточенная, либо как распределенная. Математические модели распределенных систем – это дифференциальные уравнения в частных производных, интегральные уравнения или обыкновенные уравнения с запаздывающим аргументом. Число степеней свободы распределенной системы бесконечно, и требуется бесконечное число данных для определения ее состояния.

По энергетическому признаку динамические системы делятся на консервативные и неконсервативные. Консервативные системы характеризуются неизменным во времени запасом энергии. В механике их называют гамильтоновыми. Для консервативных систем с  $n$  степенями свободы определяется гамильтониан системы  $H(p, q)$ , где  $q_i$  – обобщенные координаты,  $p_i$  – обобщенные импульсы системы,  $i = 1, 2, \dots, n$ . Гамильтониан полностью характеризует динамическую природу системы и с физической точки зрения в большинстве случаев представляет собой ее полную энергию. Эволюция во времени консервативных систем описывается уравнениями механики Гамильтона. [5]

Динамические системы с изменяющимся во времени запасом энергии называются неконсервативными. Системы, в которых энергия уменьшается во времени из-за трения или рассеяния, называются диссипативными. В соответствии с

этим системы, энергия которых во времени нарастает, называются системами с отрицательным трением или отрицательной диссипацией. Такие системы можно рассматривать как диссипативные при смене направления отсчета времени на противоположное.

## 5 Резюме

Определение динамической системы и приведены примеры динамических систем, описываемых обыкновенными дифференциальными уравнениями. Такие динамические системы могут иметь четыре типа решений: состояние равновесия, периодическое движение, квазипериодическое движение и хаотическое. Этим типам решений соответствуют аттракторы системы в виде устойчивого равновесия, предельного цикла, квазипериодического аттрактора (р-мерного тора) и хаотического (или странного) аттрактора. Важным является то, что простейшие типы квазипериодических и хаотических аттракторов могут реализовываться в динамических системах с размерностью фазового пространства не менее трех.

## Список литературы

1. Аносов Д.В. Динамическая система // Математическая энциклопедия. М.: Сов. энциклопедия, 1979.
2. Анищенко В.С. Сложные колебания в простых системах. М.: Наука, 1990.
3. Лихтенберг А., Либерман М. Регулярная и стохастическая динамика. М.: Мир, 1984.
4. Шустер Г. Детерминированный хаос. М.: Мир, 1988.
5. Неймарк Ю.И., Ланда П.С. Стохастические и хаотические колебания. М.: Наука, 1987.
6. Лоскутов А.Ю., Михайлов А.С. Введение в синергетику. М.: Наука, 1990.

УДК 543.08

## МАТЕМАТИЧЕСКАЯ МОДЕЛЬ НАКОПЛЕНИЯ ТЯЖЕЛЫХ МЕТАЛЛОВ ИЗ ЗАГРЯЗНЕННОЙ ПОЧВЫ В ВЕГЕТАТИВНЫХ ОРГАНАХ РАСТЕНИЙ

Найзабаева<sup>1</sup> Л., Нуржанов<sup>1</sup> Ч., Pidlisnyuk<sup>2</sup> V., Сатымбеков<sup>1</sup> М.

<sup>1</sup>Институт информационных и вычислительных технологий КН МОН РК,  
Алматы, Казахстан

<sup>2</sup>Jan Evangelista Purkyně University, Usti nad Labem, Czech Republic

**Аннотация.** В данной статье приведены результаты регрессионного анализа с применением метода пошаговой регрессии, направленные на установление математической зависимости накопления тяжелых металлов вегетативных органах растений в зависимости от содержания их в почве. Исходными переменными служили валовое содержание тяжелых металлов (As, Pb, Cd, Zn, Co, Ni, Cu, Cr, Ba, Sr, Mn) в почве и в вегетативных органах растений. На основании расчета, получены эмпирическое уравнение регрессии накопления тяжелых металлов в вегетативных

органах растений, произрастающих на загрязненной тяжелыми металлами почве, с учетом величины достоверности аппроксимации. Значения коэффициента детерминации близки к значению 1, что свидетельствует о статистической достоверности результатов исследования, а значение критерия Фишера подтверждают статистическую надежность уравнений в целом. Установлено, что процесс перемещения тяжелых металлов в системе «почва – растение» носит линейный характер. Данные расчеты свидетельствуют о том, что высокопродуктивный биоэнергетический вид *Miscanthus x giganteus* является фиторемедиантом, обладает способностью аккумулировать тяжелые металлы в вегетативных органах из загрязненной почвы и может быть использован в программах фиторемедиации.

**Ключевые слова:** регрессионный анализ, тяжелые металлы, растение, почва

Технология фиторемедиации является альтернативой современным инженерным технологиям, и рассматривается как экономичный и «зеленый» метод использования растений для извлечения или деградации загрязняющих веществ в почве [1, 2]. Современные инженерные технологии очистки загрязненных земель/воды/воздуха от ксенобиотиков чрезвычайно энергоемки и требуют больших капиталовложений. Интерес к фиторемедиации из года в год растет по причине не только наличия большого количества загрязнённых неорганическими и органическими ксенобиотиками и маргинальных земель, но и из-за рентабельности метода [3]. Залогом успешной очистки почв, с помощью фиторемедиации является правильный подбор растений. Растения должны быть высокопродуктивными и адаптированными к почвенно-климатическим условиям [1].

Последние годы в качестве перспективного фиторемедианта для загрязненных тяжелыми металлами почв в программах фиторемедиации рассматривают непродовольственный, высокопродуктивный, биоэнергетический вид *Miscanthus x giganteus* [4-6]. Наиболее важным качеством данного вида является способность произрастать на маргинальной, деградированной земле в течение 20 лет, а положительным качеством, характеризующим данного вида, как фиторемедианта, является технологичность его переработки. Биомасса *M.giganteus* используется для производства твердого биотоплива, целлюлозных отходов и биоэтанола. Коммерциализованы технологии создания топливных гранул, в виде пеллет из стеблей *M. Giganteus* во многих странах ЕС и США [7].

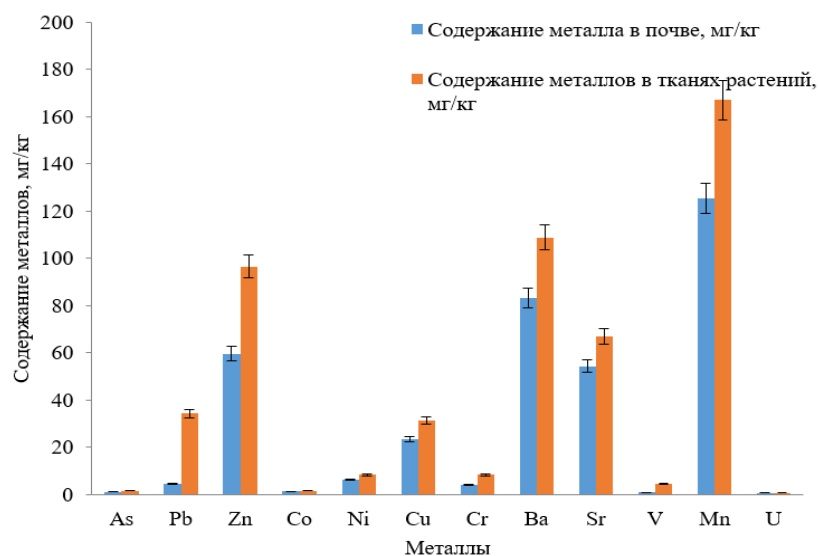
Для эффективности создания оптимальной схемы фиторемедиации разрабатываются математические модели и компьютерные программы. Математическое моделирование как инструмент современной технологии очистки загрязненных земель, ускоряет выбор оптимальной стратегии фиторемедиации для конкретных загрязненных регионов. Однако эти исследования недостаточны. Недостаточность моделирования фиторемедиационной технологии связано с тем, что восстановление загрязненной почвы - это процесс комплексный, в котором участвуют физические, биологические и химические факторы, поэтому сложно охватить сразу все аспекты моделирования [8].

Внимание ученых, в основном, обращено на моделирование параметров, определяющие фиторемедиационную значимость растительного организма. Например, в статье K.Sung с коллегами [9] с помощью модели Themodel моделировали движение органических веществ через почву и воду, а также поглощение органических химических веществ. В работе В. Motesharezadeh с коллегами [10] использовали модели Mass-Hoffman и Номее для моделирования фиторемедиации

разных типов почв, загрязненных кадмием в условиях засоления, с помощью подсолнечника. Далее S. Mathur [11] впервые разработала численную модель для прогнозирования поглощения металлов из почв, с использованием соответствующей экстракции воды и функций поглощения растворенного вещества. ID F.Lugli с коллегами [12] используя численную модель Hydrus-моделировали фиторемедиацию воды, загрязненной ионами  $Cd^{2+}$ ,  $Pb^{2+}$  и  $Zn^{2+}$ . (с учетом типа культуры и плотности посева, системы полива растений, образования почвенной корки и длины корневой системы). Создаются также различные типы модели, учитывающие влияние динамики микробной биомассы на деградацию загрязняющих веществ в корневой зоне [9, 13, 14]; модели, связанные с учетом агротехники и с внесением хелатирующих агентов для усиления детоксикации загрязнителя [12, 15, 16].

В данной статье приведены результаты, направленные на установление математической зависимости накопления тяжелых металлов вегетативных органах растений в зависимости от содержания их в почве. Анализ взаимоотношения двух изменчивых величин с помощью метода регрессии может дать ценные результаты в практическом отношении. Изучение данной зависимости могут служить основой для обоснования использования биоэнергетического вида *M.giganteus* в качестве фиторемедианта для использования в технологии очистки загрязненных земель.

Наши исследования основывались на методах математической теории планирования экспериментов, которые позволяют получить математическую модель накопления тяжелых металлов в вегетативных органах в зависимости от их от содержания в почве. При построении математических моделей использовали регрессионный анализ с применением метода пошаговой регрессии экспериментальных данных, проведенных в пакете EXCELL. Экспериментальные данные получены в результате исследования Чешской Республики, Украины и Казахстана в рамках проекта НАТО G4687 «New Phytotechnology for Cleaning Contaminated Military Sites». Основная цель математического расчета оценить способность *M.giganteus*, выращенных на загрязненной почве из территории бывшего военного горниона (Маили, Балхашский район, Алматинская область), к восстановлению почв (рисунок 1).



**Рисунок 1 – Экспериментальные данные о валовом содержании тяжелых металлов в почве и вегетативных органах *M.giganteus***

Исходя из полученных экспериментальных данных, была построена математическая модель, в основу которой легло дифференциальное уравнение. Математическое описание процессов сводится к нахождению и исследованию функциональной зависимости, т.е. функции отклика

$$y = f(x_1, x_2 \dots x_n)$$

Предварительно провели нормировку каждой переменной на диапазон разброса ее значений, чтобы исключить влияние на результаты анализа величин измерения. Нормировку данных провели по следующей формуле:

$$Y = \frac{X_i - \min_i}{\max_i - \min_i}$$

где  $X_i$  – исходные данные вектора признака  $i$ -го ( $1 \leq i \leq 12$ );  $Y_i$  – безразмерные значения от 0 до 1;  $\max$ ,  $\min$  – максимальные и минимальные значения признака.

Результаты нормированных экспериментальных накоплений тяжелых металлов в вегетативных органах растений представлены на рисунке 2, рисунке 3, рисунке 4.

На основании расчета, получено эмпирическое уравнение регрессии накопления тяжелых металлов в вегетативных органах растений, произрастающих на загрязненной тяжелыми металлами почве, с учетом величины достоверности аппроксимации.

Зависимость концентрации тяжелых металлов 1 класса опасности (As, Pb, Zn) в вегетативных органах мискантуса в зависимости от содержания их в почве представлено следующими уравнениями:

Эмпирическое уравнение регрессии накопления Zn в вегетативных органах растений:

$Y = 0,0022x^3 - 0,0448x^2 + 0,3284x - 0,369$ , с коэффициентом детерминации  $R^2 = 0,9515$ , Значение критерии Фишера  $F = 21,16$ .  $F > F_{кр}(0,05) = 4,07$ .

Эмпирическое уравнение регрессии накопления Pb в вегетативных органах растений:

$y = 0,012x^2 - 0,056x + 0,046$ , с коэффициентом детерминации  $R^2 = 0,9191$ . Значение критерии Фишера  $F = 30,64$ .  $F > F_{кр}(0,05) = 4,07$ .

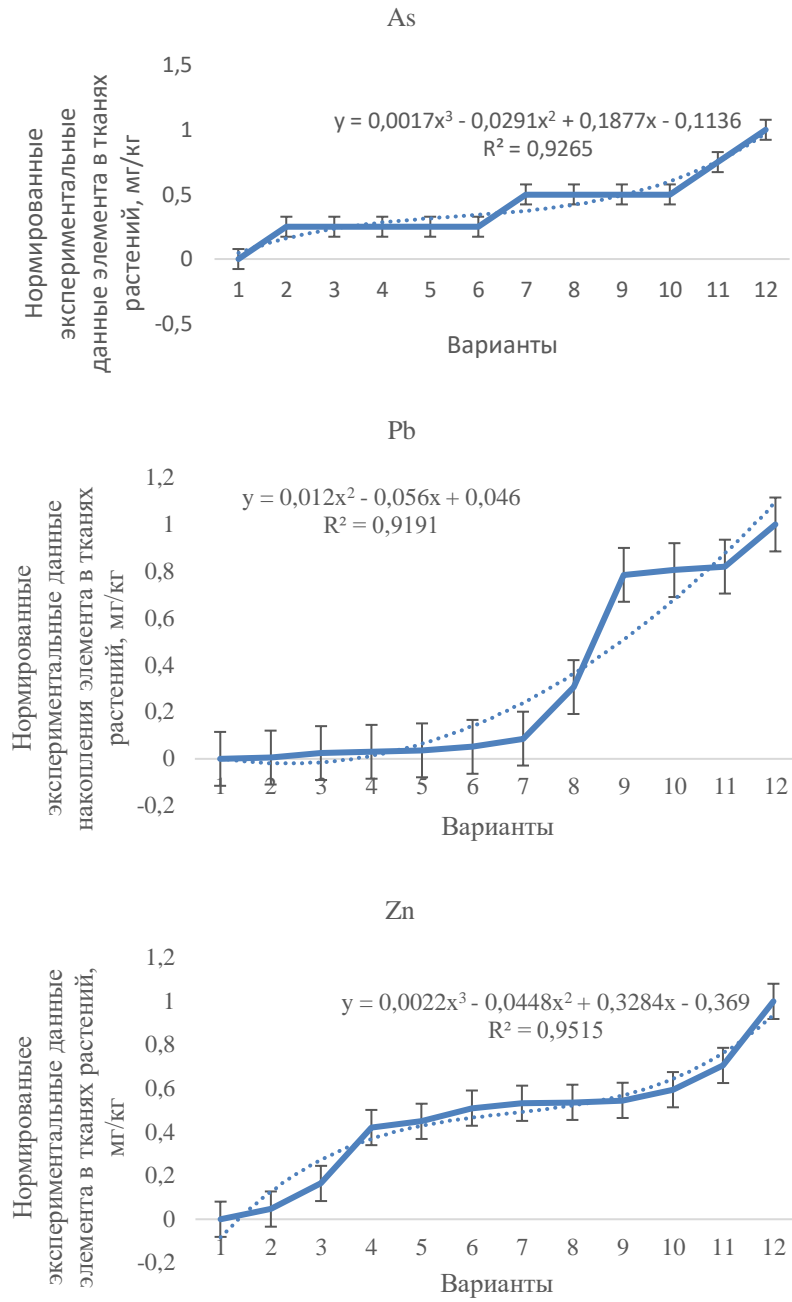
Эмпирическое уравнение регрессии накопления As в вегетативных органах растений:

$y = 0,0017x^3 - 0,0291x^2 + 0,1877x - 0,1136$ , с коэффициентом детерминации  $R^2 = 0,9265$ . Значение критерии Фишера  $F = 49,92$ .  $F > F_{кр}(0,05) = 4,07$ .

Зависимость концентрации тяжелых металлов 2 класса опасности (Co, Ni, Cu, Cr) в вегетативных органах мискантуса в зависимости от содержания их в почве представлено следующими уравнениями:

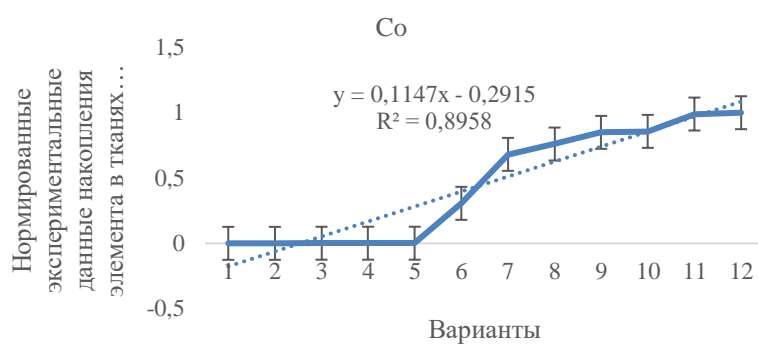
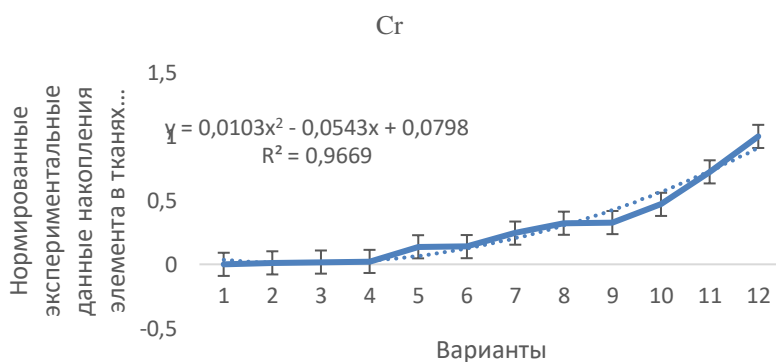
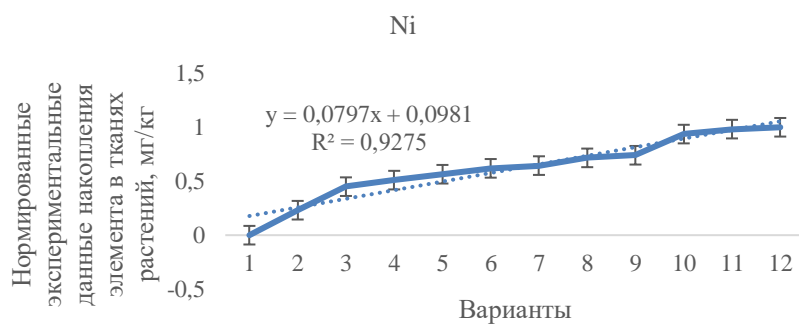
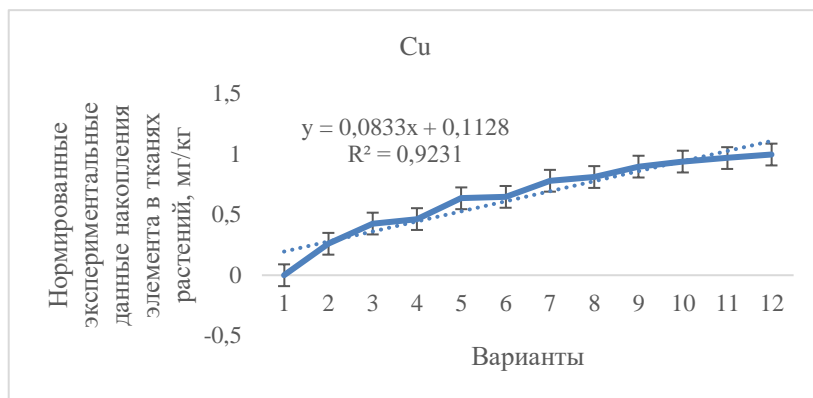
Эмпирическое уравнение регрессии накопления Co в вегетативных органах растений:

$y = 0,1147x - 0,2915$ , с коэффициентом детерминации  $R^2 = 0,8958$ , Значение критерии Фишера  $F = 95,94$ .  $F > F_{кр}(0,05) = 4,96$ .

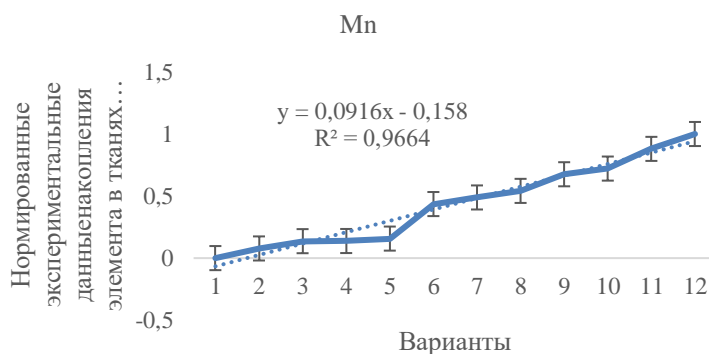
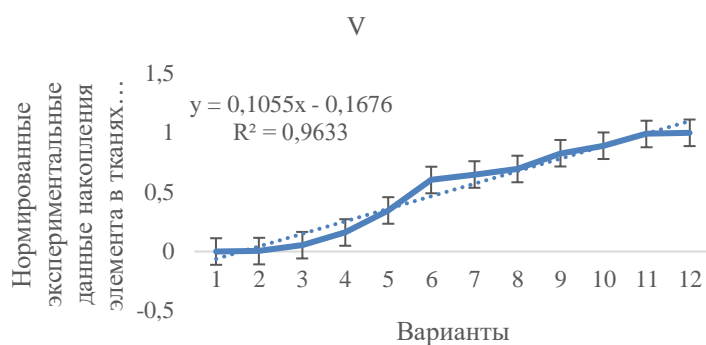
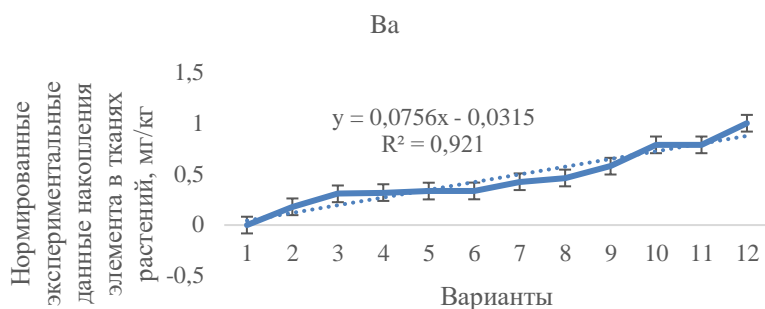
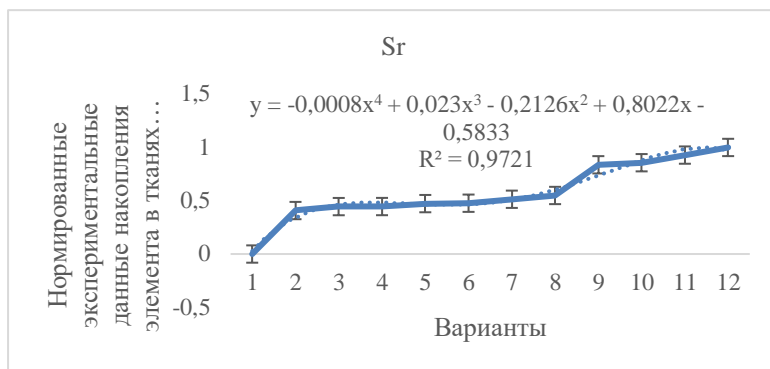


**Рисунок 2 – Нормированные экспериментальные данные накопления тяжелых металлов 1 класса опасности в вегетативных органах *Miscanthus x giganteus***





**Рисунок 3 – Нормированные экспериментальные данные накопления тяжелых металлов 2 класса опасности в вегетативных органах *Miscanthus x giganteus***



**Рисунок 4 – Нормированные экспериментальные данные накопления тяжелых металлов 3 класса опасности в вегетативных органах *Miscanthus x giganteus***

Эмпирическое уравнение регрессии накопления Ni в вегетативных органах растений:

$y = 0,0797x + 0,0981$ , с коэффициентом детерминации  $R^2 = 0,9275$ . Значение критерии Фишера  $F = 137,95$ .  $F > F_{кр}(0,05) = 4,96$ .

Эмпирическое уравнение регрессии накопления Cu в вегетативных органах растений:  $y = 0,0833x + 0,1128$ , с коэффициентом детерминации  $R^2 = 0,9231$ . Значение критерии Фишера  $F = 120,02$ .  $F > F_{кр} (0,05) = 4,96$ .

Эмпирическое уравнение регрессии накопления Cr в вегетативных органах растений:

$y = 0,0103x^2 - 0,0543x + 0,0798$ , с коэффициентом детерминации  $R^2 = 0,9669$ . Значение критерии Фишера  $F = 78,13$ .  $F > F_{кр} (0,05) = 4,07$ .

Зависимость концентрации тяжелых металлов 3 класса опасности (Ba, Mn, V, Sr) в вегетативных органах мискантуса в зависимости от содержания их в почве представлено следующими уравнениями:

Эмпирическое уравнение регрессии накопления Ba в вегетативных органах растений:

$y = 0,0756x - 0,0315$ , с коэффициентом детерминации  $R^2 = 0,9210$ . Значение критерии Фишера  $F = 126,50$ .  $F > F_{кр} (0,05) = 4,96$ .

Эмпирическое уравнение регрессии накопления V в вегетативных органах растений:

$y = 0,0916x - 0,158$ , с коэффициентом детерминации  $R^2 = 0,9633$ . Значение критерии Фишера  $F = 272,57$ .  $F > F_{кр} (0,05) = 4,96$ .

Эмпирическое уравнение регрессии накопления Mn в вегетативных органах растений:

$y = 0,0916x - 0,158$ , с коэффициентом детерминации,  $R^2 = 0,9664$ . Значение критерии Фишера  $F = 72,29$ .  $F > F_{кр} (0,05) = 4,07$ .

Эмпирическое уравнение регрессии накопления Sr в вегетативных органах растений:

$y = 0,0015x^3 - 0,0276x^2 + 0,2153x - 0,0671$ , с коэффициентом детерминации  $R^2 = 0,8981$

Значение критерии Фишера  $F = 15,45$ .  $F > F_{кр} (0,05) = 4,12$ .

Таким образом, данные расчеты свидетельствуют о том, что высокопродуктивный биоэнергетический вид *Miscanthus x giganteus* является фиторемедиантом, обладает способностью аккумулировать тяжелые металлы в вегетативных органах из загрязненной почвы и может быть использован в программах фиторемедиации. При изучении зависимости накопления 11 тяжелых металлов в вегетативных органах от содержания их в почве установлено, что коэффициенты детерминации статистически достоверны, они близки к значению 1. Значение критерии Фишера говорят о статистической надежности уравнений в целом.

### Список используемой литературы

1 Tsao D.T. Overview of phytotechnologies //In Advances in Biochemical Engineering/Biotechnology, Volume 78: Phytoremediation [Scheper, T., Tsao D. T., Editors]. Springer-Verlag, Berlin, Germany. – 2003. – P.1-50.

2 Prasad M.N. Practical use of plants for restoration soil polluted by metals // *Physiology plant*. – 2003. –Vol. 50, N 1. – 2014. – P. 764-777.

3 Baker A.J.M., McGrath, S.P., Reeves, R.D. Metal hyperaccumulator plants: a review of the ecology and physiology of a biochemical resource for phytoremediation of metal polluted soils // Contaminated soil and water. Boca-Raton, Fl., USA: Lewis Publishers. – 2000. – P. 85-107.

4 Pogrzeba M., Krzyzak J., Sas-Nowosielska A. Environmental hazards related to *Miscanthus x giganteus* cultivation on heavy metal contaminated soil // Environmental Heavy

Metal Pollution and Effects on Child Mental Development/ Risk Assessment and Prevention Strategies. Springer Science/ Business Media B.V. – 2011. – P. 213-225.

5 Pidlisnyuk B., Erickson L., Kharchenko S., Stefanovska T. Sustainable Land Management: Growing *Miscanthus* in Soils Contaminated with Heavy Metals // Journal of Environmental Protection, Special Issue in Environmental Remediation. – 2014. – Vol. 5, – P. 723-730.

6 Nsanganwimana F., Pourrut B., Mench M., Douay F. Suitability of *Miscanthus* species for managing inorganic and organic contaminated land and restoring ecosystem services. A review // Journal of Environmental Management. – 2014. – Vol. 143, –P. 123-134.

7 Heaton, E.A., Dohleman, F.G., Miguez, A.F., Juvik, J.A., Lozovaya, V., Widholm, J., Zobotin, O.A., Mcisaac, G.F., David, M.B., Voigt, T.B., Boersma, N.N., Long, S.P. *Miscanthus*: a promising biomass crop // *Advances in Botanical Research*, – 2010. – Vol. 56, –P. 75-137.

8 Thoma G.J., Lam T.B. Wolf D.C. A Mathematical Model of Phytoremediation for Petroleum-Contaminated Soil: Model Development // International Journal of Phytoremediation: – 2003. – Vol.5. No. 1. – P. 41–55.

9 Zung K.S., Corapcioglu M.Y., Drew M.C., Munster C.L. Plant contamination by organic pollutants in phytoremediation // Journal of Environmental Quality. – 2001. – Vol. 120 (1). – P.57-67.

10 Motesharezadeh B., Navabzadeh M., Liyaghat A. M. Modeling Phytoremediation of Cadmium Contaminated Soil with Sunflower (*Helianthus annuus*) Under Salinity Stress // International Journal of Environmental Research. – 2016. – Vol. 10(1). – P.109-118.

11 Mathur S. Modeling Phytoremediation of Soils // Practice Periodical of Hazardous Toxic and Radioactive Waste Management. – 2004. –Vol. 8.– P.286-297.

12 Lugli F., Mahler C.F. Phytoremediation of Metals: A Numerical Analysis // International Journal of Phytoremediation. – 2015. – Vol.17, N 3. –P. 242-248.

13 Водопьянов В.В. Математические модели и методы анализа восстановления биосистем, подверженных антропогенным воздействиям (на примере восстановления нефтезагрязненных почв) – автореферат диссертации на соискание ученой степени доктора технических наук. 2008. Уфа. – 48 с.

14 Narayanan M., Tracy J.C., Davis L., Erickson L. Modeling the fate of toluene in a chamber with alfalfa plants. 1 Theory and modeling concepts // Journal Hazardous Substance Research. – 1995. – 5a (1). –P.1-30.

15 Korostina O.A., Kuyukina M.S., Ivshina I.B., Kostina L.V., Osipenko M.A., Nyashin Y.I. Modelling of processes of filtration solutions of salts of heavy metals in soil and biosurfactant // Russian Journal of Biomechanics. – 2009. – V. 13, № 3 (45). – P. 32-43.

16 Korostina O.A., Kuyukina M.S., Ivshina I.B., Kostina L.V., Osipenko M.A., Nyashin Yu.I. The use of software for solving problems related to the purification of soil from oil and heavy metals through the biosurfactant // Russian Journal of Biomechanics. – 2011, – Vol. 15, N 3 (53). – P. 86-92.

# ЛЮБОЙ ФОРСИНГ-ТИП РЕАЛИЗУЕТСЯ В НЕКОТОРОЙ ФОРСИНГ-МОДЕЛИ И СЕМАНТИЧЕСКИЙ КРИТЕРИЙ E-ФОРМУЛ

Нуртазин А.Т., Хисамиев З.Г., Досанбай С.П.

Институт информационных и вычислительных технологий КН МОН РК  
e-mail: AbyeNurtazin@mail.ru , KhisamievZ@mail.ru

*Аннотация.* Вводится новое понятие форсинг типа и доказывается что любой форсинг тип реализуется в подходящей форсинг модели. Доказывается необходимое и достаточное семантические условие для формулы быть эквивалентной некоторой E-формуле или  $\forall$ -формуле относительно данной теории. В качестве непосредственного следствия получен критерий Робинсона о модельно полной теории.

## 1. Любой форсинг-тип реализуется в некоторой форсинг-модели

Ранее в [1] автор показал, что экзистенциальные типы, реализующиеся в экзистенциально замкнутых структурах максимальны. Этот факт обусловил их важность при изучении этого класса структур. Из [2] следует, что элементарная теория любого форсинг-компаньона полна, а экзистенциальные типы, реализуемые в форсинг-структурах не только максимальны, но и имеют единственные расширения до полных. В [3] автор назвал такие экзистенциальные типы форсинг-типами. Там же было доказано, что в случае существования простая модель над кортежом, реализующим форсинг-тип, является форсинг-моделью, и любой форсинг-тип можно реализовать в некоторой структуре, являющейся моделью форсинг-теории. Очевидно, ценность введённого там понятия значительно возрастёт, если будет доказано, что любой форсинг-тип может быть реализован в некоторой форсинг-модели. Именно доказательству этого факта посвящена предлагаемая статья.

Как уже отмечалось метод вынуждения Пола Коэна [4] для теории моделей впервые был адаптирован и использован Абрахамом Робинсоном [2]. Следует отметить, что определение Робинсона является "калькой" первоначального определения Коэна и носит явно выраженную синтаксическую форму. Между тем это определение сделано более естественным одним из классиков теории моделей выдающимся французским логиком Роландом Фрэсе в [5], где он использовал фундаментальное и чисто семантическое понятие класса всех конечных структур, вложимых данную. Теперь мы можем ввести фундаментальное семантическое понятие, часто используемое во многих современных исследованиях по теории моделей и играющее основную методическую роль в предлагаемой работе.

*Определение 1.*

Классом Фрэсе в данной счётной реляционной сигнатуре  $\Sigma$  называется любой абстрактный, наследственно замкнутый и имеющий свойство совместного вложения класс конечных  $\Sigma$  структур  $\mathfrak{F}$ .

Заметим, что в частности классом Фрэсе оказывается любой класс  $\mathfrak{F}(\mathcal{M}) = \text{Age}(\mathcal{M})$ . Более того, легко видеть, что для любого класса Фрэсе  $\mathfrak{F}$  найдётся структура  $\mathcal{M}$  такая, что выполняется равенство:  $\mathfrak{F} = \mathfrak{F}(\mathcal{M})$ .

*Определение 2.*

Мы называем две  $\Sigma$  структуры  $\mathcal{M}$  и  $\mathcal{N}$  компаньонами, если совпадают их классы Фрэсе:  $\mathfrak{F}(\mathcal{M}) = \mathfrak{F}(\mathcal{N})$ .

Класс всех компаньонов данной  $\Sigma$  структуры  $\mathcal{M}$  мы называем компаньон-классом  $C(\mathcal{M})$ . Так как с другой стороны данный компаньон-класс целиком определяется классом Фрэсе  $\mathfrak{F}$  любой содержащейся в нём структуры, мы также обозначаем его  $C(\mathfrak{F})$  и называем компаньон-классом структур, построенных из  $\mathfrak{F}$ .

Теперь нашей целью является семантическое определение вынуждения. Так как используемое здесь фундаментальное понятие класса Фрэсе является чисто семантическим, то и исходному классическому определению Коэна - Робинсона мы также придадим эквивалентную ему максимально семантическую форму. При этом мы считаем, что все рассматриваемые здесь логические формулы имеют пренексную форму, в которой все кванторы расположены в начале формулы. Известно, что тогда число следующих один за другим блоков одноимённых кванторов в начале такой формулы называется её кванторной сложностью. Теперь вынуждение формул естественно определяется индукцией по их кванторной сложности.

*Определение 3.*

Для структуры  $\mathcal{A} = \langle A; \Sigma \rangle$  из класса Фрэсе  $\mathfrak{F}$ , лежащего в ней кортежа

$\mathbf{a}$  и  $\sigma$  формулы  $\varphi(\mathbf{x})$  отношение " $\mathcal{A} \Vdash^{\mathfrak{F}} \varphi(\mathbf{a})$  в классе  $\mathfrak{F}$ " определяется индукцией по кванторной сложности  $q(\varphi)$  формулы  $\varphi$ .

1.  $q(\varphi) = 0$ .

$$\mathcal{A} \Vdash^{\mathfrak{F}} \varphi(\mathbf{a}) \Leftrightarrow \mathcal{A} \models \varphi(\mathbf{a}) \quad (\varphi(\mathbf{a}) \text{ истинна в } \mathcal{A}).$$

2.  $q(\varphi) = k + 1$  и  $\varphi(\mathbf{x}) = \exists x^1 \psi(\mathbf{x}, x^1)$

$$\mathcal{A} \Vdash^{\mathfrak{F}} \varphi(\mathbf{a}) \Leftrightarrow \text{в } A \text{ найдется кортеж } \mathbf{a}^1, \text{ что } \mathcal{A} \Vdash^{\mathfrak{F}} \exists x^1 \psi(\mathbf{a}, \mathbf{a}^1)$$

3.  $q(\varphi) = k + 1$  и  $\varphi(\mathbf{x}) = \forall x^1 \psi(\mathbf{x}, x^1)$

$$\mathcal{A} \Vdash^{\mathfrak{F}} \varphi(\mathbf{a}) \Leftrightarrow \text{не существует расширения } \mathcal{B} \text{ структуры } \mathcal{A} \text{ вынуждающей формулу } \exists x^1 \neg \psi(\mathbf{a}, x^1)$$

В приведённом определении "вынуждения" используется индукция по числу блоков кванторов в пренексном виде вынуждаемой формулы. Между тем можно предположить, что чаще может быть удобным эквивалентное приведённому определению вынуждения по числу  $Q(\varphi)$  связанных переменных в пренексном виде вынуждаемой формулы. Если это слегка модернизированное "вынуждение" обозначить  $\Vdash^*$ , то выполняется:

*Лемма 1.*

Для любой структуры  $\mathcal{A}$  из класса Фрэсе  $\mathfrak{F}$  и формулы  $\varphi(\mathbf{a})$  выполняется

$$\mathcal{A} \Vdash^* \varphi(\mathbf{a}) \Leftrightarrow \mathcal{A} \Vdash^{\mathfrak{F}} \varphi(\mathbf{a}).$$

Пусть  $p(\mathbf{x})$  -- максимальный совместный экзистенциальный тип в компаньон-теории  $T^c$ . Тогда для фиксированного  $\mathbf{a}$  и произвольных  $\mathbf{b}$  и конечных  $\sigma \subseteq \Sigma$  класс структур  $\mathfrak{F}(p) = \{ \langle \mathbf{ab}; \sigma \rangle : \exists y \mathcal{D}_{\mathfrak{F}}(\mathbf{x}, y) \in p(\mathbf{x}) \}$  оказывается подклассом  $\mathfrak{F}$  и обладает

свойством совместного вложения. Напомним, что тип  $p(\mathbf{x})$  называется форсинг-типом, если для любой формулы  $\varphi(\mathbf{a})$  в классе  $\mathfrak{F}(p)$  найдётся структура  $\mathcal{B}$ , вынуждающая в  $\mathfrak{F}$  формулу  $\varphi(\mathbf{a})$  или её отрицание. Так как класс  $\mathfrak{F}(p)$  обладает свойством совместного вложения, то для любой формулы  $\varphi(\mathbf{x})$  в этом классе найдётся конечная структура, вынуждающая в  $\mathfrak{F}$  либо формулу  $\varphi(\mathbf{a})$ , либо её отрицание. Это замечание оправдывает обозначение  $(p)$  для единственного полного расширения типа  $p(\mathbf{x})$  в форсинг-теории  $T^f$ .

Важную роль в доказательстве теоремы будет играть следующая

*Лемма.*

Для любой форсинг-типа от  $n$  переменных  $p(\mathbf{x})$  и принадлежащей  $(p)$  формулы вида  $\exists y \psi(\mathbf{x}; y)$  найдётся форсинг-расширение  $q(\mathbf{x}; y)$  типа  $p(\mathbf{x})$ , содержащее формулу  $\exists y \psi(\mathbf{x}; y)$ .

Для доказательства возьмём последовательность

$\psi_0(\mathbf{x}; y) \equiv \psi(\mathbf{x}; y), \dots, \psi_n(\mathbf{x}; y), \dots$  всех формул, совместных с типом  $p(\mathbf{x})$ . Тогда по определению форсинг-типа найдётся кортеж  $\mathbf{abb}$  и конечная сигнатура  $\sigma_0$  такие, что в классе  $\mathfrak{F}$  выполняется  $\langle \mathbf{abb}; \sigma_0 \rangle \Vdash \neg \psi(\mathbf{a}, b)$ .

Продолжая далее по индукции, считаем, что после  $n$ -го шага определены формулы

$$\psi_0(\mathbf{x}; y) \equiv \psi(\mathbf{x}; y), \dots, \psi_n(\mathbf{x}; y) \text{ и структуры}$$

$$\langle \mathbf{abb}^0; \sigma_0 \rangle \Vdash \neg \psi(\mathbf{a}, b), \dots, \langle \mathbf{abb}^n; \sigma_n \rangle \Vdash \neg \psi(\mathbf{a}, b),$$

вынуждающие предложения или их отрицания:  $\psi_0(\mathbf{x}; y) \equiv \psi(\mathbf{x}; y), \dots, \psi_n(\mathbf{x}; y)$  в классе  $\mathfrak{F}$ .

Тогда проверяем, вынуждает ли структура  $\langle \mathbf{abb}^n; \sigma_n \rangle$  формулу  $\forall y \neg \psi_{n+1}(\mathbf{a}, y)$  и в этом случае полагаем  $\langle \mathbf{abb}^{n+1}; \sigma_{n+1} \rangle = \langle \mathbf{abb}^n; \sigma_n \rangle$ . В противном случае найдётся расширение  $\langle \mathbf{abb}^{n+1}; \sigma_{n+1} \rangle$  структуры  $\langle \mathbf{abb}^n; \sigma_n \rangle$ , вынуждающее формулы  $\mathcal{D}_{\sigma_n}(\mathbf{a}, b, b^n)$  и  $\psi_{n+1}(\mathbf{a}, b)$ . Непосредственная проверка показывает, что экзистенциальный тип  $q(\mathbf{x}; y)$  кортежа  $\mathbf{ab}$  опять оказывается форсинг-типом.

Лемма доказана.

Воспользовавшись доказанной леммой теперь достаточно просто задать индуктивный процесс построения канонической структуры, которая оказывается форсинг-моделью теории  $T^f$ .

## Литература

1. Нуртазин А.Т. Счетные экзистенциально замкнутые модели универсально аксиоматизируемых теорий // Математические труды. – Новосибирск, Институт математики СО РАН им. Соболева С.Л., 2015. – т.18, №1. – С.48-97.
2. Barwise J., Robinson A. Completing theories by forcing// Ann. Math. Logic.– 1970.–№ 2.– Р. 119-142.
3. Нуртазин А.Т. Вынуждение в структурах и классах Фрэсе// Алгебра и логика, 2018.– № 4.

4. Cohen P.J., Set Theory and the Continuum Hypothesis.– New-York., Benjamin, 1966.– 467 p.
5. Fraisse R. Sur quelques classifications des systemes de relations// Publ. scient. de l'univ. d'Algers, 1955.– A1.– P.35-182.
6. Ершов Ю.Л., Палютин Е.А., Тайманов А.Д. Теория моделей.– Справочная книга по математической логике, Часть 1.– М "Наука", 1982.– 492с.
7. Нуртазин А.Т. Некоторые свойства экзистенциально замкнутых компаньонов// Алгебра и логика, 2018.– № 2.
8. Robinson A. On the Metamathematics of Algebra.– Amsterdam: North – Holland, 1951.– 540с.

## 2. Семантический критерий E-формул

Пусть  $T$  – произвольная теория сигнатуры  $\sigma$ ,  $\varphi(\bar{x})$  – произвольная формула сигнатуры  $\sigma$ ,  $\bar{x}=(x_1, \dots, x_n)$  кортеж переменных содержащих все свободные переменные формулы  $\varphi$ .

*Определение 1.*

Говорим, что формула  $\varphi(\bar{x})$   $T$ -устойчива относительно расширений, если для любых моделей  $\mathcal{M}, \mathcal{N}$  теории  $T$ , таких что  $\mathcal{M}$  подмодель  $\mathcal{N}$  выполняется  $\varphi(\mathcal{M}) \subseteq \varphi(\mathcal{N})$ , где для модели  $\mathcal{A}=\langle A, \sigma \rangle$ ,  $\varphi(\mathcal{A})=\{\bar{a} \in A \mid \mathcal{A} \models \varphi(\bar{a})\}$ .

*Определение 2.*

Говорим что формулы  $\varphi$  и  $\psi$   $T$ -эквивалентны, если выполнено  $T \models \varphi \sim \psi$ .

*Теорема 1.*

Формула  $\varphi(\bar{x})$   $T$ -устойчива относительно расширений если и только если  $\varphi(\bar{x})$   $T$ -эквивалентна некоторой E-формуле.

Доказательство. Достаточность. Очевидно, что E-формулы  $T$ -устойчивы относительно расширений.

Необходимость. Пусть  $\{\varphi_i(\bar{x}) \mid i \in I\}$  – множество всех E-формул сигнатуры  $\sigma$ , таких что  $\varphi(\bar{x})$  является логическим следствием в теории  $T$  из каждой  $\varphi_i(\bar{x})$ . Покажем, что некоторое  $\varphi_0(\bar{x})$  является логическим следствием  $\varphi(\bar{x})$  в теории  $T$ . Допустим противное, тогда для любого  $i \in I$  множество предложений  $T \cup \{\varphi(\bar{a}), \neg \varphi_i(\bar{a})\}$  в сигнатуре  $\sigma \cup \{a\}$  совместно. Отсюда и из того, конъюнкция конечного подмножество предожений из  $\{\neg \varphi_i(\bar{a}) \mid i \in I\}$  эквивалентна некоторому предложению  $\neg \varphi_j(\bar{x}), j \in I$ , следует что множество предложений  $T \cup \{\varphi(\bar{a})\} \cup \{\neg \varphi_i(\bar{a}) \mid i \in I\}$  локально совместно, и следовательно, оно совместно. Тогда, пусть  $\mathcal{M}$  модель теории  $T$ , в которой найдется кортеж  $\bar{a} \in \mathcal{M}$  такой, что для любого  $i \in I$  выполнено  $\mathcal{M} \models \varphi(\bar{a}) \& \neg \varphi_i(\bar{a})$ . Рассмотрим теперь множество предложений  $T \cup D(\mathcal{M})$ , где  $D(\mathcal{M})$  диаграмма модели  $\mathcal{M}$ . По выбору  $\bar{a} \in \mathcal{M}$  выполнено  $\mathcal{M} \models \varphi(\bar{a})$ . В силу  $T$ -устойчивости относительно расширений формулы  $\varphi(\bar{x})$  имеем  $T \cup D(\mathcal{M}) \models \varphi(\bar{a})$ . Тогда,  $T \cup D(\bar{a}, \bar{b}) \models \varphi(\bar{a})$  для некоторого  $\bar{b}=(b_1, \dots, b_m) \in \mathcal{M}$ . Отсюда,  $\mathcal{M} \models \varphi_0(\bar{a}) \& \varphi(\bar{a})$ , где  $\varphi_0(\bar{x}) = \exists y D(\bar{x}, y)$  необходимость а вместе с ней теорема доказана.



Двойственным образом можно рассмотреть понятие  $T$ -устойчивости относительно подмоделей.

*Определение 3.*

Говорим, что формула  $\varphi(\bar{x})$   $T$ -устойчива относительно подмоделей, если для любых моделей  $\mathcal{M}, \mathcal{N}$  теории  $T$ , таких что  $\mathcal{M}$  подмодель  $\mathcal{N}$  выполняется  $\varphi(N) \cap M \subseteq \varphi(M)$ .

*Предложение.*

Формула  $\varphi(\bar{x})$  сигнатуры  $\sigma$   $T$ -устойчива относительно подмоделей тогда и только тогда, когда  $\neg\varphi(\bar{x})$   $T$ -устойчива относительно расширений.

Доказательство. Необходимость. Пусть  $\mathcal{M}, \mathcal{N}$  модели теории  $T$ ,  $\mathcal{M}$  подмодель  $\mathcal{N}$  и пусть  $\bar{a} = (a_1, \dots, a_n) \in \neg\varphi(M)$ . Допустим, что  $\bar{a} \notin \neg\varphi(N)$ , тогда  $\bar{a} \in \varphi(N) \cap M$ , отсюда в силу  $T$ -устойчивости  $\varphi(\bar{x})$  относительно подмоделей имеем  $\bar{a} \in \varphi(M)$ . Противоречие, необходимость доказана. Достаточность. Пусть  $\mathcal{M}, \mathcal{N}$  модели теории  $T$ ,  $\mathcal{M}$  подмодель  $\mathcal{N}$  и пусть  $\bar{a} \in \varphi(N) \cap M$ . Допустим, что  $\bar{a} \notin \varphi(M)$ , тогда  $\bar{a} \in \neg\varphi(M)$ , отсюда в силу  $T$ -устойчивости  $\neg\varphi(\bar{x})$  относительно расширений имеем  $\bar{a} \in \neg\varphi(N)$ . Противоречие, предположение доказано.

*Теорема 2.*

Формула  $\varphi(\bar{x})$   $T$ -устойчива относительно подмоделей если и только если  $\varphi(\bar{x})$   $T$ -эквивалентна некоторой  $\forall$ -формуле.

Доказательство. Достаточность. Очевидно, что  $\forall$ -формулы  $T$ -устойчивы относительно подмоделей. Необходимость. Пусть  $\varphi(\bar{x})$   $T$ -устойчива относительно подмоделей, тогда из предыдущего предложения следует, что  $\neg\varphi(\bar{x})$   $T$ -устойчива относительно подмоделей. По из теоремы 1 следует, что  $\neg\varphi(\bar{x})$   $T$ -эквивалентна некоторой  $\exists$ -формуле. Отсюда,  $\varphi(\bar{x})$   $T$ -эквивалентна некоторой  $\forall$ -формуле. Теорема доказана.

Следствием теорем является следующая

*Теорема 3 (Робинсон [1],[2]).* Теория  $T$  модельно полна тогда и только тогда, когда любая формула  $\varphi(\bar{x})$   $T$ -эквивалентна некоторой  $\forall$ -формуле и некоторой  $\exists$ -формуле.

Доказательство. Достаточность. Пусть  $\varphi(\bar{x})$  – произвольная теория сигнатуры  $\sigma$ , по условию она  $T$ -эквивалентна некоторой  $\forall$ -формуле  $\psi(\bar{x})$  и некоторой  $\exists$ -формуле  $\theta(\bar{x})$ . Пусть  $\mathcal{M}, \mathcal{N}$  модели теории  $T$ ,  $\mathcal{M}$  подмодель  $\mathcal{N}$  и пусть  $\bar{a} \in M$ . Тогда имеем  $\mathcal{M} \models \varphi(\bar{a}) \Leftrightarrow \mathcal{M} \models \theta(\bar{a}) \Rightarrow \mathcal{N} \models \theta(\bar{a}) \Rightarrow \mathcal{N} \models \varphi(\bar{a})$ .  
 $\mathcal{M} \models \varphi(\bar{a}) \Leftrightarrow \mathcal{M} \models \psi(\bar{a}) \Rightarrow \mathcal{M} \models \psi(\bar{a}) \Rightarrow \mathcal{M} \models \varphi(\bar{a})$ .

Таким образом,  $\mathcal{M}$  элементарная подмодель  $\mathcal{N}$ . Достаточность доказана. Необходимость. Если теория модельно полна, то любая формула  $\varphi(\bar{x})$   $T$ -устойчива относительно расширений и подмоделей, следовательно  $\varphi(\bar{x})$   $T$ -эквивалентна некоторой  $\forall$ -формуле и некоторой  $\exists$ -формуле. Необходимость а вместе с ней теорема доказана.

## **Литература**

1. Кейслер Х.Д. Основы теории моделей/ Справочная книга по математической логике, 1982. – Ч.І. – с.55–108.
2. Robinson A. Complete Theories/ Amsterdam: North-Holland, 1956.

## **ИССЛЕДОВАНИЯ СТАТИСТИЧЕСКИХ СВОЙСТВ РАЗРАБОТАННОГО ГЕНЕРАТОРА ПСЕВДОСЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ**

**Нысанбаева С.Т., Капалова Н.А., Дюсенбаев Д.С., Алгазы К.Т.**

*Институт информационных и вычислительных технологий МОН РК  
Лаборатория информационной безопасности  
e-mail: sultasha1@mail.ru, Kapalova@ipic.kz, dimash\_dds@mail.ru,  
kunbolat@mail.ru*

***Аннотация.** В работе рассмотрен генератор псевдослучайных последовательностей, построенный на мультипликативной операции, и исследованы статистические свойства сгенерированных последовательностей. Приведены также результаты анализа статистических характеристик сгенерированных последовательностей, полученные по графическим и оценочным статистическим тестам и аналитико-вычислительными методами.*

### **Введение**

Генераторы псевдослучайных последовательностей (ПСП) являются неотъемлемым элементом любой системы защиты. Под псевдослучайной последовательностью будем понимать последовательность, полученную с помощью детерминированного алгоритма. Сгенерированная детерминированным алгоритмом последовательность не является случайной. Убедиться в том, что некоторая последовательность чисел обладает свойствами случайной последовательности можно либо при помощи специальных статистических тестов, либо аналитико-вычислительными методами. Именно от свойств сгенерированных последовательностей зависят надежность и безопасность процессов сбора, обработки, передачи и хранения информации, а также управления объектами. В связи с этим к формируемым для криптографических приложений последовательностям предъявляются самые жесткие требования.

Под генерацией последовательности случайных чисел подразумевается получение битовой последовательности, состоящей из двоичных знаков 0 и 1.

В настоящее время существуют различные пакеты тестов, позволяющих сделать выводы о свойствах рассматриваемых последовательностей. Среди тестов наиболее известными являются тесты Diehard, NIST, Д. Кнута и др.

*Принципы построения генераторов ПСП.* Качественный генератор ПСП, ориентированный на использование в системах защиты информации, должен удовлетворять следующим требованиям [1]:

- криптографическая стойкость;

– хорошие статистические свойства, ПСП по своим статистическим свойствам не должна отличаться от истинно случайной последовательности;

– большой период формируемой последовательности: например при шифровании для преобразования каждого элемента входной последовательности необходимо использовать свой элемент псевдослучайной гаммы;

– эффективная аппаратная и программная реализация.

При использовании криптостойкого генератора ПСП три следующие задачи для противника должны быть вычислительно неразрешимы:

– определение  $(i-1)$ -го элемента  $g_{i-1}$  последовательности на основе известного фрагмента гаммы  $g_i g_{i+1} g_{i+2} \dots g_{i+b+1}$  конечной длины  $b$ ;

– определение  $(i+1)$ -го элемента  $g_{i+1}$  последовательности на основе известного фрагмента гаммы  $g_{i-b+1} \dots g_{i-2} g_{i-1} g_i$  конечной длины  $b$ ;

– определение ключевой информации по известному фрагменту гаммы конечной длины.

Справедливо также следующее утверждение: непредсказуемый влево генератор ПСП является криптостойким. Криптоаналитик, знающий принцип работы такого генератора, имеющий возможность анализировать фрагмент  $g_i g_{i+1} g_{i+2} \dots g_{i+(t-1)}$  выходной последовательности, но не знающий используемой ключевой информации, для определения предыдущего выработанного элемента последовательности  $g_{i-1}$  не может предположить лучшего способа, чем выбор по жребию.

В рамках другого подхода к построению качественного генератора ПСП предлагается свести задачу построения криптографически сильного генератора к задаче построения статистически безопасного генератора ПСП, который должен удовлетворять следующим требованиям [1]:

– ни один статистический тест не обнаруживает в ПСП каких-либо закономерностей, иными словами, не отличает эту последовательность от истинно случайной;

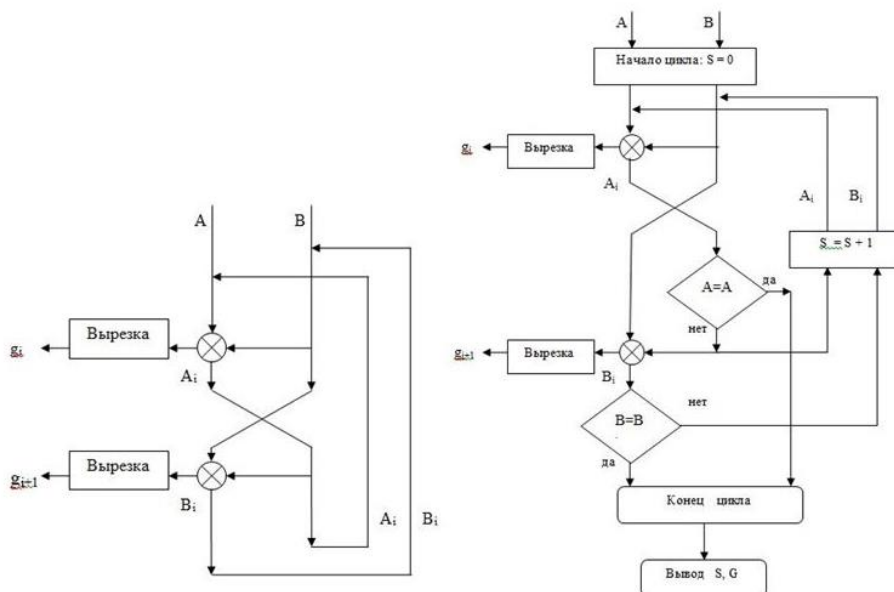
– нелинейное преобразование  $F_k$ , зависящее от секретной информации (ключа  $k$ ), используемое для построения генератора, обладает свойством «размножения» искажений, – все выходные (преобразованные) элементы возможны и равновероятны независимо от исходного элемента;

– при инициализации случайными значениями генератор порождает статистически независимые ПСП.

### 1. Алгоритм генерации псевдослучайных последовательностей

Выбираются два произвольных числа  $A$  и  $B$  и находится их произведение. Из полученного результата  $A_1 = A \cdot B$  вырезается заданное количество любых подряд идущих разрядов, которое назовем вырезкой. Первая вырезка сохраняется как первый элемент гаммы  $g_1$ . Затем  $A_1$  умножается на  $B$  и из полученного  $B_1 = A_1 \cdot B$  выделяется аналогично второй элемент гаммы  $g_2$ . На этом завершается первый шаг создания части ключевой гаммы в виде  $\{g_1, g_2\}$ . Второй шаг описанной процедуры выполняется аналогично:  $A_1$  умножается на  $B_1$  и из полученного  $A_2 = A_1 \cdot B_1$  вырезается, как указано выше, элемент  $g_3$ . Далее перемножаются  $A_2$  и  $B_1$ , из

результата  $B_2 = A_2 \cdot B_1$  вырезается  $g_4$  и т.д. (рисунок 1). В итоге циклично вырабатывается ключевая гамма необходимой длины  $G = \{g_1, g_2, g_3, g_4, \dots\}$  [3].



**Рисунок 1 – а) схема алгоритма генерации ПСП, б) схема генератора с проверкой на периодичность**

В алгоритме есть возможность задать входные параметры генератора в соответствии с требованиями к степени безопасности: длину генерируемого элемента (вырезки), местоположения производимой вырезки и длину гаммы (число произведенных вырезок).

Для генератора схема проверки повторяемости элементов гаммы  $(g_i, g_{i+1}, \dots)$ ,  $i = 1, 2, \dots$  показана на рисунке 1(б). После каждого цикла исходные данные  $A$  и  $B$  сверяются с текущими данными  $A_i$  и  $B_i$ ,  $i = 1, 2, \dots$ , соответственно. В случае совпадения этих пар чисел процесс генерации останавливается с выводом сообщения о количестве выполненных циклов по  $i$ , так как это совпадение фиксирует образование цикличности элементов последовательности, начиная с первого.

## 2. Оценка смещения вероятности появления нуля или единицы

Сгенерированные последовательности проверялись на подборке тестов Д. Кнута. Результаты приведены в разделе 3. Рассмотрим результаты операции умножения и вычислим вероятность появления «0» или «1» в каждом разряде. Пусть  $A = a_n a_{n-1} a_{n-2} \dots a_2 a_1 a_0$ ,  $B = b_m b_{m-1} b_{m-2} \dots b_2 b_1 b_0$  и их произведение  $C = c_{n+m-1} c_{n+m-2} \dots c_2 c_1 c_0$  ( $C = A * B$ ). Тогда

$$\begin{aligned}
 c_0 &= a_0 b_0 \\
 c_1 &= a_0 b_1 + a_1 b_0 \\
 c_2 &= a_0 b_2 + a_1 b_1 + a_2 b_0 \\
 &\dots \dots \dots \\
 c_{n+m-1} &= a_n b_m
 \end{aligned} \tag{1}$$

Будем вычислять вероятности принятия «0» и «1» для  $c_i, i = 0, 1, \dots, n + m - 1$ , что позволит выяснить подходящий интервал для вырезки  $g$  и равномерность распределения последовательности. В общем случае для произведения двух чисел с длинами  $n$  и  $m$ , получаем таблицу размером  $n \times m$ . Например, для двух чисел, состоящих из двух битов каждое, получаем таблицу размером  $4 \times 4$  (таблица 1).

**Таблица 1. Умножение чисел, состоящих из двух битов**

	00	01	10	11
00	0000	0000	0000	0000
01	0000	0001	0010	0011
10	0000	0010	0100	0110
11	0000	0011	0110	1001

Как видно из таблицы 1, вероятности принятия  $c_i = \langle 1 \rangle$  следующие:

$$P(c_0 = 1) = \frac{1}{4}, P(c_1 = 1) = \frac{3}{8}, P(c_2 = 1) = \frac{3}{16}, P(c_3 = 1) = \frac{1}{16}$$

Теперь вычислим эти вероятности, воспользовавшись представлением (1). При этом необходимо учесть разрядный перенос. Поэтому запишем выражение (1) следующим образом:

$$\begin{aligned} c_0 &= a_0 b_0 \\ c_1 &= a_0 b_1 \oplus a_1 b_0 \\ c_2 &= a_0 b_2 \oplus a_1 b_1 \oplus a_2 b_0 \oplus R_1, \text{ где } R_1 = a_0 b_1 a_1 b_0 \\ &\dots\dots\dots \\ c_{n+m-1} &= a_n b_m \oplus R_{n+m-2} \end{aligned}$$

Далее вычислим вероятность того, что  $c_i$  принимает значение «1» в конечных выражениях.

$$P(c_0 = 1) = P(a_0 b_0 = 1) = P(a_0 = 1)P(b_0 = 1) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}.$$

$$\begin{aligned} P(c_1 = 1) &= P(a_0 b_1 \oplus a_1 b_0 = 1) = P(a_0 b_1 = 0, a_1 b_0 = 1) + P(a_0 b_1 = 1, a_1 b_0 = 0) = \\ &= \frac{3}{4} \cdot \frac{1}{4} + \frac{1}{4} \cdot \frac{3}{4} = \frac{3}{8}. \end{aligned}$$

$$\begin{aligned} P(c_2 = 1) &= P(a_1 b_1 \oplus a_0 b_0 a_1 b_1 = 1) = P(a_1 b_1 = 0, a_0 b_0 a_1 b_1 = 1) + \\ &+ P(a_1 b_1 = 1, a_0 b_0 a_1 b_1 = 0) = P(a_1 b_1 = 0)P_{a_1 b_1 = 0}(a_0 b_0 a_1 b_1 = 1) + \\ &+ P(a_1 b_1 = 1)P_{a_1 b_1 = 1}(a_0 b_0 a_1 b_1 = 0) = 0 + \frac{1}{4} \cdot \frac{3}{4} = \frac{3}{16}. \end{aligned}$$

$$P(c_3 = 1) = P(a_0 b_0 a_1 b_1 = 1) = P(a_0 = 1)P(b_0 = 1)P(a_1 = 1)P(b_1 = 1) = \frac{1}{16}.$$

Как видно, вероятности  $c_i$ , найденные в обоих случаях, одинаковы. Однако, по мере роста длины чисел, строить таблицы подобные таблице 1, становится затруднительным. После определенных значений это выходит за рамки возможностей

вычислительных машин. Поэтому для получения соответствующих вероятностей выбирается второй способ.

Запишем  $c_i$  в общем виде:  $c_i = \bigoplus \sum_{k+l=i} a_k b_l$

Если,  $P(a_i = 0) = P(a_i = 1) = P(b_i = 0) = P(b_i = 1) = \frac{1}{2}$  то, понятно, что  $P(a_k b_l = 0) = \frac{3}{4}$  и  $P(a_k b_l = 1) = \frac{1}{4}$ , здесь  $1 \leq k \leq n, 1 \leq l \leq m$ .

Сначала рассмотрим случай, когда разряд не переносится. Тогда можно использовать одну из следующих формул:

$$P(\bigoplus \sum_{k,l} a_k b_l = 0) = \frac{1}{2} + 2^{n-1} \prod_{i=1}^n \varepsilon_i,$$

$$P(\bigoplus \sum_{k,l} a_k b_l = 0) = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} C_n^{2k} p^{2k} q^{n-2k}.$$

Полученные значения по обеим формулам одинаковы.

Рассмотрим случай, когда длины А и В равны 16 бит, тогда получим следующие результаты:

0 :- 0,75  
 1 :- 0,375  
 2 :- 0,5625  
 3 :- 0,46875  
 4 :- 0,515625  
 5 :- 0,4921875  
 6 :- 0,50390625  
 7 :- 0,498046875  
 8 :- 0,5009765625  
 9 :- 0,49951171875  
 10 :- 0,500244140625  
 11 :- 0,4998779296875  
 12 :- 0,50006103515625  
 13 :- 0,499969482421875  
 14 :- 0,500015258789063  
 15 :- 0,499992370605469  
 16 :- 0,500015258789063  
 17 :- 0,499969482421875  
 18 :- 0,50006103515625  
 19 :- 0,4998779296875  
 20 :- 0,500244140625  
 21 :- 0,49951171875  
 22 :- 0,5009765625  
 23 :- 0,498046875  
 24 :- 0,50390625  
 25 :- 0,4921875  
 26 :- 0,515625  
 27 :- 0,46875  
 28 :- 0,5625  
 29 :- 0,375  
 30 :- 0,75

Смещение вероятности от 0,5 для интервала получаемое согласно алгоритму генерации с 12-го по 19-й биты составляет 0,0001-0,0002.

Теперь рассмотрим случаи с разрядным переносом. Тогда,

$c_i = \bigoplus \sum_{k+l=i} a_k b_l \oplus R_{i-1}$ , где  $R_{i-1}$  значение, получаемое из младших разрядов.

Вероятность переноса разрядов вычисляется по формуле:

$$P(R_{i-1} = 1) = C_{i-1}^2 \left(\frac{1}{4}\right)^2 \left(\frac{3}{4}\right)^{i-3} + C_{i-1}^3 \left(\frac{1}{4}\right)^3 \left(\frac{3}{4}\right)^{i-4} + \dots + C_{i-t}^{2t} \left(\frac{1}{4}\right)^{2t} \left(\frac{3}{4}\right)^{i-3t} + \\ + C_{i-t}^{2t+1} \left(\frac{1}{4}\right)^{2t+1} \left(\frac{3}{4}\right)^{i-3t-1}$$

Как показано в примере, в случае с разрядным переносом значения зависят от предыдущих разрядов. Вероятность вычисляется по формуле условной вероятности. Учитывая вышесказанное, вычисляем вероятность появления битов в каждом разряде для умножения двух 16-битных чисел.

- 0 :- 0,25
- 1 :- 0,375
- 2 :- 0,5546875
- 3 :- 0,478515625
- 4 :- 0,507568359375
- 5 :- 0,497698783874512
- 6 :- 0,500541388988495
- 7 :- 0,499943965813145
- 8 :- 0,499967868275249
- 9 :- 0,500025807115716
- 10 :- 0,499989382267968
- 11 :- 0,500002292975503
- 12 :- 0,500000558034211
- 13 :- 0,499999017675514
- 14 :- 0,500000715106658
- 15 :- 0,499999591625646
- 16 :- 0,500000209075493
- 17 :- 0,499999588443418
- 18 :- 0,500000553402205
- 19 :- 0,499999204665269
- 20 :- 0,500001327650118
- 21 :- 0,499998081047055
- 22 :- 0,500001185053007
- 23 :- 0,500005354598003
- 24 :- 0,49997269635432
- 25 :- 0,500072756039334
- 26 :- 0,499901445600516
- 27 :- 0,49981475594015
- 28 :- 0,501909335733366
- 29 :- 0,49143905376191
- 30 :- 0,529313072562218

31 :- 0,41473388671875  
32 :- 0,28125  
33 :- 0

Здесь, смещение вероятности от 0,5 для интервала с 12-го по 19-й биты составляет 0,0001-0,0002.

В обоих случаях отклонение вероятности появления «0» или «1» от 0,5 очень малы. При выборе больших чисел в качестве А и В (более 12 бит) получаем ПСП с положительной характеристикой.

### 3. Результаты тестов Д. Кнута

Исследования статистической безопасности генераторов псевдослучайных чисел, предназначенных для решения задач криптографической защиты информации.

Построена эффективная система оценки качества генераторов ПСП. Используются наиболее известные подборки оценочных тестов:

- одна из самых первых систем для оценки свойств ПСП – «DIEHARD».
- подборка тестов Д. Кнута – математические основы, на которых строятся практически все системы оценки качества ПСП.
- специализированная система, разработанная Национальным институтом стандартов и технологий (НИСТ) США, ориентированная на исследование криптографических генераторов ПСП.
- система CRYPT-X
- графические тесты, используемые при оценке статистических свойств генераторов ПСП.

Результаты тестирования на подборке тестов Д. Кнута приведены в таблице 2.

Таблица 2. Результаты тестов Д. Кнута

№п.п	Название теста	Хи-квадрат	Результаты
1	Проверка 0 и 1	0,0162	Прошел
2	Проверка несцепленных серий	7,60	Прошел
3	Посимвольная проверка	222,44	Прошел
4	Проверка интервалов	1,93	Прошел
5	Проверка комбинаций	1,57	Прошел
6	Тест собирания купонов	3,13	Прошел
7	Проверка перестановок	15,07	Прошел
8	Проверка на монотонность	8,8	Прошел
9	Проверка корреляции	Корреляции нет	Прошел
10	Проверка линейной сложности	29,015	Прошел

### Заключение

Результаты статистического тестирования показали, что исследуемые последовательности, сгенерированные с помощью предложенного алгоритма, проходят все тесты из подборки Д. Кнута.

В качестве значений  $A_i$  и  $B_i$  во втором и последующих этапах можно использовать вырезку заданной длины из произведения  $A_{i-1}$  на  $B_{i-1}$  на участке между 8 и  $n+m-9$  разрядами. По мере приближения вырезки к центру двоичного представление произведения, отклонение от 0,5 вероятности появления «0» или «1»



стремится к нулю. Поэтому оптимальным размещением вырезки является центральная часть двоичного представление произведения  $A_i$  и  $B_i$ .

Как сказано выше, одним из требований к качеству генераторов ПСП является криптографическая стойкость. Зная любой фрагмент последовательности, выдаваемой генератором, злоумышленник не должен иметь эффективной возможности найти начальное значение, загруженное в генератор и не должен иметь возможности получить достоверную информацию о предыдущих или последующих элементах последовательности. Результаты по данному направлению будут предметом рассмотрения следующей статьи.

### **Литература**

1. Иванов М. А., Чугунков И. В., Теория, применение и оценка качества генераторов псевдослучайных последовательностей. - М.: К-ОБРАЗ, 2003. - 136 с.
2. Бияшев Р.Г., Нысанбаева С.Е., Капалова Н.А. Разработка и исследование алгоритма генерации псевдослучайных последовательностей // Вестник КазНУ. Сер. «Мат., мех., информат.» - 2008.- № 4 (59), - С. 138-143.
3. Капалова Н.А., Нысанбаева С.Е. Анализ статистических свойств алгоритма генерации псевдослучайных последовательностей: Матер. X Междунар. науч.-практ. конф. // Информационная безопасность. Ч. 2. – Таганрог: Изд-во ТТИ ЮФУ, 2008. - С. 169-172.

## **ДОСТИЖЕНИЕ КОНСЕНСУСА В СОВРЕМЕННЫХ ДЕЦЕНТРАЛИЗОВАННЫХ СИСТЕМАХ**

**Нысанбаева С.Е., Нюсупов А.Т.**

*Институт информационных и вычислительных технологий КН МОН РК,  
Казахстан  
e-mail: sultasha1@mail.ru*

### **Аннотация.**

*В данной работе проведен обзор алгоритмов для достижения консенсуса в децентрализованных системах и распределенном реестре, которые обеспечивают высокий уровень безопасности при обработке и передачи информации. Рассмотрены основные понятия децентрализованных систем и их разновидности. Особое внимание уделено распределённому реестру. Показаны классические алгоритмы для достижения консенсуса, среди которых выделены алгоритмы на основе Blockchain технологии. Отмечены преимущества и недостатки децентрализованных систем. Показана важность применения подобных систем для повышения безопасности передачи данных в современных информационных сетях и системах.*

**Ключевые слова.** *децентрализованное сети, распределённый реестр, консенсус, blockchain*

## **Введение**

В современном мире широко используются различные децентрализованные системы для обеспечения отказоустойчивости, безопасности, целостности и доступности при распределённой обработке информации. Под децентрализованными, одноранговыми или пиринговыми (peer-to-peer, P2P) подразумеваются системы, в которых отсутствуют узлы, управляющие работой других узлов [1]. В отличие от систем вида «клиент-сервер» каждый узел данной сети выполняет роль как клиента, так и сервера. Данный подход позволяет предотвратить возможное несанкционированное изменение конфиденциальных данных пользователей на стороне сервера.

Одноранговые системы разрабатываются для решения различных задач и проблем. Большинство P2P систем предназначены для:

- обмена информацией;
- создания криптовалютных систем;
- распределенного вычисления.

Одним из распространённых видов децентрализованных систем при обмене информацией между пользователями являются Torrent системы. В подобных системах множество узлов участвуют в отправке и принятии данных. Узлы между собой соединяются посредством треккера – специализированного сервера, который работает по протоколу НТТР (Hyper Text Transfer Protocol). В нем хранятся IP адреса клиентов и хэш-суммы для идентификации клиентов данной сети. Torrent системы позволяют обмениваться информацией различного вида (документы видео, фото и т.п.).

В настоящее время активно развиваются также криптовалютные системы, которые используются пользователями для выполнения финансовых операций. В основе подобных систем лежит технология распределенного реестра. Данная технология позволяет хранить, передавать и получать доступ к информационным ресурсам в общей базе данных. Реестр – это распределенная цифровая книга, в котором применяются криптографические алгоритмы для контроля процессов создания и передачи цифровых активов [2]. В реестре важным является достижение консенсуса для устранения рисков двойных трат криптовалюты и повышения отказоустойчивости и конфиденциальности данных. Консенсус является важной составляющей для поддержания целостности распределённого реестра.

## **Алгоритмы консенсуса в распределённом реестре**

Существует большое количество алгоритмов для достижения консенсуса в распределённом реестре. Теория консенсуса в распределённых системах сформировалась в 80-х годах XX века, когда американский ученый Лесли Лампорт предложил алгоритм [3] для решения «задачи византийских генералов» (задача взаимодействия нескольких удаленных узлов, которые получили команды из главного центра). Данное решение используется для частного случая, в котором не учитывается динамическое изменение узлов сети. Решение задачи для общего случая стало возможным с появлением в 2008 году технологии распределенного реестра – цепочки блоков (Blockchain [4]), в составе децентрализованной платежной системы Bitcoin [5].

Наиболее известными алгоритмами консенсуса являются: Paxos [6], Chandra–Toucg [7], Raft [8], и алгоритмы, входящие в технологию Blockchain [9]. Paxos [6] – это гибкий и отказоустойчивый протокол консенсуса, который используется в приложениях для согласования общих значений между распределенными участниками. В данном протоколе применяется три класса агентов: поставщик (proposer), получатель (acceptor) и обучающийся (learner). Принцип работы алгоритма

основывается на взаимосвязи агентов, в котором предложенные поставщиками значения, обрабатываются получателями с целью принятия или отклонения конкретного предложения. Все принятые решение запоминаются обучающимся агентом. При взаимодействии агентов обеспечивается безопасность консенсуса и гарантируется выбор только одного значения, а также сохраняется целостность данных при передаче. Chandra–Toueg [7] – алгоритм для решения задач консенсуса в ненадежных распределённых системах. Данный алгоритм распознает неудачные процессы с помощью модуля детектора отказов. Модуль содержит лист предполагаемых процессов, которые могут быть нарушены. Процесс заносится в лист подозрительных в случае, когда количество ошибок превышает установленную норму. В результате ошибочные процессы исключаются из работы системы, что способствует достижению целостности взаимодействия всех процессов в системе. Raft [8] является гибким алгоритмом достижения консенсуса, который отличается от Paxos значительно упрощенной структурой при сохранении схожей эффективности. Основным преимуществом Paxos является его понятность в освоении и легкость создания практических распределённых систем. Алгоритм основан на изменениях состояний узлов сети в заданной последовательности, что гарантирует надёжную доставку сигналов. Устойчивость системы достигается благодаря механизму выбора приоритетного узла – лидера. Если лидер долгое время не принимает и отправляет сигналы, то приоритетным становится другой узел, который определяется путем голосования.

Актуальным и современным решением достижения консенсуса в децентрализованных системах являются алгоритмы, лежащие в основе технологии распределённого реестра Blockchain. К наиболее известным относятся: подтверждение выполнения работы (Proof of Work, PoW [9]) и подтверждение доли (Proof of Stake, PoS [10]). Принцип достижения консенсуса PoW основывается на выполнении сложной задачи вычисления, результаты которого позволяют участвовать в валидации совершенных транзакций (операция обмена информационными ресурсами). За проведённые вычисления полагается вознаграждение, которое мотивирует пользователей принимать участие в распределённом консенсусе. Потенциальной уязвимостью алгоритма является возможность завладения злоумышленником 51% вычислительных ресурсов всей сети, с помощью которой он может изменять истории транзакций в свою пользу. Аналогичным алгоритмом является PoS, в котором вместо подтверждения работы используется доказательство владения заданным количеством ценных информационных ресурсов, как правило, электронной валюты. Подтверждение наличия таких ресурсов дает право участвовать в доле консенсусе, равной от общей доли всех участников сети.

### **Заключение**

Большинство из основных алгоритмов достижения консенсуса в децентрализованных системах и распределённых реестрах обеспечивают высокую степень безопасности в области хранения и передачи информации и характеризуются конфиденциальностью, целостностью и доступностью данных. Рассмотрены важные аспекты развития технологии распределённого реестра, которая позволяет значительно улучшить существующие системы финансовых институтов, оптимизировать бизнес процессы и создать прочную основу для проведения электронных платежей.

Достоинствами алгоритмов распределённого консенсуса в децентрализованных системах являются:

- обеспечение конфиденциальности, целостности и доступности данных при передаче информации;
- возможность создания эффективных отказоустойчивых систем;
- отсутствие двойных трат информационных ресурсов в криптовалютной системе.

Таким образом, децентрализованные системы на основе различных алгоритмов консенсуса имеют высокий потенциал развития в будущем, так как заложенные в них методы и технологии могут гибко подстраиваться и модифицироваться под различные условия.

### **Список литературы**

1. Raval S. Decentralized Applications: Harnessing Bitcoin's Blockchain Technology. – O'Reilly Media, Inc., 2016. – 118 с.
2. Pinna A., Ruttenberg W. Distributed Ledger Technologies in Securities Post-Trading Revolution or Evolution? // ECB Occasional Paper. – 2016. – №172. – С. 1-35.
3. Lamport L., Shostak R., Pease M. The Byzantine generals problem // ACM Transactions on Programming Languages and Systems (TOPLAS). – 1982. – Т. 4. – №. 3. – С. 382-401.
4. Dhillon V., Metcalf D., Hooper M. Blockchain Enabled Applications: Understand the Blockchain Ecosystem and How to Make it Work for You. – Apress: Berkeley, 2017. – 218 с.
5. Nakamoto S. Bitcoin: A peer-to-peer electronic cash system. – 2008. – С. 1-10.
6. Meling H., Jehl L. Tutorial summary: Paxos explained from scratch // International Conference on Principles of Distributed Systems. – Springer, 2013. – С. 1-10.
7. Hossain M. A. Performance Analysis of Chandra-Toueg Consensus Algorithm with S Failure Detector. – Kent State University, 2000. – С. 1-3.
8. Ongaro D., Ousterhout J. K. In search of an understandable consensus algorithm // USENIX Annual Technical Conference. – 2014. – С. 305-319.
9. Arati B. Understanding blockchain consensus models // Persistent Systems Ltd, Tech. Rep. – 2017. – с. 14.
10. Bentov I., Gabizon A., Mizrahi A. Cryptocurrencies without proof of work // International Conference on Financial Cryptography and Data Security. – Springer, 2016. – С. 142-157.

## **КРИПТОГРАФИЧЕСКАЯ ЗАЩИТА В АВТОМАТИЗИРОВАННЫХ СИСТЕМАХ**

**Нысанбаева С.Е., Усатова О.А.**

*Институт информационных и вычислительных технологий КН МОН РК,  
Республика Казахстан  
e-mail: guoa\_olgq@mail.ru*

***Аннотация.** Рассматриваются различные подходы, используемые при разработке средств криптографической защиты цифровой информации.*

Высокая степень автоматизации, широкое внедрение компьютерных систем в различные сферы человеческой деятельности делают автоматизированные системы обработки данных чрезвычайно уязвимыми по отношению к деструктивным воздействиям, ставят общество в зависимость от степени безопасности используемых информационных технологий. Вследствие этого важной характеристикой любой автоматизированной системы независимо от ее сложности и назначения становится безопасность циркулирующей в ней информации. Это также относится и к отдельным людям и группам людей, которые используют возможность скрывать свои реальные намерения для получения конкурентного преимущества и снижения уязвимостей.

Защита конфиденциальной информации от несанкционированного доступа (НСД) является одной из самых древних. Отметим некоторые из факторов, которые определяют трудоемкость решения задач защиты информации:

- увеличение объемов информации, накапливаемой, хранимой и передаваемой с помощью компьютерной техники;
- появление новых транспортных средств и носителей передачи данных;
- расширение круга пользователей, имеющих доступ к ресурсам компьютерных систем;
- усложнение режимов функционирования технических средств компьютерной системы;
- повсеместное распространение новых инфокоммуникационных технологий;
- увеличение количества технических средств и связей в автоматизированных системах обработки данных.

Чтобы уменьшить последствия от воздействия НСД, необходимо создание и использование системы безопасности. Цель создания такой системы - предупреждение последствий умышленных и случайных деструктивных воздействий, следствием которых могут быть разрушение, модификация или утечка информации. В этом случае эффективная система безопасности должна обеспечивать:

- методы обеспечения физической безопасности компонентов системы;
- ограничение доступа;
- разграничение доступа;
- разделение доступа (привилегий) – разрешение доступа только при одновременном предъявлении полномочий всех членов группы;
- криптографическое преобразование информации и реализованные на его основе криптографические протоколы.

Среди множества организационных, методологических и технических мер особое место занимают методы криптографической защиты информации, без которой практически не обходится ни одна автоматизированная система. Криптография – это наука защиты информации путем ее преобразования в нечитаемый вид, способ хранения и передачи данных, позволяющий лишь уполномоченным лицам (процессам) читать и обрабатывать их. Первые методы криптографии использовались людьми при письме на дереве или камне, которые передавались другим людям, имевшим все необходимое для расшифрования сообщений. На сегодняшний день криптография применяется в потоках двоичного кода инфокоммуникационных систем.

В настоящее время существует целый ряд алгоритмов шифрования данных, основными из которых являются:

- подстановка, перестановка, замена;
- процедуры сети Фейстеля;

- симметричные (на пример криптосистемы DES, AES, IDEA);
- асимметричные (на пример криптосистемы RSA, DSA, Эль-Гамала, ECDSA, ГОСТ Р 34.10-2012, обмен ключами Диффи — Хелмана) [1].

Асимметричные алгоритмы шифрования характеризуются тем, что для зашифровывания информации используют один ключ (открытый), а для расшифровывания - другой (секретный). Эти ключи различны и не могут быть получены один из другого. Данное свойство асимметричных систем шифрования позволяет в какой-то мере решить проблему распределения ключей между пользователями, которое является основным недостатком симметричных систем. Для того чтобы гарантировать защиту данных, к асимметричным системам шифрования предъявляется два важнейших требования [2]:

- преобразование открытого текста должны быть необратимыми и исключать его восстановление на основе открытого ключа;
- определение секретного ключа на основе открытого должно быть невозможным для современного уровня развития вычислительных средств.

Однако следует отметить, что алгоритмы асимметричных криптосистем более трудоемки и медленней по сравнению с симметричными алгоритмами, что на практике рационально их использовать там, где объем зашифрованной информации незначителен, но очень важен. Это свойство асимметричных шифров и определило области применения таких систем — как правило, с их помощью осуществляется, распределение ключей, аутентификация пользователей за счет применения электронной подписи [3].

Современные инфотелекоммуникационные системы характеризуются высокой скоростью передачи информации. Использование симметричных криптографических алгоритмов позволяет быстрее зашифровывать и расшифровывать цифровые данные. Симметричные алгоритмы шифрования основаны на том, что отправитель и получатель информации используют чаще всего один и тот же ключ. Этот ключ должен храниться в секрете и передаваться способом, исключающим его перехват [4].

Симметричные криптографические алгоритмы делятся на две основные группы: блочные и поточные шифры. Блочные шифры производят криптографические преобразования с блоками открытой информации. Как правило, рекомендуемая длина блока определяется в стандартах шифрования. Применение процедур перемешивания и рассеивания символов позволяет обеспечить равномерность распределения символов (битов) в зашифрованном блоке данных. Применение простейших операций перестановок, подстановок и циклических сдвигов позволяет реализовать данные процедуры с использованием микропроцессорных вычислительных устройств.

Блочные алгоритмы шифрования являются более медленными по сравнению с поточными алгоритмами. Поточные шифры - алгоритмы побитного или побайтного шифрования. Эти системы обеспечивают более высокое быстродействие процесса зашифрования и расшифрования, однако они не обладают достаточной степенью стойкости [3].

Автоматизация и цифровизация приводят к росту угроз несанкционированного доступа к информации, и как следствие, к необходимости постоянной поддержки и развития систем защиты. Повышение производительности вычислительной техники и появление новых видов атак на шифры ведет к понижению стойкости известных криптографических алгоритмов. Это приводит к необходимости создания и использования различных подходов при разработке криптографических средств обеспечения информационной безопасности [5].

В Институте информационных и вычислительных технологий активно занимаются вопросами защиты информации. Разработка криптографических средств защиты цифровой информации проводится сотрудниками лаборатории Информационной безопасности. Ими разработаны системы шифрования и электронной цифровой подписи с использованием непозиционных полиномиальных систем счисления (НПСС, ее синонимы - модулярная арифметика, системы счисления в остаточных классах). Проводится так же исследование надежности разработанных криптосистем различными методами криптоанализа. В лаборатории осуществляется программная реализация созданных криптосистем, в настоящее время начаты работы по проектированию и разработке программно-аппаратных средств их реализации.

### **Литература**

1. Бабаш, А. В. История криптографии. Часть I / А.В. Бабаш, Г.П. Шанкин. - М.: Гелиос АРВ, 2016. - 240 с.
2. Панасенко С.П. Алгоритмы шифрования. Специальный справочник/ С.П. Панасенко.- Санкт-Петербург «БХВ - Петербург», 2009.- 576с.
3. Бабенко, Л. К. Современные алгоритмы блочного шифрования и методы их анализа / Л.К. Бабенко, Е.А. Ищукова. - М.: Гелиос АРВ, 2015. - 376 с.
4. Рябко Б.Я. Криптографические методы защиты информации/ Рябко Б.Я., Фионов А.Н. М.: Горячая линия – телеком, 2005.- 229с.
5. Шаньгин, В. Ф. Информационная безопасность и защита информации / В.Ф. Шаньгин. - Москва: Огни, 2016. - 551 с.

## **КОНЦЕПЦИЯ ЦИФРОВОЙ ИНТЕЛЛЕКТУАЛЬНОЙ МЕДИЦИНЫ**

**Пак А.А., Джаксылыкова А.Б.**

*Институт информационных и вычислительных технологий КН МОН РК,  
Казахстан*

*e-mail: aa.pak83@gmail.com, aselya17.89@mail.ru*

***Аннотация.** Согласно государственной программе «Цифровая модернизация здравоохранения» необходимыми услугами инновационного Казахстана, являются услуги «смарт-медицины», а именно дистанционная профилактика и лечение, «электронной медицины». В свою очередь предиктивное моделирование на основе электронных записей здравоохранения (ЭЗЗ) определяющий драйвер развития «смарт-медицины», что в перспективе позволит улучшить качество жизни. Разработка предиктивных статистических моделей обычно требует извлечение некоторых рекомендаций в качестве предсказываемой переменной из нормализованных ЭЗЗ, что является трудозатратным процессом, который также приводит к значительной потере информации о пациенте. Мы считаем, что рассматриваемый в настоящей работе подход может быть использован для того, чтобы создать точные масштабируемые предикторы для широкого круга медицинских сценариев, причем предикторы могут быть дополнены объяснениями, которые подчеркивают доказательства диагноза в карте пациента.*

## 1. Введение

Феномен электронной медицины (ЭМ) связан с возможностью оцифровки данных о здоровье пациента, которые далее могут быть обработаны и проанализированы с целью улучшения ухода и диагностики. По факту, объем регулярно собираемых данных о состоянии здоровья пациентов обладает экспоненциальным ростом и является одним из самых быстрорастущих сегментов цифровой вселенной [1]. К сожалению, большая часть этой информации просто хранится на серверах и не используется для создания интеллектуальных информационных систем, которые могут использовать клиницисты для улучшения оказания помощи и повышения уровня жизни пациентов. В мировом научном сообществе существует гипотеза, что синтез машинного обучения и медицинских данных может обеспечить значительные выгоды не только для безопасности и качества ухода за пациентами, но и для экономики государств [2, 3, 4, 5].

Несмотря на богатство и потенциал имеющихся данных, существуют определенные сложности в создании интеллектуальных информационных систем в приложении к медицине. Общеизвестным мнением является тот факт, что 80% усилий в создании аналитических моделей является предварительная обработка (извлечение, трансформация, загрузка и очистка) наборов данных [6, 7], что серьезно ограничивает масштабируемость прогностических моделей.

Дополнительной проблемой разработки является вопрос защиты персональных данных [13], ведь чаще всего доступные базы представлены в ненормализованном виде с указанием ФИО и место жительства пациента, что подпадает под закон РК от 21 мая 2013 № 94-V «О персональных данных и их защите».

Другим важным локальным прикладным аспектом разработки подобного рода приложений является разнородность самих данных, обусловленная большим количеством локальных игроков на рынке CRM решений для медицины [14-16].

Далее количество потенциальных предикторов в электронных медицинских записях (EHR) может исчисляться в тысячах, особенно для общедоступных данных. Традиционные подходы к моделированию подобной сложности путем выбора ограниченного числа входных переменных зарекомендовали себя малоэффективными [8].

Однако, было показано что разработки в области глубокого обучения и искусственных нейронных сетей могут позволить решить многие из этих проблем и «разблокировать» информацию в EHR [9, 10, 11, 12]. Эти архитектуры известны своей способностью обрабатывать большие объемы относительно грязных данных, включая ошибки в ярлыках и большого количества входных переменных (ака проклятие размерности). Ключевым преимуществом является тот факт, что исследователям обычно не нужно указывать, какие потенциальные предикторы модель должна использовать при обучении. Более того, технологии, реализующие такое преимущество, выделяют в отдельное направление известное как обучение представлением (Representation Learning, RL).

В настоящей работе будет рассмотрена оптимальное хранилище медицинских данных в приложении к задачам разработки предикторов широкого спектра медицинских сценариев.

Работа имеет следующую структуру, в разделе 2 представлено резюме государственной программы «Цифровой модернизации здравоохранения». В разделе 3 рассмотрена унифицированная модель представления данных для обработки истории болезни пациента с критическим анализом.



## 2. Цифровая модернизация здравоохранения

Документ «Архитектура интероперабельности» выпущенный РГП на ПХВ «Республиканский центр электронного здравоохранения» Министерства здравоохранения Республики Казахстан от 2017 представляет собой основной архитектурный концепт программы «Цифровая модернизация здравоохранения». Основным лейтмотивом является понятие интероперабельности, иными словами, приведение разнородных информационных компонент здравоохранения к общему знаменателю, с целью унификации и гармонизации их взаимодействия. Для обеспечения интероперабельности взаимодействующих информационных подсистем на пространстве электронного здравоохранения Республики Казахстан все компоненты должны быть способны обмениваться информацией, используя стандартизированные методы, форматы, стандарты. В концепции представлена архитектура и модель интероперабельности, а также определён набор стандартов, а именно CDA R2, HL7 v3, HL7 FHIR. XML, eXML, призванных обеспечить синтаксическую совместимость данных, и общий подход к решению задачи обеспечения интероперабельности взаимодействующих информационных подсистем [19].

Основным архитектурным элементом концепции является сущность «Электронного паспорта здоровья» — это центральный механизм сбора, обработки, хранения и обмена медицинской информацией между участниками системы здравоохранения, включающими поставщиков медицинских услуг, пациентов и медицинские информационные системы. Элементарными информационными сущностями являются медицинские документы, приведенные в таблице 1.

**Таблица 1. Реестр документов, используемый в «Архитектура интероперабельности» [18]**

№	Шаблоны документов	Clinical Document Templates
1	Медицинское свидетельство о рождении	Birth Certificate
2	Медицинское свидетельство о смерти	Death Certificate
3	Медицинское свидетельство о перинатальной смерти	Perinatal Death Certificate
4	Карта учета материнской смертности	Maternal Death Certificate
5	Карта учета родившегося живым, мертворожденного и умершего ребенка в возрасте до 5 лет	Liveborn, Stillborn and Deceased Child of up to 5 years Certificate
6	Минимальный набор медицинских данных о пациенте	Patient Summary
7	Направление	Referral
8	Неструктурированный документ	Unstructured Document
9	Согласие на прикрепление	Consent On Attachment
10	Согласие на процедуру	Consent On Procedure
11	Отказ	Dissent
12	Выписка	Discharge Summary
13	Полномочия	Mandate

14	Рецепт	Prescription
15	Отпуск лекарственных средств	Dispense
16	Отказ от госпитализации	Hospitalization Dissent
17	Дополнение к направлению	Referral Addendum
18	План лечения	Care plan
19	Консультация специалиста	Consultation Note

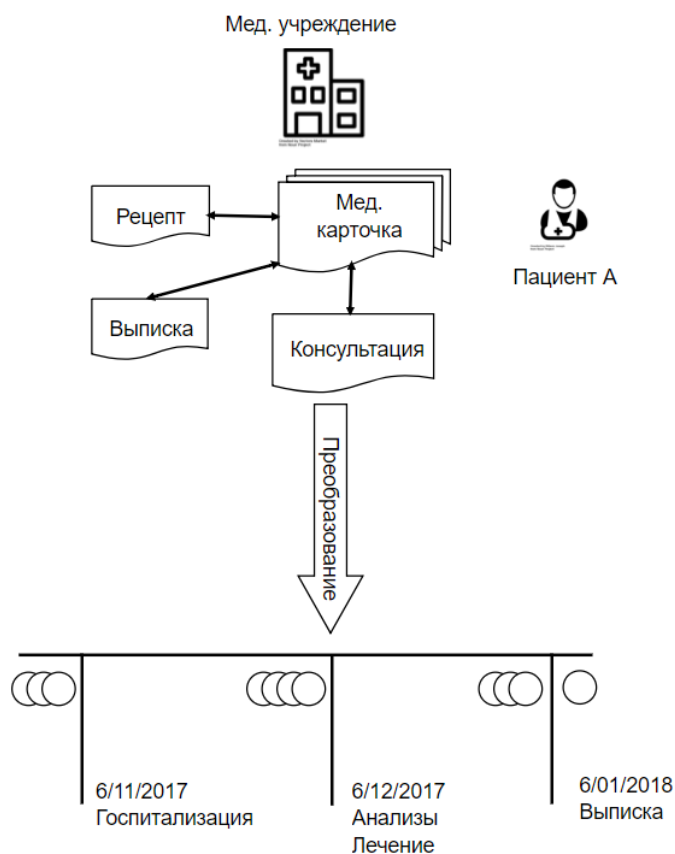
Документы, представленные в таблице 1, имеют различные мета-поля, которые описывают дату создание, дату изменения, авторов и др., помимо метаинформации, каждый документ декларирует некоторое событие в истории болезни пациента, например, проведенный анализ или процедуру, назначение лекарств или получение результатов после консультации у специалиста и др. Далее каждое событие имеет свой свойства, которые могут быть описаны либо количественной переменной, либо категориальной переменной. Все количественные переменные имеют собственные шкалы значений, например, температура живого человек принадлежит интервалу 35.9 – 37.2 С, кровяное давление 80-140 мм. рт. ст., частота сердечных сокращений в покое 60-90 уд./мин., тоже верно и для других измеряемых физиологических, биохимических показателей. Прием лекарств также можно описать с помощью таких свойств как длительность, дозировка, наименование. Жалобы пациента можно свести к бинарным или категориальным признакам.

Все вышеизложенные события и признаки содержатся в неструктурированных или плохо-структурированных документах. В следующем разделе мы опишем формат данных необходимый для унификации и возможности применения методов машинного обучения к медицинским сценариям.

### 3. Представление данных и обработка

Мы разработали единую структуру данных, которая может использоваться для прогнозирования всех возможных медицинских сценариев, тем самым мы предлагаем решить проблему ручного создания наборов данных для каждого нового прогнозирования. Этот подход представляет собой ЭЗЗ во временном порядке: данные организованы по пациентам и по времени. Наиболее подходящим представлением событий в истории болезни пациента является стандарт FHIR [19].

Использование свободного текста для прогнозирования позволяет получить новый уровень объяснения прогнозов. Клиницисты исторически не доверяли нейросетевым моделям из-за их непрозрачности. Предложенная архитектура данных может визуализировать, какие факты модель «рассматривала» для каждого отдельного диагноза, эту информацию может использовать клиницист, чтобы получить интуицию при принятии решения, иными словами, для клинициста важно при принятии решения понимать на основании каких фактов экспертная система рекомендует тот или иной исход. Тем не менее, существуют и другие возможные методы интерпретации моделей глубокого обучения [20-21]. Мы представляем концепцию преобразования данных в рамках «Архитектуры интероперабельности» [18] для увеличения доступности медицинских данных в исследовательских целях.



**Рисунок 1. Концептуальная диаграмма архитектуры преобразование неструктурированных данных**

### Литература

1. The Digital Universe: Driving Data Growth in Healthcare. <https://www.emc.com/analyst-report/digital-universe-healthcare-vertical-report-ar.pdf> (Accessed: 02-05-2018)
2. D. W. Bates et al. “Big data in health care: using analytics to identify and manage high-risk and high-cost patients”. In: Health Aff. 33.7 (2014), pp. 1123–1131.
3. J. Larry Jameson and Dan L Longo. “Precision medicine—personalized, problematic, and promising”. en. In: N. Engl. J. Med. 372.23 (June 2015), pp. 2229–2234.
4. Harlan M Krumholz. “Big data and new knowledge in medicine: the thinking, training, and tools needed for a learning health system”. en. In: Health Aff. 33.7 (July 2014), pp. 1163–1170.
5. Ravi B Parikh, J Sanford Schwartz, and Amol S Navathe. “Beyond Genes and Molecules - A Precision Delivery Initiative for Precision Medicine”. en. In: N. Engl. J. Med. 376.17 (Apr. 2017), pp. 1609–1612.
6. Gil Press. Cleaning Big Data: Most Time-Consuming, Least Enjoyable Data Science Task, Survey Says. <https://www.forbes.com/sites/gilpress/2016/03/23/data-preparationmost-time-consuming-least-enjoyable-data-science-task-survey-says/>. Accessed: 02-05-2018. Mar. 2016.
7. Steve Lohr. “For Big-Data Scientists, ‘Janitor Work’ Is Key Hurdle to Insights”. In: The New York Times (Aug. 2014).

8. Benjamin A Goldstein et al. “Opportunities and challenges in developing risk prediction models with electronic health records data: a systematic review”. en. In: J. Am. Med. Inform. Assoc. 24.1 (Jan. 2017), pp. 198–208.
9. Andrea Frome et al. “DeViSE: A Deep Visual-Semantic Embedding Model”. In: Advances in Neural Information Processing Systems 26. Ed. by C J C Burges et al. Curran Associates, Inc., 2013, pp. 2121–2129.
10. Varun Gulshan et al. “Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs”. In: JAMA 316.22 (Dec. 2016), pp. 2402–2410.
11. Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. “Deep learning”. en. In: Nature 521.7553 (May 2015), pp. 436–444.
12. Yonghui Wu et al. “Google’s Neural Machine Translation System: Bridging the Gap between Human and Machine Translation”. In: (Sept. 2016). arXiv: 1609.08144 [cs.CL].
13. Куракин А.С., Алгоритм деперсонализации персональных данных. Научно-технический вестник информационных технологий, механики и оптики, 2012, №6 (82), 130-135 cc
14. <http://crmavi.com/> (Accessed: 02-05-2018)
15. <http://medelement.kz> (Accessed: 02-05-2018)
16. <https://www.medesk.ru/medcrm/> (Accessed: 02-05-2018)
17. Цифровая модернизация здравоохранения. <http://www.rcrz.kz/docs/prezent2.pdf> Доступен: 08-06-2018
18. Архитектура интероперабельности [https://www2.ezdrav.kz/images/platform/Architectura\\_interoperabelnosti.pdf](https://www2.ezdrav.kz/images/platform/Architectura_interoperabelnosti.pdf) Доступен: 08-06-2018
19. Health Level 7. FHIR Specification Home Page (2017). Available at: <http://hl7.org/fhir/> (Accessed 02-05-2018).
20. Suresh, H. et al. Clinical intervention prediction and understanding using deep networks. arXiv [cs.LG] (PMLR, Los Angeles, CA, USA, 2017).
21. Avati, A. et al. Improving palliative care with deep learning. arXiv [cs.CY] (2017).

## **МЕТОДЫ АППРОКСИМАЦИИ РЕШЕНИЯ СИСТЕМ УРАВНЕНИЙ И НЕРАВЕНСТВ**

**Посыпкин М.А., Ахметжанов М.А., Намазбаев Б.Д.**

*Институт информационных и вычислительных технологий КН МОН РК,  
Казахстан  
e-mail: [mposypkin@gmail.com](mailto:mposypkin@gmail.com)*

***Аннотация.** В работе рассматривается задача аппроксимации множеств решений систем нелинейных уравнений и неравенств. Приведена задача аппроксимации робота параллельной структуры, которая сводится к решению системы уравнений. Показано, как такая система может быть сведена к системе неравенств меньшей размерности. Приводятся алгоритмы решения систем указанного вида. Доказаны свойства и оценки точности получаемых аппроксимаций.*

## Введение

Во многих задачах на практике возникает необходимость решения систем равенств и неравенств. В частности, такие задачи часто встречаются в робототехнике при построении аппроксимации рабочей области робота. Приведем пример такой задачи.

Рассмотрим робот 3-RPR [1]. В названии 3 означает число степеней свободы, а буквы R,P,R специфицирует типы подвижных соединений. Данный робот, представленный на рисунке 1, представляет собой планарный манипулятор с шестью ротационными шарнирами для закрепления штанг и тремя призматическими двигателями. Двигатели изменяют длину штанг, соединяющих неподвижное основание с углами треугольной плоской платформы, тем самым управляя ее движением и ориентацией. В дальнейшем будем считать основание и платформу равносторонними треугольниками.

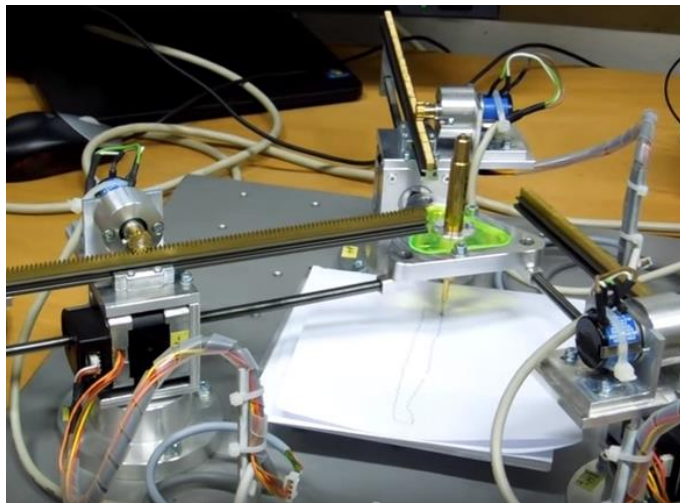


Рисунок 1. Робот-манипулятор типа 3-RPR

Кинематическая схема робота представлена на рисунке 2. Положение манипулятора характеризуется параметрами, перечисленным в таблице 1.

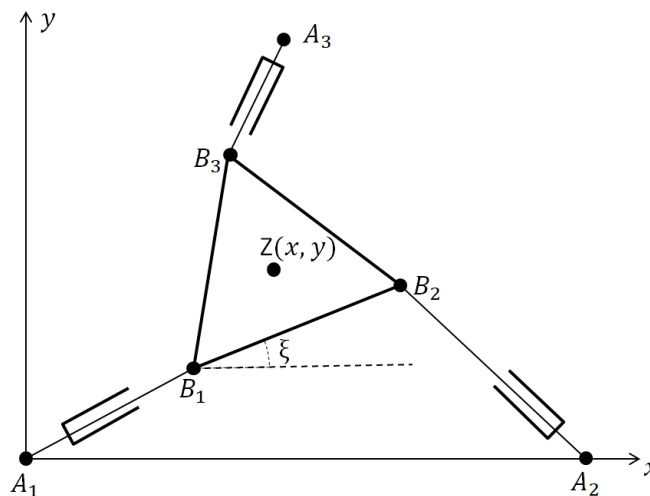


Рис. 2. Кинематическая схема робота 3-RPR

**Таблица 1. Параметры, характеризующие положение робота 3-RPR**

Параметр	Описание
$x, y$	координаты центр подвижной платформы
$\xi$	угол поворота платформы относительно нейтрального положения
$l_1, l_2, l_3$	длины штанг $\overline{A_1B_1}, \overline{A_2B_2}, \overline{A_3B_3}$ соответственно

Обозначим фиксированные координаты точек  $A_i$  в системе координат  $Oxу$  через  $x_i^A, y_i^A, i = 1, 2, 3$ . Фиксированные координаты точек  $B_i$  в подвижной системе координат, связанной с платформой, обозначим через  $\hat{x}_i, \hat{y}_i, i = 1, 2, 3$ . Предполагается, что в начальном положении оси подвижной системы координат параллельны осям системы  $Oxу$ . Величина  $\xi$  определяет угол поворота подвижной системы координат относительно начального положения против часовой стрелки. Тогда координаты точки  $B_i$  в системе  $Oxу$  задаются формулами:

$$\begin{aligned} x_i^B &= x + \hat{x}_i \cos \xi - \hat{y}_i \sin \xi, \\ y_i^B &= y + \hat{x}_i \sin \xi + \hat{y}_i \cos \xi. \end{aligned} \quad (1)$$

Уравнения связи для робота 3-RPR определяются соотношениями  $|A_iB_i| = l_i, i = 1, 2, 3$ . Запишем уравнения связи для данного робота, используя введенные переменные:

$$(x_i^B - x_i^A)^2 + (y_i^B - y_i^A)^2 = l_i^2, \quad i = 1, 2, 3.$$

Подставляя выражения (1), получим:

$$(x + \hat{x}_i \cos \xi - \hat{y}_i \sin \xi - x_i^A)^2 + (y + \hat{x}_i \sin \xi + \hat{y}_i \cos \xi - y_i^A)^2 = l_i^2, i = 1, 2, 3. \quad (2)$$

Таким образом, получена система из трех нелинейных уравнений относительно шести переменных, перечисленных в таблице 1. Рабочая область определяется уравнениями (2) и системой интервальных ограничений, задающих диапазон  $[l_i^{\min}, l_i^{\max}]$  изменения длин штанг  $l_i$ . Также целесообразно задать границы изменения переменных  $(x, y, \xi)$ . Без ограничения общности можно положить:

$$\begin{aligned} \xi &\in [0, 2\pi], \\ x &\in \left[ \max_{i=1,2,3} (x_i^A - l_i^{\max}) + \frac{r}{2}, \min_{i=1,2,3} (x_i^A + l_i^{\max}) - \frac{r}{2} \right], \\ y &\in \left[ \max_{i=1,2,3} (y_i^A - l_i^{\max}) + \frac{r}{2}, \min_{i=1,2,3} (y_i^A + l_i^{\max}) - \frac{r}{2} \right], \end{aligned}$$

где  $r$  – расстояние от центра до углов платформы.

Итоговая система имеет следующий вид:

$$\begin{aligned} (x + \hat{x}_i \cos \xi - \hat{y}_i \sin \xi - x_i^A)^2 + (y + \hat{x}_i \sin \xi + \hat{y}_i \cos \xi - y_i^A)^2 &= l_i^2, i = 1, 2, 3, \\ l_i &\in [l_i^{\min}, l_i^{\max}], i = 1, 2, 3, \\ \xi &\in [0, 2\pi], \end{aligned}$$

$$\begin{aligned} x &\in \left[ \max_{i=1,2,3} (x_i^A - l_i^{max}) + \frac{r}{2}, \min_{i=1,2,3} (x_i^A + l_i^{max}) - \frac{r}{2} \right], \\ y &\in \left[ \max_{i=1,2,3} (y_i^A - l_i^{max}) + \frac{r}{2}, \min_{i=1,2,3} (y_i^A + l_i^{max}) - \frac{r}{2} \right]. \end{aligned} \quad (3)$$

На практике интерес представляет не все множество решений системы (3) в пространстве  $R^6$ , а его проекция на множество координат выходного звена. В рассматриваемом случае либо множество достижимых координат  $x, y$  центра платформы  $Z$ , либо множество всех возможных троек  $(x, y, \xi)$ . Также определенный интерес может представлять диапазон достижимых углов  $\xi$ .

Непосредственное использование системы (3) для аппроксимации рабочей области может быть нереализуемо по причине высокой вычислительной сложности, обусловленной размерностью задачи. Заметим, система (3) содержит три уравнения и шесть параметров. Можно упростить систему, сведя ее к системе неравенств, исключив переменные  $l_i, i = 1, 2, 3$ :

$$\begin{aligned} (x + \hat{x}_i \cos \xi - \hat{y}_i \sin \xi - x_i^A)^2 + (y + \hat{x}_i \sin \xi + \hat{y}_i \cos \xi - y_i^A)^2 - (l_i^{max})^2 &\leq 0, i = 1, 2, 3, \\ (l_i^{min})^2 - (x + \hat{x}_i \cos \xi - \hat{y}_i \sin \xi - x_i^A)^2 - (y + \hat{x}_i \sin \xi + \hat{y}_i \cos \xi - y_i^A)^2 &\leq 0, i = 1, 2, 3, \\ \xi &\in [0, 2\pi], \\ x &\in \left[ \max_{i=1,2,3} (x_i^A - l_i^{max}) + \frac{r}{2}, \min_{i=1,2,3} (x_i^A + l_i^{max}) - \frac{r}{2} \right], \\ y &\in \left[ \max_{i=1,2,3} (y_i^A - l_i^{max}) + \frac{r}{2}, \min_{i=1,2,3} (y_i^A + l_i^{max}) - \frac{r}{2} \right]. \end{aligned} \quad (4)$$

Приведенный пример показывает, что важная практическая задача аппроксимации рабочей области робота может быть сведена к нахождению решения системы равенств либо неравенств с интервальными ограничениями на переменные. Далее в работе мы рассмотрим разработанные подходы к решению указанных задач.

### Аппроксимация множества решений системы уравнений

Рассмотрим систему уравнений общего вида:

$$\begin{aligned} g_i(x) &= 0, i = 1, \dots, m, \\ x &\in [a, b], \end{aligned} \quad (5)$$

где  $g_i(x): R^n \rightarrow R$  – непрерывные отображения, а  $[a, b] = \{x \in R^n: a_i \leq x_i \leq b_i, i = 1, \dots, n\}$  –  $n$ -мерный параллелепипед, задающий интервалы изменения параметров. Пусть  $X$  – множество всех решений системы (5).

Система (5) эквивалентна следующей системе:

$$\begin{aligned} \phi(x) &= 0, i = 1, \dots, m, \\ x &\in [a, b], \end{aligned}$$

где  $\phi(x)$  – свертка функций  $g_i(x), i = 1, \dots, m$ . В качестве возможных вариантов свертки можно взять любую норму вектора в пространстве  $R^m$ . Например можно положить  $\phi(x) = \sum_{i=1}^m g_i^2(x)$ .

Для нахождения аппроксимации множества  $X$  применяется алгоритм **CovEq**, результатом работы которого будет список параллелепипедов  $A$ , образующих аппроксимацию.

*Алгоритм CovEq*

```

L := {[a, b]}
while L ≠ ∅ do
  take P ∈ L, L := L \ {P}
  m := minx∈P φ(x), M := maxx∈P φ(x)
  if m ≤ 0 ≤ M then
    if d(P) ≤ δ then
      A := A ∪ {P}
    else
      {P1, P2} = divide P
      L := L ∪ {P1, P2}
    endif
  endif
endif
done

```

Принцип работы алгоритм **CovEq** состоит в том, что исходный параллелепипед делится на параллелепипеды меньшего размера до тех пор, пока диаметр параллелепипеда не станет меньше заданной величины  $\delta$ . Обрабатываемые параллелепипеды заносятся в список  $L$ . Деление происходит всегда вдоль максимального ребра, что обеспечивает конечность алгоритма. При этом, параллелепипеды, заведомо не содержащие точек из множества решений системы (5), отбрасываются. Для этого выполняется величин  $m = \min_{x \in P} \phi(x)$ ,  $M = \max_{x \in P} \phi(x)$  с нулем. Если  $m > 0$  или  $M < 0$ , то параллелепипед  $P$  не содержит ни одной точки из  $X$  и может быть исключен из дальнейшего рассмотрения.

**Утверждение 1.** Множество  $A$ , построенное в результате работы алгоритма **CovEq**, содержит все множество  $X$  и расстояние по Хаусдорфу между множествами  $A$  и  $X$  не превосходит  $\delta$ :

$$X \subseteq A, \quad (6)$$

$$h(A, X) \leq \delta. \quad (7)$$

**Доказательство.** По построению, отсеvu подвергаются только параллелепипеды, не содержащие точек из  $X$ , поэтому справедливо включение (6). Для любого параллелепипеда  $P$  из совокупности  $A$  справедливо  $m \leq 0 \leq M$ . В силу компактности  $P$  и непрерывности функции  $\phi(x)$  найдется такая точка  $x_0 \in P$ , что  $\phi(x_0) = 0$ . То есть  $x_0 \in X$ . Так как диаметр параллелепипеда  $P$  не превосходит  $\delta$ , то все точки  $P$  находятся в шаре радиуса  $\delta$  с центром в точке  $x_0$ . Тем самым (с учетом включения (6)) доказана вторая часть утверждения – неравенство (7).

К сожалению, точное нахождение величин  $m, M$  возможно только в относительно простых случаях. В большинстве случаев используются оценки  $\hat{m}, \hat{M}$ :  $\hat{m} \leq m \leq M \leq \hat{M}$ . В качестве примеров можно рассмотреть интервальные [2] или



Липшицевы оценки [3]. При этом, алгоритм **CovEq** заменяется алгоритмом **CovEqEst**, полученным из **CovEq** заменой величин  $m, M$  на  $\hat{m}, \hat{M}$ . Справедливо следующее утверждение относительно алгоритма **CovEqEst**.

**Утверждение 2.** Множество  $A$ , построенное в результате работы алгоритма **CovEqEst**, содержит все множество  $X$ :  $X \subseteq A$ .

Утверждение 2 доказывается по аналогии с утверждением 1. Оценить расстояние по Хаусдорфу между множествами  $X$  и  $A$  без дополнительных предположений в отношении функций  $g_i(x), i = 1, \dots, m$  не представляется возможным.

### Решение систем неравенств

Система неравенств в общем виде имеет следующих вид:

$$\begin{aligned} g_i(x) &\leq 0, i = 1, \dots, m, \\ x &\in [a, b], \end{aligned} \quad (8)$$

где  $g_i(x): R^n \rightarrow R$  – непрерывные отображения, а  $[a, b] = \{x \in R^n: a_i \leq x_i \leq b_i, i = 1, \dots, n\}$  –  $n$ -мерный параллелепипед, задающий интервалы изменения параметров. Пусть  $X$  – множество всех решений системы (8).

Систему (8) можно заменить эквивалентной системой:

$$\begin{aligned} \psi(x) &\leq 0, \\ x &\in [a, b], \end{aligned} \quad (9)$$

где  $\psi(x) = \max_{i \in \{1, \dots, m\}} g_i(x)$ .

Для приближения множества  $X$  применяется алгоритм **CovIneq**, результатом работы которого будет списки параллелепипедов  $A, B$ , аппроксимирующих внутренность и границу множества  $X$  соответственно.

### Алгоритм CovEq

$L := \{[a, b]\}$

**while**  $L \neq \emptyset$  **do**

  take  $P \in L, L := L \setminus \{P\}$

$m := \min_{x \in P} \psi(x), M := \max_{x \in P} \psi(x)$

**if**  $M < 0$  **then**

$A := A \cup \{P\}$

**else if**  $m \leq 0$  **then**

**if**  $d(P) \leq \delta$  **then**

$B := B \cup \{P\}$

**else**

$\{P_1, P_2\} = \text{divide } P$

$L := L \cup \{P_1, P_2\}$

**endif**

**endif**

**done**

По аналогии с утверждением 1, доказывается следующее утверждение.

**Утверждение 3.** Для множеств  $A$  и  $B$ , построенных в результате работы алгоритма **CovIneq**, справедливы следующие соотношения:

$$A \subseteq \text{Int } X \subseteq X \subseteq A \cup B, \quad (10)$$

$$\partial X \subseteq B, \quad (11)$$

$$h(A \cup B, X) \leq \delta. \quad (12)$$

По аналогии со случаем равенств точные значения максимума и минимума можно заменить на оценки  $\hat{m}, \hat{M}$ :  $\hat{m} \leq m \leq M \leq \hat{M}$ . Аналогично алгоритм **CovIneq** заменяется алгоритмом **CovIneqEst**, полученным из **CovIneq** заменой величин  $m, M$  на  $\hat{m}, \hat{M}$ . Справедливо следующее утверждение относительно алгоритма **CovIneqEst**.

**Утверждение 3.** Для множеств  $A$  и  $B$ , построенных в результате работы алгоритма **CovIneqEst**, справедливы следующие соотношения:

$$A \subseteq \text{Int } X \subseteq X \subseteq A \cup B, \quad (13)$$

$$\partial X \subseteq B. \quad (14)$$

### Заключение

Рассмотрена задача аппроксимации множества решения системы равенств или неравенств. Приведен практический пример, который может быть сформулирован в одной из этих форм. Достоинством формулировки в виде системы неравенств является меньшая размерность пространства решений, но получение такой формализации требует существенных усилий от разработчика, связанных с необходимостью ручного преобразования исходной системы уравнений.

Разработано два алгоритма для получения аппроксимации множеств решений систем уравнений и неравенств. Доказаны свойства получаемых аппроксимаций и полученных оценки их точности. В дальнейшем планируется проведение вычислительных экспериментов с помощью указанных алгоритмов для широкого спектра задач робототехники.

### Литература

1. Merlet J. P. Parallel robots. – Springer Science & Business Media, 2006. – Т. 128.
2. Moore R. E. Methods and applications of interval analysis. – Siam, 1979. – Т. 2.
3. Sergeyev Y. D., Kvasov D. E. Global search based on efficient diagonal partitions and a set of Lipschitz constants //SIAM Journal on Optimization. – 2006. – Т. 16. – №. 3. – С. 910-937.

# МОДЕЛИ И МЕТОДЫ СОЗДАНИЯ ИНФОРМАЦИОННЫХ СИСТЕМ ПОДДЕРЖКИ НАУЧНО-ОБРАЗОВАТЕЛЬНОЙ ДЕЯТЕЛЬНОСТИ, ИНТЕГРИРОВАННЫХ В ОТКРЫТОЕ СЕМАНТИЧЕСКОЕ ПРОСТРАНСТВО

Самбетбаева М.А., Тусупов Д.А., Федотов А.М., Батырханов А.Г.  
Институт информационных и вычислительных технологий КН МОН РК  
e-mail: Madina\_jgtu@mail.ru

***Аннотация.** Информационные системы (ИС) для поддержки научно-образовательной деятельности оперируют с различного рода информацией, как например, публикации, электронные документы, электронные коллекции, онтологические описания, массивы данных, логические описания и др. Как правило, эти ресурсы, востребованные разными группами исследователей, оказываются недоступными из-за невозможности их поиска и идентификации. Семантические связи между информационными ресурсами повышают их ценность и предоставляет дополнительные возможности для информационного поиска и идентификации. Данные, интегрированные в открытое семантическое пространство, представляют собой совокупность знаний о некоторой предметной области в виде семантической структуры, на основе которой становятся возможными качественно новые наукометрические измерения и исследования структурных свойств корпуса научных знаний.*

***Ключевые слова.** Информационная система, онтология предметной области, тезаурус, управляемый словарь, система управления информационными ресурсами, система хранения информационных ресурсов, классификация и систематизация информационных ресурсов, семантическое пространство.*

В классической монографии [3], изданной ВИНТИ, на основе выделения двух типов информационных потребностей: потребности в сведениях об источниках необходимой научной информации и потребности в самой необходимой научной информации говорится, что для удовлетворения информационных потребностей первого предназначены информационные системы, получившие название документальных, второго типа – фактографических. В настоящее время наиболее востребованным средством информационного обеспечения научно-образовательной деятельности становятся ИС, сочетающие возможности информационных систем обоих названных типов и позволяющие удовлетворять информационные потребности квалифицированного пользователя в соответствии со схемой «документ – рассуждение – факт» [4, 5]. Заметим, что эта схема соответствует RDF схеме связанных данных [6, 7]. Будем называть такие ИС интеллектуальные информационные системы (ИИС). Важным этапом процесса функционирования ИИС является извлечение из текстов документов содержащихся в них *фактов*, то есть, в наиболее общем смысле, «особого рода предложений, фиксирующих эмпирическое знание» [8]. Подход разработанный в концепции Семантик WEB [9] здесь оказывается неприменимым из-за того, что рассматриваемые ИС работают с документами произвольной структуры [10].

Предполагаемое исследование ориентировано на разработку модели и созданию прототипа распределенной ИИС поддержки научно-образовательной деятельности с учетом связывания данных и морфологии казахского языка, унифицирующей процесс обмена результатами исследований [11-13].

Модель ИИС поддержки научно-образовательной деятельности, на основе анализа информационных потребностей пользователей, должна обеспечить функциональность [28]:

Авторам работы не известны разработки ИИС, которые обеспечивают такую функциональность.

В настоящее время существуют достаточно мощные информационные системы, которые в той или иной степени удовлетворяют информационные потребности пользователей, например, euroCRIS, eLibrary, Информика [14], ИСТИНА [15], MathNET [16], Socionet [17], сайты электронных библиотек: Библиотека Конгресса, Мировая цифровая библиотека. Самым продвинутым, с точки зрения нашего исследования, можно назвать проект «Европеана» [18], но как все библиотечные проекты он является чисто документальным. С точки зрения предоставления сервисов можно отметить систему «Соционет» [17] и разрабатываемый пакет «Либмета» [19]. Однако основными недостатками большинства систем являются ограниченность возможностей проведения аналитической работы с ресурсами и обеспечения интеграции ресурсов как внутри каждой из систем, так и с внешними системами (часто не учитываются международные стандарты и рекомендации, низкая интероперабельность). Это крайне неудобно в сфере научно-образовательной деятельности.

Одна из главных задач интеграции в открытое семантическое пространство в состоит в том, что необходимо установить связи между конкретными научными фактами (например, «что означает термин кибернетика» или «кто автор данной статьи») и сущностями информационной системы (персоны, факты, данные, документы, публикации, ключевые термины и т.п.). Поэтому исследование принципов и разработка алгоритмов автоматизации научно-информационного процесса, включая алгоритмы связывания и идентификации данных, является весьма актуальной проблемой.

Задача идентификации объектов реального мира, упоминаемых в слабоструктурированных документах, решается с использованием управляемого словаря (тезауруса). Разработка специализированных тезаурусов актуальна сама по себе как развитие и систематизация понятийного аппарата предметной области (в данном случае информатики), так и для логического поиска информации, как средство формирования поисковой потребности, формулирования поисковых предписаний, адекватного автоматического индексирования, систематизации и классификации документов.

Однако ограничением для большинства информационных систем поддержки научно-образовательной деятельности является их распределенность и гетерогенность. Типичный подход к решению указанной проблемы предполагает либо организацию информационного обмена между системами и, либо создание нового, обобщающего ресурса. Такой подход решает проблему интероперабельности только на синтаксическом и структурном уровнях, и не может считаться достаточно эффективным в сфере научно-образовательной деятельности. Существующие к настоящему моменту методы интеграции информационных ресурсов относятся к структурным методам и не решают проблемы семантической интероперабельности.

В рамках нашего исследования планируется рассмотреть задачи автоматического связывания и идентификации данных в систему интеграции данных, позволяющей формировать, хранить и предоставлять унифицированный доступ к данным. Идентификация документов включает проверку на дубликаты или почти дубликаты. В

работе будут разработаны алгоритмы поиска дубликатов документов и аналогичных документов (семантические дубликаты).

В ИС поддержки научно-образовательной деятельности, так называемые семантические дубликаты, дубликаты по смыслу играют большую роль. Существующие алгоритмы настроены, чтобы находить полные или почти дубликаты, или заимствования. Поэтому необходима разработка алгоритма поиска аналогичных документов (дубликатов по смыслу).

Использование формата Связанных Открытых Данных LOD (Linked Open Data), помимо поддержки технологий публикации, предусматривает использование онтологий предметных областей с использованием семантических запросов к данным. Онтология задает описание связей между сущностями информационной системы (персоны, факты, данные, документы, публикации, ключевые термины и т.п.).

Важным этапом практической реализации предлагаемых в исследовании подходов должна стать реализация алгоритмов синтаксического и семантического анализа текстов с целью извлечения фактов.

Поскольку публикация научных данных в нашей стране осуществляется на трех языках, то предполагаемый семантический поиск в системе должен осуществляться с учетом морфологии казахского, русского и английского языков, причем на основе трёхязычного тезауруса с учетом лексических эквивалентов. Авторам исследования не известны работы, которых анализировались документы одновременно на русском, английском и казахском языках.

**Значимость исследования** в национальном и международном масштабе высока, поскольку тематика исследования актуальна и вносит большой теоретический и практический вклад в информационные системы для поддержки научно-образовательной деятельности оперируют с различного рода информацией, как например, публикации, электронные документы, электронные коллекции, онтологические описания, массивы данных, логические описания и др. Как правило, эти ресурсы, востребованные разными группами исследователей, оказываются недоступны из-за невозможности их поиска и идентификации. Таким образом, разработка ИС, унифицирующей процесс обмена результатами научных исследований, будет способствовать эффективному взаимодействию территориально и организационно распределенных групп исследователей, что подчеркивает актуальность исследования в целом.

**Использование.** Создание модели информационной системы поддержки научно-образовательной деятельности, унифицирующей процесс обмена результатами научных исследований, работающей с данными и документами, снабженной адаптивными сервисами конвертации наукоемких ресурсов, включающая подмодели описания, представления, автоматического связывания и их взаимодействия, с тонкими поисковыми и классификационными механизмами, призванные удовлетворению информационных потребностей исследователей весьма актуальная задача, вносящая существенный вклад в развитие данной научной области.

### **Методы и модель исследования**

На сегодняшний день технологической основой управления доступом к распределенным информационным ресурсам являются технологии, основанные на трех технологических «китах»: WEB-технологии, модель каталогов LDAP [13, 20-22] и модель протокола Z39.50 [23-25]. Эти технологические составляющие реализуют основные функции распределенной ИИС, как системы, основанной на знаниях. Отметим, что стандарт Z39.50 определяет службу и протокол типа клиент/сервер для

информационного поиска. Для реализации функций управления (управление доступом к распределенной информации и распределенное управление информацией) более всего подходят технологии, основанные на модели протокола LDAP, т.к. именно на основе LDAP сегодня проще всего организовать технологию «единой точки доступа» к информационным ресурсам, включая идентификацию, аутентификацию и авторизацию пользователей в распределенных информационных системах. Немаловажным обстоятельством при этом является тот факт, что LDAP основан на идеологии распределенного хранения информации (деревьев каталогов) на фоне глобальной идентификации всех элементов каталогов, содержит внутри себя определения механизмов и процедур репликаций данных между различными серверами и очень хорошо поддерживается разработчиками прикладного и системного программного обеспечения. Последнее позволяет достаточно просто переходить от локального к распределенному управлению информационными системами и контролю доступа к их ресурсам.

Для осуществления целей исследования при работе с удаленными хранилищами документов и репозиториями выбран протокол Z39.50 (или протоколы SRW/SRU [26]) и в случае их отсутствия у архива открытого доступа - протокол OAI-PMH [27]. К тому же, ни один другой стандарт (именно стандарт ISO 23950:1998 [23], причем международный, а не рекомендация) не содержит полного набора необходимых правил, регламентирующих сетевой доступ к базам данных, не допускает абстрагирование от конкретных схем и структур данных при поиске информации и при ее извлечении из хранилищ и предоставления пользователю в необходимых ему форматах.

Информационная модель разрабатываемой ИИС [29-31] должна быть многоуровневой и включать в себя следующие компоненты: хранилище данных (репозиторий), сервер метаданных, сервер приложений (диспетчера), системы поддержки контролируемых словарей-классификаторов, предназначенных для систематизации ресурсов. Сервер метаданных используется для динамической генерации схем баз данных и ведения служебных баз данных, в которых хранятся данные, обеспечивающие поддержку стандартных функций системы, динамически определяемые отношения между документами и динамическое распределение ресурсов системы. Сервер приложений используется для поддержки пользовательских сервисов и поддержки работы алгоритмов, например, морфологического анализа или процедур кластеризации и классификации. Система управления информационными ресурсами требует создания единой справочной системы о конфигурации и содержании собственно ИС. Наличие такой справочной системы является необходимым условием для успешного построения системы управления распределенными информационными ресурсами [32, 33].

Концептуальная модель ИИС [34] разрабатывается на основе эталонной модели OAIS RM [35]. Разработка концептуальной модели ИИС основана на онтологии предметной области, совместимой с уже существующими. Модель ИИС оперирует со следующими сущностями: авторы (персоны, действующие лица, организации и др. субъекты деятельности, включая компьютерные приложения), документы, публикации, словарные статьи, ключевые термины, данные и др. объекты деятельности, включая факты – особый вид документа, понимаемый как: характеристика сущности, описываемой в онтологии информационной системы, представляемая как единичное значение данных [10, 30].

Основу онтологии составляет создание управляемых словарей (тезауруса) с учетом полноты представленной информации, то есть онтология рассматривается как

общая часть модели предметной области и тезауруса, связывающая знания о мире со знаниями о языке, причем полноценный тезаурус невозможен без онтологии, поскольку она, пусть даже в простейшей форме, является скелетом всякой системы данных и/или знаний. Тезаурус можно интерпретировать как запас семантической информации, содержащейся в документах на данную тему, то есть как описание структуры знаний.

В рамках исследования будет использован разработанный по стандарту ISO 25964 в ИВТ СО РАН трехязычный тезаурус информатике и информационным технологиям. Для реализации тезауруса была выбрана схема данных Zthes. На данный момент компоненты реализации тезауруса позволяют просматривать, редактировать и добавлять термины тезауруса в систему через web-формы а также импортировать и экспортировать термина в виде XML, RDF, DTD файлов с возможным преобразованием записей в схемы данных SKOS, MARC, MODS. Тезаурус содержит более 20 тыс. терминов на русском, казахском и английском языках, и позволяет проводить индексирование документов сразу на трех языках. На основе алгоритма Портера разработан морфологический анализатор для казахского языка. Анализатор и словарь были внедрены в библиотеку phpMorphu. Связи и знания в онтологии представлены в виде RDF триплетов, используется механизм глобальной уникальной идентификации URI, язык представления моделей данных в форме объектов и их взаимосвязей RDF (его расширения RDFSchema и SKOS) и язык веб-онтологий OWL, а также SPARQL - протокол доступа и язык поисковых запросов к источникам данных RDF.

Одним из возможных вариантов эффективного доступа к информационным ресурсам является классификация документов. При использовании фасетной классификации объекты классифицируются одновременно по нескольким независимым друг от друга признакам (фасетам). Применительно к электронным документам в качестве фасетов выступают элементы метаданных. Опишем простейшую формальную модель классификации документов с использованием метаданных.

Обозначим множество метаданных  $\{M^j, 1 \leq j \leq n\}$ .

В произвольном документе  $d$  элементы  $M^j$  могут присутствовать один или несколько раз или отсутствовать.

Пусть  $D$  - множество документов, цитируемых в информационной системе,  $D = \{d_i: i \in I\}$ , где  $I$  – некоторая нумерация документов из  $D$ .

Документ  $d$  принадлежит множеству  $D$  с индексом  $i$ , т.е.

$$d_i = (m_i^{1,k_1}, m_i^{2,k_2}, \dots, m_i^{n,k_n}),$$

где запись  $m_i^{j,k_j}$  означает, что метаданные  $M^j$  встречается в документе  $d_i$   $k_j$ -раз, если  $k_j = 0$ , тогда метаданные  $M^j$  отсутствует в документе  $d_i$ .

Пусть  $d_i \in D$ .

Предложим метод (процедуру) переопределения индексов из множества  $I$ .

Пусть  $d_i = (m_i^{1,k_1}, m_i^{2,k_2}, \dots, m_i^{n,k_n})$ . Определим индекс  $i^*$  для документа  $d_i$  следующим образом:  $i^* = c(k_1, \dots, k_j, \dots, k_n)$ , где  $c$  – канторовская функция нумерации кортежей неотрицательных целых чисел.

Через функцию  $h$  обозначим процедуру переопределения индексов из множества  $I$ , т.е.  $i^* = h(i)$ .

Пусть  $I^* = \{i^* \mid i^* = h(i), i \in I\}$ .

В общем случае,  $I^* \neq I$ .

Переопределим запись документов  $d_i = (m_i^{1,k_1}, m_i^{2,k_2}, \dots, m_i^{n,k_n})$  следующим образом:  $\overline{d_{i^*}} = (k_1, \dots, k_j, \dots, k_n)$ .

В этом случае, мы можем рассматривать множество векторов  $D^* = \{\overline{d_{i^*}}(k_1, \dots, k_n) : i \in I\}$  вместо множества документов  $D = \{d_i : i \in I\}$ . Определим для векторов  $\overline{a}$  и  $\overline{b}$  из  $D^*$  операции  $\otimes$  и  $\oplus$  следующим образом:

$$\overline{a} \otimes \overline{b} = \overline{d_{i^*}}(k_1^*, \dots, k_j^*, \dots, k_n^*),$$

где  $k_j^* = \min \{k_j^{\overline{a}}, k_j^{\overline{b}}\}$  для  $j = \overline{1, n}$  и  $i^* = c(k_1^*, \dots, k_n^*)$ .

$$\overline{a} \oplus \overline{b} = \overline{d_{i'}}(k_1', \dots, k_j', \dots, k_n'),$$

где  $k_j' = \max \{k_j^{\overline{a}}, k_j^{\overline{b}}\}$  для  $j = \overline{1, n}$  и  $i' = c(k_1^1, \dots, k_n^1)$ .

Определим на множестве индексов  $I^*$  отношение частичного порядка  $\ll$  следующим образом: Пусть  $i_0 = c(k_1^0, k_2^0, \dots, k_n^0)$ ,  $i_1 = c(k_1^1, k_2^1, \dots, k_n^1)$ .

Тогда  $i_0 \ll i_1 \stackrel{df}{\Leftrightarrow} k_1^0 \leq k_1^1 \& \dots \& k_n^0 \leq k_n^1$ , где отношение  $\leq$  естественный порядок на целых числах.

Будем писать для векторов  $\overline{a}$  и  $\overline{b} \in D^*$ ,  $\overline{a} \ll \overline{b} \Leftrightarrow i_{\overline{a}} \ll i_{\overline{b}}$ , где  $i_{\overline{a}}$ ,  $i_{\overline{b}}$  – индексы векторов  $\overline{a}$  и  $\overline{b}$  соответственно.

Из введенных выше операций  $\otimes$  и  $\oplus$  следуют следующие соотношения:

1.  $\overline{a} \otimes \overline{b} \ll \overline{a}$  и  $\overline{a} \otimes \overline{b} \ll \overline{b}$ ;
2.  $\overline{a} \ll \overline{a} \oplus \overline{b}$  и  $\overline{b} \ll \overline{a} \oplus \overline{b}$ .

Таким образом, на множестве  $D^*$  мы можем ввести отношение частичного порядка  $\ll$ , т.е. мы получим структуру частичного порядка  $\mathbf{D}^* = \langle D^*, \ll \rangle$ .

Пусть  $\overline{a}_i = (k_1, \dots, k_n)$ , определим «вес» вектора  $\overline{a}_i$  следующим образом:

$$w(\overline{a}_i) = \sum_{j=1}^n k_j.$$

Рассмотрим множество векторов  $D^* \setminus \{\overline{0}\}$ , где  $\overline{0}$  – нулевой вектор. Для произвольных векторов  $\overline{a}$  и  $\overline{b}$  определим меру «близости» векторов:

$$\rho(\overline{a}, \overline{b}) = 1 - \frac{w(\overline{a} \otimes \overline{b})}{w(\overline{a} \oplus \overline{b})},$$

где  $w(\overline{a} \oplus \overline{b}) \neq \overline{0}$ . Если  $w(\overline{a} \oplus \overline{b}) = \overline{0}$ , тогда вектора  $\overline{a}$  и  $\overline{b}$  несравнимы (т.е. нет смысла говорить о «близости» векторов).

Свойство:  $0 \leq \rho(\overline{a}, \overline{b}) \leq 1$ .

Если  $\rho(\overline{a}, \overline{b}) = 0$ , тогда вектора  $\overline{a}$  и  $\overline{b}$  одинаковы.

Если  $\rho(\overline{a}, \overline{b}) = 1$ , тогда вектора  $\overline{a}$  и  $\overline{b}$  несравнимы.

Пусть  $\rho(\overline{a}, \overline{b}) = \varepsilon$ , тогда, чем меньше  $\varepsilon$ , тем вектора  $\overline{a}$  и  $\overline{b}$  более близки.

Теперь все размышления о документах можно вести на языке векторов и обратно.



Для решения задач тематической классификации предлагается дополнительно описывать категории и ключевые понятия словарными статьями, которые наряду с документами подвергаются весовому индексированию. В экспериментах, проведенными нами по классификации авторефератов диссертационных работ, такой подход показал очень хорошие результаты. В частности, такой подход позволяет решить проблему «полисемии», которая в большом количестве присутствует терминологической базе информатики.

При разработке алгоритма поиска документов «по аналогии» [10] при решении задачи кластеризации предлагается использовать формализм, основанный на использовании классов толерантности. Поисковое предписание, определяющее набор классификационных признаков с указанием сочетаний значений этих признаков при помощи логических операций, определяет конкретное ядро толерантности на множестве документов, которое и выдается пользователю в качестве ответа на его информационный запрос. Кроме того, на множестве классов толерантности можно также, в свою очередь, ввести отношение толерантности, при этом толерантными считаются классы, имеющие хотя бы один общий документ. .

Алгоритмизация научно-информационных процессов предполагает их точное описание на информационно-логических и алгоритмических языках, предназначенных для представления соответствующих алгоритмов, включая алгоритмы связывания и идентификации данных. Задача идентификации объектов реального мира, упоминаемых в слабоструктурированных документах решается с использованием управляемого словаря (тезауруса). В случае уникальности имен это не вызывает проблем. В противном случае, как например, с персонами, выступающих в роли авторов публикаций, необходимо учитывать различные дополнительные признаки, по которым производится идентификация. Идентификацию авторов можно уточнить путем отнесения статьи к списку работ автора.

Идентификация документов включает проверку на дубликаты или почти дубликаты. Здесь используются методы n-грамм, шинглов и семантические алгоритмы, основанные на методах классификации. Существует большое количество алгоритмов, работающих как с метаданными (каталожным описанием), так и полными текстами. В исследовании будут разработаны алгоритмы поиска дубликатов документов и аналогичных документов (семантических дубликатов).

При создании ИИС разумно следующее понимание факта: содержащаяся в тексте и метаданных документа совокупность связей между сущностями, описывается в онтологии информационной системы. Отсюда, в частности, вытекает следующее важное замечание: именно онтология ИИС определяет, что будет считаться фактом в рамках этой системы, поскольку она задает описание связей между сущностями информационной системы (персоны, факты, данные, документы, публикации, ключевые термины и т.п.). Важным этапом практической реализации предлагаемых в исследовании подходов должна стать реализация алгоритмов синтаксического и семантического анализа текстов с целью извлечения фактов.

## **Литература**

1. Bizer, C. Linked Data – the story so far / Christian Bizer, Tom Heath, Tim Berners-Lee // IJSWIS: Int. j. on Semantic Web a. inform. systems. — 2009. — Vol. 5, N 3. — P. 1—22.
2. Berners-Lee, T. Linked Data. [Design issues] [Electronic resource] / Tim Berners-Lee // W3C: World Wide Web Consortium. — Electronic data. — [Keio, Korea, Jap.; Cambridge, MA, USA; Biot, France; Beihang, China], 2006, last change: 2009.

3. Михайлов А.И., Черный А.И., Гиляревский Р.С. Научные коммуникации и информатика. М: Наука, 1976.
4. Арский Ю.М., Гиляревский Р.С., Туров И.С., Черный А.И. Инфосфера: Информационные структуры, системы и процессы в науке и обществе. М.: ВИНТИ, 1996.
5. Шокин Ю.И., Федотов А.М., Барахнин В.Б. Проблемы поиска информации. Новосибирск: Наука, 2010.
6. Klyne, G. Resource Description Framework (RDF): concepts and abstract syntax: W3C Recommendation 10 Febr. 2004 Keio, Korea, Jap. ; Cambridge, MA, USA ; Sophia Antipolis, France, 2004. — Mode of access: <https://www.w3.org/TR/2004/REC-rdf-concepts-20040210/>.
7. Alemu, G. Linked Data for libraries: benefits of a conceptual shift from libraryspecific record structures to RDF-based data models / Getaneh Alemu, Brett Stevens, Penny Ross, Jane Chandler // *New Libr. World*. — 2012. — Т. 113, N 11-12. — P. 549—570.
8. Ракитов А. Факт // *Философская энциклопедия*. М: Советская энциклопедия, 1970. Т. 5. С. 298.
9. Berners-Lee T., Hendler J., Lassila O. The Semantic Web // *Scientific American*. 2001. Vol. 284(5). P. 34-43.
10. Barakhnin V.B., Fedotov A.M. Building models of documentary and factographic retrieval in digital libraries // *Automatic Documentation and Mathematical Linguistics*. - 2014. - Vol.48. - Iss. 6. - P.296-304. - ISSN 0005-1055. - EISSN 1934-8371.
11. Федотов А. М. Концептуальные подходы к построению распределенных систем // *Труды Международной конференции по вычислительной математике МКВМ-2004*. — Новосибирск: Изд. ИВМ и МГ СО РАН, 2004. — С. 132–143.
12. Федотов А. М. Методологии построения распределенных систем // *Вычислительные технологии*. 2006, Т. 11, Избранные доклады X Российской конференции «Распределенные информационно-вычислительные ресурсы», Новосибирск 6-8 октября 2005 г. С. 3–16.
13. Шокин Ю. И., Федотов А. М., Жижимов О. Л. Технология распределенных информационных систем // *Современные информационные технологии для научных исследований: Материалы конф. Магадан, 2008*. С. 18–21.
14. Абрамов А.Г., Булакина М.Б., Сигалов А.В. Электронная библиотека портала "Единое окно доступа к образовательным ресурсам": концепция и реализация проекта, достижения и проблемы // 28-29 октября 2010 г., Санкт-Петербург / ЦГПБ им. В.В. Маяковского; ред. Сост. И.Е. Прозоров. - СПб.: Северная звезда, 2010. С. 172-178.
15. Интеллектуальная система тематического исследования научно-технической информации (ИСТИНА) / С.А. Афонин и др. Под ред. академика В.А. Садовниченко., – М.: Издательство Московского университета, 2014. – 262 с.
16. Аджиев А.С., Бездушный А.Н., Серебряков В.А. О реализации веб-системы математической информации [Электронный ресурс] / *Российский научный электронный журнал Электронные библиотеки*, - 2004. Т. 7. Вып. 1.– Режим доступа: <http://www.elbib.ru/index.phtml?page=elbib/rus/journal/2004/part1/ABS>
17. Паринов С.И., Ляпунов В.М., Пузырев Р.Л. Система Соционет как платформа для разработки научных информационных ресурсов и онлайн-сервисов // [Электронный ресурс] / *Российский научный электронный журнал Электронные библиотеки*. 2003. Т. 6, вып. 1.
18. Europeana.eu. Collections : [offic. website] [Electronic resource] / Europeana Foundations, Europ. Union. — Electronic data. — Den Haag, Netherlands, [2016].

19. Атаева О. М., Серебряков В. А., Персональная открытая семантическая цифровая библиотека LibMeta. Конструирование контента. Интеграция с источниками LOD, // Информ. и её примен., 11:2 (2017), - с. 85–100
20. Федотов А. М., Шокин Ю. И., Жижимов О. Л., Молородов Ю. И. Служба директорий LDAP как единая информационная среда // Открытое и дистанционное образование. 2007. № 4 (28). С. 31–41. – (14) .
21. Барахнин В. Б., Жижимов О. Л., Степанов Ю. Ю., Федотов А. М. LDAP-каталог организации как ядро корпоративной распределенной информационной системы // Инновационные недра Кузбасса. IT-технологии: Сб. науч. тр. Кемерово, 2008. С. 226–232.
22. Жижимов О. Л., Федотов А. М. Модели управления доступом к распределенным информационным ресурсам // Электронные библиотеки: перспективные методы и технологии, электронные коллекции: Тр. IX Всерос. науч. конф. Переславль-Залесский: Изд-во «Университет города Переславля», 2007. С. 301–304.
23. ISO 23950:1998 Information and documentation — Information retrieval (Z39.50) — Application service definition and protocol specification.
24. Жижимов О. Л., Мазов Н. А. Принципы построения распределенных информационных систем на основе протокола Z39.50. Новосибирск, 2004. 361 с.
25. Жижимов О. Л., Федотов А. М., Шокин Ю. И. Технологическая платформа массовой интеграции гетерогенных данных // Вестник НГУ Серия: Информационные технологии. - 2013. - Том 11, Выпуск № 1. - С. 24–41. - ISSN 1818-7900.
26. SRU — Search/Retrieval via URL / The Library of Congress. — USA .
27. The Open Archives Initiative Protocol for Metadata Harvesting. Protocol Version 2.0 of 2002-06-14. `Document Version 2015-01-08. <http://www.openarchives.org/OAI/2.0/openarchivesprotocol.htm>
28. Шокин Ю.И., Федотов А.М., Жижимов О.Л. Технологии создания распределенных информационных систем для поддержки научных исследований // Вычислительные технологии. - 2015. - Т.20. - № 5. - С.251-274. - ISSN 1560-7534.
29. Федотов А.М., Барахнин В.Б., Жижимов О.Л., Федотова О.А. Информационная модель электронной библиотеки по научному наследию // - 2014. - Санкт-Петербург: ФГБУ «Президентская библиотека имени Б. Н. Ельцина». - С.175-202. - ISBN 978-5-905273-51-3.
30. Федотов А.М., Барахнин В.Б., Жижимов О.Л., Федотова О.А. Модель информационной системы для поддержки научно-педагогической деятельности // Вестник Новосибирского государственного университета. Серия: Информационные технологии. - 2014. - Т.12. - № 1. - С.89-101. - ISSN 1818-7900. - EISSN 2410-0420.
31. Шокин Ю.И., Федотов А.М., Жижимов О.Л., Федотова О.А. Система управления электронными библиотеками в ИРИС СО РАН // В.А. Серебрякова. - 2014. - Т.1. - Москва: Вычислительный центр РАН. - С.11-39. - ISBN: 978-5-19601-103-6.
32. Жижимов О.Л., Федотов А.М., Юданов Ф.Н. Модель управления информационными ресурсами организации // Вестник Новосибирского государственного университета. Серия: Информационные технологии. - 2010. - Т.8. - № 4. - С.81-95. - ISSN 1818-7900. - EISSN 2410-0420.
33. Шокин Ю.И., Федотов А.М., Жижимов О.Л., Федотова О.А. Эволюция информационных систем: от WEB-сайтов до систем управления информационными ресурсами // Вестник Новосибирского государственного университета. Серия: Информационные технологии. - 2015. - Т.13. - № 1. - С.117-134. - ISSN 1818-7900. - EISSN 2410-0420.

34. Федотов А.М., Абделиева М.Н., Байдавлетов А.Т., Бапанов А.А., Самбетбаева М.А., Федотова О.А. Концептуальная модель научно-образовательной информационной системы // Вестник Новосибирского государственного университета. Серия: Информационные технологии. - 2015. - Т.13. - № 3. - С.52-67. - ISSN 1818-7900. - EISSN 2410-0420.

35. ISO 14721:2012 Reference model for an Open archival information system (RM OAIS) / Recommended Practicle: CCSDS 650.0-M-2 (Magenta Book). June 2012. Available at: <http://public.ccsds.org/publications/archive/650x0m2.pdf>.

## **КОГНИТИВНАЯ SMART–ТЕХНОЛОГИЯ ДЛЯ ПЕРСОНАЛИЗИРОВАННОГО ДИСТАНЦИОННОГО ОБУЧЕНИЯ ЛЮДЕЙ С ОСЛАБЛЕННЫМ ЗРЕНИЕМ ИНЖЕНЕРНЫМ СПЕЦИАЛЬНОСТЯМ В НЕФТЕГАЗОВОЙ ОТРАСЛИ**

**Самигулина Г.А., Лукманова Ж.С.**

*Институт информационных и вычислительных технологий КН МОН РК,  
Казахстан  
e-mail: galinasamigulina@mail.ru*

***Аннотация.** Исследования посвящены актуальной проблеме подготовки профессиональных инженерных кадров для работы на современном сложном высокотехнологичном промышленном оборудовании фирмы Honeywell в нефтегазовой отрасли. Целью работы является разработка и внедрение инновационной персонализированной когнитивной Smart-технологии дистанционного обучения людей с ослабленным зрением в лабораториях коллективного пользования с использованием когнитивно-визуальных схем для динамической подачи учебной информации в зависимости от типа центральной нервной системы обучающихся (холерики, сангвиники, флегматики и меланхолики) и особенностей зрения (миопии или гиперметропии). Обработка многомерных данных осуществляется на основе модифицированных алгоритмов искусственных иммунных систем и других подходов искусственного интеллекта. Так как проблемы со зрением в той или иной мере возникают у более чем восьмидесяти процентов населения, а дистанционное обучение подразумевает большое количество проведенного времени за компьютером, что приводит к дополнительной нагрузке на зрение и ухудшению восприятия информации, то очевидна перспективность данных исследований. В дальнейшем предлагаемая когнитивная Smart-технология может применяться при разработке мнемосхем для управления сложными техническими, технологическими процессами и осуществления информационной поддержки деятельности человека с учетом индивидуальных психофизиологических особенностей восприятия и осознания текущей информации.*

Прогресс в области создания нового сложного высокотехнологичного промышленного оборудования, а также сложности, связанные с его эксплуатацией требуют разработки новых эффективных подходов для подготовки высокопрофессиональных инженерных кадров и применения новых инновационных технологий дистанционного обучения.

Характерной особенностью современного дистанционного обучения (ДО) является разработка и внедрение интеллектуальных технологий, рост инноваций и переход на новые формы персонализированного обучения. Широкое распространение получили Smart – технологии, которые предполагают применение когнитивного подхода при организации процесса обучения [1, 2]. Особенно в последнее время вызывают интерес когнитивные Smart-технологии ДО для людей с ослабленным зрением [3].

Актуальны разработки интеллектуальных информационных технологий дистанционного обучения (ДО) с использованием различных нетрадиционных биоинспирированных подходов на основе нейронных сетей, генетических алгоритмов, искусственных иммунных систем [4], алгоритмов роевого интеллекта и т.д.

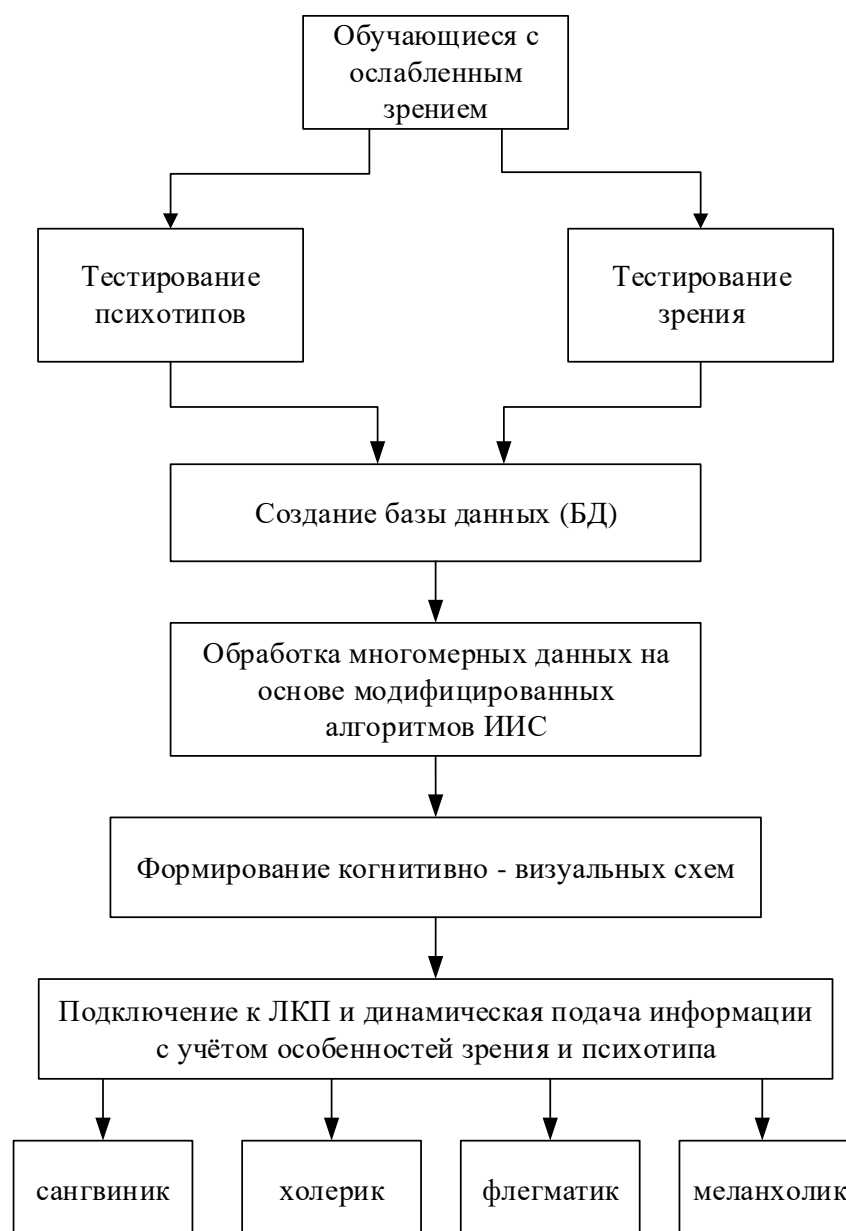
Предлагаемые исследования являются продолжением цикла работ, посвященных дистанционному обучению инженерным специальностям. В работе [5] представлена процедура сбора данных с реального объекта промышленной автоматизации с использованием оборудования Schneider Electric и организации информационного обмена данными для интеллектуальной системы управления на основе Искусственной Иммунной Системы (ИИС). Применяется технология OPC (OLE for Process Control), которая предназначена для обмена данными с системой измерения и управления, а также может использоваться при организации дистанционного доступа к реальному промышленному оборудованию в лабораториях коллективного пользования для обучения. Исследования [6] посвящены разработке специализированной мультиагентной системе дистанционного обучения для людей с ограниченными возможностями зрения на основе нейронных сетей и нейро-нечеткой логики.

Целью данной работы является разработка инновационной когнитивной Smart-технологии [7] персонализированного дистанционного обучения инженерным специальностям нефтегазовой отрасли людей с ослабленным зрением в лабораториях коллективного пользования (ЛКП) на современном промышленном оборудовании фирмы Honeywell (с использованием платформы Exregion PKS) на основе применения различных модифицированных алгоритмов искусственных иммунных систем и других подходов искусственного интеллекта, а также когнитивных визуальных технологий динамической подачи учебной информации в зависимости от психотипа обучающегося и особенностей зрения.

На рисунке 1 приведена структурная схема когнитивной Smart-технологии ДО для людей с ослабленным зрением. Диагностика состояния зрения осуществляется с использованием тестов, например, с помощью таблицы Сивцева [8]. Таблица Сивцева - это набор печатных знаков для определения остроты зрения, то есть шаблон для проверки зрения.

Для оценки психотипа, характеризующего личность обучающегося, используется тест определяющий темперамент по А. Белову [9]. Рассматривается четыре психотипа: сангвиник, холерик, флегматик и меланхолик. Сангвиник обладает сильным характером и гибкостью ума, легко идет на контакт. **Холерик - это вспыльчивый и эмоциональный человек, нетерпеливый и склонный к конфликтам,** но быстро усваивающий информацию, обладающий лидерскими качествами. **Флегматик** - спокойный, сдержанный, терпеливый психотип, отличается самообладанием в стрессовых ситуациях, но в непривычных условиях становится пассивным. **Меланхолики** - это чувствительные, неэнергичные, пессимистические люди, подозрительные и ревнивые, но обладающие аналитическим складом ума, способные к творчеству и умеющие доводить дело до конца.

Основное достоинство применения этого теста заключается в том, что он показывает процентное соотношение всех четырех типов в характере человека и какие черты преобладают. Таким образом, этот тест помогает при анализе когнитивных способностей и способствует построению адекватной модели обучающегося с ослабленным зрением. Обработка многомерных данных осуществляется на основе модифицированных алгоритмов искусственных иммунных систем и других подходов искусственного интеллекта. В зависимости от психотипа обучающегося и состояния зрения формируются когнитивно-визуальные схемы для динамической подачи обучающей информации и работы с оборудованием в ЛКП.

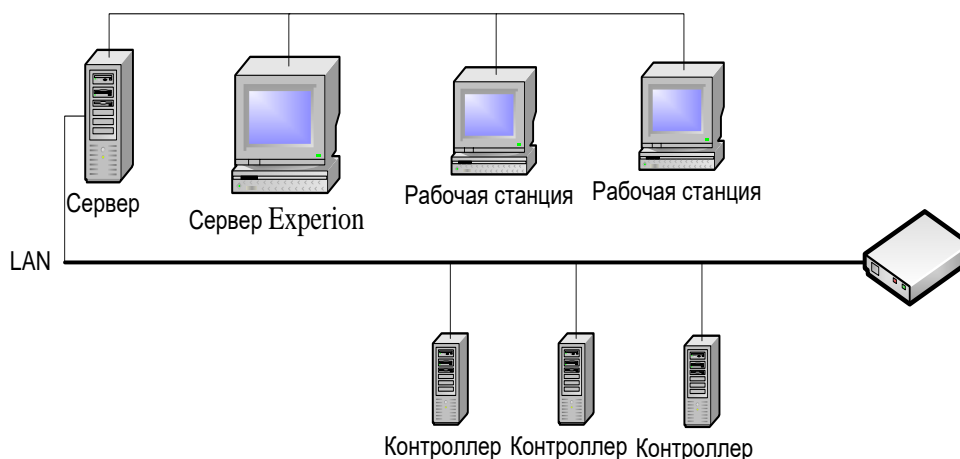


**Рис.1. Когнитивная SMART-технология ДО**

При поддержке компании Honeywell в Satbaev University создан учебный лабораторный центр для обучения и доступа к реальному промышленному оборудованию, которое применяется в нефтегазовой отрасли [10]. В данной

лаборатории представлено современное промышленное оборудование на основе распределенной системы управления - Experion PKS, которая широко используется на ряде предприятий, таких как: ТОО «Тенгизшевройл» и Карачаганак Петролеум Оперейтинг. Использование платформы Experion PKS позволяет создать единый комплекс, который обеспечивает прямой доступ к каждой рабочей станции, любому контроллеру и системам сбора, хранения и обработки данных о ходе производственного процесса.

На рисунке 2 представлена модель ЛКП с оборудованием компании Honeywell на базе Experion PKS.



**Рис.2. Модель ЛКП с оборудованием фирмы Honeywell**

Доступ к информации осуществляется через Internet на базе запатентованной фирмой Honeywell технологии HMIWeb, которой снабжены все программные средства верхнего уровня. Система Experion PKS поддерживает промышленный интерфейс OPC (Ole for Process Control) и технологии COM/DCOM, OLE, ActiveX, которые используются в наиболее распространенных промышленных и полевых сетях.

Работа выполняется по гранту Комитета Науки Министерства Образования и Науки Республики Казахстан (2018–2020 гг.), по теме «Разработка когнитивной Smart – технологии для интеллектуальных систем управления сложными объектами на основе подходов искусственного интеллекта».

### **Литература**

1. You, Sik Hong. Smart Education System / Sik Hong You // Journal of the Institute of Internet, Broadcasting and Communication. – 2013. – Vol. 13. – P. 255-260.
2. Yousif, Jabar H., Saini, Dinesh Kumar and Uraibi, Hassan S. Artificial Intelligence in E-learning-Pedagogical and Cognitive Aspects / Jabar H. Yousif, Dinesh Kumar Saini and Hassan S. Uraibi // Proc. of the World Congress on Engineering. London, U.K. – 2011. – Vol 2. – P. 452-457.
3. Bhowmick, A., Hazarika, S.M. An insight into assistive technology for the visually impaired and blind people: state-of-the-art and future trends / A. Bhowmick, S. M. Hazarika. // Journal on Multimodal User Interfaces, 2017. -Volume 11, Issue 2. – P.149–172.
4. Lee, Chin-Ling., Huang, Chih-Hui., Lin, Cheng-Jian. Test-Sheet Composition Using Immune Algorithm for E-Learning Application / Chin-Ling Lee, Chih-Hui Huang,

Cheng-Jian Lin. // New Trends in Applied Artificial Intelligence. – Berlin: Springer–Verlag, 2007. – P. 823-833.

5. Samigulina, G.A., Samigulina, Z.I. Intelligent System of Distance Education of Engineers, based on Modern innovative Technologies / G.A. Samigulina, Z.I. Samigulina // Proceedings of the II International Conference on Higher Education Advances, HEAd'16. J. Procedia- Social and Behavioral Sciences. – Valencia, Spain: Elsevier, 21-23 June 2016. - № 228. – P. 229-236.

6. Samigulina, G. A., Shayakhmetova, A. S., Nyussupov, A. Innovative intelligent technology of distance learning for visually impaired people / G.A. Samigulina, A.S. Shayakhmetova, A. Nyussupov // J. Open engineering. – 2017. - №7. - P. 444-452.

7. Самигулина, Г.А., Самигулина, З.И. Разработка инновационной когнитивной SMART – технологии дистанционного обучения современному промышленному оборудованию в рамках концепции цифроватизации промышленности Казахстана / Г.А. Самигулина, З.И. Самигулина // Материалы круглого стола «Цифроватизация промышленности - основа четвертой промышленной революции». – Алматы, 20 апреля 2018. – С. 88-91.

8. <https://www.vseozrenii.ru/proverka-zreniya/duohromnyj-test/> (05.06.2018).

9. <https://experimental-psychic.ru/test-na-temperament-formula-po-belovu/> (05.06.2018).

10. Техническая документация: Руководство по планированию серверной и клиентской части фирмы Honeywell. – 2008. – 37с.

## **ДИАГНОСТИКА ПРОМЫШЛЕННОГО ОБОРУДОВАНИЯ НА ОСНОВЕ МОДИФИЦИРОВАННОГО АЛГОРИТМА РОЯ ЧАСТИЦ И ПОДХОДА ИСКУССТВЕННЫХ ИММУННЫХ СИСТЕМ**

**Самигулина Г.А., Масимканова Ж.А.**

*Институт информационных и вычислительных технологий КН МОН РК,  
Казахстан  
galinasamigulina@mail.ru, masimkanovazh@gmail.com*

**Аннотация.** *Статья посвящена разработке Smart-технологии диагностики промышленного оборудования для нефтегазовой отрасли. С бурным развитием цифровых технологий актуально применение подходов искусственного интеллекта. Построение оптимального набора данных осуществляется на основе модифицированного алгоритма роя частиц с весом инерции (IWPSO) для дальнейшего прогнозирования технического состояния промышленного оборудования с использованием искусственных иммунных систем. Приведены результаты моделирования модифицированного алгоритма роя частиц на базе реальных производственных данных нефтегазовой компании ТенгизШевройл (на основе точных замеров показаний с датчиков установки У300).*

В настоящее время актуально внедрение инновационных цифровых Smart-технологий в нефтегазовую отрасль для поиска, разведки и разработки месторождений на основе применения последних достижений искусственного интеллекта. Методы



искусственного интеллекта успешно применяются в крупных нефтегазовых компаниях (КазМұнайГаз, ТенгизШевройл, Chevron, Total, Shell, Лукойл) для решения задачи оптимизации производства, размещения скважин, прогнозирования разведки нефти и газа. Особенно широкое применение получили алгоритмы роевого интеллекта (РИ), такие как алгоритмы муравьиной и пчелиной колоний, роя частиц, кукушки, серых волков и др. Достоинствами алгоритмов РИ являются высокая точность и простота программной реализации. Правильный подбор значений весовых коэффициентов алгоритмов улучшает способность к глобальному и локальному поиску, решает проблему преждевременной сходимости.

Большой интерес представляют гибридные системы на основе модифицированных алгоритмов роя частиц. В работе [1] предлагается совместное использование алгоритма роя частиц (particle swarm optimization, PSO) и метода опорных векторов с использованием метода наименьших квадратов (LS-SVM) для прогнозирования добычи нефти и газа. Результаты моделирования показывают, что предложенный алгоритм имеет хорошую сходимость, высокую точность прогнозирования и скорость обучения. Статья [2] посвящена применению алгоритма PSO и метода опорных векторов (SVR) для формирования оптимального набора гиперпараметров с использованием реальных промышленных наборов данных, полученных при добыче нефти с четырех различных нефтяных скважин. Гибридная PSO-SVR модель сравнивается с моделями на основе метода случайного поиска (RAND-SVR) и метода проб и ошибок (TE-SVR). В результате исследования доказывається эффективность PSO-SVR модели при построении оптимального набора параметров.

В работе [3] используются искусственные нейронные сети (ANN) и алгоритм PSO для прогнозирования образования конденсата в зависимости от выделяемого газа (Condensate-to-gasratio, CGR) при добыче нефти. В работе [4] предлагается алгоритм PSO для прогнозирования оптимального набора параметров бурения. Применение предложенного метода позволяет минимизировать общую стоимость работы скважин.

Разработаны различные модифицированные алгоритмы роя частиц, такие как Guaranteed Convergence PSO (GCP SO), modified PSO (MPSO), inertia weight PSO (IWPSO), Fully informed PSO (FIP SO), Adaptive PSO (APSO) и др. В исследовании [5] рассматривается модифицированный алгоритм роя частиц с весом инерции (IWPSO), в котором с увеличением количества итераций уменьшается вес инерции. По сравнению с классическим алгоритмом роя частиц данный алгоритм улучшает поиск оптимальных значений путем решения задачи преждевременной сходимости. В статье [6] анализируется влияние веса инерции на эффективность работы алгоритма IWPSO. В результате исследования отмечается, что использование алгоритма роя частиц с предложенным весом инерции позволяет быстро обрабатывать большой объем данных и обеспечивает лучшее решение.

Актуально использование подхода искусственных иммунных системы (ИИС) для решения широкого круга задач в нефтегазовой отрасли. В статье [7] используется ИИС для обнаружения ошибок в работе оборудования при разработке нефтегазовых месторождений. В работе [8] используются ИИС для прогнозирования добычи углеводородов из скважин. Результаты исследования показывают, что использование ИИС повышает точность прогнозирования.

Проведенный литературный обзор доказывает актуальность применения подходов искусственного интеллекта в нефтегазовой отрасли, в частности, решения задачи прогнозирования на основе алгоритмов РИ [9] и искусственных иммунных систем [10].

**Постановка задачи.** Необходимо разработать Smart-технологии диагностики промышленного оборудования на основе модифицированного алгоритма роя частиц и подхода искусственных иммунных систем (ИИС) на базе реальных производственных данных нефтегазовой компании ТенгизШевройл.

Для решения поставленной задачи разработан следующий алгоритм:

Шаг 1. Формирование базы данных (БД) производственной информации (суточных замеров с датчиков промышленного оборудования).

Шаг 2. Построение оптимального набора данных на основе модифицированного алгоритма роя частиц [11].

Шаг 3. Решение задачи распознавания образов и прогноз на основе подхода искусственных иммунных систем.

Шаг 4. Диагностика технического состояния промышленного оборудования и принятие решения.

Оптимальный набор производственных данных формируется на основе алгоритма роя частиц. Ниже приведен классический алгоритм роя частиц. Пусть позиция частицы в  $D$ -мерном пространстве представляется следующим образом [12]:

$$X_i = (x_{i1}, x_{i2}, \dots, x_{iD}), \text{ а скорость частицы: } V_i = (v_{i1}, v_{i2}, \dots, v_{iD}).$$

Для выбранной фитнес-функции задается размер популяции, количество итераций и весовые коэффициенты алгоритма. При каждой итерации определяются локальное лучшее ( $p_{id}$ ) и глобальное лучшее ( $p_{gd}$ ) значения. Локальное лучшее значение рассматривается как наилучшее положение частицы в пространстве поиска. Глобальное лучшее значение является лучшим положением всех частиц в популяции. Далее частицы меняют скорость (1) и положения (2) по формулам:

$$v_{id}^{k+1} = wv_{id}^k + c_1r_1(p_{id} - x_{id}^k) + c_2r_2(p_{gd} - x_{id}^k), \quad (1)$$

$$x_{id}^{k+1} = x_{id}^k + v_{id}^{k+1}, \quad i = 1, 2, \dots, N; d = 1, 2, \dots, D, \quad (2)$$

где  $v_{id}$  – скорость частицы,  $k$  – количество итераций,  $w$  – вес инерции,  $c_1, c_2$  – коэффициенты ускорения,  $r_1, r_2$  – равномерно распределенные случайные числа в интервале  $[0,1]$ , которые используются для сохранения разности популяции,  $p_{id}$  – локальное лучшее значение,  $p_{gd}$  – глобальное лучшее значение,  $N$  – размер популяции.

Выбор дескрипторов выполняется с использованием алгоритма CFS (Correlation-based feature selector, CFS), предложенного в работе [13]. Фитнес-функция представляется следующим образом:

$$M_s = \frac{k\bar{r}_{sf}}{\sqrt{k + k(k-1)\bar{r}_{ff}}} \quad (3)$$

где  $s$  – подмножество, которое содержит  $k$  дескрипторов,  $\bar{r}_{sf}$  – средняя корреляция «дескриптор-класс» ( $f \in s$ ),  $\bar{r}_{ff}$  – средняя корреляция между «дескриптор-дескриптор».

Далее сравниваются значения фитнес-функций [14]:

–если  $M(x_{id}^{k+1}) \leq M(p_{id})$ , то установить  $p_{id} = x_{id}^{k+1}$ .

–если  $M(p_{id}) \leq M(p_{gd})$ , то установить  $p_{gd} = p_{id}$ .

Основным недостатком описанного алгоритма является быстрая сходимость.

Для решения проблемы преждевременной сходимости применяется модифицированный алгоритм роя частиц с весом инерции (IWPSO). При итерации вес инерции частицы меняется от большого значения к меньшему [12] по формуле:

$$w = w_{\max} - \frac{w_{\max} - w_{\min}}{\text{iter}_{\max}} k, \quad (4)$$

где  $w$  – вес инерции,  $w_{\max}$  – начальный вес инерции,  $w_{\min}$  – конечный вес инерции,  $\text{iter}_{\max}$  – максимальное количество итераций,  $k$  – текущая итерация. При  $w \geq 1$  скорость частицы увеличивается и позволяет частицам лучше исследовать пространство поиска. При  $w < 1$  частицы замедляются и находят локальные значения.

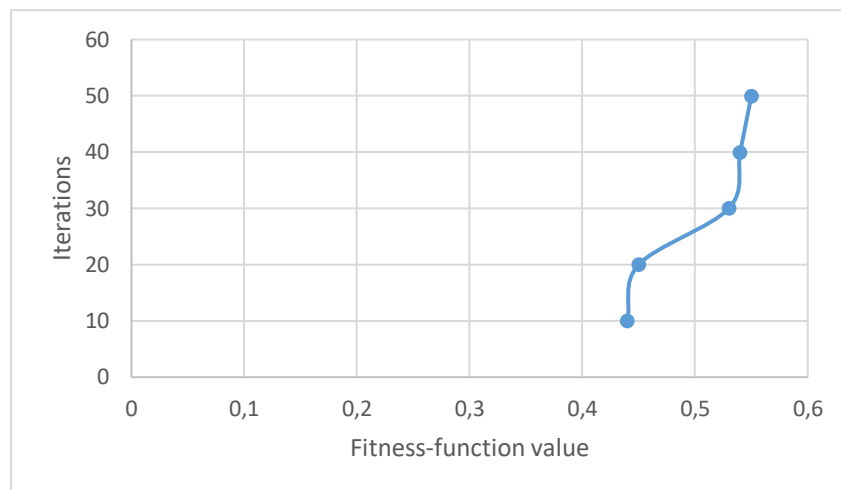
### Результаты моделирования

Исследования проводились на базе крупнейшего промышленного предприятия Казахстана по разработке и добыче нефти ТенгизШевройл [15]. Завод состоит из нескольких комплексно - технологических линий и установок различного назначения. В качестве сложного объекта управления рассматривается Установка 300, предназначенная для очистки нефтяных газов от кислых компонентов и состоящая из нескольких агрегатов (абсорбера высокого давления, абсорбера низкого давления, регенератора амина, охладителя амина и др.). Фрагмент базы данных (БД) суточных замеров показаний с датчиков установки У300 (например, LIC31002 – уровнемер буйковый, FT31005 – преобразователь разности, TT31020 – преобразователь температуры и т.д.) представлен в таблице 1.

**Таблица 1. Фрагмент БД суточных замеров датчиков У300**

№	LIC31002	FT31005	PDT31001	TT31020	PDT31007	...
1	65	85,563	5,733	107,204	54,593	...
2	63,9	57,315	5,906	102,336	58,006	...
3	66,1	96,944	4,692	110,13	47,273	...
4	65,4	85,257	5,556	108,655	54,822	...
5	65,7	85,844	6,093	107,803	54,517	...
6	65,9	87,093	5,206	109,279	57,028	...
7	64,3	83,714	6,14	106,187	52,405	...
...	...	...	...	...	...	...
800	64,9	86,189	5,649	107,698	55,314	...

Для выделения информативного набора данных на основе модифицированного алгоритма роя частиц с весом инерции используется программный пакет Yagriz, разработанный на языке программирования Python. При моделировании заданы следующие параметры: количество частиц в рое  $N=10$ , количество итераций  $k=10$ , коэффициенты ускорения  $c_1 = 1$  и  $c_2 = 2$ , вес инерции  $w = 1$ . В результате моделирования из 19 дескрипторов были выбраны 8 информативных дескрипторов. Результаты моделирования зависимости значений фитнес-функций от количества итераций представлены на рисунке 1.



**Рис. 1. Моделирование модифицированного алгоритма роя частиц с весом инерции**

Значение фитнес-функции показывает информативность отобранного набора дескрипторов и меняется от 0 до 1. В таблице 2 представлены результаты моделирования при разном количестве итераций.

**Таблица 2. Результаты моделирования**

Модифицированный алгоритм роя частиц	Количество итераций	Размер популяции	Значение фитнес-функции	Набор информативных дескрипторов
Алгоритм роя частиц с весом инерции	10	10	0.44	8
	20	20	0.45	8
	30	30	0.53	7
	40	40	0.54	7
	50	50	0.55	6

При итерации равной 50 и популяции равной 50 достигается оптимальное значение фитнес-функции и выбирается 6 информативных дескрипторов.

Работа выполнена по гранту КН МОН РК по теме: "Разработка когнитивной Smart-технологии для интеллектуальных систем управления сложными объектами на основе подходов искусственного интеллекта" (2018-2020 гг.).

### **Литература**

1. Qiao, Y., Peng, J., Ge, L., Wang, H. Application of PSO LS-SVM forecasting model in oil and gas production forecast // *Cognitive Informatics & Cognitive Computing*. – IEEE, 2017. – DOI: 10.1109/ICCI-CC.2017.8109791 (дата обращения: 30.05.2018).
2. Akande, K.O., Owolabi, T.O., Olatunji, S.O., Abdul Raheem, A. A hybrid particle swarm optimization and support vector regression model for modelling permeability prediction of hydrocarbon reservoir // *Journal of Petroleum Science and Engineering*. – 2017. – Vol. 150. – С. 43-53.
3. Zendehboudi, S., Ahmadi, M.A., James L., Chatzis I. Prediction of Condensate-to-Gas Ratio for Retrograde Gas Condensate Reservoirs Using Artificial Neural Network with Particle Swarm Optimization // *Energy Fuels*. – 2012. – Vol. 26 (6). – P. 3432–3447.

4. Self, R., Atashnezhad, A., Hareland, G. Reducing Drilling Cost by finding Optimal Operational Parameters using Particle Swarm Algorithm // Proc. of Society of Petroleum Engineers Source SPE Deepwater Drilling and Completions Conference. URL: <https://www.researchgate.net/publication/304998993> (дата обращения: 04/03/2018).
5. Bin, J., Zhigang, L., Xingsheng, G. A dynamic inertia weight particle swarm optimization algorithm // Chaos, Solitons and Fractals. – 2008. – Vol. 37. – P. 698-705.
6. Umapathy, P., Venkateshaiah, C., Arumugam, M.S. Particle Swarm Optimization with Various Inertia Weight Variants for Optimal Power Flow Solution // Discrete Dynamics in Nature and Society. – 2010. – URL: <http://dx.doi.org/10.1155/2010/462145> (дата обращения: 15.05.2018).
7. Araujo, M., Aguilar, J., Aponte, H. Fault Detection System in Gas Lift Well based on Artificial Immune System. URL: <https://pdfs.semanticscholar.org/8a90/4a9938416dd50b9ee2c7d2a5090b7c468046.pdf> (дата обращения: 05.06.2018).
8. Panja, P., Velasco, R., Pathak, M., Deo, M. Application of artificial intelligence to forecast hydrocarbon production from shales // Petroleum. – 2018. – Vol. 4. – P. 75-89.
9. Самигулина, Г.А., Масимканова, Ж.А. Обзор современных методов роевого интеллекта для решения задачи выделения информативных признаков при создании новых лекарственных препаратов // Проблемы информатики. – Новосибирск, 2016. – №2. – С.50-61.
10. Samigulina, G.A. Immune network modeling technology for complex objects intellectual control and forecasting system: monograph. - USA: Science Book Publishing House, 2015. -172 p.
11. Samigulina, G.A., Massimkanova, Zh.A. Multi-agent system for forecasting based on modified algorithms of swarm intelligence and immune network modeling // Proc. Inter. conf. «Smart Innovation, Systems and Technologies: Agents and Multi-Agent Systems: Technologies and Applications 2018». – Australia: Springer, 2018 (в печати).
12. Mu, A., Cao, D., Wang, X. A Modified Particle Swarm Optimization Algorithm // Natural Science. - 2009. - Vol.1, No.2. - P. 151-155.
13. Hall, M.A. Correlation-based Feature Selection for Machine Learning // Thesis of Doctoral dissertation. - The University of Waikato, 1999. - 198 p.
14. Ahmad, I. Feature Selection Using Particle Swarm Optimization in Intrusion Detection // International Journal of Distributed Sensor Networks. – 2015. – Vol. 11. – URL: <https://doi.org/10.1155/2015/806954> (дата обращения: 28.03.2018).
15. Технологический регламент на технологический процесс очистки углеводородных газов на установке 300 // ТОО «ТенгизШевройл». – 2016 г. – 146 с.

## **ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА УПРАВЛЕНИЯ СЛОЖНЫМИ ОБЪЕКТАМИ НА ОСНОВЕ АЛГОРИТМА КЛОНАЛЬНОГО ОТБОРА**

**Самигулина З.И., Алдибекова А.Н.**

*Институт информационных и вычислительных технологий  
КН МОН РК, Казахстан  
[zarinasamigulina@mail.ru](mailto:zarinasamigulina@mail.ru), [aitkul.aldibekova@gmail.com](mailto:aitkul.aldibekova@gmail.com)*

***Аннотация.** Статья посвящена разработке интеллектуальной системы управления сложными объектами с использованием последних достижений в области искусственного интеллекта на базе современных распределенных систем управления предприятиями (PCU). Рассмотрены основные механизмы управления на основе PCU Honeywell DCS, а так же принципы передачи и хранения данных для работы с прогнозируемыми событиями. Предложена архитектура интеллектуальной системы на основе алгоритма искусственных иммунных систем (клонального отбора) для интеграции с Honeywell DCS.*

В настоящее время последние тенденции в области развития высокотехнологичных систем автоматизированного управления сложными объектами направлены на внедрение интеллектуальных технологий на базе современных методов искусственного интеллекта (ИИ). Новейшие программно-аппаратные комплексы позволяют повсеместно внедрять элементы искусственного интеллекта для анализа и прогнозирования данных с измерительных приборов посредством обработки архивной информации SCADA систем и распределенных систем управления (DCS, Distributed Control System).

Для разработки подобных комплексов хорошо зарекомендовал себя подход искусственных иммунных систем (ИИС). В последнее время к ИИС проявляют всё больший интерес во многих областях: вычислительной технике, автоматизации и управлении, математическом моделировании, медицине, дистанционном обучении и т.д.

Искусственные иммунные системы представляют собой раздел вычислительного интеллекта на основе принципов иммунологии. Первые исследования появились в конце 80-х (Bersini H., Varela F.), начале 90-х годов (Farmer J., Packard H., Perelson A.) для машинного обучения и автоматизированного решения проблем. Работы были посвящены теории иммунной сети и направлены на компьютерное обучение, проблемы управления и оптимизации [1]. Следующим этапом развития ИИС стало применение в области защиты компьютерной безопасности, классическими работами в данном направлении являются исследования ученых Forrest S., Hofmeyr A., Somayaji A. и Kephart J.

Современные ИИС подразделяются на несколько основных направлений: клональный отбор [2,3], негативный отбор [4], алгоритмы на основе иммунной сети [5-7]. Данные методы широко используются в области управления сложными объектами. Например, работа [8] посвящена исследованиям в сфере проектирования системы управления двигателем постоянного тока на основе алгоритма клонального отбора (КО). Авторами разработана интеллектуальная система ПИД-регулирования с настройкой параметров в реальном времени с помощью КО. Исследования [9] направлены на разработку системы идентификации на базе клонального отбора и адаптивного ПД регулятора для динамической системы с неопределенностью. Работа [10] посвящена созданию алгоритма управления автономными подводными аппаратами на базе КО. Подход искусственных иммунных систем с непрерывным обучением [11] используется для диагностики нарушений напряжения в электрических распределенных системах. В работе [12] алгоритм негативный отбор применяется для диагностики неисправностей ветряных турбин. Исследования [13] посвящены разработке системы агентов обнаружения неисправностей для химических процессов котельной с использованием динамического симулятора завода на базе негативного отбора. Авторами в работе [14] представлен мультиагентный алгоритм искусственной иммунной сети для оценки эффективности дистилляционной установки.

Исследования [15] посвящены применению подходов эволюционной оптимизации для системы управления электростанцией с использованием иммунно-инспирированных алгоритмов. Алгоритмы оптимизации реализованы в среде MATLAB, а работа электростанции смоделирована в Dynsim. Авторами описан вычислительный интерфейс между данными компонентами.

Таким образом, анализ литературы доказывает актуальность применения искусственных иммунных систем для управления сложными объектами.

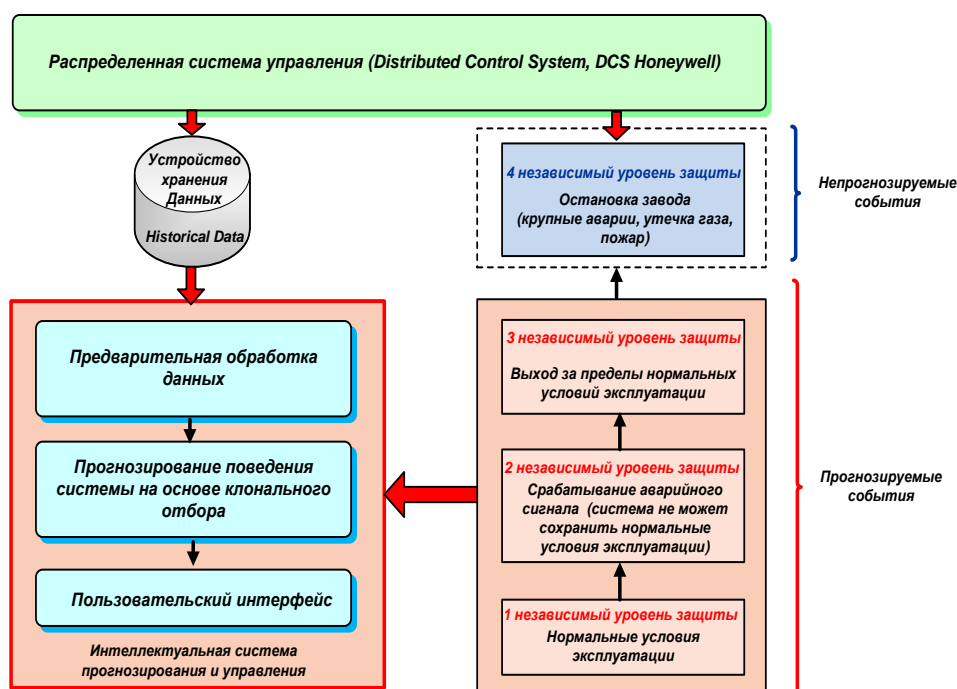
Постановка задачи исследования формулируется следующим образом: необходимо разработать интеллектуальную систему управления сложными объектами на основе алгоритма искусственных иммунных систем (клонального отбора) для анализа и прогноза состояния объектов промышленной автоматизации на базе распределенной системы управления Honeywell DCS.

Для решения поставленной задачи рассмотрим основные принципы управления технологическими установками на заводах ТенгизШевройл. Завод имеет центральный пункт управления (CCR, Central Control Room), где с помощью пультов DCS отдаются команды [16]. Система управления Honeywell состоит из узлов, особых компонентов сети, включающих в себя пульты оператора GUS, устройства хранения данных (HM, History Module), архивный модуль (ARM, Archive Reply Module), устройство управления, контроллер (HPM, High-Performance Process Manager), вычислительную аппаратуру (AM, Application Module) и т.д. Передача данных между узлами осуществляется с помощью двух сетей связи, работающих по разным протоколам: локальная сеть управления (LCN, Local Communications Networks) и универсальная сеть управления (UCN, Universal Control Network). С помощью модуля сетевого интерфейса (NIM, Network Interface Module) данные передаются между сетями LCN и UCN.

К локальной сети управления подсоединен исторический модуль и база данных истории процесса, причем скорость передачи данных составляет 5 миллионов бит в секунду. Для обеспечения безопасности применяется резервированная пара коаксиальных кабелей. Для разработки интеллектуальной Smart системы управления на базе алгоритмов искусственного интеллекта интересен исторический модуль (HM), который имеет жесткий диск для хранения данных. Основными его функциями являются: сбор и хранение отобранных данных, хранение программ и программного обеспечения, хранение основных операционных систем. А так же база данных истории процесса (PHD), которая позволяет собирать данные с систем управления, обеспечивает долговременные исторические данные и может представлять информацию другим компьютерным системам на базе операционной системы Microsoft WinServer 2003 или 2008.

На рисунке 1 представлена укрупненная схема архитектуры интеллектуальной системы управления и прогнозирования на основе алгоритма клонального отбора для распределенной системы управления Honeywell. Управление технологическим процессом делится на несколько независимых уровней защиты, которые представляют собой прогнозируемые и непрогнозируемые события. Первый уровень защиты рассчитан на нормальные условия эксплуатации, обеспечивается за счет работы DCS системы управления, которая следит за параметрами технологического процесса и меняет их в автоматическом режиме в случае возникновения необходимости. Второй уровень независимой защиты рассчитан на возникновение аварийного сигнала, при котором система не может сохранить нормальные условия эксплуатации. В таком случае оператор принимает решение по устранению проблемы. Третий уровень независимой защиты срабатывает, когда система вышла за пределы нормальных

условий эксплуатации, происходит контролируемый останов. К четвертому уровню независимой защиты относятся непрогнозируемые события, такие как крупная авария, утечка газа, возникновение пожара, при котором система не может предотвратить происшествие, но может минимизировать ущерб.



**Рис. 1** Архитектура интеллектуальной системы прогнозирования и управления на основе алгоритма искусственных иммунных систем (клонального отбора)

Интеллектуальная система прогнозирования и управления работает с сегментом прогнозируемых событий. Информация о состоянии объектов по сети LCN поступает на устройство хранения данных НМ, далее осуществляется анализ и прогноз данных на основе алгоритма клонального отбора.

В основе алгоритма клонального отбора лежит следующий принцип. Генерируется популяция из  $N$  антител, каждая популяция определяет случайное решение для процесса оптимизации. Далее строится новая популяция-кандидат за счет отбора на каждой итерации лучших антител. Следующим этапом новые антитела оцениваются и добавляются к исходной популяции. Наихудшие антитела предыдущего поколения заменяются на новые, случайно созданные [2]. Достоинством данного подхода является возможность решения задач в условиях неопределенности параметров.

Таким образом, предложенная интеллектуальная система прогнозирования и управления на основе алгоритма искусственных иммунных систем (клонального отбора) применяется для управления прогнозируемыми событиями на базе распределенной системы управления фирмы Honeywell и может быть использована при разработке многофункциональной искусственной иммунной системы, состоящей из подсистем, реализующих основные механизмы и алгоритмы функционирования иммунной системы человека для оценки состояния и прогноза поведения интеллектуальной системы.



*Работа выполнена по гранту КН МОН РК по теме: "Разработка когнитивной Smart-технологии для интеллектуальных систем управления сложными объектами на основе подходов искусственного интеллекта" (2018-2020 гг.).*

### **Литература**

1. Cutello, V., Nicosia G. An immunological approach to combinatorial optimization problems // *Lecture Notes In Computer Science*. – Springer-Verlag, 2002. – P. 361-370.
2. Castro, L., Von Zuben, F. Learning and optimization using the clonal selection principle // *IEEE Transactions on Evolutionary Computation*. – 2002. – Vol.6. – P. 239-251.
3. Cort, N., Coello C. Multiobjective optimization using ideas from the clonal selection principle // *Lecture Notes in Computer Science*. – Springer, 2003. – P. 158-170.
4. Ji, Z., Dasgupta D. Revisiting negative selection algorithms // *Evolutionary Computation*. – 2007. – Vol. 15. - №2. – P. 223-251.
5. Castro, L. Timmis, J. *Artificial Immune Systems: A New Computational Intelligence Approach*. - Springer, 2002. – 380 p.
6. Watkins, A., Timmis, J., Bogges, L. Artificial immune recognition system (AIRS): An immune-inspired supervised learning algorithms // *Genetic Programming and Evolvable Machines*. – 2004. – Vol.5. - №3. – P. 291 – 317.
7. Tarakanov, A. Formal peptide as a basic agent of immune networks: from natural prototype to mathematical theory and applications // *Int. Workshop of Central Eastern Europe on Multi-agent Systems*. - 1999.
8. Yang, D., Gu, T., Zhou, H., Zeng, J., Jiang, Z. Design of DC Motor Control System Based on Clonal Selection Algorithm // *Advanced Manufacturing Systems, ICMSE 2011*. – 2011. – Vol. 201-203. – P. 2343-2348.
9. Hu, J., Yin, J., Li, T. Clonal Selection Identification and Adaptive PD Control for Uncertain Dynamical System // *Fifth International Conference on Natural Computation*. – 2009. – Vol.4. – P. 27-31.
10. Lee, J., Roh, M., Lee, Jn., Lee, D. Clonal Selection Algorithms for 6-DOF PID Control of Autonomous Underwater Vehicles // *Artificial Immune Systems*. – 2007. – P. 182-190.
11. Lima, F., Lopes, M., Lotufo, A., Minussi, C. An artificial immune system with continuous-learning for voltage disturbance diagnosis in electrical distribution systems // *Expert systems with application*. – Elsevier, 2016. – P. 1-12.
12. Alizadeh, E., Meskin, N., Khorasani, K. A Negative Selection Immune System Inspired Methodology for Fault Diagnosis of Wind Turbines // *IEEE Transactions on Cybernetics*. – 2017. – P. 1-15.
13. Kimura, N., Takeda, Y., Tsuge, Y. Agent Based Fault Detection System for Chemical Processes using Negative Selection Algorithm // *Advances in Science, Technology and Engineering Systems Journal*. – 2018. – Vol. 3. - №2. – P. 90-98.
14. Shi, X., Qiu, F. A Multi-agent Artificial Immune Network Algorithm for the Tray Efficiency Estimation of Distillation Unit // *Chinese Journal of Chemical Engineering*. – Elsevier, 2012. – Vol. 20. - №6. – P. 1148-1153.
15. Al-Sinbol, G., Perhinschi, G., Bhattacharyy, D. Evolutionary Optimization of Power Plant Control System Using Immunity-Inspired Algorithms // *International review if chemical engineering*. – 2017. – Vol.9. - №1. - P. 1-8.
16. Автоматизация. Система TotalPlant Solution. Локальная сеть управления LCN. Обслуживание // Honeywell. – 123 с.

# К ВОПРОСУ СОЗДАНИЯ РН С УЛУЧШЕННЫМИ ЭКОЛОГИЧЕСКИМИ ХАРАКТЕРИСТИКАМИ ЗА СЧЕТ ИЗВЛЕЧЕНИЯ ЭНЕРГЕТИЧЕСКИХ РЕСУРСОВ, НАХОДЯЩИХСЯ В НЕВЫРАБАТЫВАЕМЫХ ОСТАТКАХ ТОКСИЧНОГО КОМПОНЕНТА ТОПЛИВА В БАКАХ И МАГИСТРАЛЯХ ОС, ЗА СЧЁТ ИХ ПРИНУДИТЕЛЬНОГО ИСПАРЕНИЯ

Суйменбаев Б.Т.<sup>1</sup>, Трушляков В.И.<sup>2</sup>, Ермолдина Г.Т.<sup>1</sup>,  
Бапышев А.М.<sup>1</sup>

<sup>1</sup>Институт информационных и вычислительных технологий МОН РК,

<sup>2</sup>Омский государственный технический университет

e-mail: [bts49@mail.ru](mailto:bts49@mail.ru)

**Аннотация.** Предложена концепция снижения негативного техногенного воздействия одноразовых отработавших ступеней (ОС) в районах падения (РП) отделяющихся частей ракет-носителей (РН) с улучшенными экологическими и тактико-техническими характеристиками. Основу научно-методического обеспечения снижения техногенного воздействия пусков РН составляют рекомендации создаваемой информационно-аналитической системы района падения (ИАС-РП) для разработки технологических, схемных и проектно-конструкторских решений (ПКР) эксплуатантом РН для обеспечения снижения негативного техногенного воздействия ОС в РП. Обсуждаются решаемые задачи ИАС-РП, схема взаимодействия ИАС-РП с разработчиками РН, методы выбора проектно-конструкторских параметров автономной бортовой системы спуска ОС, обеспечивающей повышение тактико-технических характеристик и снижения техногенного воздействия пусков РН.

**Ключевые слова:** техногенное воздействие, районы падения, информационная аналитическая система, отработавшая ступень ракеты-носителя, неиспользуемые остатки топлива, пожаровзрывобезопасность, манёвры, автономная бортовая система спуска.

## Введение

Снижение техногенного воздействия ракетно-космической деятельности на окружающую среду является одной из актуальных проблем практической космонавтики. Это касается как загрязнения околоземного космического пространства крупногабаритным космическим мусором в виде верхних ступеней ракет космического назначения (РКН), так и отчуждения значительных по площади районов падения на поверхности Земли для нижних ступеней РКН [1].

Если проблема крупногабаритного космического мусора в околоземном космическом пространстве верхними ступенями РКН носит интернациональный характер, то проблема районов падения нижних ступеней РКН носит локальный характер и касается в основном эксплуатантов РКН и собственников территорий, где располагаются эти районы падения.

Использование на РКН маршевых жидкостных ракетных двигателей (ЖРД) дополнительно усугубляет эту проблему за счёт:

— наличия остатков топлива в баках верхних ступеней РКН может приводить к взрывам на орбитах (к настоящему времени уже случились взрывы более 160 ступеней);

— взрывов при спуске вторых ступеней РКН уже при входе в атмосферу, что приводит к существенному расширению площади разлёта фрагментов ступени и, соответственно, площади выделяемого района падения;

— увеличения площадей районов падения нижних ступеней РКН вследствие наличия в баках остатков топлива (дополнительные возмущения на угловое движение ступени со стороны остатков топлива в баках);

— проливов остатков топлива в районах падения нижних ступеней, что кроме механического разрушения приводит ещё и к химическому загрязнению почвы, вод, к пожарам.

В качестве одного из возможных направлений снижения техногенного воздействия РКН с ЖРД на окружающую среду на защищаемых орбитах в околоземном космическом пространстве была предложена реализация повторного запуска маршевого ЖРД для управляемого спуска вторых ступеней после выполнения ими своей миссии, используя невыработанные запасы кислорода и водорода.

Однако это предлагаемое решение неприменимо для спуска с траекторий выведения нижних ступеней РКН, падающих в зоны отчуждения на поверхности Земли. Это чрезвычайно важно для Казахстана, имеющей преимущественно районы падения на собственной территории, в отличие от США, Франции, Японии, имеющих районы падения ступеней в акваториях Мирового океана.

Повторный запуск маршевого ЖРД ступени РКН представляет собой достаточно сложную научно-техническую задачу, решённую в настоящее время, но для её практической реализации на разработанных и эксплуатируемых ЖРД требуется вложение значительных дополнительных средств, проведение больших объёмов дополнительных наземных экспериментов, и, как следствие, снижение достигнутой полётной надёжности.

В качестве альтернативного варианта реализации энергетических ресурсов, находящихся в невыработанных остатках топлива, предлагается разработка активной бортовой системы спуска (АБСС) ступеней РКН, которая начинает своё функционирование после отделения ступени [2].

Основными системами АБСС являются:

— система газификации, предназначенная для извлечения невыработанных остатков жидкого топлива, находящихся в неопределённом фазовом состоянии и граничном положении в топливных баках после выключения маршевого ЖРД. Испарение невыработанных остатков топлива предложено осуществлять путем подачи в топливные баки горячих газообразных продуктов сгорания специальных само-горящих газогенерирующих составов (ГГС).

— система реализации газифицированных продуктов в баках на основе подачи газифицированных продуктов из каждого бака в специальный газовый ракетный двигатель (ГзРД) или газореактивную систему стабилизации.

Наличие многоступенчатости ракет космического назначения (РКН) и невыработываемых остатков жидких компонентов ракетного топлива (КРТ) в баках ступеней после выключения маршевого ЖРД привели к тому, что в результате ракетно-космической деятельности кроме прочих проблем, связанных с проектированием, изготовлением и эксплуатацией, возникли две проблемы в части экологического воздействия пусков на окружающую среду:

- появление в защищаемых областях околоземного космического пространства (ОКП) потенциально взрывоопасного крупногабаритного космического мусора в виде орбитальных ступеней с остатками топлива в баках;

- наличие значительных площадей районов падения, выделяемых под отделяющиеся части ступеней, в частности, отработанные ступени (ОС), хвостовые отсеки (ХО), головные обтекатели (ГО).

Первая проблема имеет международный уровень и в настоящее время для разработчиков РКН сформулированы конкретные требования к организационным и проектно-конструкторским мероприятиям, позволяющим снизить поступление в защищаемые области ОКП потенциально взрывоопасного крупногабаритного космического мусора в виде ОС, а также последствиям при неуправляемом спуске в атмосферу [3-7].

Вторая проблема имеет преимущественно российскую специфику и связана, прежде всего, с наличием районов падения ОС, ХО, ГО на территории страны, в отличие от других стран, эксплуатирующих РКН и имеющих районы падения в акватории Мирового океана. Наличие районов падения на территории хозяйствующих субъектов приводит к необходимости проведения работ в части выделения этих районов, а также по ликвидации последствий после каждого пуска. Как следствие – значительные организационные и финансовые затраты.

В ряде работ рассмотрены различные варианты решения первой проблемы, например, путём повторного запуска маршевого ЖРД для управляемого спуска ОС после выполнения своей миссии в течение 1 – 2 витков. Для орбит, находящихся в ОКП до высоты 2000 км, рекомендован перевод ОС на орбиту с баллистическим сроком существования не более 25 лет с постепенным входом её в атмосферу и сгоранием в верхних слоях, а для более высоких орбит – перевод на орбиту захоронения со сроком существования свыше 1000 лет. Рассмотрена активная бортовая система спуска (АБСС), обеспечивающая либо перевод ОС на орбиту со сроком существования менее 25 лет, либо управляемый спуск в течение 1 – 2 витков [8-13].

Наличие многоступенчатости РКН и влияние этого фактора на окружающую среду заключается в необходимости выделения значительных территорий под районы падения. Подробно рассмотрены составляющие затрат, связанные с выделением и эксплуатацией районов падения, подходы по снижению их площадей. Основное внимание при снижении площадей районов падения уделено повышению точностных характеристик приборного состава системы управления, учёту сезонных факторов состояния атмосферы при расчёте программ выведения и т.д.

Величины остатков КРТ в топливных баках на момент выключения маршевого ЖРД в идеальном случае (когда все возмущающие факторы при заправке РКН, полёте действуют на увеличение фактических остатков топлива в баках) составляют до 3% от начальной заправки и являются результатом проведения значительного объёма работ по их снижению в различных направлениях, в том числе:

- повышение точности измерения
- уровней заливок КРТ в баках;
- оптимизация конструкции заборных устройств КРТ в баках и разработки системы сброса остатков КРТ после выключения маршевого ЖРД (пример модернизации ЖРД на первой ступени РКН «Протон-М»);
- модернизация системы управления топливом для обеспечения опережающей выработки наиболее токсичного КРТ и доработки ЖРД для безаварийного выключения по окончании одного из КРТ;

- использование терминального управления для движения РКН по «дожигающим» траекториям и разработка соответствующих алгоритмов;
- минимизация гарантийных остатков КРТ и введение индивидуальных заправок КРТ для каждой полезной нагрузки и т.д.

Все эти разработки требуют значительных затрат средств и времени на этапах НИР и ОКР и приводят, как правило, к усложнению конструкции, повышению сложности алгоритмов управления и, соответственно, к снижению тактико-технических характеристик РКН [14-16].

Проведённые исследования показали, что применение АБСС в составе ОС исключает необходимость проведения этих разработок в процессе проектирования, т.к. предлагаемые технология, схемные и проектно-конструкторские решения позволяют эффективно извлечь невырабатываемые остатки КРТ, а энергетический ресурс, заключённый в них, использовать для снижения техногенного воздействия на окружающую среду в различных направлениях, в том числе и для снижения количества и площадей районов падения.

Сравнительный анализ альтернативных вариантов решения данной проблемы показал высокую эффективность и потенциальные возможности АБСС для повышения тактико-технических характеристик РКН, снижения техногенного воздействия на окружающую среду и повышения технического уровня РКН, в частности, для последующего перехода к мягкой посадке ОС.

## **1 Существующие научно-методические подходы снижения влияния пусков РН в районах падения отделяющихся частей**

В настоящее время установлен следующий регламент ликвидации последствий техногенного воздействия пусков РН в РП космодрома Байконур.

1. Комплекс работ по устранению экологических последствий падения ОС в РП включает в себя:

- а) визуальная засечка района падения ОС методом триангуляции;
- б) оперативное реагирование на возникновение пожаров, последствий взрывов;
- в) поиск фрагментов ОС, документирование, инвентаризация заводских номеров двигателей и др. элементов;
- г) сбор и вывоз фрагментов ОС и т.д.

2. Комплекс работ по детоксикации почвенного покрова

Детоксикация почвенного покрова проводится по разработанным методикам на основе данных многолетних исследований «НИЦ «Фарыш-Экология», в частности, разработаны две технологии детоксикации почвенного покрова, загрязненного компонентами ракетного топлива:

- технологический регламент на детоксикацию почв, загрязненных НДМГ и продуктами его химической трансформации комбинированным методом;
- технологический регламент микробиологической детоксикации почв, загрязненных ракетным топливом – керосином, основанный на деструкции керосина аэробными формами аборигенных штаммов почвенных микроорганизмов.

## **2 Разработка научно-методического обеспечения снижения техногенного воздействия пусков РН в РП**

В целях снижения техногенного воздействия пусков РН с ЖРД в районах падения (РП) представляется целесообразным решение следующих задач:

1. Формирование необходимого объема информации по РП для разработки проектно-конструкторских решений и мероприятий для РН, пуск которой будет

осуществлён в этот РП. Эта задача решается путём разработки информационно-аналитической системы района падения (ИАС-РП)

2. Разработка проектно-конструкторских решений и мероприятий непосредственно для РН, пуск которой будет осуществлён в этот РП. Эта задача решается путём разработки активной бортовой системы спуска (АБСС) отработавшей ступени (ОС) на основе информации от ИАС-РП

3. Создание системы «ИАС-РП – разработчик РН»

4. Предложения по корректировке жизненного цикла РН и содержания отдельных его этапов

Для реализации поставленных задач необходима разработка научно-методического обеспечения снижения техногенного воздействия пусков РН в районах падения космодрома Байконур, которое должно:

- максимально использовать существующие наработки по системе экологического мониторинга космодрома (СЭМК)

- включать в себя современные методы повышения экологической безопасности РН, реализация которых возможна и доступна разработчикам и эксплуатантам РН, районов падения

- включать механизмы оценки влияния последствий пуска РН с конкретными тактико-техническими характеристиками на конкретный район падения и их экономическую оценку

- формировать рекомендации по изменению координат точек падения отработавших ступеней (ОС), предложения по изменению проектно-конструкторских параметров ОС, предотвращающих взрывы ОС, кардинальное снижение площадей районов падения.

### **3 ИАС-РП как элемент существующей системы экологического мониторинга космодрома Байконур**

В силу специфических особенностей ракетно-космической деятельности степень воздействия на компоненты биосферы различна в пространственно-временном масштабе. Поэтому в имеющейся информационно-аналитической системе (ИАС) космодрома особое внимание уделяется определению степени антропогенной трансформации экосистем и их компонентов под воздействием ракетно-космической деятельности, а также определению вклада неспецифического антропогенного воздействия на экосистемы.

Существующая ИАС состоит из следующих подсистем:

- подсистема подготовки исходных данных (ИД), архивации и документирования;

- подсистема планирования, учета и контроля, анализа и регулирования;

- подсистема математического и геоинформационного моделирования;

- подсистема обработки информации, анализа и принятия решений;

- подсистема нормативно-правового обеспечения.

ИАС космодрома, основанная на геоинформационных технологиях, решает ряд задач, связанных со сбором и обработкой информации о характере и видах воздействия РКД на окружающую среду; степени изменения свойств компонентов экосистем; определение масштабов аварий и источников загрязнения; прогнозирование последствий аварий; обеспечение государственных органов управления информацией для анализа и прогнозирования экологической ситуации в районах осуществления РКД.

ИАС-РП и взаимодействие с разработчиком РН и разработчиком системы экологического мониторинга космодрома (СЭМК) представляется следующим образом:

1) Информации от разработчика РН:

- технические характеристики РН, в том числе: тип РН, масса конструкции, остатки топлива в баках и магистралях

- оценка запасов топлива на борту ОС для манёвра в заданный район падения

- возможность манёвра ОС для изменения РП

- возможности воздушного взрыва ОС

- площади разлёта фрагментов ОС в планируемом РП

- наличие других возможных РП и точек прицеливания

- возможность наземного взрыва и т.д.

2) Информация от СЭМК по текущему РП:

- обобщённая характеристика возможных РП и точек прицеливания

- динамика облака рассеивание компонентов топлива

- возможность пожара в РП

- предложения по другим РП и точкам прицеливания

- стоимость работ по устранению экологического ущерба и т.д.

Основные идеи при организации взаимодействия ИАС-РП и СЭМК:

1. Использование существующей информации из блоков СЭМК (ИАС, геоинформационной системы, системы наблюдения)

2. Проведение совместно с ИАС СЭМК математического моделирования для получения оценок техногенного воздействия на окружающую среду для различных сценариев развития ситуации при пуске РН с ЖРД и конкретными техническими характеристиками РН, РП, в том числе:

- загрязнения воздушного пространства

- загрязнения почвогрунтов в РП

- Возможности пожаров и их последствий

3. Проведение совместно с ИАС СЭМК математического моделирования для получения оценок по затратам средств на устранение последствий пуска РН с ЖРД в указанный район падения для различных сценариев

4. Выбор вариантов изменения (смещения) РП, точек прицеливания с учётом наличия АБСС и т.д.

Разработка научно-методического обеспечения снижения техногенного воздействия пусков РН в РП предполагает решение следующих основных задач.

1. Создание ИАС-РП, как элемента существующей ИАС космодрома Байконур, для сбора данных, касающихся паспортных характеристик РП и характеристики конкретной ОС, падающей в этот РП, в том числе: расчётные и фактические остатки жидкого топлива в баках и магистралях (на момент выключения ЖРД), динамики спуска ОС на высотах ниже 20 км, параметры высотных взрывов ОС (при их наличии), разбросы фрагментов ОС, факт возгорания растительности и т.д.

2. Создание замкнутой системы «ИАС-РП – РН», формирующей рекомендации для эксплуатанта РН на основе комплексной оценки экологических последствий падения ОС с соответствующими характеристиками, в конкретные точки падения выбранного РП. Эти рекомендации являются исходными данными для разработки комплекса мероприятий по снижению техногенного влияния пуска РН по конкретному району падения, в том числе:

а) обеспечения пожаровзрывобезопасности ОС;

б) возможности изменения координат точек падения ОС для минимизации экологического ущерба в конкретном РП;

в) изменению программ выведения РН;

г) манёвру ОС и т.д.

3. Создание научно-методического обеспечения для разработки автономных бортовых систем спуска (АБСС) ОС, обеспечивающих снижение техногенного воздействия пусков РН с маршевыми ЖРД в районах падения при одновременном повышении тактико-технических характеристик РН в целом (расширение диапазона наклонов орбит выведения полезных нагрузок, возможное повышение массы выводимого полезного груза, снижение стоимости пуска РН и т.д.).

Достижение указанных показателей обеспечивается за счёт использования следующих основных перспективных технологий:

а) создание ИАС-РП и использование оперативной информации о прогнозируемом техногенном воздействии ОС с соответствующими техническими характеристиками (остатки топлива в баках и магистралях после выключения ЖРД) и параметрами входа (скорость движения центра масс, относительно центра масс) и т.д. на конкретный РП,

б) разработки рекомендаций ИАС-РП по возможности выбора другого РП или координат точки падения ОС в существующем РП;

в) практически полного извлечения энергетических ресурсов, находящихся в невырабатываемых остатках жидкого топлива в баках и магистралях ОС за счёт их испарения;

г) использования полученных энергетических ресурсов для совершения манёвров ОС на атмосферном участке траектории спуска с использованием АБСС.

На рисунке представлены преимущества разработки автономной бортовой системы спуска отработавшей степени РН.



Рис. Разработка автономной бортовой системы спуска отработавшей степени РН



## **Выводы**

Таким образом, на основании проведенного анализа:

1. Сформулирована концепция снижения негативного техногенного воздействия в районах падения космодрома Байконур для одноразовых неспасаемых ОС, основанная на оперативных рекомендациях созданной ИАС-РП к разработчику РН и созданию автономной бортовой системы спуска ОС, которая позволит обеспечить:

- а) пожаровзрывобезопасность ОС,
- б) минимизацию площади падения ОС,
- в) возможность манёвра для изменения местоположения района падения ОС.

2. Разработаны предложения по созданию методики проектирования ИАС-РП для оценки техногенного воздействия пуска РН на выбранный район падения, интегрированной в общую ИАС космодрома Байконур.

3. Предложено использование АБСС для обеспечения кардинального снижения техногенного воздействия пусков РН с маршевыми ЖРД в районах падения ОС. Разработаны предложения по созданию автономной бортовой системы спуска ОС, обеспечивающей кардинальное снижение техногенного воздействия пусков РН с маршевыми ЖРД в районах падения ОС на основе модернизации программ управления выводением РН на активном участке траектории выведения, программ управления движением ОС на атмосферном участке траектории спуска для многократного снижения площади района падения (до размера круга с радиусом ~ 50 – 100 м) с практически полной выработкой жидких остатков топлива.

*Исследования проведены при поддержке гранта Минобрнауки РК (№AP05131162 от 02.02.2018), гранта Минобрнауки РФ (14.577.21.0157), контракта ГК Роскосмоса (НИР «Магистраль»).*

## **Литература**

1. Экологические проблемы и риски воздействий ракетно-космической техники на окружающую природную среду. Справочное пособие, под общей редакцией Адушкина В.В., Козлова С.И., Петрова А.В. - М: Изд. «Анкил». – 2000. - с.640.

2. Trushlyakov, V., Zharikov, L., Lempert, D. Development of solid fuel gas generating compositions to ensure non explosiveness of spent orbital stages of liquid rocket of space launch vehicles//Eurasian Chemico-Technological Journal. – Vol.19, No1 (2017). - P.63-70, - <http://ect-journal.kz/index.php/ect01/article/view/504>.

3. IADC-04-06, Rev 5.5 May 2014 Support to the IADC Space Debris Mitigation Guidelines - <http://www.iadc-online.org/Documents/IADC-04-06%20Support%20to%20IADC%20Guidelines%20rev5.5.pdf>, - <http://www.ntv.ru/novosti/1664219> (<https://news.mail.ru/incident/30087239/?frommail=1>

4. Blue Origin. URL: - <https://www.blueorigin.com> (дата обращения: 19.09.2016).

5. SpaceDev - // Space Nevada Corporation. URL: <http://www.spacedev.com> (дата обращения: 19.09.2016).

6. Analiz tsenovoi politiki SpaceX pri ispol'zovanii mnogorazovykh raket - // Kosmicheskaya lenta. 2016. 26 aprelya. URL: <http://kosmolenta.com/index.php/863-2016-04-26-resuability-economy> (data obrashcheniya: 19.09.2016).

7. [http://e-notabene.ru/ik/article\\_20451.html](http://e-notabene.ru/ik/article_20451.html)

8. Шатров, Я.Т., Баранов Д.А., Суйменбаев, Б.Т., Трушляков, В.И. Povysheniye pozharovzryvobezopasnosti pri ekspluatatsii otrabotavshikh stupeney raket-nositeley s zhidkostnymi raketnymi dvigatelyami [Fire and explosion safety improvement during the

launch vehicle worked-off stages with liquid propulsion engine operation]//Pozharovzryvobezопасnost —Fire and Explosion Safety, 2016. - Vol.25, no.4. - pp.30–42. DOI: 10.18322/PVB.2016.25.04.30-42.

9. V. Trushlyakov, Ya. Shatrov, B. Sujmenbaev, D. Baranov. The designing of launch vehicles with liquid propulsion engines ensuring fire, explosion and environmental safety requirements of worked-off stages// Acta Astronautica. – 2017. – Vol. 131. – pp. 96-101.

10. Б.А. Титов, С.А. Рычков. Уменьшение размеров районов падения отработавших блоков ракеты-носителя типа “Союз” при преднамеренном членении их конструкции. /Вестник СГАУ. - №1, 2007.

11. Патент №0002572014 от 27.12.2015. Способ уменьшения районов падения отработанных ракетных блоков первой ступени ракетоносителя при их параллельном соединении.

12. Ю.Л. Кузнецов, Д.С. Украинцев. Анализ влияния схемы полёта ступени с ракетно-динамической системой спасения на энергетические характеристики двухступенчатой ракеты-носителя среднего класса. /Вестник СГАУ Т. 15, - № 1, 2016г.

13. V. Trushlyakov, Ya. Shatrov, B. Sujmenbaev, D. Baranov. The designing of launch vehicles with liquid propulsion engines ensuring fire, explosion and environmental safety requirements of worked-off stages// Acta Astronautica. – 2017. – Vol. 131. – pp. 96-101.

14. Суйменбаев, Б.Т., Ахмеджанов, А.С., Караданов, Т.К. Математическое моделирование турбулентной диффузии загрязняющих веществ в атмосферу от мгновенного высоко-температурного источника большой мощности //Гидрометеорология и экология, №3, 2006.

15. Краснов, В.М., Суйменбаев, Б.Т. Sprayed rocket fuel in the atmosphere, dimensions of the cloud and its flight trajectory Geophysical Research Abstracts, - V3. -2001.

16. Суйменбаев Б.Т., Краснов В.М. Некоторые аспекты последствий аварий ракеты-носителя “Протон” над территорией Казахстана //Экологические системы и приборы. - №2. - 2001.

## ПРОЛЕГОМЕНЫ К ЧЕТВЕРТОМУ ЗАКОНУ ДИАЛЕКТИКИ

**Сулейменов И.Э., Пак И.Т., Бакиров А.С.**

***Аннотация.** Показано, что осуществление прорывных работ в области искусственного интеллекта невозможно без кардинальной модернизации методологической основы исследований в данной области. Показано, что такая модернизация может быть осуществлена только с применением аппарата объективной диалектики. Последовательное его применение к объектам, так или иначе связанным с представлениями об интеллекте, приводит к формулировке принципа диалектической симметрии, который, по-видимому, может рассматриваться как четвертый закон диалектики. Первые практические шаги по внедрению предложенной концепции в практику связаны с созданием новой парадигмы образования, которая оперирует, в том числе, представлениями о том, что «образование», понимаемое как общественная институция, не сводимо к индивидуальному обучению. Образование в современных условиях можно и нужно*

*рассматривать, в первую очередь, как инструмент воздействия на общественное сознание, точнее, его подсистемы.*

**Ключевые слова:** *объективная диалектика, сущности информационной природы, диалектика информации, законы диалектики, высшее образование*

Проблематика искусственного интеллекта (ИИ) становится одной из основных направлений исследований, связанных со становлением цифровой эпохи [1]. Наиболее дискуссионным вопросом, возникающим при обсуждении сущности ИИ, очевидно, является вопрос о природе интеллекта как такового, вопрос, который можно сформулировать предельно кратко: «Что же такое интеллект?». Уместно подчеркнуть, что в настоящее время этот вопрос приобретает не отвлеченно-теоретическое, но непосредственное практическое значение, в том числе, в связи с разработкой инструментов, позволяющих адекватно управлять информационными потоками в обществе.

Ответ на него неоднократно пытались дать с самых различных позиций [2,3], но, по нашему мнению, все эти попытки обладают вполне определенным недостатком, связанным с тем, что понятие «интеллект» неразрывно связано с понятием «информация», последовательной трактовки которого в литературе до самого последнего времени предложено не было.

Действительно, можно заведомо утверждать, что «интеллект» представляет собой некую сущность сугубо информационной природы. Так, сознание человека появляется в результате обмена сигналами между нейронами головного мозга, каждый из которых выполняет сравнительно простые функции. Принято говорить, что сознание представляет собой системное свойство, но само по себе это утверждение только констатирует очевидный на сегодняшний день факт – оно не сводимо к свойствам отдельных нейронов, это – новое качество, которое появляется в сложной системе в силу наличия разветвленной системы связей между ее элементами.

Разумеется, можно говорить о том, что в данной ситуации представляет собой проявление одного из законов объективной диалектики – закона перехода количества в качество. Но, далеко не очевидно, *какая конкретно* сущность здесь испытывает количественно-качественный переход.

В данной работе предпринимается попытка показать, что человеческое сознание есть результат процессов самоорганизации, только опосредованно связанных с материальным носителем, это – *самоорганизация неких сущностей, имеющих сугубо информационную природу*. Подчеркнем, что в настоящее время уже можно с уверенностью говорить о том, что во многих сложных системах различной природы возникают нетривиальные информационные сущности, которые только опосредованно связаны с материальными носителями информации. Такие информационные сущности обладают своим собственным поведением, причем часто проявления их свойств затруднительно обнаружить, изучая поведение элементов системы по отдельности. Примером нетривиальной информационной сущности такого рода является бюрократия – результат процессов *информационной* самоорганизации в административных системах. Как показано в [4] бюрократия представляет собой некий отдаленный аналог биологического организма, обладающий выраженным аналогом инстинкта самосохранения и собственным «целеполаганием», характер которого, как правило, не имеет ничего общего с теми задачами, для которых изначально создавался административный аппарат.

Может показаться несколько удивительным, но весьма наглядной иллюстрацией к возникновению нетривиальных информационных сущностей в коммуникационных

оболочках сложных систем, служит поэзия. Для поэтических текстов в полной мере характерно то, что Лотман именовал «внетекстовыми структурами», т.е. некое новое качество, которое не связано с прямой передачей информации через слова естественного языка. В этом смысле допустимо говорить о том, что слова поэтического текста сами образуют особую коммуникационную оболочку, в которой содержится нечто иное, что невозможно увидеть, истолковывая текст буквально. Именно это обстоятельство и пытались положить (разумеется, не используя язык современной теории информации) в основу своих философских построений теоретики символизма; так Вяч. Иванов утверждал, что поэзия есть «тайнопись неизреченного» [5]. Символисты полагали, что ценность стихотворной речи состоит в «недосказанности», «утаенности смысла».

Изучение процессов самоорганизации сущностей информационной природы представляет собой нетривиальную задачу, так как они являются ненаблюдаемыми (в том смысле, который в данный термин вкладывает объективная диалектика). *Любой* мыслимый эксперимент, так или иначе, может предоставить сведения только об изменениях состояния материальных тел – носителей информации, судить о процессах самоорганизации информационных сущностей можно, следовательно, только по *косвенным* данным (в особенности это относится к ситуациям, когда информация является распределенной, как это имеет место в нейронных сетях).

Точнее, при изучении поведения сущностей *информационной* природы на первый план выходит именно *методология*.

Действительно, нужно иметь *теоретический* инструмент, который позволит адекватно *истолковать* большой массив экспериментальных данных и привести их в форму, позволяющую *судить* о том, что именно происходит в коммуникационной оболочке сложной системы. Основой для разработки такого рода инструментов может быть, очевидно, только методология, которая применительно к теории информации развита заведомо недостаточно, поскольку до самого недавнего времени отсутствовало адекватное истолкование самого понятия «информация» как такового.

Как показано в [6], дать такое истолкование можно *только* средствами философии, поскольку понятие информации относится к числу базовых. Точнее данное понятие можно и нужно рассматривать именно как *категорию* объективной диалектики. В частности, это означает, что изучение проблем, связанных с понятием «интеллект», в частности, наиболее фундаментальных проблем искусственного интеллекта, нельзя вести иначе, кроме как с привлечением аппарата объективной диалектики.

Отсюда – необходимость создания собственной методологической основы теории искусственного интеллекта, принципиально отличающейся от той, что используется в настоящее время. Фактически, методология работ, ведущихся в данной области во всем мире, не отличается от методологии, которую используется при разработке классических вычислительных устройств. Очевидно, рано или поздно придется выйти за ее рамки, так как системы ИИ необходимо должны представлять собой нечто качественно иное, нежели существующие цифровые устройства.

Как будет ясно из дальнейшего, развитие собственной методологической основы для работ в этом направлении с непреложностью приводит к положению, которое вполне можно рассматривать как четвертый закон диалектики.

По состоянию исследований на сегодняшний день его можно кратко сформулировать как «*принцип диалектической симметрии*».

Рассмотрим основания для выдвижения данного принципа.

Как отмечается в [6], информация – одно из самых фундаментальных понятий, широко используемых как в естественных, так и в гуманитарных науках. Поэтому говорить об «определении» информации в смысле школьной дефиниции представляется неоправданным. Дать определение в таком контексте – значит раскрыть содержание одного термина, используя другие. Пользуясь этим методом, рано или поздно придется прийти к понятиям, которые уже нельзя будет выразить через другие.

Чтобы этот круг был разорван, должны существовать определенные базовые категории, которые нельзя раскрыть через какие-либо другие. Объективная диалектика находит выход из положения, определяя такие категории *через противопоставление*.

Существует целый ряд парных диалектических категорий: содержание и форма, количество и качество и т.д. Им нет определения, для них и не может существовать «дефиниции» именно в силу их базового, фундаментального характера.

Именно к таким понятиям относятся, в том числе, категории материи и информации [6].

Любопытно отметить, что до самого последнего времени философская трактовка понятия «материя» испытывала те же самые трудности, что и в случае анализа категории информации (это, впрочем, представляется достаточно объяснимым в силу существования теснейшей связи между упомянутыми фундаментальными понятиями.)

Чтобы показать это, напомним «ленинское определение материи»:

«Материя есть философская категория для обозначения объективной реальности, которая дана человеку в ощущениях его, которая копируется, фотографируется, отображается нашими ощущениями, существуя независимо от них».

Сходные по структуре определения используются и в настоящее время. В постсоветских учебниках по философии можно найти, например, такие определения как «С точки зрения диалектики, материя есть объективная реальность – причина, основа содержания и носитель всего многообразия мира» [7]. Или там же: «Материя – это не реальная возможность всех форм, а действительное их бытие».

По своей сути (точнее по логической конструкции) указанные определения практически не отличаются от определений информации, рассмотренных выше. Это подчеркивается и во многих общедоступных ресурсах, которые в настоящее время широко используют студенты для изучения философии: ленинское определение, скорее, гносеологично, чем онтологично, ибо сказать, что материя – объективная реальность, – это в содержательном плане ничего не сказать<sup>1</sup>.

Трудно сказать почему, но во всей материалистической диалектике категория материи занимала особое, т. е. не имеющее пары, место. Понимая, что базовые категории не могут быть определены никаким иным образом, кроме как через противопоставление, специалисты по данному вопросу оставляли категорию «материи» (в этом смысле) в стороне.

Затруднения, о которых говорилось выше, снимаются автоматически, если признать, что «материя» и «информация» представляют собой именно парные диалектические категории. Их, так же, как и другие, следует определять (и можно определить) только через противопоставление. Существует точка зрения, в соответствии с которой информация противопоставляется энергии [8]. Однако

---

<sup>1</sup> <http://filosofedu.ru/index.php/otvety-na-voprosy-po-teorii-filosofii/559-materija-i-ejo-sistemnaja-organizacija>

последовательно применяя аппарат объективной диалектики, ее нельзя признать состоятельной, так как движение рассматривается как форма существования материи.

Природа объективной реальности двойственна, что в том или ином смысле признавалось и признается всеми философскими школами (здесь уместно вспомнить, например, о широких для того времени дискуссиях относительно универсалий, в которых столкнулись точки зрения номиналистов и реалистов).

Любой объективно существующий объект – атом, молекула или нечто макроскопическое – содержит (сформулируем это так, за неимением лучшего термина) как материальную, так и информационную составляющие. Эта информация проявляется и на микроуровне, и на макроуровне, и на всех промежуточных уровнях.

Предложенный подход, который можно назвать принципом дуализма материи и информации, автоматически расставляет все по местам. Материя – философская категория, но это не значит, что ее конкретные разновидности нельзя измерять в соответствующих физических единицах. Материя может выступать в самых различных ипостасях, и для измерения их характеристик в каждом конкретном случае, вообще говоря, требуются свои методы.

Точно также дело обстоит и с информацией. В некоторых областях существуют методы ее достоверного измерения, в каких-то – еще нет. Однако каковы бы не были успехи в конкретной области, связанной с изучением информации той или иной разновидности (например, биологической), вряд ли оправдано придавать им абсолютное значение.

Однако при использовании сформулированного выше противопоставления, на первый взгляд, исчезает конструктивность определения информации; оно становится слишком общим, чтоб им можно было пользоваться в тех же самых целях, в которых, например, проводится различие между микро- и макроинформацией. Преодолеть данное затруднение можно, вводя категорию отчужденной информации [6].

Конструктивность определения информации, даваемой через противопоставление, выражается также в том, что на этой основе можно сформулировать принцип диалектической симметрии. (Как известно, многие важные научные результаты были получены именно на основании соображений симметрии, например, принцип Онсагера.)

Сходным образом, принципы симметрии могут быть применены и к противопоставлению материи и информации, рассматриваемых как диалектические категории. Из этого принципа, прежде всего, вытекает, что подобно тому, как может существовать бесконечное разнообразие форм материи, точно так же может существовать и бесконечное разнообразие форм информации. Материя может быть косной, а может и быть живой, следовательно, симметричные понятия должны существовать применительно к информации; это создает дополнительные возможности для изучения нетривиальных информационных сущностей, возникающих в сложных системах, философскими методами.

Далее, именно принцип диалектической симметрии, связанный с дуальностью материи и информации, позволяет адекватно истолковать соотношение того, что именуется «энергия» и того, что именуется «информация». Как известно, существуют многочисленные спекуляции относительно вопроса о том, может ли информация обладать «энергетикой» и тому подобное. Обратимся к таблице 1.

**Таблица 1. К иллюстрации принципа диалектической симметрии**

Вещество	↔	Энергия
Информация	↔	?

«Фактор X», обозначенный в данной таблице знаком вопроса, пока не имеет названия и, по-видимому, даже не обсуждался в научной литературе (во всяком случае, не обсуждался широко). Однако, суждения, связанные с этим фактором, уже бытуют в массовом сознании. Так, принято говорить о положительной или отрицательной «энергетике» человека, об энергии психического», встречаются и представления о том, что информация неким образом может конвертироваться в энергию и/или рассматриваться как ее форма.

Анализировать весь спектр таких суждений заведомо не имеет смысла, поскольку за ними не стоит ничего, кроме неких интуитивных соображений, остающихся недоказуемыми.

Можно поступить проще – вспомнить генезис понятия «энергия» в его современной трактовке и попытаться действовать – или наметить план действий – по аналогии.

Как утверждают поясняющие статьи в общедоступных справочных ресурсах, понятие «кинетическая энергия» было впервые использовано не так давно, в 1829 году (работы Г. Кориолиса). Понятие «потенциальная энергия» является еще более поздним – 1853 год (У. Ренкин); по историческим меркам совсем недавно.

Первоначально оба эти понятия появились как результат математического исследования уравнений, описывающих движение материальной точки. То, что впоследствии стало формулой для кинетической энергии – произведение массы на квадрат скорости – появляется как первый интеграл при решении дифференциального уравнения второго порядка, в которое явно не входит переменная времени. Сходным образом обстоит дело и с потенциальной энергией – соответствующую функцию ввели просто для того, чтобы как-то назвать результат неявного интегрирования векторной функции, описывающей поле сил.

Именно это обстоятельство и имел в виду лорд Кельвин, заявивший в 1881 году перед слушателями:

*The very name energy, though first used in its present sense by Dr Thomas Young about the beginning of this century, has only come into use practically after the doctrine which defines it had ... been raised from mere formula of mathematical dynamics to the position it now holds of a principle pervading all nature and guiding the investigator in the field of science.*

Отголоски дискуссий середины 19-го века, участники которых – с разных позиций – пытались дать истолкование элегантным математическим построениям, приводящим к формулам, интерпретированным позднее как выражение законов сохранения, можно проследить и по философской литературе – достаточно вспомнить Анти-Дюринг Ф. Энгельса. Вспомнить об этих дискуссиях здесь уместно по вполне определенной причине – понятие «энергия» в его современном виде родилось из математических выкладок и только потом получило и физическое, и философское

истолкование. В массовое сознание оно проникло гораздо позже – где-то на рубеже 19-го и 20-го веков.

Следовательно, есть все основания поставить вопрос о построении теории, в которую будет входить информация, *обладающая способностью к воздействию* – параллели с ранними представлениями о «живой силе» и «кинетической энергии» просматриваются достаточно отчетливо.

По-видимому, именно такого рода соображения в обозримом будущем будут положены в основу количественного описания поведения нетривиальных информационных структур, порождаемых сложными системами, которые, в том числе, оказывают все более выраженное влияние на человеческое общество. Эти же соображения позволят пролить свет на давний вопрос об «энергетике информации» на корректной основе, что является еще одним аргументом в пользу конструктивности определения информации, даваемого методами объективной диалектики.

Принцип диалектической симметрии можно применить не только для внесения ясности в вопрос об «энергетике информации». Из него, в частности, непосредственно вытекает также существование того, что выше было названо нетривиальными информационными сущностями. Обратимся к Таблице 2.

**Таблица 2. Иерархия уровней организации материи и информации**

Уровни организации материи	Уровни организации информации
<p><b>Неорганический уровень</b> Подуровни: - микромир: субмикрорэлементарный – микроэлементарный – ядерный – атомарный – молекулярный уровни - макромир: макроскопические тела - мегамир: планеты – звездно-планетные комплексы – галактики – метagalактики</p>	<p>Неотчужденная информация, информация, отчуждаемая в результате физико-химических взаимодействий</p>
<p><b>Биологический уровень</b> Подуровни: - уровень биологических макромолекул – клеточный уровень – микроорганизменный уровень – уровень органов и тканей – уровень организма в целом – популяционный уровень – биоценозный уровень – биосферный уровень</p>	<p>Относительно обособленная информация, обладающая собственным поведением, самоорганизующаяся информация</p>
<p><b>Социальный уровень</b> Подуровни: - уровень индивидов – уровень семьи – уровень коллективов – уровень социальных групп – уровень национальностей и наций – уровень государств и систем государств (цивилизаций) – общечеловеческий (ноосферный) уровень</p>	<p>Информация, осознающая самое себя, высшие информационные сущности различной природы</p>

Данная таблица призвана подчеркнуть следующее обстоятельство. Подобно тому, как материя может выступать в самых различных проявлениях, так и



информация не может рассматриваться как нечто однородное (в частности, нет и не может быть завершеного понимания информации, относящегося ко всем ее проявлениям). Из принципа диалектической симметрии однозначно вытекает, что подобно тому как существуют иерархические уровни, отвечающие диалектической категории материи, так существуют и уровни, отвечающие категории информации.

В частности, речь идет о нетривиальных информационных сущностях, возникающих и развивающихся в коммуникационном пространстве, упоминавшийся выше. В гуманитарной литературе до сих пор широко используется термин «общественное сознание». Истолкование этого термина, доступное в философских словарях и справочниках, де-факто полностью построено на отрицаниях. Одно из них звучит, например, так

Общественное сознание – не простая совокупность индивидуальных сознаний, а целостное образование, обладающее внутренней структурой, которая включает в себя различные уровни (теоретическое, обыденное сознание, идеология, общественная психология) и формы сознания (политическое, правовое, этическое, эстетическое, научное, религиозное, философское) [9].

Или

Общественное сознание – отражение общественного бытия; совокупность коллективных представлений, присущих определённой эпохе. Оно отражает в сущности и само состояние конкретного общества. Общественное сознание противопоставляется индивидуальному сознанию как то общее, что содержится в сознании каждого человека как члена общества. но не с сайта [10].

По существу, все известные определения общественного сознания так или иначе построены на отрицании. Ключевым в цитированном и аналогичных ему определениях является тот факт, что общественное сознание не сводится к совокупности индивидуальных сознаний, что это – нечто иное. Однако, такого рода утверждения не позволяют ни в малейшей степени раскрыть природу этой сущности.

Первый шаг в данном направлении был сделан в монографии [11], где была предложена нейросетевая модель ноосферы. В рамках данной модели общество рассматривалось как аналог нейронной сети, причем индивидам ставились в соответствие отдельные нейроны, а межличностным коммуникациям – связи между нейронами сети.

Данная аналогия, помимо прочего, позволяет утверждать, что информационная сущность, именуемая общественным сознанием, в действительности обладает весьма сложным строением. Это вытекает из того, что ее порождает разветвлённая система коммуникаций, которая сама обладает очень сложным строением. Более того, есть все основания утверждать, что общественное сознание образовано совокупностью различных информационных сущностей, особенности которых определяются особенностями порождающих коммуникаций. Так родовые и семейно-клановые связи порождают информационные сущности одного типа, профессиональные коммуникации – другого и т.д.

Как одну из сущностей, пребывающих и развивающихся в общественном сознании, можно рассматривать любую из научно-технических парадигм. Напомним, что термин «парадигма» вошел в широкое употребление после выхода в свет знаменитой книги американского историка науки Томаса Куна [12]

Философская энциклопедия предлагает следующее толкование данного термина. Научная парадигма – совокупность научных достижений, признаваемых всем научным сообществом в тот или иной период времени и служащих основой и образцом для новых научных исследований. Впрочем, цитируемая статья вносит весьма

существенное уточнение: «К настоящему времени понятие парадигмы еще не получило точного значения».

По существу, Томас Кун говорил о том, что отдельная теория или отдельное научное достижение не может владеть умами настолько, чтобы полностью задавать вектор развития той или иной отрасли знания, это может сделать только определенная совокупность таких теорий. Одна из них может быть опровергнута или модифицирована, но в целом структура знания остается неизменной, соответственно не видоизменяется и вектор, задающий направление развития.

Несколько утрируя, можно сказать, что научная парадигма – это основа той идеологии (корректнее говорить о системе научных взглядов), которую разделяет подавляющее большинство ученых, работающих в данной области и которая официально кодифицирована текстами, признаваемыми классическими или каноническими.

Для наглядности можно провести аналогию между научной парадигмой и любым из вероучений. В средние века в среде теологов (схоластов) протекали жаркие дискуссии, сталкивались самые различные точки зрения, но все они оставляли неизменной первооснову – христианское вероучение в католической версии. Напротив, участники диспутов искренне полагали, что истина, рождающая в их спорах, служит только его укреплению.

Следовательно, парадигма – это то, что делает систему научных взглядов (при всей их пестроте и разнообразии) неким единым целым; она же формирует вполне определенный понятийный аппарат, определенный дискурс, который позволяет представителям научного сообщества ощущать себя единым целым, противопоставляя себя иным прочим.

Парадигма задает некое общепринятое видение мира, она очерчивает круг проблем, которые признаются значимыми и требующими решения. Одновременно она определяет допустимые методы решения этих проблем и т.д. Все это вместе создает определенную научную традицию, а также комплементарные ей формальные и неформальные институции (к последним относятся, в том числе, и различные неписанные правила, которых придерживается большинство ученых).

Обратим внимание на следующее обстоятельство: научная парадигма существует и развивается во многом подобно живому организму; отдельные ее носители (конкретные ученые) покидают этот мир (как могут отмирать отдельные клетки биологического организма), но рассматриваемая информационная сущность *как целое* продолжает существовать.

Более того, парадигма как таковая только опосредованно связана с памятью отдельного ученого, скорее ее существование следует соотносить с коллективной памятью, точнее с тем фрагментом глобальной коммуникационной среды, который связан с научными исследованиями и разработками. Не вызывает также сомнений, то научная парадигма представляет собой объект сугубо информационной природы, причем обладающий выраженным собственным поведением. Аргументом в пользу рассмотрения парадигмы как особой информационной сущности, обладающей своим собственным поведением, является и то, что соответствующая информация является распределенной, причем здесь вполне уместна аналогия с нейронной сетью. Потеря отдельной части коммуникационной сети не сказывается на ее функционировании как целого.

Не вызывает сомнений, что в современных условиях развитие информационных сущностей, порождаемых коммуникациями в обществе, претерпевают выраженные трансформации, что, в первую очередь, обусловлено развитием

телекоммуникационных технологий. В данном отношении многие авторы, например [13,14], говорят о становлении новой медиареальности.

Отсюда вытекает следующее важнейшее обстоятельство, имеющее непосредственное практическое значение. Высшее образование в современных условиях де-факто имеет дело не с индивидуальностями, как это было во времена Яна Амоса Коменского. Оно взаимодействует с нетривиальными информационными сущностями, поведение которых, в частности, определяет насколько эффективным будет обучение. Упрощая, студент легко усвоит те положения, которые зафиксированы окружающей его коммуникационной средой, тогда как усвоение сведений, не попадающих в данную категорию, будет крайне затруднено. Они де-факто будут отторгнуты как «неактуальные», «неинтересные» или «не имеющие отношения к повседневности». Непонимание этого обстоятельства, попытка преодолеть многочисленные кризисные явления в современной высшей школе в рамках парадигмы, восходящей к Коменскому, только усугубляют негативные тренды.

Соображения такого рода говорят о том, что методологическая проработка вопросов, так или иначе связанных с понятием «интеллект», имеет значение, в том числе, и с точки зрения преодоления кризисных явлений, присущих современной высшей школе. Модернизация высшего образования РК, необходимость которой подчеркивалась в Послании Президента РК от 10 января 2018 г., требует трансформации базовых подходов к организации высшего образования, комплементарных существенным изменениям, протекающим в современном обществе. Для этого, в свою очередь, требуется адекватное истолкование всей системы понятий, так или иначе связанных с категорией «информация». Формулировка принципа диалектической симметрии, предложенная в данной работе, позволяет сделать это, избегая многочисленных противоречий, которыми изобиловали ранее предложенные трактовки понятия «информация» и «интеллект».

Разумеется, говорить о том, что на данный момент данный принцип получил достаточное обоснование, равно как и о том, что ему целесообразно присвоить статус закона диалектики, на данном этапе исследований преждевременно (именно поэтому в заглавии работы был использован термин «пролегомены»). Однако, уже не вызывает сомнений, что многообразие сущностей информационной природы, для которых пока не имеется достаточного терминологического аппарата, не может не быть в том или ином смысле симметричным многообразию проявлений сущностей, связанных с категорией «материя». Ввиду очевидной фундаментальности такой симметрии, она должна выражаться теми или иными законами диалектики, окончательная формулировка которых представляется делом будущего.

*Работа выполнена в рамках РННТД коммерциализации 0218-17-ГК «Разработка и реализация новой комплексной информационной образовательно-профорориентационной технологии»*

### **Список литературы**

1 Никитина Е. А. Искусственный интеллект: философия, методология, инновации //Философские проблемы информационных технологий и киберпространства. – 2014. – №. 2.

2 Холодная М. А. СУЩЕСТВУЕТ ЛИ ИНТЕЛЛЕКТ КАК ПСИХИЧЕСКАЯ РЕАЛЬНОСТЬ? Вопросы психологии, № 5, 1990.—с. 121-128 //Вопросы психологии. – 1990. – №. 5. – С. 121-128.

- 3 Труфанов С. Н. ЧТО ТАКОЕ ИНТЕЛЛЕКТ? //Прикладная психология и психоанализ. – 2012. – №. 4. – С. 3-3.
- 4 Сулейменов И. Э., Нуртазин, А. А., Габриелян, О. А., Шалтыкова, Д. Б., Тасбулатова, З. С., & Панченко, С. В. Бюрократия с точки зрения теории самоорганизации //Образовательные ресурсы и технологии. – 2017. – №. 2 (19).
- 5 Михайлова М. Ю. Актуальные проблемы изучения семантики невыразимого //Вестник Башкирского университета. – 2015. – Т. 20. – №. 3.
- 6 Габриелян О. А., Мун, Г. А., Панченко, С. В., Сулейменов, И. Э. Диалектика информации //Ученые записки Крымского федерального университета имени В.И. Вернадского. Философия. Политология. Культурология. – 2015. – Т. 1. – №. 2. – С. 137-148.
- 7 Спиркин, А.Г. Философия : Учебник : 2-е изд. / А.Г. Спиркин. – М.: Гардарики, 2006. – 736 с.
- 8 Kirschenmann, P. K. Information and Reflection: on Some Problems of Cybernetics and How Contemporary Dialectical Materialism Copes with Them / P.K. Kirschenmann – Springer Science & Business Media, 2012. – 240 p.
- 9 Сергеева Т. Б. Словарь-справочник по философии для студентов лечебного, педиатрического и стоматологического факультетов //Ставрополь: СтГМА. – 2005.
- 10 Мамардашвили М. К. Анализ сознания в работах Маркса //Вопросы философии. – 1968. – Т. 6. – С. 14-25.
- 11 Сулейменов И. Э., Григорьев П. Е. Физические основы ноосферологии //Алматы—Симферополь. – 2008. – С. 158.
- 12 Томас К. Структура научных революций //Исследователь/Researcher. – 2010. – №. 1-2.
- 13 Савчук В. В. Медиареальность. Медиа субъект. Медиа философия //Медиа философия II. Границы дисциплины. – 2009. – С. 226-241.
- 14 Кузнецова Е. И. Медиа реальность как коммуникативный медиум //Медиа философия. – 2009. – Т. 2. – №. 2. – С. 67-79.

## **ОСНОВНЫЕ НАПРАВЛЕНИЯ МОНИТОРИНГА ТЕРРИТОРИИ СИНЬЦЗЯН-УЙГУРСКОГО АВТОНОМНОГО РАЙОНА КНР В ПРЕДЕЛАХ БАССЕЙНОВ ТРАНСГРАНИЧНЫХ РЕК И ЗАВИСИМЫХ ТЕРРИТОРИЙ НА БАЗЕ СПУТНИКОВОЙ ИНФОРМАЦИИ**

**Терехов А.Г.**

*Институт информационных и вычислительных технологий КН МОН РК,  
Казахстан  
e-mail: aterekhov1@yandex.ru*

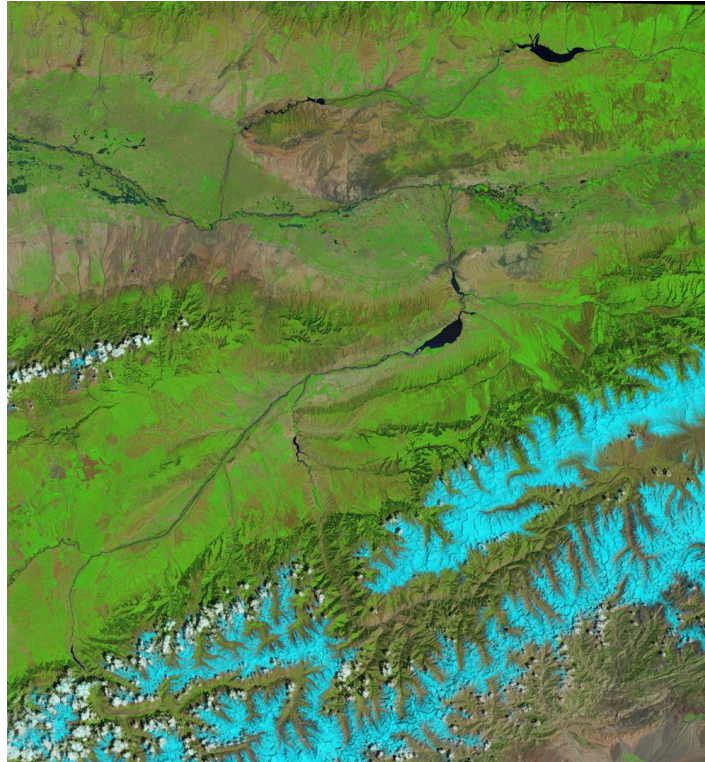
***Аннотация.** Работа посвящена организации и проведению спутникового мониторинга объектов гидротехнической инфраструктуры в пределах бассейна трансграничной (КНР-Казахстан) реки Иле. Обеспеченность водой территории Казахстана сильно зависит от расположенных выше по течению реки территорий*

*СУАР КНР. При этом важны параметры землепользования и воднотранспортной инфраструктуры СУАР КНР. Создана и тестируется система оперативного мониторинга запасов воды в доминирующих водохранилищах в китайском секторе бассейна р. Иле (Капшагайское [р.Текес] и Жарынтайское [р.Каиш]). Разработана спутниковая методика оценки расхода воды в реках региона, рассмотрены особенности регистрации расхода воды на гидропостах в районе границы КНР-Казахстан (Саньдаохэцзы (КНР). Пристань Добын (Казахстан), «164 км» (Казахстан). Показано, что в период 2006-2016 гг. система регистрации расхода воды в р.Иле на территории Казахстана имела значительные проблемы в точности измерений. Спутниковая методика оценки расхода воды в р.Иле может способствовать улучшению ситуации с достоверностью измерений расхода воды в реке.*

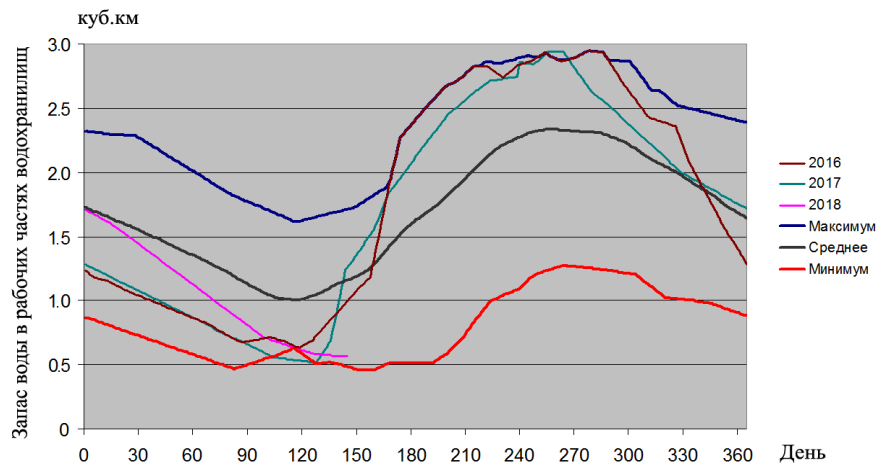
Бассейн трансграничной крупной реки (например р.Иле) является единым водохозяйственным «механизмом». Трансграничный характер территории бассейна создаёт проблемы доступности гидрологической информации с сопредельной территории. Её недостаток снижает эффективность управления водными ресурсами в нижней части бассейна [1]. В случае бассейнов трансграничных (КНР – Казахстан) рек Кара-Ертис и Иле, ситуация усугубляется аридностью климата и как следствие, значительной вариативностью сезонных объёмов речного стока на фоне её климатического дефицита и связанных с этим проблем с ирригацией пахотных земель [2]. Кроме того, рост потребления воды на территории СУАР КНР [3], включая развитие межбассейновой системы переброски речного стока, направленной на обводнения внутренних территорий за счёт трансграничных водных ресурсов, создаёт реальные угрозы водной безопасности значительной части Казахстана [4]. В том числе, озеру Балхаш, как конечному водному объекту бассейна р. Иле [5-8]. Отдельной проблемой становится согласование измерительных систем и регистрируемых объёмов расхода воды на гидропостах обеих стран в районе границы.

Отсутствие полноты объективной оперативной информации о состоянии и водопотреблении территории СУАР КНР в рамках бассейнов основных трансграничных рек (р. Кара-Ертис и р. Иле) и зависимых от их водных ресурсов территорий Джунгарской равнины послужило предпосылкой к проведению данных работ. Другой важной задачей является разработка спутниковых технологий оценки расхода воды в трансграничных реках, которая позволяет рассматривать их водный режим, как в оперативном режиме, так и в исторической перспективе (на глубину архива спутниковой съёмки).

В настоящее время создан комплекс программ обработки спутниковой информации в задаче оценки запасов воды в двух доминирующих водохранилищах в долине реки Иле (КНР) [9]. Этот инструмент позволяет решать ряд задач связанных с гидрологией сопредельной территории КНР [10-14]. С 2018 года в комплекс исходных спутниковых данных дополнительно к LANDSAT 7,8 (разрешение 30 м) были включены снимки с европейской орбитальной системы SENTINEL-2 аппараты А и В, с пространственным разрешением 20 метров, что позволило увеличить частоту посещения спутником территории водохранилищ и уменьшить период обновления данных до 2 дней. На рис.1, 2 показана территория долины реки Иле (КНР) с водохранилищами и дана оперативная информация по текущим и запасам воды в рабочих частях в сравнении с сезонами 2016, 2017 гг. и нормами (минимум, среднее, максимум) периода 2006-2016 гг.



**Рис. 1. Спутниковый снимок LANDSAT-8 разрешение 30 м, за 4 июля 2015 года. Китайская часть долины реки Иле с двумя доминирующими региональными водохранилищами: Капшагайское (р.Текес) и Жарынтайское (р. Каш)**



**Рис. 2. Спутниковые оценки суммарных запасов воды в рабочих частях Капшагайского (р. Текес, КНР) и Жарынтайского (р.Каш, КНР) водохранилищ. Построено по данным LANDSAT-7,8; SENTINEL-2A по состоянию на конец мая 2018 года**

Задача спутниковой оценки расхода воды в трансграничных реках территории КНР и Казахстана имеет большое значение для формирования оперативных баз гидрологических данных. Оперативные базы гидрологической информации верхней, китайской части трансграничного бассейна необходимы для эффективного

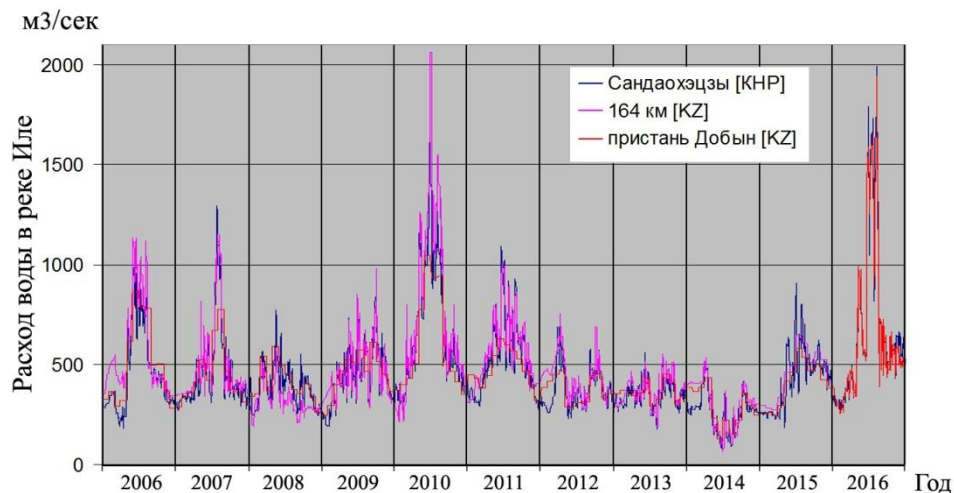
администрирования водным хозяйством среднего и нижнего течения реки, относящегося к территории Казахстана. В случае реки Иле, протекающей по плоской, слабо наклоненной равнине, сложенной, по большей части, легко размываемыми песчаными отложениями, возникают серьезные проблемы в точности измерения расхода воды. Меандрирующее русло реки в период многоводья заливает поймы, открывает и закрывает протоки. Все это приводит к существенной доле неопределенности в оценках расхода воды.

Методика определения расхода воды по спутниковым снимкам имеет большое преимущество перед стандартными наземными измерениями наличием возможности вернуться к историческим данным и соответственно, анализировать их с помощью различных методов обработки. Информация о наполненности русла водой зарегистрированная спутником доступна всегда. В наземных измерениях состояние реки мгновенно, измеренная информация остается, но провести проверку измерения повторно, позже невозможно, время не имеет обратного хода.

Задача стыковки данных по расходу воды в реке Иле в районе границы КНР-Казахстан, между постами КНР и Казахстан, а также между постами Казахстана, является ключевой для работы двухсторонних (КНР-Казахстан) соглашений по делению водных ресурсов трансграничной реки Иле в будущем, но в настоящий момент эта проблема далека от решения. Несмотря на отсутствие заметных притоков у реки Иле в районе границы с КНР, три гидропоста этого региона (Саньдаохэцзы (КНР), пристань Добын (Казахстан) и «164 км» (Казахстан)) см. рис.3 регистрируют заметно различающиеся данные, рис.4-6.

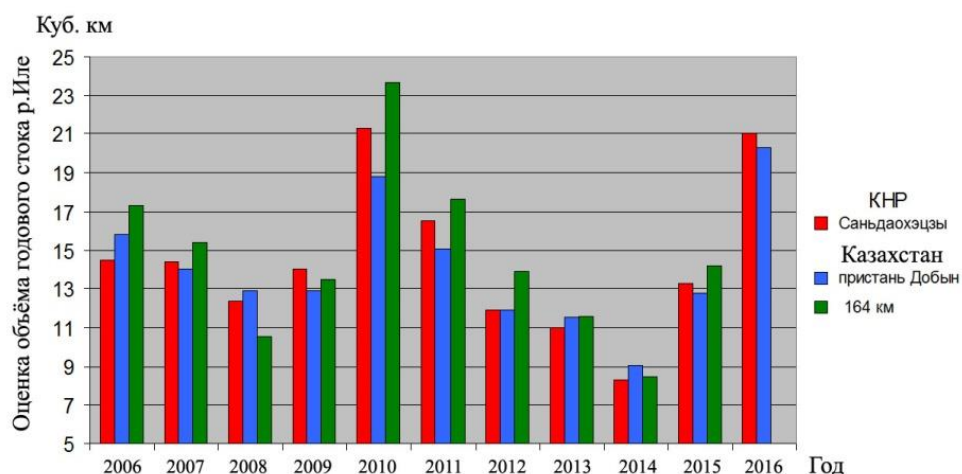


**Рис. 3. Картосхема расположения гидропостов в бассейне трансграничной реки Иле в районе границы КНР-Казахстан. Построено по снимку SENTINEL-2A на 7 июня 2018 г.**



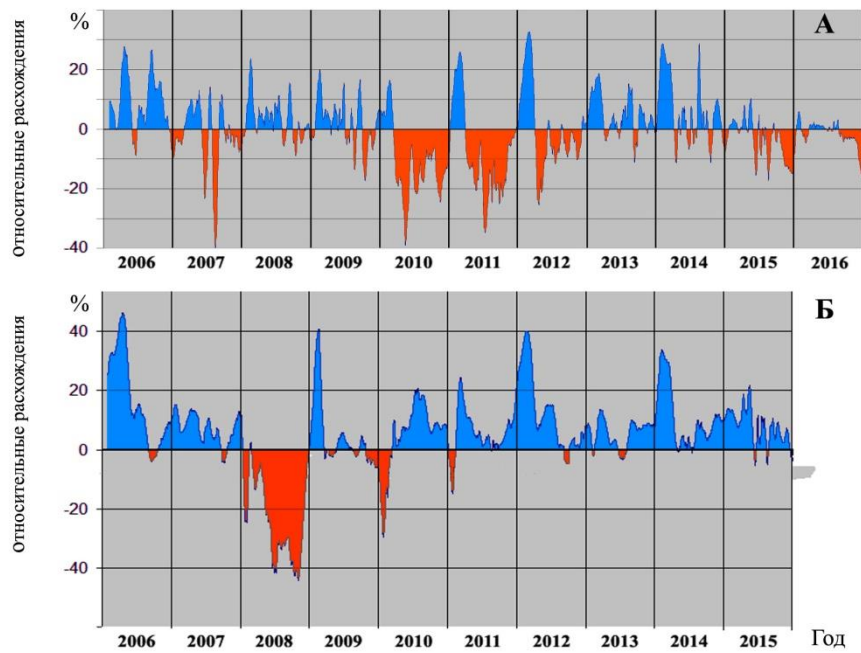
**Рис. 4. Зарегистрированные расходы воды в реке Иле (среднесуточно) на гидропостах в районе границы КНР-Казахстан в период 2006-2016 гг.**

Наиболее надежными представляются данные гидропоста Сандаохэцзы (КНР) с частично обустроенным в настоящий момент створом (бетонные дамбы), рис.7. Но у этих данных есть два недостатка. Во-первых, это данные другого государства, с которым происходит деление ограниченных водных ресурсов трансграничной реки. Т.е. нет уверенности в гарантированной доступности и объективности этой информации. Во-вторых, соотношение между уровнем воды в реке и её расходом воды на этом посту представляет собой достаточно консервативную кривую, не имеющую существенных межгодовых вариаций. И это несмотря на нестабильность гидрологического створа (формирование–размывание отмелей), возможного влияния ледовых условий в холодный период и режимов многоводья, когда окружающая гидропост пойма заливается водой, образуя 2-3 километровые подтопления, рис.8-10.

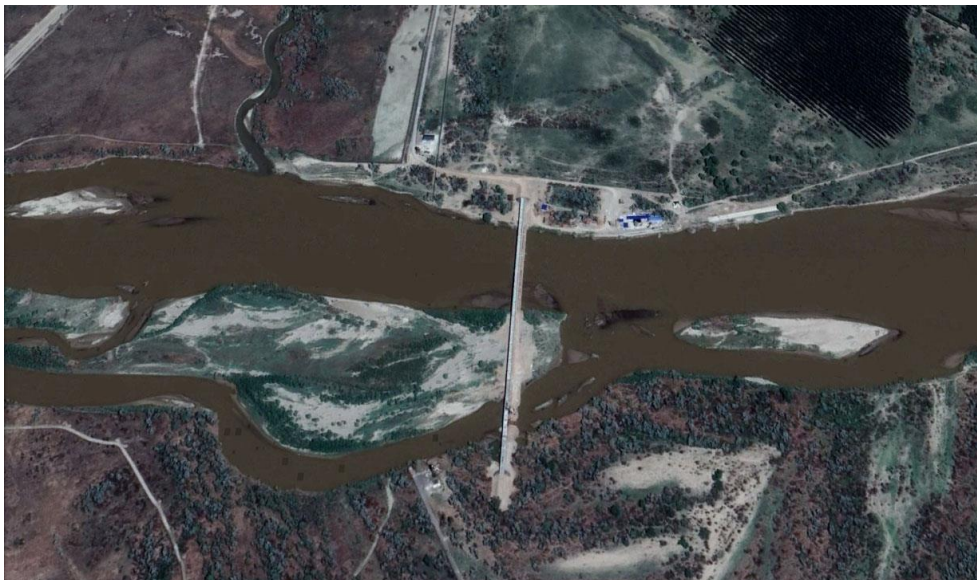


**Рис. 5. Объемы годового стока на р. Иле зарегистрированные на гидропостах в районе границы КНР-Казахстан в период 2006-2016 гг.**





**Рис. 6. Относительные расхождения в измеренных величинах расхода воды на р. Иле (плавающее 30-дневное окно) на гидропостах в районе граница КНР-Казахстан в период 2006-2016 гг. А) Пристань Добын – Саньдаохэцзы; Б) «164 км» - Саньдаохэцзы**



**Рис. 7. Состояние русла р. Иле в створе гидропоста Саньдаохэцзы (КНР) на 8 мая 2018 года по данным субметровой спутниковой съёмки (снимок из системы Google Earth)**

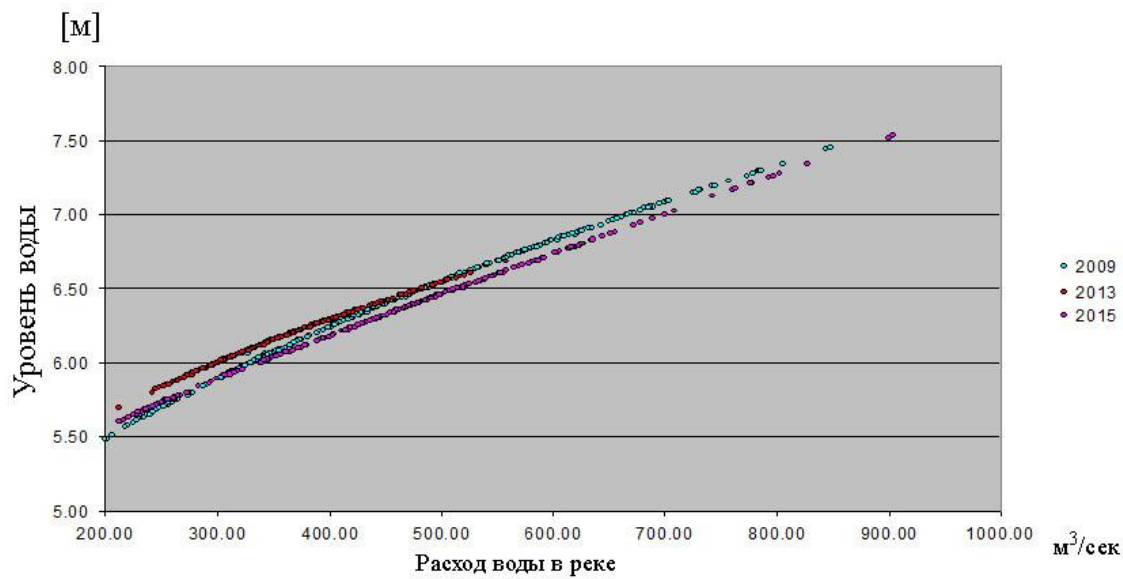


Рис. 8. Соотношение между расходом и уровнем воды в створе гидропоста Сандаохэцзы (КНР) в сезонах 2009, 2013 и 2015 гг. (суточные данные)

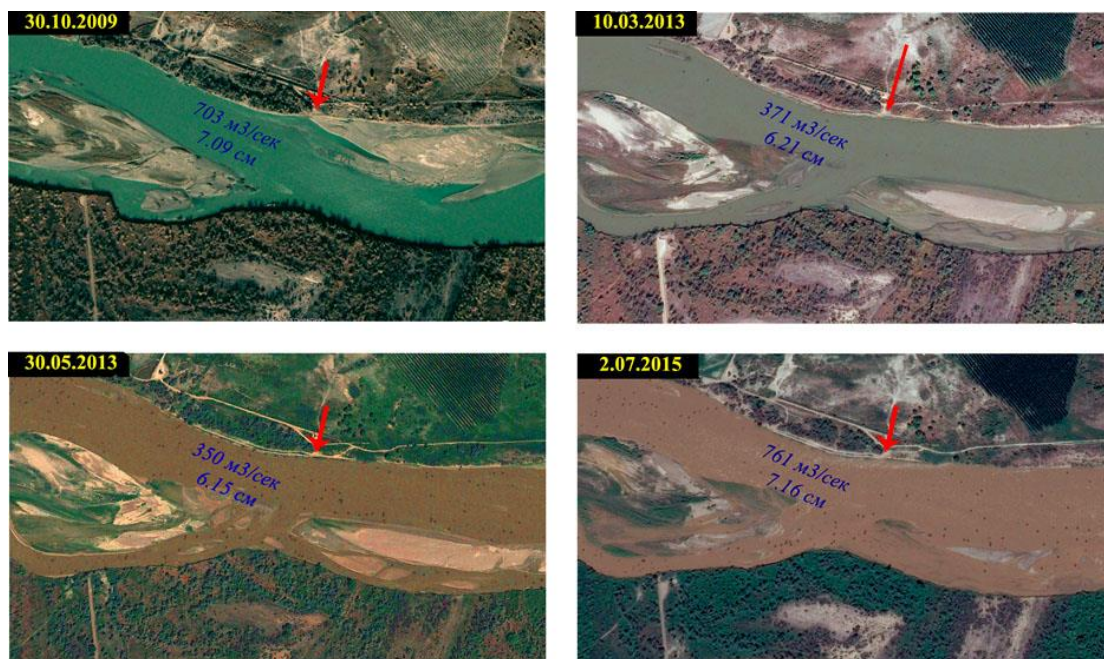
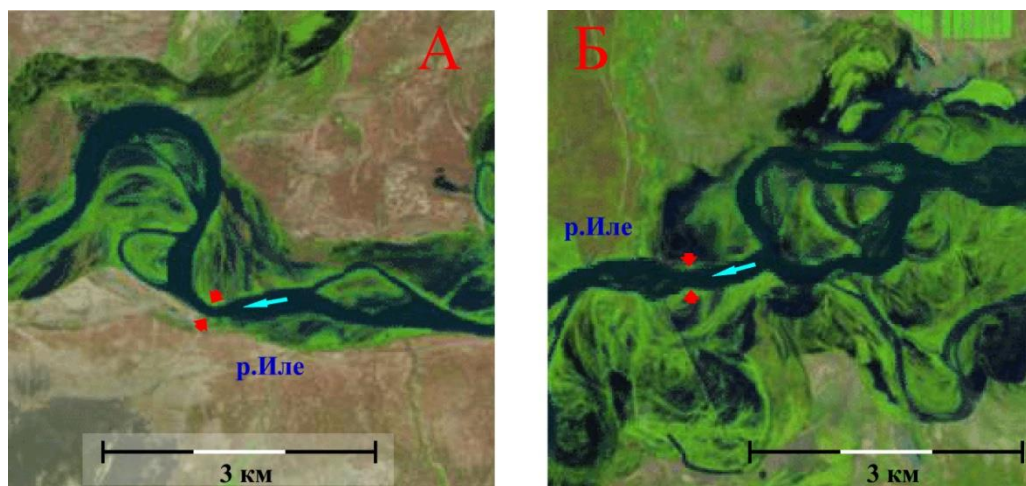
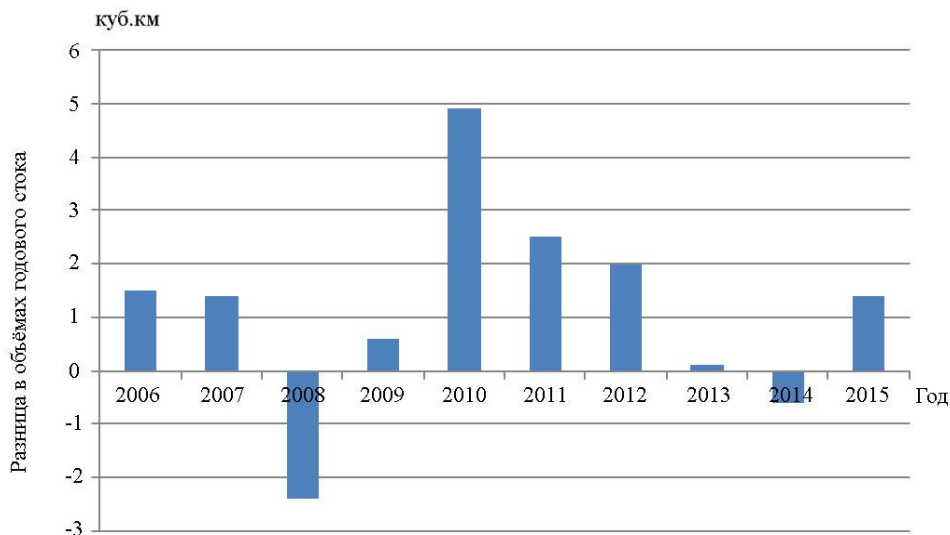


Рис. 9. Состояние русла р. Иле в створе гидропоста Сандаохэцзы (КНР) в сезонах 2009, 2013 и 2015 гг. по данным субметровой спутниковой съёмки (Quickbird)



**Рис. 10. Состояние русла р. Иле в створе гидропостов «Пристань Добын» [А] (расход воды 1640 куб.м/сек) и «Саньдаохэцзы (КНР)» [Б] (1730 куб.м/сек) на 17 июля 2016 года по данным спутниковой съёмки SENTINEL-2A (разрешение 20 м)**

Различия между данными по расходу воды регистрируемыми на гидропостах Казахстана (Пристань Добын – «164 км») недопустимо велики. В объёмах годового стока в период 2006-2015 гг. различие находилось в пределах от [-2.4] до [+4.9] млрд. куб.м, рис.11.



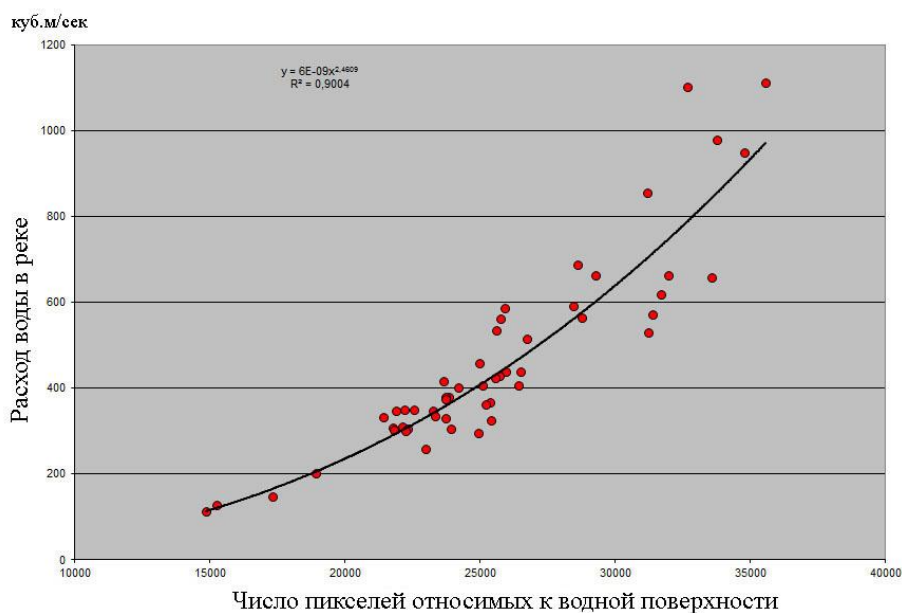
**Рис. 11. Разница в объёмах годового стока р.Иле, зарегистрированных на гидропосту «164 км» против данных поста «Пристань Добын»**

Если учесть, что между этими постами нет сколько-нибудь заметных притоков в реку Иле ситуация с системой организации контроля над расходом воды в реке представляется неудовлетворительной.

Проведение спутникового мониторинга над расходом воды в реках возможно только на основе данных среднего пространственного разрешения (10-30 м). Спутниковые данные субметрового разрешения помимо их высокой стоимости посещают территорию слишком редко. Частота залёта спутников среднего разрешения систем LANDSAT 7,8 и SENTINE L-2 (А,Б) позволяет обеспечивать обновление данных, примерно раз в 2-3 дня, т.е. формировать квазисуточные базы данных. Хотя серьёзной помехой остается облачный покров.

Разрешения спутникового снимка в 10-30 метров недостаточно для регистрации позиции русла реки, с точностью необходимой для оценки расхода воды, но такие спутниковые данные позволяют регистрировать различные уровни наполненности водой долины реки в целом. Был разработан метод обработки снимков 10-20 километровой участка долины реки Иле с целью получить численные характеристики по наполненности водой прирусловых территорий. Предполагалось, что эта характеристика должна иметь однозначную, функциональную связь с расходом воды в реке и таким образом возможно построение эмпирической калибровочной кривой между спектральными характеристиками прирусловой территории на тестовом участке и расходом воды в ней.

На основе суточных данных по расходу воды в р.Иле на гидропосту «164 км» и спутниковых снимков системы LANDSAT 5,7,8 в период 1997-2014 гг. (62 снимка) спутниковая методика оценки была опробована, рис.12. В целом, результаты показывают работоспособность такого подхода.



**Рис. 12. Взаимосвязь между наполненностью русла р. Иле водой (число пикселей на территории тестового участка прирусловой территории отнесённых к водной поверхности) и расходом воды. Построено на основе спутниковых снимков LANDSAT 5,7,8 периода 1997-2014 гг. и среднесуточных расходов воды по гидропосту «164 км» в день пролета спутника**

Таким образом, спутниковый мониторинг водных объектов позволяет оценивать запасы воды в водохранилищах китайского сектора бассейна реки Иле в оперативном режиме с периодом обновления данных от 2 дней. В задаче спутникового определения

расхода воды в реке Иле, снимки среднего разрешения (10-30 метров) спутниковых систем LANDSAT 7,8 и SENTINEL-2 (A,B), с периодом возврата от 2 дней, позволяют строить схемы оценки на основе спектральных характеристик прирусловых территорий. Эта методика опирается на физические характеристики территории, никак не связанные с аппаратно-методической основой измерения расхода воды на гидропостах. Это обстоятельство может позволить улучшить понимание существующих проблем в измерениях расхода воды в трансграничной реке Иле и содействовать улучшению системы наземных измерений.

### Литература

1. Мальковский И. М., Толеубаева Л. С. Трансграничный бассейн озера Балкаш: сценарии сбалансированного водопользования / Материалы II Всероссийской научной конференции «Водные и экологические проблемы Сибири и Центральной Азии». –Барнаул, 2014. – С. 337-343.
2. World Bank program, ECAEXT - project «Irrigation in Central Asia».-[http://siteresources.worldbank.org/ECAEXT/Resources/publications/Irrigation-in-Central-Asia/Irrigation\\_in\\_Central\\_Asia-Full\\_Document-English.pdf](http://siteresources.worldbank.org/ECAEXT/Resources/publications/Irrigation-in-Central-Asia/Irrigation_in_Central_Asia-Full_Document-English.pdf). (дата обращения: 03.09.2017).
3. А. Абдурахим. Общая оценка земельных ресурсов / Исследования аридных районов. Урумчи, 2000, N 1. - С.40-51.
4. Медеу А.Р., Мальковский И.М., Исаков Н.А., Толеубаева Л.С. /Водная безопасность Республики Казахстан: проблемы и решения. -Алматы, 2012. -200 с.
5. И.И. Скоцеляс Актуальные проблемы гидрометеорологии озера Балхаш и Прибалхашья/ -СПб: Гидрометеиздат, 1995. - 269 с
6. А.Б. Самакова Проблемы гидроэкологической устойчивости в бассейне озера Балкаш. / - Алматы, Издательство "КАГАНАТ", 2003. -579 с.
7. Турсунова А.А. , Мырзахметов А.Б. Водные ресурсы Иле-Балкашского бассейна с учетом международных принципов совместного использования. /Материалы международной научной конференции «Европейская наука XXI века». - Пшемысль, Польша. 2012. С.23-34.
8. Куджибаева Г. Актуальные вопросы по использованию водных ресурсов в области взаимоотношения трансграничных водных объектов / Ж. Наука и новые технологии. 2013. № 4, С.55-56
9. Терехов А.Г., Долгих С.А. Геоинформационная система оперативной оценки объёма запаса воды в водохранилищах китайского сектора бассейна реки Иле/ в кн. Водные ресурсы Центральной Азии и их использование. Матер. Межд. научно-практ. конф. посвященной подведению итогов объявленного ООН десятилетия «Вода для жизни». г. Алматы, 22-24 сентября 2016 г., книга 1. С.170-175.
10. Терехов А.Г. Спутниковые оценки параметров межгодового контр регулирования стока реки Иле с помощью водохранилищ в китайском секторе бассейна в период 2007-2016 годов/ Гидрометеорология и Экология 2017. №3 (86).- С.81-92.
11. Терехов А.Г., Долгих С.А. Спутниковый мониторинг гидрологического режима водохранилища на реке Текес (китайская часть бассейна реки Иле) / Вестник КНУ (серия географическая) – 2015, № 1 (40), -С.143–147.
12. Терехов А.Г., Пак И.Т, С.А.Долгих Данные LANDSAT 5,7,8 и ЦМР в задаче мониторинга гидрологического режима Капчагайского водохранилища на реке Текес (китайская часть бассейна реки Иле) / Ж. Современные проблемы дистанционного зондирования Земли из космоса.2015, Т.12, № 6, -С. 174-182.

13. Терехов А.Г., Калимолдаев М.Н., Пак И.Т. Компьютерное моделирование и спутниковые данные в задачах мониторинга некоторых гидрологических параметров в бассейнах трансграничных рек на примере китайской части бассейна реки Иле/ ВЕСТНИК КазНУ (Серия: математика, механика, информатика). 2016. №3/1(90). - С.75-86.

14. Терехов А.Г. Потенциал влияния водохранилищ китайского сектора бассейна реки Иле на режим её водности на границе КНР-Казахстан / Гидрометеорология и экология. 2016. №4 (83). -С.63-70.

## ТЕХНОЛОГИИ ОПРЕДЕЛЕНИЯ ГРАММАТИЧЕСКИХ КАТЕГОРИЙ КАЗАХСКОГО ЯЗЫКА

Тусупова М.Д.<sup>1</sup>, Мурзин Ф.А.<sup>1,3</sup>, Еримбетова А.С.<sup>1,2</sup>

<sup>1</sup>Новосибирский государственный университет, Россия

<sup>2</sup>Институт информационных и вычислительных технологий КН МОН РК,  
Казахстан

<sup>3</sup>Институт систем информатики им. А.П. Ершова СО РАН, Россия  
e-mail: [m.tussupova@gmail.com](mailto:m.tussupova@gmail.com), [murzin@academ.org](mailto:murzin@academ.org), [aigerian@mail.ru](mailto:aigerian@mail.ru)

**Аннотация.** Казахский язык является агглютинативным, номинативным языком, представляющим тюркскую группу языков. В данной работе описаны технологии определения существительных, прилагательных, местоимений, глаголов и наречий, а также процесс нахождения корня слова для заданного исходного слова. В качестве корпуса был взят корпус казахского языка – KazKorpus.

Казахский язык является агглютинативным языком, образование слов и новых форм слов осуществляется последовательным присоединением аффиксов. Поэтому набор меток должен быть разработан таким образом, чтобы максимально точно фиксировать эту информацию. Определение метки части речи, основанной на правилах, является самым старым подходом, в котором используются рукописные правила для меток (тегов). Тегеры на основе правил зависят от словаря или словаря, чтобы получить возможные теги для каждого помеченного тега. Ручные правила используются для определения правильного тега, когда слово имеет более одного возможного тега. [1]

В качестве корпуса был взят корпус казахского языка KazKorpus. В настоящее время размер первичного корпуса превышает 135 миллионов слов и содержит более 400 000 документов, разделенных на пять основных жанров:

- Литературный: включает казахские литературные произведения, в том числе романы, рассказы, стихи и т. д.;

- Официальное: включает в себя главным образом официальные статуты, приказы, акты и другие юридические документы, подготовленные правительственными организациями в течение 2009 и 2012 годов;

- Научный: включает книги, научные монографии, диссертации, статьи и эссе из различных областей, таких как информатика, биология, химия и т.д.;

- Публицистика (средства массовой информации): содержит периодические издания и статьи из онлайн-источников, т. Е. Газет и журналов, опубликованных за последние десять лет;

- Неофициальный: включает документы с казахскими текстами, извлеченными из популярных блогов, начиная с 2009 года.

Основными источниками данных были интернет-сайты, а также оцифрованные книги, диссертации и статьи из публичных и личных библиотек, которые помогали сбрасывать различные типы шума, такие как тексты на других языках, веб-баннеры и т. Д. На рост потребовалось около 7 месяцев до его текущего размера, а процесс сбора данных все еще продолжается. В таблице 1 представлено количественное описание корпуса. [2]

**Таблица 1. Количественное описание корпуса KazKorpus**

Жанр	Количество документов	Количество слов	Количество уникальных слов
Литература	8 255	7 733 456	423 445
Публицистика	404 884	79 302 154	951 659
Официальные	25 302	44 670 856	335 264
Научные	527	2 227 878	153 877
Неформальные	6 110	1 337 953	162 074
ВСЕГО	445 078	135 272 297	1 365 202

На рисунке 1 показана работа теггера, основанного на правилах, а так же работа стеммера. Так как видно, что части речи не всегда определяются корректно, либо же не определяются, было решено нами добавить UnigramTagger.

	word	stem	pos
0	қарағандыда	қарағанды	[[қарағандыда, NN]]
2	бабалар	баба	[[бабалар, NN]]
3	сезі	сез	[[сезі, None]]
5	фолькорлық	фолькорлы	[[фолькорлық, ADJ]]
6	фестивалі	фестивал	[[фестивалі, None]]

**Рис. 1 Определение части речи с использованием теггера, основанного на правилах**

UnigramTagger реализует простой статистический алгоритм тегирования: для каждого токена он присваивает тег, который наиболее вероятно относится к типу этого токена. Прежде чем UnigramTagger можно использовать для пометки данных, он должен быть обучен на учебном корпусе. Он использует этот корпус, чтобы определить, какие теги наиболее распространены для каждого слова. UnigramTaggers обучаются с использованием метода обучения, который принимает тегированный корпус. После обучения UnigramTagger тег можно использовать для пометки немаркированных корпусов. UnigramTagger присваивает тегу по умолчанию None слову, тип которого не встречался обучающих данных.

На рисунках 2, 3 видна работа UnigramTagger на основе двух корпусов: KazKorpus и словарь Link Grammar Parser.

	word	stem	pos	pos_learn
0	қарағандыда	қарағанды	[[қарағандыда, NN]]	[[қарағанды, None]]
2	бабалар	баба	[[бабалар, NN]]	[[баба, None]]
3	сөзі	сөз	[[сөзі, None]]	[[сөз, NN]]
5	фолькорлық	фолькорлы	[[фолькорлық, ADJ]]	[[фолькорлы, None]]
6	фестивалі	фестивал	[[фестивалі, None]]	[[фестивал, NN]]

Рис. 2 Применение UnigramTagger на KazKorpus

	word	tag	pos_learn
0	адам	NN	[[адам, NN]]
1	қыз	NN	[[қыз, NN]]
2	бала	NN	[[бала, NN]]
3	ит	NN	[[ит, NN]]
4	мысық	NN	[[мысық, NN]]

Рис. 3 Применение UnigramTagger на словаре Link Grammar Parser

Производительность UnigramTagger сильно зависит от качества его обучающего набора. В частности, если набор упражнений слишком мал, он не сможет достоверно оценить наиболее вероятный тег для каждого слова. Производительность также пострадает, если набор для обучения значительно отличается от текстов, которые мы хотим пометить. [3] В случае рассмотренных рисунков видно, что качество определения частей речи на разных корпусах отличается. Причиной тому – UnigramTagger обучался на словаре Link Grammar Parser.

### Литература

1. <http://language.worldofcomputing.net/pos-tagging/rule-based-pos-tagging.html>
2. <http://kazcorpus.kz/klcweb/kk/>
3. [https://www.nltk.org/\\_modules/nltk/tag.html](https://www.nltk.org/_modules/nltk/tag.html)



# ЧИСЛЕННАЯ ОЦЕНКА ГЕНЕРАЛИЗАЦИИ В ТЕКСТАХ СМИ НА РУССКОМ ЯЗЫКЕ С УЧЕТОМ ИСТОЧНИКОВ ТОНАЛЬНОСТИ: ЛЕКСИЧЕСКИЙ ПОДХОД

Уалиева И.М., Красовицкий А.М., Мусабаев Р.Р.,  
Мейрамбеккызы Ж.

*Институт информационных и вычислительных технологий КН МОН РК  
i.ualiyeva@ipic.kz*

## **Аннотация**

*В статье представлен метод численной оценки генерализации текста с учетом его тональности. Предлагаемый метод основан на лексическом подходе. Особенностью метода является сравнение численной оценки генерализации текста с его экспертной оценкой. В статье описывается процесс наполнения словарей, показана оценка работы метода на базе расширенного словаря.*

## **Введение**

В современных условиях развития общества недостоверность СМИ и их влияние на социум представляют собой одну из важных проблем. Особое беспокойство вызывает усиленное негативное воздействие СМИ на мнение и сознание граждан. Наряду с этим возникает необходимость в анализе публикаций на их достоверность и наличие манипуляции. Существует множество стилистических, лингвистических способов для создания текстов в манипуляционном ключе, однако в нашем исследовании нами отобраны два критерия: негативная тональность и генерализация (обобщенность). Генерализация[1] представляет собой обобщенный способ описания события, необоснованное утверждение, которое может содержать скрытую ложь. Метод генерализации позволяет журналистам без углубленного понимания вопроса и проблемы создавать определенную поверхностную картину события либо объекта, при этом со скрытым навязанным посылом. Согласно исследованиям, в основе предубеждений лежит генерализация[9]. Примером генерализации могут служить статьи, где в обобщенной и негативной форме может быть выражена авторская позиция: «На самом деле этот союз - сугубо политический, так как страны, примыкающие к нему, автоматически становятся политическими марионетками Москвы»[24] (об экономическом союзе). Такие выводы автор статьи делает не на основе конкретного исследования, а в силу своего журналистского опыта. Генерализованные высказывания помимо обобщенного характера имеют негативную тональность: *Кремль контролирует телевидение административно и отчасти финансово, а Останкино контролирует Кремль суггестивно, символически. Оно компостирует мозги Кремлю.* Генерализованный метод передачи информации может служить для журналиста методом передачи субъективного отношения отдельного лица в качестве «общего объективного мнения». [7]

Целью исследования является разработка метода вычисления генерализации в корреляции с тональностью и авторской позицией. Данный подход авторов и разработанное ими программное приложение поможет в дальнейшем автоматизированным способом выявлять генерализованные публикации, которые негативно воздействуют на читателя.

## Описание программного приложения

Для проведения исследования по распознаванию ключевых словоформ генерализации, тональности и авторской позиции, авторами были разработаны специальная программа на C# с подключением SQLite и Regular Expressions. В программу включены лемматизатор, словарь с морфологическими признаками с использованием распознающих правил, выделения выражений генерализации и расчета степени генерализации. Программный инструментарий имеет следующие шаги:

- парсинг текста;
- лемматизация;
- подключение словарей и лингвистических правил;
- морфологический разбор, распознавание синтаксических элементов;
- применение на них правил;
- суммирование весов правил, что в результате дает оценку степени генерализации.

Для всего текста программа выдает вес генерализации текста. Генерализация также может определяться с учетом тональности, авторства, слов-фактов.

## Постановка задачи

Пусть имеется текст публикации, каждая публикация имеет свою степень генерализации. Под генерализацией будем рассматривать свойство автора обобщать, делать выводы, основанные на наблюдениях за отдельными явлениями, в ходе сложного взаимодействия суждений, рассуждений, абстрагирования, воображения и решения задач. Требуется получить численную оценку генерализации текста публикации *gen*.

## Обзор подходов к решению задачи

Задача поиска генерализации непрерывно связана с задачей тональности, так как экспрессия оказывает наибольшее влияние на читателя [5]. Чтобы получить численную оценку генерализации текста можно применить множество подходов [16]. Предлагаемые исследователями подходы предложены для определения тональности текста. К примеру, Vinodhini и Chandrasekaran[20], выделили следующие наиболее популярные подходы для анализа тональности текстов:

- субъективный лексический подход – когда составляется список слов, где каждому слову присваивается оценка, указывающая на характер слова: положительное, отрицательное или нейтральное;
- *n*-граммы – когда на основе обучающих данных составляются юниграммы, биграммы, триграммы или их комбинации для дальнейшей классификации текста;
- машинное обучение – когда для извлечения информации из текста и обучения модели используют алгоритмы машинного обучения.

Исследователи (Maynard, Funk[11], D'Andrea[8], Thakkar[17]) предлагают схожие подходы для классификации тональности: (i) машинное обучение, (ii) лексический подход и (iii) гибридный подход.

Подход, основанный на машинном обучении, использует методы классификации, чтобы классифицировать текст. Документы разбиваются на два множества: обучающую и тестовую выборки. Обучающая выборка используется для обучения, тогда как тестовая выборка нужна для проверки того, как хорошо алгоритм произвел классификацию.

В лексическом подходе исследователи выделяют следующие методы:

- метод на основе словаря: первоначально создается словарь ключевых слов (вручную, на основе шаблонов или готовые словари), далее идет сопоставление слов из текста со словарями [11];

- метод, основанный на машинном обучении, объединяет лингвистический подход и машинное обучение. В этом подходе выявленные исследователями лингвистические признаки интегрируются в алгоритмы машинного обучения;

- метод, основанный на текстовых корпусах: из корпуса выявляются слова, и на основе этих слов делается заключение о тональности [15];

- ансамблевые методы [21]: повышают точность классификации путем комбинирования массивов отдельных обучающих выборок. Техники подхода: баггинг, бутстрапинг.

В гибридном подходе комбинация машинного обучения и подхода, основанного на лексике, имеет потенциал для улучшения качества определения тональности [13]. В комбинированном подходе сочетается тщательно разработанная лексика и мощность алгоритмов машинного обучения.

Лексический подход для определения тональности текста, с набором слов в словари [16], [18] был применен авторами для численной оценки генерализации. Ограничением применения машинного обучения является отсутствие большого объема текстового корпуса для его обработки. Применив лексический подход, авторы могут обойти такого рода ограничение, сосредоточившись на сборе словарей, осмыслении экспертами тестовых данных.

### **Предлагаемый авторами подход к решению задачи**

Пусть  $T$  – входящий текст в виде последовательности слов,  $K = \{k_1, k_2, \dots, k_l\}$  – это множество ключевых слов или фраз, которые используются для численной оценки генерализации текста. Слова или фразы из  $K$  назовем элементами  $K$ .  $F = \{f_1, f_2, \dots, f_m\}$  – множество элементов, которые выражают факт,  $O = \{o_1, o_2, \dots, o_p\}$  – множество элементов, выражающих мнение автора публикации. Заметим, что  $O$  представлено вводными оборотами. Численная оценка генерализации  $gen(T, K)$  будет вычисляться по частоте вхождений элементов текста, входящих в множество  $K$ :

$$gen(T, K) = \frac{|T \cap K|}{|T|},$$

$T \cap K$  – последовательность элементов  $T$ , из которой удалили все вхождения из  $K$ ,  $|T|$  – число элементов в тексте.

Расширенная оценка генерализации будет вычисляться по частоте вхождений слов множества  $S$ , входящих в множества  $K, F, A, O$ . Предполагается, что слова и фразы, входящие в множества  $F, A, O$  ослабляют генерализацию.

Вес генерализации предложения с учетом ключевых элементов

Легко заметить, что слова, фразы или вводные обороты из разных словарей имеют свой вес. Рассмотрим текст как последовательность предложений  $S = s_1, s_2, \dots, s_q$ , где  $q$  – количество предложений в тексте. Установим для каждого  $w_i \in s_r, 1 \leq r \leq q$ ,  $w_i$  – это подстрока  $s_r$  и  $w_i \in K$  вес  $weight(w_i) = 1$  для случая, если в предложении только находится только один элемент из ключевого словаря  $K$ . В случае, если число элементов из  $K$  будет больше одного, то вес генерализации

предложения  $weight(s_r)$  усилятся путем умножения суммы их весов на коэффициент генерализации  $\alpha$ :

$$weight^K(s_r) = \begin{cases} weight(w_i), w_i \in s_r, w_i \in K \\ \alpha \sum_{i=2}^l weight(w_i), w_i \in s_r, w_i \in K \end{cases}$$

где  $l$  – количество встречающихся слов, фраз или вводных оборотов в тексте из множества  $K$  в предложении  $s_r$ .

Вес генерализации предложения с учетом слов-модификаторов.

В случае ослабления генерализации будем также рассматривать текст как последовательность предложений  $S = s_1, s_2, \dots, s_q$ , где  $q$  – количество предложений в тексте. Рассчитаем вес  $weight^M(s_r)$  предложения  $s_r$  по следующей формуле:

$$weight^M(s_r) = \prod_{i=1}^m weight(w_i), w_i \in F, A, O, w_i \in s_r,$$

где  $m$  – количество встречающихся слов, фраз или вводных оборотов в тексте из множеств слов-модификаторов  $F, A, O$ . Установим первоначальный вес  $weight(w_i)$ , равным 0,75 для ослабления генерализации для случая, когда в предложении только находится только одно слово-модификатор.

Таким образом, вес генерализации каждого предложения в тексте с учетом включения словарей ключевых слов и словарей ослабления генерализации, будет рассчитываться по следующей формуле:

$$weight(s_r) = \begin{cases} weight^K(s_r), w_i \in K, s_r \cap (F \cup A \cup O) = \emptyset \\ weight^K(s_r) * weight^M(s_r), w_i \in K, F, A, O, \text{ для любого } w_i \in s_r \\ 0, w_i \in F, A, O, s_r \cap K = \emptyset \end{cases}$$

Численная оценка генерализации с учетом интенсификации  $gen(S, K, F, A, O)$  будет вычисляться как сумма весов генерализации каждого генерализованного предложения к общему числу предложений  $S$  в тексте:

$$gen(S, K, F, A, O) = \frac{\sum_r weight(s_r)}{|S|}.$$

Достоверность численной оценки генерализации будет основываться на экспертных оценках. Таким образом, для достоверности результата численной оценки текста нужно минимизировать разность между значениями, полученными экспертами  $\overline{gen}$ , и предсказанными значениями  $gen$ :

$$\sum (\overline{gen}_j - gen_j)^2 \rightarrow \min, \text{ где } j - \text{индексы статей всего текстового корпуса.}$$

Оптимизация идет за счет настройки весов ключевых элементов, а также за счет весов словарей тональности.

Рассмотрим подробнее на примерах. Рассчитаем вес для предложения «В Казахстане падает общий уровень жизни: снижаются качество образования, медицинских услуг и культура». В предложении три слова и фраза из множества ключевых слов  $K$ : «Казахстан», «падает», «общий уровень жизни», «снижаются».

Коэффициент генерализации  $\alpha$  установим, равным 3. Рассчитаем вес генерализации предложения по формуле (3). Тогда вес будет равным  $3 * (1 + 1 + 1 + 1) = 12$ .

Другое предложение «*Все как один: взрослые и дети, соблюдали жесткую дисциплину, призывали укреплять единство группы*» содержит три ключевых слова из множества  $K$ : «все как один», «взрослые», «дети» и два слова-модификатора, выражающих мнение автора из множества  $O$ : «жесткий» и «единство группы». Слова-модификаторы, выражающие мнение автора, ослабляют вес предложения, следовательно, их первоначальный вес равен 0,75. Тогда общий вес этих слов-модификаторов будет равен произведению  $0,75 * 0,75$ , что в итоге даст 0,56. Три ключевых слова дадут вес, равный 9 (девять). В итоге, согласно вышеприведенной формуле, общий вес генерализации будет равным  $9 * 0,56 = 5,04$ . На примере явно видно, что наличие слов-модификаторов значительно ослабляет генерализацию.

### Наполнение словарей

Основной сложностью лексического подхода является формирование словарей. Исследователи (D'Andrea et al.[8]) выделяют следующие техники для набора словарей: (i) ручной сбор словарей, (ii) с использованием текстового корпуса, (iii) с использованием шаблонов словарей или готовых словарей, при таком подходе можно собрать базовый набор слов вручную, дополнить его путем поиска новых слов в словарях RussNet или взять уже готовые словари, например, синонимов и антонимов. Для набора словарей авторы использовали ручной сбор для формирования словаря базовых элементов и готовый словарь тональности – словарь оценочных слов и выражения русского языка РуСентиЛекс[3]. РуСентиЛекс состоит из 16 000 слов и словосочетаний. Данный лексикон является упорядоченным словарем слов и словосочетаний. Слова разделены по типам тональности (негативный: *агрессия*, положительный: *прогресс*, нейтральный: *балласт*) и по источнику: оценка (opinion), чувство (feeling), факт (fact). В формировании словаря участвовали эксперты-лингвисты, которые вручную анализировали употребление каждого слова в новостном контенте. При создании словаря авторы ссылались на тезаурус русского языка РуТез.

Словарь базовых элементов был собран нами вручную, объем на данный этап более 600 слов, в нем множество базовых элементов генерализации  $K = \{k_1, k_2, \dots, k_l\}$ , где  $l$  – размер словаря. Категории базовых элементов были сформированы на основе работ исследователей Данковой[2], Смит[6], Орловой [5], Фроловой [7]. Во множество базовых элементов  $K$  вошли следующие категории слов для численной оценки генерализации:

- неконкретные глаголы, выражающие абстрактность, неконкретность событий, неопределенность действий: «казаться», «терпеть», «думать».
- неконкретные существительные: «холод», «видения», «многие».
- неконкретные местоимения: «какой-нибудь», «кое-что», «чей-нибудь».
- универсальные обобщители – это слова, не допускающие исключений: «без всякого сомнения», «ясное дело», «вечно», «все», «всегда».
- лексические единицы со значением регулярности/нерегулярности: «как правило», «обычно», «нечасто», «постоянно», «редко».
- количественные показатели: «каждый десятый», «много», «немало».
- лексические единицы с семантикой общепризнанности: «считается», «известно», «бесспорно», «несомненно».

• стилистические приемы, содержащие метафоры с обобщающим значением: «побил рекорды», «так называемое цивилизованное общество», «это наркотическая петля».

Авторами была выдвинута гипотеза, что высказанное автором, собственное мнение в виде вводных оборотов или других конструкций, а также подтвержденное событие, явление, происшествие (факт) ослабляет генерализацию, как, например, в нижеприведенных предложениях:

(1) *Виновник торжества был впечатлен тем, что горожане смогли оказать поддержку всем членам его семьи*

(2) *Строптивый и жесткий характер совсем не помогал ему налаживать отношения с коллегами*

(3) *Все как один: взрослые и дети, соблюдали жесткую дисциплину, призывали укреплять единство группы*

В первом предложении фраза «виновник торжества» из словаря тональности, обозначенная в нем как факт, ослабляет генерализацию базовых слов «горожане» и «всем». Во втором и третьем предложениях слова-мнения «строптивый», «жесткий», «группа», выражающие мнение автора, также ослабляют генерализацию предложения. Назовем такие элементы, ослабляющие генерализацию, элементами-модификаторами. Словари элементов-модификаторов были сформированы вручную на основе словаря тональности и вводных оборотов Национального Корпуса Русского Языка (НКРЯ).

### Эксперименты

Собранный словарь ключевых элементов имеет иерархическую структуру, а именно, для фраз имеется базовая форма, например, «каждый» и ее уточнение «каждый второй». Последняя имеет больший вес генерализации. Примеры применения регулярных выражений в словарях:

<code>\d+\.\?d{1,4}</code> миллионов	<code>\d+\.\?d* millions</code>
более <code>\d+\.\?d*</code> миллионов	more than <code>\d+\.\?d* millions</code>
привычны <code>\w{1,4}</code>	regular <code>\w{1,4}</code>
кажд <code>\w{1,4}</code>	each <code>\w{1,4}</code>
кажд <code>\w{1,4}</code> втор <code>\w{1,4}</code>	each <code>\w{1,4}</code> second <code>\w{1,4}</code>
кажд <code>\w{1,4}</code> трет <code>\w{1,4}</code>	each <code>\w{1,4}</code> third <code>\w{1,4}</code>

Чтобы минимизировать разницу между экспертной оценкой генерализованности текста и предсказанной программным путем, мы настраивали параметры весовых коэффициентов, такие как веса слов, фраз, вводных оборотов из ключевого словаря, веса элементов-модификаторов, коэффициент генерализации  $\alpha$ . Валидация слов, включённых в словарь, проверяется самими экспертами, на основе перекрестных оценок. В дальнейшем, валидность словарей будет дополнительно проверена с учетом значимости самих экспертов на основе метода Дельфи. Параметры весовых коэффициентов, правила, элементы словаря будут настраиваться до тех пор, пока разница между экспертной оценкой и программной оценкой не станет минимальной.

Для проведения экспериментов с численной оценкой генерализации и дальнейшего сравнительного анализа оценок, полученных экспертным и программным путем, авторы использовали публикации официальных и полуофициальных СМИ РК, которые находятся в открытом доступе. Выборка включала по 4 статей на генерализацию, позитивную тональность и негативную тональность. Всего экспертами определено пять степеней генерализации по аналогии

с коэффициентами корреляции: очень слабая генерализация, слабая генерализация, средняя генерализация, сильная и очень сильная генерализация.

Таблицы 1,2 описывают общее количество предложений в публикации, количество предложений с генерализацией, численную оценку генерализации публикации для выборки из публикаций. Публикации отобраны с учетом экспертной оценки степени генерализации. В таблице 1 приведены результаты сравнения алгоритма и экспертной оценки, полученные с использованием словаря по статье Данковой [2].

**Таблица 1**

№	Ссылка на статью	Количество предложений с генерализацией в публикации	Всего предложений	Численная оценка генерализации	Экспертная оценка генерализации
1	Дуванов	9	42	0,84/0,98	Очень сильная
2	Бомпиева	2	70	0,011/0	Очень слабая
3	Сатпаев Досым	3	46	0,087/0,036	Средняя
4	Тукпиев	33	157	0,092/0,081	Слабая

В таблице 2 приведены результаты алгоритма и экспертной оценки с использованием словаря по статье Данковой[1], Смита, Фроловой и Орловой [5-7].

**Таблица 2**

№	Ссылка на статью	Количество предложений с генерализацией в публикации	Всего предложений	Численная оценка генерализации	Экспертная оценка генерализации
1	Дуванов	36	45	0,489/1,24	Очень сильная
2	Бомпиева	25	70	0,107/0,305	Очень слабая
3	Сатпаев Досым	30	46	0,554/0,801	Средняя
4	Тукпиев	84	157	0,376/0,512	Слабая

Как видно из таблиц, наилучшее приближение получено во втором случае. Мы показали, что с расширением словарей, полученных благодаря лингвистам, растет точность распознавания генерализации с учетом тональности в тексте.

### **Заключение**

Задача получения численной оценки генерализации публикации не является тривиальной. Авторами предлагается методика численной оценки генерализации, основанная на лексическом подходе, и ее сравнении с экспертной оценкой. Минимизация расхождения между экспертной оценкой и вычисляемой алгоритмом, достигается за счет оптимизации параметров весовых коэффициентов ключевого словаря, элементов-модификаторов и коэффициента генерализации  $\alpha$ .

В словарь отбирались слова, фразы, вводные обороты генерализации, выбранные согласно работ исследователей в области психологии и лингвистики. Словари были сформированы вручную и на базе готовых словарей и вводных оборотов НКРЯ.

Методика численной оценки была апробирована на небольшом текстовом корпусе. Получены численные оценки, вычислены точность распознавания текста и корреляция между алгоритмической оценкой и экспертной. Эксперимент показал, что численная оценка имеет среднюю степень корреляции между алгоритмической оценкой генерализации публикации и ее экспертной оценкой.

По нашей программе ПЦФ предлагаемую методику численной оценки генерализации текста мы будем апробировать на большом текстовом корпусе объемом не менее 30 млн словоупотреблений, глубиной в семь лет официальных средств массовой информации РК, в котором будут храниться статьи, собранные из открытых источников.

В других работах авторы хотят провести исследования для выявления позитивной/негативной тональности публикаций, а также выявить совместное влияние генерализованности и тональности публикации на аудиторию, что позволит выявить аномальные публикации, которые можно исследовать более глубоко.

#### **Список литературы:**

1. Большой энциклопедический словарь. <http://www.classes.ru/all-russian/russian-dictionary-encycl-term-13943.htm> (дата обращения: 20.05.2018)
2. Данкова Н.С. Стратегия генерализации как средство репрезентации судебной власти (на материале российских и английских печатных СМИ) // Политическая лингвистика, N.1, 2016, – 73-81 с.
3. Лукашевич Н.В., Левчик А.В. Создание лексикона оценочных слов русского языка РуСентилекс // Труды конференции OSTIS-2016, – 377-382 с.
4. Морозова А. А. Достоверность информации в социальных сетях и критерии ее верификации (на примере «ВКонтакте») // Вестник Челябинского государственного университета, N. 6 (402), 2017, – 75-83 с. КиберЛенинка: <https://cyberleninka.ru/article/n/dostovernost-informatsii-v-sotsialnyh-setyah-i-kriterii-ee-verifikatsii-na-primere-vkontakte>(дата обращения 02.06.2018)
5. Орлова О. Г. Составление словаря стереотипов о русских: содержание и форма. // Политическая лингвистика., 2013.– N. 3 (45), – 175-182 с.
6. Смит С. Овладейте силой внушения! Добивайтесь всего, чего хотите! – Москва.: Издательство «АСТ», 2014. – 443 с.
7. Фролова И. В. 2015 О выражении субъективности в аналитических статьях качественной британской и российской прессы. //Политическая лингвистика. 2015. – N 1 (51). –, 138-145 с. <http://elar.uspu.ru/handle/uspu/1579>
8. D’Andrea A., Ferri F., Grifoni P., Guzzo T. Approaches, Tools and Applications for Sentiment Analysis Implementation. International Journal of Computer Applications (0975 – 8887) 2015. – Vol. 125, N.3.
9. Dijk T. 1984 Prejudice in Discourse. Pragmatics & Beyond, Amsterdam: Benjamins.
10. Lauesen S. User Interface Design. A Software Engineering Perspective. Addison-Wesley, 2005, – P. 604.
11. Maynard, D., and Funk, A. Automatic detection of political opinions in tweets. In: Proceedings of the 8th international conference on the semantic web, ESWC’11, . – 2011. P. 88-99.



12. Mejia A. The General in the Particular // Journ. of Philosophy of Education. 2010. Vol. 44, N. 1. – P. 93-107.
13. Mudinas A., Zhang D., Levene M. Combining lexicon and learning based approaches for concept-level sentiment analysis. In Proceedings of the First International Workshop on Issues of Sentiment Discovery and Opinion Mining (WISDOM '12). ACM, New York, NY, USA, – 2012. Article 5, P. 8. DOI= <http://dx.doi.org/10.1145/2346676.2346681>.
14. Repaka R., Pallela R., Koppula A., and Movva V. A Novel Machine Learning Approach for Aspect Based Sentiment Analysis. In Proceedings of the 9th International Workshop on Semantic Evaluation, SemEval@NAACL-HLT 2015, Denver, Colorado, USA, <http://aclweb.org/anthology/S/S15/S152126.pdf>.
15. Rice D. R., and Zorn C. 2013 Corpus-based dictionaries for sentiment analysis of specialized vocabularies //In Proceedings of NDATAD, P. 98-115.
16. Taboada M., Brooke J., Tofiloski M., Voll K., Stede M. Lexicon-Based Methods for Sentiment Analysis. Computational Linguistics, June 2011, Vol. 37, N. 2, P. 267-307.
17. Thakkar, Harsh and Patel, Dhiren. 2015. Approaches for Sentiment Analysis on Twitter: A State-of-Art study. arXiv preprint arXiv:1512.01043.
18. Turney, P.D. 2002 Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews. In: 40th Annual Meeting of the Association for Computational Linguistics (ACL 2002), Philadelphia, Pennsylvania, P. 417–424.
19. Ualiyeva I., Krassovitsky A., Mussabayev R. Generalization rate evaluation in open publication materials in Russian language. The 16th International Conference Information Technologies and Management, April 26-27, 2018, Information Systems Management Institute, Riga, Latvia .
20. Vinodhini G., and Chandrasekaran, Dr. 2012. Sentiment Analysis and Opinion Mining: A Survey. International Journal of Advanced Research in Computer Science & Technology (IJARCST), 6(2):282-292.
21. Whitehead M., and Yaeger L. 2010 Sentiment Mining Using Ensemble Classification Models. In: Sobh T. (eds) Innovations and Advances in Computer Sciences and Engineering. Springer, Dordrecht.
22. Бомпиева Ж. Казахстану нужны серьезные результаты в сфере профессионального образования. <https://tengrinews.kz/conference/240/> (дата обращения: 05.06.2018)
23. Дуванов С. Как остановить «утечку мозгов» из Казахстана. <http://www.ca-portal.ru/article:39057> (дата обращения: 01.06.2018)
24. Дуванов С. Не ходите в ЕврАзЭС. <http://www.fergananews.com/articles/8363> (дата обращения: 01.06.2018)
25. Сатпаев Д. Лингвистическая деколонизация. [https://forbes.kz/life/opinion/dosyim\\_satpaev\\_lingvisticheskaya\\_dekolonizatsiya/](https://forbes.kz/life/opinion/dosyim_satpaev_lingvisticheskaya_dekolonizatsiya/) (дата обращения: 01.06.2018)
26. Тукпиев Ж. Нечем дышать! В каких городах Казахстана самый опасный воздух? <http://www.kazpravda.kz/news/obshchestvo/nechem-dishat-v-kakih-gorodah-kazahstana-samii-opasnii-vozduh/> (дата обращения: 01.06.2018)

# ПОДХОД К СТАТИСТИЧЕСКОМУ ИССЛЕДОВАНИЮ ПРОБЛЕМ ИНФОРМАЦИОННОГО ОБЕСПЕЧЕНИЯ ИННОВАЦИОННОЙ ДЕЯТЕЛЬНОСТИ ПРЕДПРИЯТИЙ В КАЗАХСТАНЕ

<sup>1,2</sup>Утепбергенов И.Т., <sup>1,2</sup>Тойбаева Ш.Д., <sup>1,2</sup>Буранбаева А.И.,  
<sup>1</sup>Коржаспаев А.Е.

<sup>1</sup> *Институт информационных и вычислительных технологий КН МОН РК,  
Казахстан*

<sup>2</sup> *Университет Туран, Казахстан  
e-mail: i.utepbergenov@ipic.kz*

**Аннотация:** *Представлен и обоснован подход на основе сбора и обработки фактических данных для Республики Казахстан для проверки обоснованности гипотезы о том, что инновационная производительность на уровне республики и ее отдельных регионов зависит от полноты и качества информационной поддержки инноваций на протяжении всего их жизненного цикла.*

**Ключевые слова:** *управление инновациями, информационное обеспечение.*

Инновации развиваются во всех уголках мира, развитые страны находятся на стадии перехода к шестому технологическому укладу, представленному нано- и биотехнологиями, в то время как Казахстан переходит на пятый уклад, базирующийся на развитии информационных систем и телекоммуникаций.

Политика развития инновационных исследований должна основываться на достижениях исторически сложившихся научных школ, наличии научного опыта и использовании накопленной информации и знаний (Послание Президента Республики Казахстан «Стратегия «Казахстан-2050», 2012, с. 22) [3].

Одним из способов диверсификации национальной экономики и переориентации от сырьевого материала к модели промышленного обслуживания является развитие цифровых технологий. 12 декабря 2017 года Указом Правительства Республики Казахстан № 827 была утверждена Государственная программа «Цифровой Казахстан».

В соответствии с этой программой необходимо принять меры для повышения качества существующей инфраструктуры инновационного развития, а ключевым направлением развития отрасли ИКТ является обеспечение роста доли услуг в области информационных технологий. Планируется, что программные мероприятия будут осуществляться по пяти направлениям, одним из которых является «Создание инновационной экосистемы». Таким образом, инновационное и цифровое развитие рассматриваются в органическом единстве, что создает благоприятные предпосылки для реализации проектов, направленных на информационную поддержку инновационной деятельности. Одним из таких проектов является проект Института информационно-вычислительных технологий, который проводится авторами по гранту Министерства науки и образования Республики Казахстан в сотрудничестве с российскими и немецкими учеными (проект № AP05134019 «Разработка научно-методических основ и прикладных аспектов построения распределенной системы информационного обеспечения инновационной деятельности с учетом специфических особенностей каждого из этапов жизненного цикла инноваций»).

Существующие инновационные организации Казахстана для повышения конкурентоспособности придерживаются собственными разработанными внутренними этапами жизненного цикла инноваций, основываясь на базовых этапах жизненного цикла инноваций. В общедоступных публичных источниках субъектов инноваций таких, как инновационные организации, учебные заведения, исследовательские институты, описывают только характеристики стадий жизненного цикла, но не предлагают алгоритмы определения стадии жизненного цикла инновационной организации на практике. Для получения сведений от инновационных организаций для улучшения и развития инновационной деятельности научно-исследовательские институты проводят эмпирические исследования жизненных циклов организаций, и первоначальным этапом является использование достаточно громоздких анкет для определения текущей стадии жизненного цикла анализируемой компании, что не показывает общего состояния инновационной организации.

Предлагаемые подходы к реализации этого проекта предполагают обобщение и развитие опыта Казахстана, России и Европы в решении задач информационной поддержки науки и образования в отношении инноваций как современной деятельности с ее собственными специфическими характеристиками.

Проект, в частности, направлен на проверку на основе сбора и обработки фактических данных для Республики Казахстан обоснованности гипотезы о том, что инновационная производительность на уровне республики и ее отдельных регионов зависит от полноты и качества информационной поддержки инноваций на протяжении всего их жизненного цикла, с момента начала вывода инновационного продукта с рынка.

Подтверждение гипотезы позволит обосновать необходимость реинжиниринга существующих локальных систем информационной поддержки и их интеграции с созданной системой распределенной информационной поддержки инноваций.

Система предусматривает создание единой точки входа для навигации в национальном и глобальном информационном пространстве посредством предоставления метаинформации об информационных ресурсах, относящихся к задаче пользователя. Точка входа - информационный портал, где метаинформация сосредоточена как на информационных ресурсах Казахстана, так и на ресурсах других стран.

Цель данного этапа проекта - сбор данных для исследования проблем информационного обеспечения инновационной деятельности с учетом специфики отдельных этапов жизненного цикла инноваций в целях формирования конкретных проектных предложений по совершенствованию региональной системы информационного обеспечения инновационной деятельности.

Объект исследования – предприятия и организации, ведущие инновационную деятельность в различных отраслях экономики.

Предмет исследования – параметры, характеризующие информационную обеспеченность деятельности инновационных предприятия и организации.

Жизненный цикл (ЖЦ) – это совокупность связанных между собой явлений, процессов, работ, образующих заверченный круг развития в определенный период времени. Жизненный цикл инновации (ЖЦИ) представляет собой отрезок времени, в течение которого инновация обладает активной жизненной силой и приносит и производителю, и продавцу прибыль или какую-либо другую реальную выгоду. [1]

ЖЦИ принято так же называть «инновационным циклом» (ИЦ). Сущность этого понятия состоит в том, что каждая инновация реализуется по схеме, называемой

инновационным циклом, включающим различные этапы — от идеи до коммерциализации новшества.

ЖЦИ различаются по видам инноваций. Эти различия затрагивают, прежде всего, общую продолжительность цикла, продолжительность каждой стадии внутри цикла, особенности развития самого цикла, разное количество стадий. Виды и количество стадий жизненного цикла определяются особенностями той или иной инновации. Однако у каждой инновации можно определить «стержневую», то есть базовую, основу, жизненного цикла с четко выделенными стадиями. [2]

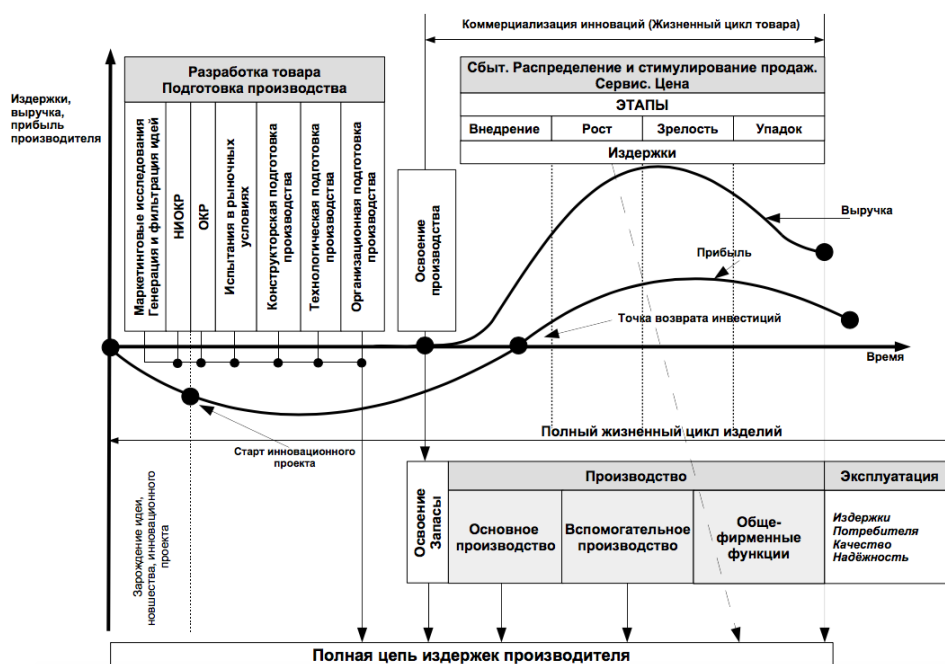


Рис. 1 Жизненный цикл инноваций

ЖЦИ состоит из семи конкретных стадий (рис.1), таких как:

1. разработка нового изделия – организация инновационного процесса, вложение капитала;
2. выход на рынок – изделие приносит прибыль в период внедрения;
3. развитие рынка – рост объема продаж изделий на рынке, анализ времени, когда новое изделие активно продается и рынок достигает насыщения этим изделием;
4. стабилизация рынка – прекратился рост объема продаж;
5. уменьшение рынка – происходит спад сбыта изделия, но спрос на данное изделие есть, значит, есть объективные предпосылки к увеличению объема продажи изделия;
6. подъем рынка – спрос существует, производитель изучает условия спроса, меняет свою кадровую и ценовую политику, применяет различные формы материального стимулирования продажи изделия, активизирует рекламу. Это позволяет увеличить объем продаж на какой-то период времени;
7. падение рынка – происходит полная реализация изделия или полное прекращение продажи изделия из-за его невостребованности у покупателя.

Современные реалии диктуют новые условия развития. Динамично развивающееся общество предопределяет ускорение всех экономических процессов.

Сегодня вопрос временных затрат в отдельных ситуациях стоит выше величины затрат материальных. Данные тенденции характерны и для развития науки, в том числе экономического анализа. Необходимость принятия оперативных управленческих решений обуславливает сроки проведения необходимых аналитических процедур, не оказывая влияния при этом на качественные результаты анализа. При осуществлении инновационной деятельности необходимость оперативного качественного анализа возрастает, так как сегодня та или иная информация может являться актуальной, а завтра – уже нет.

Инновационный процесс предприятий состоит из нескольких взаимосвязанных этапов, каждый из которых подлежит анализу. Предприятия в отдельных ситуациях даже не подозревают, что внедряя новые технологии, расширяя ассортимент, модернизируя производственный процесс, улучшая технологические процессы, они реализуют различные инновационные проекты. Эффективность их реализации оценивается лишь посредством роста показателей доходности и прибыльности, которые конечно, являются одними из важнейших показателей, но не исследуют специфику такого явления, как «инновация». С другой стороны, отдельные производители не представляют возможным внедрять инновационные технологии, в связи с отсутствием финансовых возможностей для их разработки или приобретения. Однако в большинстве своем предприятия имеют большую часть необходимых средств, но нерациональное их использования или неверное определение их количества, в виду отсутствия регулярного экономического анализа, приводит к неэффективному использованию и перераспределению средств.

Современные тенденции инновационного развития отдельных отраслей и предприятий характеризуются низким уровнем инновационной восприимчивости, несмотря на то, что экономические субъекты имеют достаточно стабильное финансовое положение, наличие свободных собственных средств, положительную динамику прибыли и уровня рентабельности.

Большинство существующих методик оценки инновационной деятельности представляют собой функционально сложный и затратный процесс. Расчет систем показателей требует большого объема запрашиваемой информации, длительного времени на ее обработку и оценку результатов анализа. Так как инновационная деятельность является специфическим направлением деятельности предприятий, то сложность и трудоемкость аналитического процесса увеличивается. Отдельный этап инновационного процесса характеризуется специфическими показателями, которые необходимо исследовать в динамике, в сравнении с общеустановленными или средними, учитывать и выявлять влияние отдельных факторов на то или иное явление или процесс.

Таким образом, все инновационные процессы взаимосвязаны и взаимообусловлены и соответственно не должны рассматриваться в отрыве друг от друга. Однако это не исключает возможности логического обособления факторов в процессе экономических расчетов. При этом степень влияния каждого фактора на изменение исследуемого объекта определяется при условии, что влияние остальных факторов принимается неизменным.

Представители инновационного бизнеса часто опираются на сведения государственной статистики. В целях обобщения и развития опыта Казахстана, России и Европы в решении задач информационной поддержки науки и образования в отношении инноваций были исследованы возможности получения статистической информации.

На государственном уровне у нас в Казахстане статистическую информацию о деятельности инновационных предприятий агрегирует комитет по статистике Министерства национальной экономики Республики Казахстан. Сбор данных об инновационной активности организаций и измерением инновационного процесса на предприятиях и организациях осуществляется с помощью статистической формы 1-инновация «Отчет об инновационной деятельности» общегосударственного статистического наблюдения РК (введена 8 ноября 2017 года, приказ №159).

На региональном и федеральном уровне статистическую информацию о деятельности инновационных предприятий агрегирует Федеральная служба государственной статистики (Росстат) и ее территориальные органы в субъектах РФ. Сбор данных об инновационной активности организаций осуществляется с помощью следующих статистических форм:

- №4-инновация «Сведения об инновационной деятельности организации»;
- №2-наука «Сведения о выполнении научных исследований и разработок»;
- №2-наука (краткая) «Сведения о выполнении научных исследований и разработок»;
- №3-информ «Сведения об использовании информационных и коммуникационных технологий и производстве вычислительной техники, программного обеспечения и оказания услуг в этих сферах»;
- №1-технология «Сведения о создании и использовании передовых производственных технологий»;
- №1-лицензия «Сведения о коммерческом обмене технологиями с зарубежными странами (партнерами)».

Непосредственным способом измерения инновационного процесса на предприятиях и организациях является форма №4-инновация.

Критерием для формирования выборки обследуемых организаций по форме федерального статистического наблюдения №4-инновация "Сведения об инновационной деятельности организаций" в соответствии с Общероссийским классификатором видов экономической деятельности (ОКВЭД) являются юридические лица и организации, относящихся к следующим видам экономической деятельности:

- Раздел А. Сельское, лесное хозяйство, охота, рыболовство и рыбоводство (за исключением охоты, рыболовства и рыбоводства);
- Раздел В. Добыча полезных ископаемых;
- Раздел С. Обрабатывающие производства;
- Раздел. Д Обеспечение электрической энергией, газом и паром; кондиционирование воздуха (за исключением торговли);
- Раздел Ф. Строительство;
- Раздел J. Деятельность в области информации и связи;
- Раздел М. Деятельность профессиональная, научная и техническая.

Сбор статистических сведений об инновационной деятельности должен соответствовать международным стандартам, описанных в Руководстве Осло (Рекомендации по сбору и анализу данных по инновациям). Сравнив основные требования к виду деятельности отчетных единиц, можно прийти к выводу, что отраслевое деление по форме №4-инновация по существу соответствует принятым Руководством Осло основным классификациям отчетных единиц и единиц наблюдения, по которым должны быть получены сведения (группы и классы ISIC, Rev. 3.1/NACE, Rev. 1.1) [4].

В то же время, существуют и различия между казахстанской, российскими формами и международными стандартами. Так, согласно Руководству, к объектам изучения инновационной статистики необходимо отнести: гостиницы, рестораны, транспорт (различных типов), почту, финансовое посредничество, операции с недвижимостью, архитектуру и прочее.

Кроме того, существует проблема неопределенности объема инновационных производств. Форму подают субъекты крупного и среднего предпринимательства, а малые предприятия и микропредприятия «остаются неохваченными». Это ведет к «перекосу» в структуре инновационных предприятий по размерам и снижению точности статистических данных. Таким образом, основная форма отчетности об инновациях далека от идеала, а следовательно, необходимо проведение дополнительных статистических исследований детальности инновационных предприятий и организаций, в том числе в вопросе удовлетворенности информационным обеспечением.

Рассмотренная сложившаяся ситуация позволяет обосновать необходимость развития и использования экспресс-методик оценки инновационной направленности предприятий, в частности к таковым относится анкетирование.

Анкета является основным инструментом опроса и представляет собой документ, содержащий структурно-организованный набор вопросов, каждый из которых связан с задачами проводимого исследования. Эта связь выражается в необходимости получения информации, отражающей характеристики изучаемого объекта. Анкетирование, как метод исследования, позволяет за короткий срок получить максимально возможный объем информации о каком-либо продукте, узнать мнения общества по каким-то вопросам и в прочих подобных случаях [5].

Метод анкетирования имеет следующие достоинства:

- оперативность получения необходимой информации;
- возможность организации массовых обследований;
- незначительная трудоемкость процедур подготовки и проведения исследований, обработки их результатов;
- отсутствие субъективного пристрастия к кому-либо из анкетиртуемых.

Наряду с достоинствами метод анкетирования обладает и недостатками, среди которых можно выделить:

- невозможность переформулирования вопросов, в связи с отсутствием личного контакта;
- не достаточная достоверность результатов, на итоги которых могут оказать влияние различных факторов.

Метод анкетного исследования активно применяется в социологических, педагогических, статистических исследованиях. В экономических исследованиях анкетирование используется недостаточно широко, в виду абстрактности своих результатов. По нашему мнению, четко выстроенные и верно сформулированные анкеты, отвечающие цели исследования и преследующие ее достижение, являются одной из составляющих информационного обеспечения инновационной деятельности предприятий. Результаты анкетирования возможностей осуществления инновационной деятельности позволят объективно оценить существующую экономическую ситуацию на предприятии и спроектировать результаты ответов на конкретную модель инновационного развития. Практическое применение анкетирования в рамках оценки инновационной направленности предприятий позволит обеспечить аппарат управления оперативной комплексной информацией о

возможных тенденциях развития инновационной деятельности на предприятии, тем самым повысив эффективность принимаемых управленческих решений относительно дальнейших мероприятий в направлении развития инновационной деятельности.

ИРН проекта №AP05134019



Республиканское государственное предприятие (РГП) на праве хозяйственного ведения (ПХВ)

"Институт информационных и вычислительных технологий"

Лаборатория инновационных и смарт технологий

Зав. лабораторией  
д.т.н., проф. Утепбергенов И.Т.

**Наименование темы проекта:** Разработка научно-методических основ и прикладных аспектов построения распределенной системы информационного обеспечения инновационной деятельности с учетом специфических особенностей каждого из этапов жизненного цикла инноваций.

### Многоаспектная анкета для инновационных и академических организаций

1. Укажите отчетный период \_\_\_\_\_
2. Укажите фактическое место осуществления инновационной деятельности организации (независимо от места регистрации организации) – область, город, район, населенный пункт \_\_\_\_\_
3. Укажите наименование организации \_\_\_\_\_
4. Укажите общую численность работников на конец отчетного периода \_\_\_\_\_ (человек)
5. Укажите область (сегмент) инновационной деятельности Вашей организации \_\_\_\_\_
6. Организационно-правовая форма Вашей организации \_\_\_\_\_
7. Укажите категорию, к которой относится Ваша организация:
  - стартап
  - малые (до 100 сотрудников)
  - средние (до 250 сотрудников)
  - крупные (более 250 сотрудников)
  - инжиниринговый центр
  - консалтинговый центр
  - ВУЗ
  - научно-исследовательский институт
  - технологическая инновационная компания (ТИС)
  - другое: \_\_\_\_\_
8. Ваша организация
  - производит инновационную продукцию
  - оказывает услуги
  - доводит продукт (услугу) до коммерческого продукта (товарного образца)
  - осуществляет ~~прототипирование~~
  - другое: \_\_\_\_\_
9. Укажите виды работ, выполняемых Вашей организацией
  - производство новой или усовершенствованной продукции
  - предоставление услуг сервисного сопровождения и обслуживания инновационной деятельности
  - выполнение НИР
  - опытно-конструкторские работы

Рис. 2 Первая страница многоаспектной анкеты



Для того чтобы не использовать сокращения вопросов многоаспектной анкеты необходимо проводить анкетирование на всех стадиях жизненного цикла инноваций. Поэтому на основе проведенного анализа статистических документов, была разработана форма многоаспектной анкеты для анкетирования инновационных предприятий. На рис. 2 представлена первая страница данной анкеты, включающей 55 вопросов, для выявления исследовательской специфики информационных потребностей новаторов на каждом этапе жизненного цикла инноваций от генерации идей до вывода продукта с рынка.

При проведении выборочного наблюдения необходимо определить минимальный объем выборочной совокупности, который будет удовлетворять требованию репрезентативности статистических данных. Изначально исследователи задают величину допустимой ошибки выборки и доверительную вероятность, а минимальный объем выборки рассчитывается исходя из требуемой точности. Объем выборки зависит от метода отбора и вида оцениваемой величины (Табл. 1).

**Таблица 1. Способы расчета минимального объема выборки для случайного отбора**

Метод отбора	Оцениваемая величина	
	для средней	для доли
Повторный	$n = \frac{t^2 \sigma^2}{\Delta^2}$	$n = \frac{t^2 w(1-w)}{\Delta^2}$
Бесповторный	$n = \frac{t^2 \sigma^2 N}{N\Delta^2 + t^2 \sigma^2}$	$n = \frac{t^2 w(1-w)N}{N\Delta^2 + t^2 w(1-w)}$

где:  $\sigma^2$  – дисперсия признака;  
 $t$  – коэффициент доверия;  
 $\Delta$  – предельная ошибка выборки;  
 $N$  – численность генеральной совокупности.

Исследование проблем информационного обеспечения инновационных предприятий Казахстана является бесповторным, поскольку каждый респондент заполняет анкету только один раз.

Важной проблемой при подготовке к исследованию остается предварительная оценка значения дисперсии ( $\sigma^2$ ). Существует несколько способов ее приблизительного оценивания:

1. Используют значение дисперсии из предыдущих исследований.
2. По правилу «трех сигм» общий размах вариации заключается в отрезке  $\bar{x} \pm 3\sigma$ , следовательно,  $\sigma$  можно получить как размах, деленный на 6.
3. Если приблизительно известно среднее значение признака, то полагают, что  $\sigma = \bar{x}/3$ .
4. При изучении альтернативного признака берется максимально возможная дисперсия 0,25.

Первые три способа определения дисперсии в предлагаемом исследовании неприменимы, поскольку в программе наблюдения отсутствуют вопросы, варианты ответа на которые измеряются в интервальной и относительной шкалах. Следовательно, учитывая то, что все вопросы анкеты дают ответы, выраженные в номинальной шкале, возможно оценивать лишь доли единиц совокупности,

обладающие определенным значением признака, а в качестве метода расчета средней применим четвертый метод. Таким образом, дисперсия имеет следующее значение:

$$\sigma^2 = w(1 - w) = 0,25$$

Для расчёта минимального объема выборки необходимо воспользоваться формулой:

$$n = \frac{t^2 w(1 - w)N}{N\Delta^2 + t^2 w(1 - w)}$$

В отличие от повторного отбора, в бесповторном необходимо знать приблизительную численность генеральной совокупности. К сожалению, органы статистики не публикуют информацию о количестве инновационных предприятий, поэтому данную величину необходимо оценить.

Таким образом, предложенный подход позволит подтвердить обоснованность гипотезы о необходимости реинжиниринга существующих локальных систем информационной поддержки и их интеграции с создаваемой системой распределенной информационной поддержки инноваций в Казахстане.

#### **Литература**

1. Инновационный менеджмент / И.Ю. Евграфова, Е.О. Красильникова. – М.: Окей- книга. - 2009. - 84 с.
2. Агарков С. А., Кузнецова Е. С., Грязнова М. О., Инновационный менеджмент и государственная инновационная политика, Академия Естествознания, 2011
3. Послание Президента Республики Казахстан «Стратегия «Казахстан-2050»: новый политический курс государства». 14 декабря 2012. Проверено 25.02.2018 из <https://strategy2050.kz/ru/multilanguage/>.
4. Анцибор А.В. Анкетирование, как метод экспресс-анализа инновационной деятельности, Евразийский союз ученых\_28.11.15\_11(20) – [Электронный ресурс].
5. [www.instat.gov.al/media/2956/nace\\_rev11.pdf](http://www.instat.gov.al/media/2956/nace_rev11.pdf).

## **ФОРМАЛЬНАЯ МОДЕЛЬ ОЦЕНИВАНИЯ КАЧЕСТВА ЭКСТРАКЦИИ И ИДЕНТИФИКАЦИИ ЗНАНИЙ ИЗ СЛАБОСТРУКТУРИРОВАННОЙ ТЕКСТОВОЙ ИНФОРМАЦИИ**

**Хайрова Н.Ф., Мамырбаев О.Ж., Избасаров Е.Ж., Мухсина К.Ж.**

*Институт информационных и вычислительных технологий КН МОН РК  
erlan\_1081@mail.ru*

*Аннотация. В данной статье рассматривается модель для определения каждого семантического поля по результатам выделения множества смысловых лингвистических единиц (полнотекстовых документов, сверхфразовых единств, абзацев и др.), соответствующих данной смысловой парадигме.*

**Ключевые слова.** Модель оценивания качества, слабоструктурированная текстовая информация.

**Введение.** Для оценки эффективности созданной информационно-лингвистической технологии экстракции и идентификации знаний из слабоструктурированной текстовой информации необходимо выделить метрики – совокупность объективно измеряемых показателей, характеризующих деятельность пользователей до и после внедрения данной технологии. К таким метрикам обычно относят время поиска пользователем информации по тому или иному вопросу и уровень знаний, извлеченных пользователями данной системы. При этом, в отличие от временных показателей, характеризующих длительность выполнения тех или иных процессов и достаточно просто поддающихся объективному измерению, метрика уровня знаний достаточно сложно поддается измерению. Тем не менее, основную ценность для социально-экономических организационных систем обычно представляют новые, скрытые, неявные и неформализованные знания, извлеченные из информационных потоков, позволяющие принимать новые нетрадиционные решения.

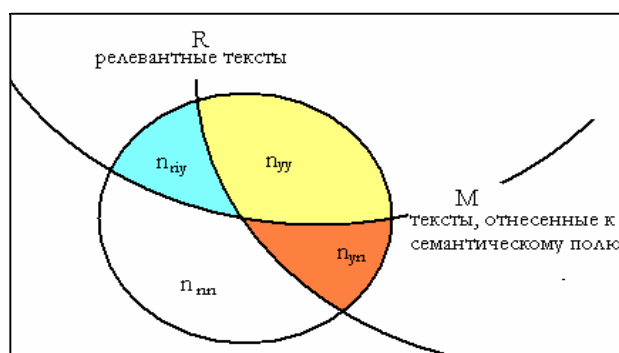
**Постановка и алгоритм решения задачи.** На сегодняшний день не существует стандартных бенчмарков для измерения качества извлечения из текстов знаний [7]. Обычно, для решения задачи измерения качества семантической классификации и информационного поиска используют *метод тестовых коллекций* [4,6], заключающийся в сравнении результатов работы исследуемой схемы на заранее определенных данных с оценками экспертов на тех же данных. В результате сравнения получается одно-двухкритериальная оценка эффективности. Поскольку понятия «эффективного извлечения знаний», «качества знаний» не определены, количественная оценка результатов работы системы нетривиальна. Традиционный подход в подобных случаях – сравнение с «эталонным» результатом, – плохо применим из-за необходимости создания эталонного ответа для каждого конкретного набора электронных документов.

Для оценки работы системы используем алгоритм, для которого выводы, сделанные системой, согласуются с мнением экспертов. Для получения интегральных показателей качества работы системы идентификации знаний в слабоструктурированных текстовых информационных потоках применима методика усредненных метрик [2].

Обычно в поисковых системах точность вычисляется как отношение правильных выданных источников к общему числу выданных источников, а полнота – как отношение числа правильных выданных источников к общему числу правильных источников, существующих в системе [3,5]. Будем использовать показатели количественной оценки эффективности поиска и классификации, утвержденные межгосударственным стандартом по информации, библиотечному и издательскому делу [1]. Такими показателями являются: коэффициент точности – *precision*, коэффициент полноты – *recall* и коэффициент аккуратности *accuracy*.

Для определения перечисленных коэффициентов необходимо для каждого семантического поля по результатам выделения множества смысловых лингвистических единиц (полнотекстовых документов, сверхфразовых единств, абзацев и др.), соответствующих данной смысловой парадигме, определить:  $n_{yy}$  – число выделенных элементов, релевантных семантическому полю с точки зрения эксперта,  $n_{yn}$  – число выделенных элементов, не релевантных, с точки зрения эксперта,  $n_{ny}$  –

количество релевантных элементов, не выделенных системой, и  $n_{nn}$  – количество нерелевантных элементов, не выделенных системой (рис.1.).



**Рис. 1. Метрики оценки эффективности**

При этом, если нет элементов, получивших с точки зрения эксперта определение *undefined*, сумма значений метрик равна количеству смысловых лингвистических элементов, поступивших на обработку:

$$D = n_{nn} + n_{ny} + n_{yn} + n_{yy},$$

где  $D$  – множество лингвистических элементов, поступивших в систему на обработку.

Отношение релевантности является субъективным, сложно определяемым и имеющим скорее психологическую природу понятием. Мы будем использовать определение релевантности [8], в котором релевантность зависит от четырех понятий *Relevance(IR, IN, C, T)*, где  $IR$  – информационный ресурс,  $IN$  – информационная потребность,  $C$  – контекст и  $T$  – время. В нашей модели информационный ресурс представлен множеством лингвистических смысловых элементов уровня связного текста, поступившим на обработку  $IR = D$ . Информационную потребность можно разделить на неосознанную (истинную потребность) эксперта в знаниях, оперируя которыми эксперт решает некоторую информационную проблему, стоящую перед ним, и осознанную (внутреннее понимание реальной потребности). Переход между неосознанной и осознанной потребностью в знаниях вносит погрешность в вычисление эффективности работы подсистемы, основанной на интеллектуальных моделях.

Релевантность, определяемая неосознанной потребностью пользователя, лежит в основе определения коэффициента пертинентности информации [1]. Осознанная потребность интеллектуального поиска знаний определяет его полноту и точность. Осознанная потребность эксперта в знаниях, необходимых для решения некоторых задач, формируется в сфере мышления, но, сформировавшись в реальном контексте предметной области  $C$  и времени  $T$ , информационная потребность  $IN$  описывается средствами естественного языка.

При определении эффективности работы системы релевантность, т.е. соответствие связного текста крупной смысловой парадигме определяется экспертом по шкале “*Relevance / irrelevant / undefined*” и показывает соответствие электронного текста некой локальной области знаний (или крупной смысловой парадигме).

Исходя из предположения, что система принимает решение о принадлежности к данной локальной области знаний каждого связного текста, вычисляем коэффициент

аккуратности как отношение правильно принятых решений системы к их общему числу:

$$accuracy = \frac{n_{yy} + n_{nn}}{n_{yy} + n_{ny} + n_{yn} + n_{nn}}$$

Ошибку вычисляем как отношение неправильно принятых системой решений к их общему числу:

$$error = \frac{n_{yy} + n_{nn}}{n_{yy} + n_{ny} + n_{yn} + n_{nn}}$$

В нашей модели полезно объединить точность и полноту в одной усредненной величине. Но при этом невозможно использовать среднее арифметическое, так как при стремящейся к нулю точности среднее арифметическое полноты и точности будет меньше 50 %. Будем опираться на среднее гармоническое точности и полноты, называемое мерой Ван Ризбергена, или *F-measure*:

$$F_1 - measure = \frac{2}{\frac{1}{precision} + \frac{1}{recall}} = \frac{2 \times precision \times recall}{precision + recall}$$

При этом используемая в предложенной модели извлечения новых знаний из поступающей в систему слабоструктурированной текстовой информации метрика должна учитывать возможность использовать различные веса  $\alpha$  для учета полноты и точности:

$$F_\beta - measure = \frac{1}{\alpha \frac{1}{precision} + \frac{(1-\alpha)}{recall}} = \frac{(\beta^2 + 1) \times precision \times recall}{\beta^2 \times precision + recall}$$

где  $\alpha \in [0,1], \beta^2 = \frac{(1-\alpha)}{\alpha}, \beta \in [0, \infty]$ . При значении коэффициентов  $\alpha = 1/2$  или  $\beta = 1$  мера придает одинаковый вес полноте и точности и получается сбалансированная  $F_1$ -мера.

Если  $0 < \beta < 1$  - большее значение при расчете уделяется точности, а при  $\beta > 1$  - больший вес приобретает полнота. Так как в принятой системе расчета эффективности релевантность является субъективной, то эта субъективность будет перенесена на точность. В отличие от поисковой машины Интернет-поиска, в информационной системе, извлекающей знания для их последующего использования, точность приобретает первостепенное по отношению к полноте значение.

**Заключение.** Исходя из вышесказанного в данной модели можно объединить точность и полноту в одной усредненной величине. Но при этом невозможно использовать среднее арифметическое, так как при стремящейся к нулю точности среднее арифметическое полноты и точности будет не меньше 50 %. При этом используемая в предложенной модели извлечения новых знаний из поступающей в систему слабоструктурированной текстовой информации метрика должна учитывать возможность использовать различные веса для учета полноты и точности.

*Работа выполнена в рамках проекта ИРН АР 05131073 Методы и модели поиска и анализа криминально значимой информации в неструктурированных и слабоструктурированных текстовых массивах.*

### **Литература**

1. ГОСТ 7.73 – 96 SU. Поиск и распространение информации. Термины и определения. – Введ. 10.10.1996; принят Межгосуд. советом СНГ по стандартизации, метрологии и сертификации. – М.:Госстандарт.
2. Кураленок И. Оценка систем текстового поиска /И. Кураленок , И. Некрестьянов // Программирование. – 2002. – N 28(4). – С.226 –242.
3. Сегалович И. В. Методы сравнительного анализа современных поисковых систем и определения объема Рунета / И. В. Сегалович, Ю. Г. Зеленков, Д. О. Нагорнов // Электронные библиотеки: перспективные методы и технологии, электронные коллекции – RCDL'2006 : труды 8-й Всерос. науч. конф.; Суздаль, Россия, 2006. [Электронный ресурс] – Режим доступа: [www.rcdl2006.unijar.ac.ru](http://www.rcdl2006.unijar.ac.ru).
4. Шабанов В. И. Метод классификации текстовых документов, основанный на полнотекстовом поиске /В. И. Шабанов , А. М. Андреев // Труды РОМИП'2003. – СПб. : НИИ Химии СПб гос. ун-та, 2003. –С.52–71.
5. Bar-Yossef Z. and Gurevich. M. Efficient search engine measurements. In C. L. Williamson, M. E. Zurko, P. F. Patel-Schneider, and P. J. Shenoy, editors. ACM. – Ban, Canada, 2007. – P.401–410.
6. Cormack G.V. A Efficient construction of large test collections / G. V.Cormack , C.R. Palmer ,C.L. Clarke // Proc. of the SIGIR'98– P.282–289.
7. Manning C., Schütze H. Foundations of Statistical Natural Language Processing. MIT Press,2000.
8. Mizzaro S. How many relevances in information retrieval? Proceeding of the Workshop «Information Retrieval and Human Computer Interaction», GIST Technical Report GR96-2 / Glasgow University. – Glasgow: The British Computer Society – P.57–60.

## **ПРИМЕНЕНИЕ ТЕХНОЛОГИИ ОТСЛЕЖИВАНИЯ ГЛАЗ В ИССЛЕДОВАТЕЛЬКИХ РАБОТАХ**

**Шокишалов Ж.М.**

*Институт информационных и вычислительных технологий КН МОН РК,  
Казахстан  
e-mail: [jas\\_moderator@mail.ru](mailto:jas_moderator@mail.ru)*

***Аннотация.** В последние несколько лет все большее внимание уделяется глазной биометрии. С одной стороны, это может быть связано с увеличением доступности глазных трекеров, т. е. устройств, способных измерять характеристики при движении глаз, такие как направление взгляда и размер зрачка. С другой стороны, функции и поведение глаз все больше и больше рассматриваются как потенциально более безопасные методы аутентификации, особенно при использовании в сочетании с традиционными методами проверки подлинности. В этой статье подробно*

*описывается технология отслеживания глаз и рассматривается ее применение в информационной безопасности.*

Трекер отслеживания движение глаз (eye tracker) - это устройство, которое позволяет обнаружить взгляд пользователя и, таким образом, найти то, что отображается, обычно на экране [1]. Движения глаз происходят как последовательности чрезвычайно быстрых саккад (длительностью менее 100 мс) и фиксации, которые обычно длятся от 100 до 600 мс. Наблюдатель должен иметь возможность записывать координаты взгляда пользователя определенное количество раз в секунду (частота дискретизации) и косвенно обнаруживать посредством правильных алгоритмов, фиксаций и саккад. Инфракрасный или ближний инфракрасный свет обычно используется для освещения глаз пользователя, а роговые и зрачковые отражения позволяют точно измерять глаза, поскольку они могут устранить смещения головы и движения глаз.

Чтобы точно зафиксировать взгляд пользователя, необходимо настроить калибровку глазного трекера. Типичный процесс калибровки требует, чтобы пользователь смотрел на круг, который последовательно отображается в предопределенных положениях на экране. Таким образом, можно найти правильное отображение между отражениями роговицы инфракрасного света на глазах пользователя и целевыми точками на экране.

Важным вопросом, касающийся трекеров отслеживания движение глаз нескольких лет назад была их очень высокая цена, редко ниже 15 000 долларов США (и часто выше 20 000 долларов США). Однако сейчас есть гораздо более дешевые устройства, стоимость которых колеблется примерно от 100 до 2000 долларов США. Даже если их точность иногда немного ниже, чем у дорогостоящих инструментов, их низкая цена делает их доступными для многих других применений.

Технологии отслеживания глаз связаны с измерением положения глаза относительно головы или точки зрения [2]. Глазные трекеры, такие как Tobii-TX / T120 / T60 или Eye- Tribe, относятся ко второй категории, они захватывают движения глаз, фокусируя инфракрасные огни в глазах пользователя и выполняя алгоритм, чтобы обнаружить точную точку, куда направлен взгляд [2].

Основными компонентами системы слежения за глазами являются [3]: датчики для захвата движения глаз, алгоритмы интерпретации потоков изображений, создаваемых датчиками, и приложений, которые предоставляют пользователю простой способ доступа и интерпретации данных. Измерения глаз, которые можно использовать в технологиях отслеживания глаз, - это фиксации, саккады и размер зрачка.

Фиксации - это периоды стабильности, при которых взгляд фиксируется на одной точке. Визуальная информация обрабатывается во время фиксаций.

Саккады - это быстрые движения глаз, которые фокусируют новую часть сцены. Во время саккад можно сделать небольшую визуальную обработку. Различные показатели используются для получения информации из необработанных данных, предоставляемых наблюдателем:

- общее количество фиксаций, которые, как полагают, отрицательно связаны с эффективностью поиска [4],
- количество фиксаций в каждой интересующей области, которое должно отражать важность указанной области. Несколько исследователей разработали инновационные методы анализа и представления этой метрики отслеживания глаз.

Wooding [5] представила «карту фиксации» для выделения наиболее часто фиксированных областей изображения,

- доля времени, затраченного на каждую интересующую область, которая может выявлять шаблоны, описывающие, как внимание пользователя привлекается к интересующей области,

- ответ ученика, который может помочь в определении когнитивной нагрузки пользователя [6]. Индекс когнитивной активности, определяемый как среднее число внезапных разрывов в сигнале в секунду за выбранный период времени, может оценить когнитивную нагрузку [7]. Его также можно использовать в эмоциональном моделировании.

- продолжительность фиксации и длина саккады, для получения информации о психическом состоянии пользователя [8],

- путь сканирования (последовательность фиксации), который может указывать на эффективность элементов в пользовательском интерфейсе [9].

Отслеживание глаз широко используется в области информационной безопасности. В статье «Setting safety precautions: a long-term research of habitation through the problems of fmri, eye and field experiments» описано исследование в области информационных систем и взаимодействия человека и компьютера. Исследование показало, что зависимость - сокращение реакции на повторное стимулирование - является серьезной проблемой и ведет к предупреждению о безопасности. С помощью нескольких экспериментов авторы решили две проблемы: привыкание в нейробиологическом аспекте, и когда привыкание отрицательно сказывается на фактическом соответствии безопасности.

«Мы добавили к прошлым исследованиям, изучая привыкание в продольном направлении, как с помощью fMRI, так и с отслеживанием глаз, а также с помощью полевого эксперимента с соблюдением безопасности. Наши результаты показывают, что привыкание происходит с течением времени на уровне нейробиологического уровня. Мы также продемонстрировали, что подверженность повторным предупреждениям приводит к уменьшению приверженности безопасности. Наконец, мы показали, что полиморфные предупреждения эффективны в снижении привыкания с течением времени, что проявляется как в качестве внимания на нейробиологическом уровне, так и в фактической приверженности безопасности» [10].

Исследования также проводились по привычке пользователя в повседневной жизни. Обычно пользователи привыкли закрывать окна предупреждения, не читая их, при установке сторонних программ, тем самым нарушая простые правила безопасности и конфиденциальности персонального компьютера. Авторы статьи «Look Before You Authorize: Using Eye-Tracking To Enforce User Attention Towards Application Permissions» пытались заставить пользователей активировать кнопку согласия только после прочтения диалогового окна, прежде чем нажимать на него [11].

В работе «Preventing Lunchtime Attacks: Fighting Insider Threats With Eye Movement Biometrics» рассматривается возможность поведения глаз пользователей как новой биометрической основы.

«В контролируемом эксперименте мы проверяем, как время и близость задачи влияют на движение глаз и стабильность элементов, а также то, как различные подмножества функций влияют на производительность классификатора. Эти подмножества функций могут использоваться для адаптации биометрического движения глаз к различным методам аутентификации и моделях угроз. Наши результаты показывают, что биометрия движения глаз обеспечивает надежную и стабильную идентификацию и аутентификацию пользователей. Мы исследуем



различные подходы, в которых злоумышленник может попытаться использовать внутренние знания, чтобы имитировать законного пользователя. Наши результаты показывают, что, хотя это предварительное знание измеримо, оно не увеличивает вероятность успешного олицетворения. Чтобы определить временную стабильность наших функций, мы повторяем эксперимент дважды в течение двух недель. Результаты показывают, что мы можем достоверно аутентифицировать пользователей в течение всего периода. Мы показываем, что решение классификации зависит от всех функций и подражания, некоторые из них недостаточно, чтобы обмануть классификатор. Мы подробно обсудим преимущества и ограничения нашего подхода и дадим практическую информацию об использовании этой биометрии в реальной среде» [12].

Отслеживание глаз пользователей в области информационной безопасности предоставляют дополнительную полезную информацию для анализа и выявления уязвимостей в системах управления электронными документами.

Для проведения наших экспериментальных работ был выбран трекер «Eye Tribe». Несмотря на скромную цену, в определенных ситуациях он имеет очень удовлетворительную точность и может давать хорошие выходные данные.

В планируемой автоматизированной среде для работы с электронными документами введение слежения за глазами будет эффективным для следующих целей:

- Анализ пользователя при работе с компьютером для создания личного профиля пользователя в качестве конкретного метода идентификации. Если поведение пользователя изменится, система снова запросит дополнительный метод идентификации (электронный ключ, пароль).

- Дополнительные метаданные при создании, изменении, отправке и получении электронного документа. Метаданные помогут влиять на конфликтную ситуацию в системе.

- Анализ электронного документа для повышения безопасности и целостности. Анализ повлияет на улучшение структуры электронного документа с помощью тепловых карт на документе. Поможет понять, какие части часто интересны пользователям.

Таким образом, подводя итоги, можно сказать, что применение технологии отслеживания глаз может повлиять на электронный рабочий процесс. Анализируя поведение пользователей, вы можете понять, какие аспекты представляют интерес для пользователей, какие части документа более восприимчивы к интересам злоумышленников.

### **Литература**

1. A.T. Duchowski, Eye Tracking Methodology – Theory and Practice, second ed., Springer-Verlag, London, 2007.
2. Andrew Duchowski. 2007. Eye tracking methodology: Theory and practice. Vol. 373. Springer Science & Business Media, Berlin, Germany.
3. Tobii. accessed June, 2017. Tobii eye tracker. <http://www.tobii.com/group/about/this-is-eye-tracking/>. (accessed June, 2017).
4. Alex Poole and Linden J Ball. 2006. Eye tracking in HCI and usability research. Encyclopedia of human computer interaction 1 (2006), 211–219.
5. David S Wooding. 2002. Fixation maps: quantifying eye-movement traces. In Proceedings of the 2002 symposium on Eye tracking research & applications. ACM, New York, NY, 31–36.

6. Eckhard H Hess. 1965. Attitude and pupil size. Scientific american (1965).
7. Sandra P Marshall. 2002. The index of cognitive activity: Measuring cognitive workload. In Human factors and power plants, 2002. proceedings of the 2002 IEEE 7th conference on. IEEE, Los Alamitos, CA, 7–7.
8. Han-Chin Liu and Hsueh-Hua Chuang. 2011. An examination of cognitive processing of multimedia information based on viewers' eye movements. Interactive Learning Environments 19, 5 (2011), 503–517.
9. Anthony Hughes, Todd Wilkens, Barbara M Wildemuth, and Gary Marchionini. 2003. Text or pictures? An eyetracking study of how people view digital video surrogates. In International Conference on Image and Video Retrieval. Springer, Berlin, Germany, 271–280
10. Duchowski A T. Eye tracking techniques[M]//Eye Tracking Methodology. Springer, Cham, 2017: 49-57.
11. Hess C W, Muri R, Meienberg O. Recording of horizontal saccadic eye movements: methodological comparison between electro-oculography and infrared reflection oculography[J]. Neuro-ophthalmology, 1986, 6(3): 189-197.
12. Young L R, Sheena D. Survey of eye movement recording methods [J]. Behavior research methods & instrumentation, 1975, 7(5): 397-429.

## МОДЕЛИРОВАНИЕ ТРАНСПОРТНОЙ ИНФРАСТРУКТУРЫ ГОРОДСКОГО ОБЩЕСТВЕННОГО ТРАНСПОРТА

<sup>1</sup>Яворский В.В., <sup>1</sup>Ахмедиярова А.Т., <sup>2</sup>Байдикова Н.В., <sup>2</sup>Чванова А.О.

<sup>1</sup> Институт информационных и вычислительных технологий КН МОН РК

<sup>2</sup> Карагандинский государственный индустриальный университет

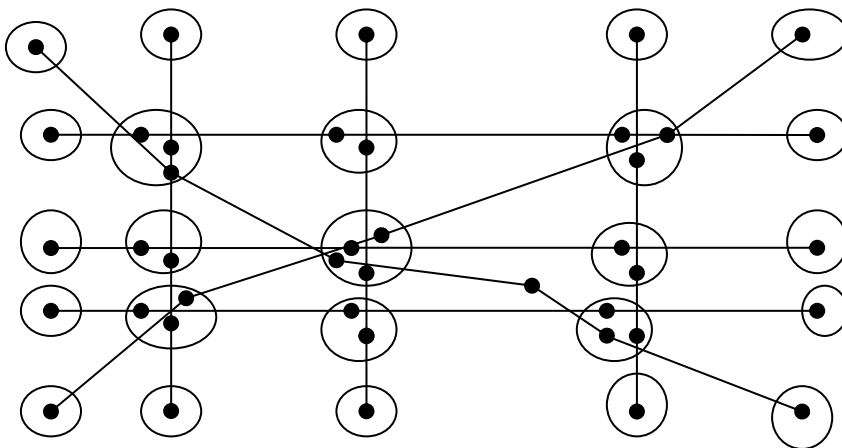
e-mail: yavorskiy-v-v@mail.ru

***Аннотация.** Рассмотрены возможности моделирования передвижения пассажиров городского пассажирского транспорта на основе гиперсетей. Показано, что равновесное состояние соответствует минимальным затратам времени на передвижение пассажиров и определяет распределение пассажиров по путям следования в соответствии со стратегией выбора пути*

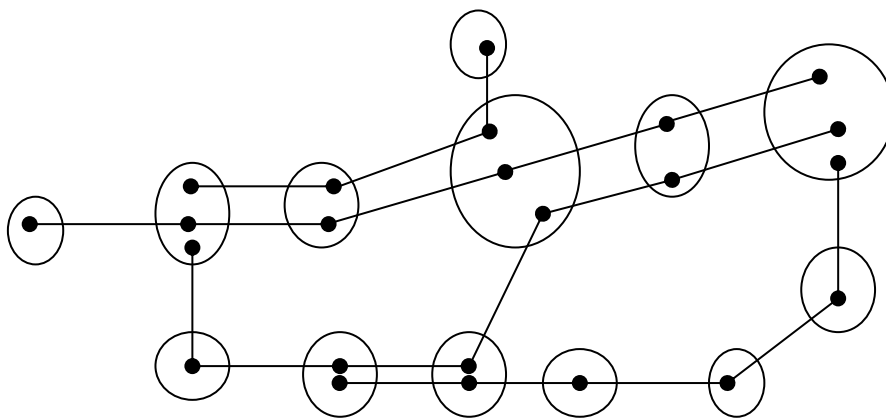
Возможность решения задачи моделирования процессов транспортировки пассажиров на гиперсети городского общественного транспорта (ГОТ) зависит от вида изучаемой гиперсети. При этом под гиперсетью ГОТ будем понимать совокупность маршрутов в  $k = \overline{1, n}$  с заданным количеством транспорта на каждом маршруте  $\vec{r} = \{r_k\}$ . Будем также считать, что перевозка пассажиров на гиперсети осуществляется между транспортными районами (ТР) города, в которых имеется один обобщенный пункт обслуживания (остановка) маршрута, в случае, если маршрут проходит через ТР. Информация о потенциальных пассажиропотоках между ТР города задана в виде матрицы корреспонденций пассажиропотоков  $\|\lambda_{ij}\|$ .

Предлагается классифицировать гиперграфы [1] сетей ГОТ с позиций последующего анализа на три основных вида. Если маршруты, составляющие

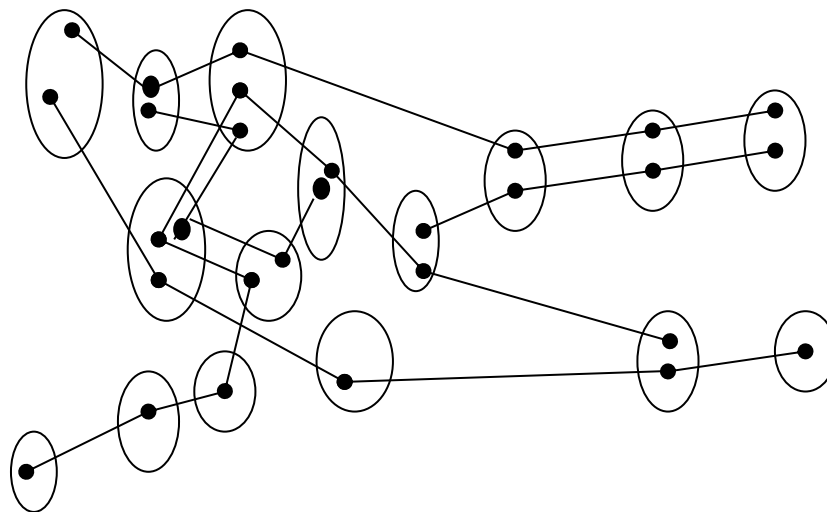
гиперсеть ГОТ, не имеют общих пассажиропотоков, будем называть их гипергиперсетью независимых маршрутов (рис. 1). Если маршрут имеет общие пассажиропотоки, но на каждой остановке их можно разделить пропорционально потоку поступающих на остановку потоков обслуживания, который определяется после рассмотрения процесса обслуживания на предыдущих остановках маршрута, то ГОТ будем называть гипергиперсетью с разделяющимися пассажиропотоками (рис. 2). Все остальные гиперсети ГОТ будем называть сложными (рис. 3).



**Рис. 1** Гиперсеть независимых маршрутов



**Рис. 2** Гиперсеть маршрутов ГОТ с разделяющимися зависимыми пассажиропотоками



**Рис. 3 Сложный тип сетей ГОТ**

Для моделирования процессов перевозки пассажиров на сетях ГОТ первых двух видов могут быть непосредственно использованы модели процессов на отдельных маршрутах. При этом для моделирования процессов на гиперсети независимых маршрутов необходимо отдельно исследовать эти процессы для каждого маршрута. Для получения параметров пассажиропотоков и характеристик транспортного обслуживания гиперсети в целом необходимо сложить полученные для отдельных маршрутов параметры и характеристики.

При моделировании процессов пассажирских перевозок на гиперсети маршрутов ГОТ с разделяющимися пассажиропотоками можно исходить из следующего положения. Если пассажиры одной остановки обслуживаются несколькими маршрутами, то общий пассажиропоток при прочих равных условиях распределяется между маршрутами пропорционально потоку поступающих свободных мест на маршрутах. Это предположение делает необходимым совместное исследование тех групп маршрутов, которые обслуживают одинаковые корреспонденции. Само аналитическое описание группы связанных маршрутов проводится аналогично описанию одного маршрута. При этом расчет параметров и характеристик необходимо начинать с конечных остановок маршрута, на которых поток поступающих на остановку свободных мест для каждого из маршрутов определен. Находя часть общего пассажиропотока, обслуживаемого маршрутом, и добавляя к нему индивидуальный поток пассажиров, получаем потенциальный поток пассажиров на остановку маршрута, что позволяет с использованием описанной модели маршрута рассчитать интенсивность потока свободных мест, поступающих на следующую остановку данного маршрута, и т.д. Такое моделирование процессов может быть осуществлено для сетей, у маршрутов которых нет общих пассажиропотоков во встречных направлениях (кольцевые остановки).

Описанные методы анализа процессов перевозки пассажиров могут быть использованы для сравнительно несложных гиперсетей, анализ же сложных гиперсетей ГОТ с учетом многих факторов, влияющих на перевозки пассажиров, возможен лишь с использованием моделей массового поведения пассажиров при перевозках. При этом формируется потокораспределение на всей гиперсети.

Наиболее распространенной методикой определения потокораспределения являются энтропийные модели [2, 3], моделирующие массовое поведение с учетом

некоторых априорных предпочтений. Имеются также методы, в которых более детально исследуется самоорганизация и психофизические закономерности поведения пассажиров [4-7].

Предлагается модель, которая, с одной стороны, является представителем второго направления, так как в ней используются психофизические закономерности выбора пути следования, а также закономерности процессов обслуживания, а с другой стороны, как модель первого направления, где ищется равновесное распределение пассажиропотоков.

Стратегия выбора пассажирами  $\xi$ -го пути следования из  $i$ -го района в  $j$ -й из всей совокупности возможных путей может быть задана вероятностью

$$P_{\xi}^{ij} = \frac{1}{N} \left\{ 1 + a_t \ln \left[ \frac{\prod_{\xi=1}^N T_{\xi}^{ij}}{(T_{\xi}^{ij})^N} \right] + a_q \ln \left[ \frac{\prod_{\xi=1}^N q_{\xi}^{ij}}{(q_{\xi}^{ij})^N} \right] \right\}, \quad (1)$$

где  $N \geq 2$  – число возможных путей;  $T_{\xi}^{ij}$ ,  $q_{\xi}^{ij}$  – соответственно время передвижения и число посадок на  $\xi$ -ом пути следствия;  $a_t, a_q$  – коэффициенты относительного влияния времени проезда и числа пересадок на выбор пути.

Имеет место следующее условие:  $0 \leq P_{\xi}^{ij} \leq 1$ . В качестве множества  $H_{ij}$  обычно достаточно рассматривать пути, отличающиеся от кратчайшего между ТР  $i$  и  $j$  не более чем на 15 - 20% и не имеющие более двух-трех пересадок. По результатам обследований [49] получена следующая пропорция для коэффициентов:  $a_t : a_q = 0,632 : 0,328$ . Значения коэффициентов можно найти из условия, что для пути  $\xi$ , отличающегося от кратчайшего на 15% и имеющего три пересадки (четыре посадки в ТЕ,  $q_{\xi} = 4$ ),  $P_{\xi}^{ij} = 0$ :

$$a_t = 1,28 \left( \frac{1}{n-1} \right); a_q = 0,592 \left( \frac{1}{n-1} \right). \quad (2)$$

Анализ процессов перевозки пассажиров удобно проводить, задавая гиперсеть маршрутов графом, структура которого позволяет рассчитывать пассажиропотоки с использованием стратегии выбора пассажирами путей передвижения. Предлагается представлять гиперсеть маршрутов ГОТ графом  $L(Z, W)$  (рис. 4).

Множество вершин графа состоит из двух типов вершин  $i = \overline{1, n}$ , соответствующих рассматриваемым транспортным районам, и  $i_k$ , соответствующих остановкам различных маршрутов  $k = \overline{1, m}$ , проходящих через район  $i$ . Множество дуг  $W$  графа состоит из четырех типов дуг: дуг типа  $(i, i_k)$ , длины которых  $t^{no}(j, j_k)$  равны времени подхода и ожидания пассажиров, желающих осуществить посадку в ТЕ  $k$ -го маршрута; дуг типа  $(i_k, j_k)$ , длины которых  $t^{mp}(i_k, j_k)$  задают время передвижения по соответствующему перегону  $k$ -го маршрута; дуг типа  $(i, j)$ , длины которых  $t^{nu}(i, j)$  задают время пешего передвижения между ТР  $i$  и  $j$ ; дуг пересадок  $(i_k, i)$ , длины

которых  $t(i_k, i)$  равны дополнительным затратам времени, которые возникают у пассажиров при пересадках.

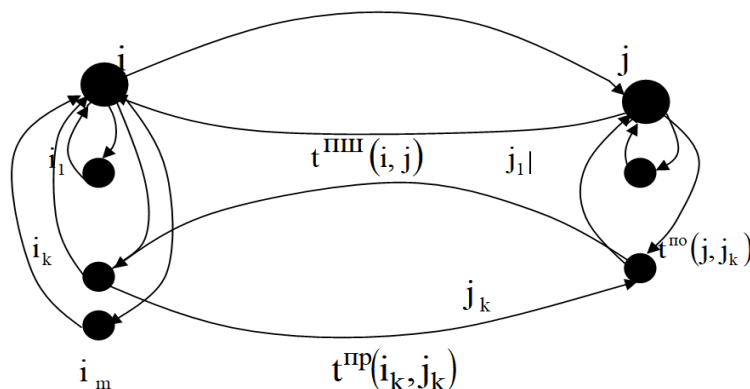


Рис. 4 Граф маршрутной гиперсети

Длины дуг типа  $(i, j)$  и  $(i_k, i)$ , соответствующие пешим переходам и пересадкам при моделировании процессов пассажиров, будем считать постоянными. Например, длительность пешего перехода можно оценить по формуле

$$t^{nu}(i, j) = \frac{l(i, j)}{V_n}, \quad (3)$$

где  $l(i, j)$  и  $V_n$  - соответственно среднее расстояние между ТР  $i$  и  $j$  и скорость пешего передвижения.

Можно также считать постоянной и длины дуг типа  $(j, j_k)$ . Время подхода к остановке и ожидания посадки  $t^{no}(i, i_k)$  можно оценить по формуле

$$t^{no}(i, i_k) = \frac{1}{V_n} \left( \frac{1}{4\sigma_{ik}} + \frac{a}{4} \right) + \frac{\tau_k}{2}, \quad (4)$$

где  $a$  и  $\tau_k$  - соответственно среднее расстояние между остановками на маршрутах ГОТ можно оценить и интервал между транспортными единицами (ТЕ) на маршруте  $k$ ;  $\sigma_{ik}$  - плотность маршрута  $k$  в ТР  $i$ .

Будем называть идеальным потокораспределением такое распределение пассажиров по дугам графа  $L(Z, W)$ , которое рассчитано в соответствии с (1) в предположении, что обслуживание пассажиров происходит без отказов. При этом время передвижения

$$t^{np}(i_k, i_j) = \frac{l(i, j)}{V_T}, \quad (5)$$

где  $V_T$  - средняя скорость сообщения транспорта.

Идеальное потокораспределение полностью соответствует интересам пассажиров, осуществляющих поездки, однако оно не учитывает процессов, которые возникают при их обслуживании. Для получения реального потокораспределения предлагается использовать имитационную модель, являющуюся упрощением описанной модели процессов перевозки на маршруте.

Пусть  $\Pi(i_k, j_k)$  – интенсивность пассажиропотоков на перегонах маршрута  $k$  в идеальном потокораспределении, а  $q_k$  – интенсивность потока свободных мест на маршруте:

$$q_k = \frac{r_k V_T \delta_k}{2l_k}, \quad (6)$$

где  $l_k$  – длина маршрута  $k$ .

Будем определять  $\Pi^\delta(i_k, j_k)$  – интенсивность потока обслуженных пассажиров  $k$ -го маршрута на перегоне  $(i_k, j_k)$  и  $\Pi^{\text{отказ}}(i_k, j_k)$  – интенсивность отказов, используя следующую модель:

1. Если  $\tau_k \leq t_{\text{ож}}^M$  и  $q_k \geq \Pi(i_k, j_k)$ , то  
 $\Pi^\delta(i_k, j_k) = \Pi(i_k, j_k)$ ,  $\Pi^{\text{отказ}}(i_k, j_k) = 0$ . (7)

2. Если  $\tau_k \leq t_{\text{ож}}^M$  и  $q_k < \Pi(i_k, j_k)$ , то  
 $\Pi^\delta(i_k, j_k) = q_k$ ,  $\Pi^{\text{отказ}}(i_k, j_k) = \Pi(i_k, j_k) - q_k$ . (8)

3. Если  $\tau_k > t_{\text{ож}}^M$  и  $\tau_k q_k \geq \Pi(i_k, j_k) t_{\text{ож}}^M$ , то  
 $\Pi^\delta(i_k, j_k) = \Pi(i_k, j_k) \frac{t_{\text{ож}}^M}{\tau_k}$ ,  $\Pi^{\text{отказ}}(i_k, j_k) = \Pi(i_k, j_k) \left(1 - \frac{t_{\text{ож}}^M}{\tau_k}\right)$ . (9)

4. Если  $\tau_k > t_{\text{ож}}^M$  и  $\tau_k q_k < \Pi(i_k, j_k) t_{\text{ож}}^M$ , то  
 $\Pi^\delta(i_k, j_k) = q_k$ ,  $\Pi^{\text{отказ}}(i_k, j_k) = \Pi(i_k, j_k) - q_k$ . (10)

Таким образом, с использованием модели (1)-(10) может быть описан процесс обслуживания пассажиров при идеальном потокораспределении. Так как пассажиры ГОТ имеют информацию о процессе обслуживания, который их ожидает, то они, естественно, не будут совершать передвижения по тем идеальным для них путям, на которых наблюдается большое число отказов, то есть будет происходить перераспределение идеальных потоков и приспособление передвижения к реальной пропускной способности маршрутов гиперсети. Для того, чтобы учесть это, используем следующую процедуру пересчета временных параметров гиперграфа  $L(Z, W)$ . Естественно предположить, что часть пассажиров, получивших отказ на перегоне маршрута, будут вынуждены совершить по этому перегону пешее передвижение. В соответствии с этим произведём перерасчет времени передвижения по перегонам маршрутов  $(i_k, j_k) \in W$ :

$$t^{np'}(i_k, j_k) = \frac{1}{\Pi(i_k, j_k)} \left[ \Pi^{\text{отказ}}(i_k, j_k) \frac{l(i, j)}{V_T} + \Pi^\delta(i_k, j_k) \frac{l(i, j)}{V_n} \right]. \quad (11)$$

Такое изменение параметров графа  $L(Z, W)$ , разумеется, приведет к изменению путей передвижения  $H_{ij}$  между транспортными районами  $i, j \in Z$ . Для этих путей вновь рассчитаем вероятности  $P_{\xi}^{ij}$  по формуле (11). Проведем теперь с использованием новых вероятностей  $P_{\xi}^{ij}$  расчет нового потокораспределения. Для этого потокораспределения также будем изучать процесс обслуживания и т. д. Описанную процедуру необходимо повторять до тех пор, пока не сформируются неизменные  $\bar{H}_{ij}$  и соответствующие им  $\{P_{\xi}^{ij}\}$ ,  $\{T_{\xi}^{ij}\}$ .

Распределение потоков пассажиров по дугам графа  $L(Z, W)$  в равновесном состоянии  $\bar{\Pi}(h, e)$ ,  $(h, e) \in W$  учитывает стратегию выбора пассажирами путей следования и реальные процессы обслуживания на гиперсети ГОТ.

В целом для предлагаемой модели можно считать справедливым вывод, сделанный для модели, учитывающей ограничения на провозную способность транспорта [8], который заключается в том, что сходимость имеет место, если функции  $\{P_{\xi}^{ij}\}$  и  $\{T_{\xi}^{ij}\}$  являются непрерывными.

Соответствующее равновесному состоянию потокораспределение  $\bar{\Pi}(i, j)$  будем называть реальным потокораспределением на гиперсети ГОТ. Величина же

$$T = \sum_{(i, j) \in W} t(i, j) \bar{\Pi}(i, j), \quad (12)$$

где  $t(i, j)$  – временной параметр дуги  $(i, j) \in W$ , определяет суммарные затраты времени пассажиров на передвижение в городе.

Показано, что равновесное состояние соответствует минимальным затратам времени на передвижение пассажиров. В предлагаемой модели равновесное состояние определяет распределение пассажиров по путям следования в соответствии со стратегией выбора пути.

Процедура пересчета временных характеристик дуг графа маршрутной гиперсети ГОТ и потоков на них предусматривает использование для нахождения равновесного потока метода возможных направлений, причем на каждом шаге для отыскания подходящего направления нужно решать задачу распределения потоков в соответствии со стратегией (11).

### Литература

1. Попков В.К. Математические модели связности. / В.К. Попков – Новосибирск: Изд. ИВМиМГ СО РАН, 2006. - 409 с.
2. Вильсон А.Д. Энтропийные методы моделирования сложных систем. / А.Д. Вильсон – М.: Наука, 1978. – 248с.
3. Попков Ю.С. Системный анализ и проблемы развития городов. / Ю.С. Попков – М.: Наука, 1983.-512с.
4. Гольц Г.А. Транспорт и расселение. / Г.А. Гольц – М.:Наука, 1991.-248с.
5. Зенгбуш М.В., Белинский А.Ю., Дынкин А.Г. Пассажиропотоки в городах. / М.В. Зенгбуш, А.Ю. Белинский, А.Г. Дынкин – М.: Транспорт, 1984.-135с.
6. Яворский В.В. Моделирование распределения пассажиропотоков на маршрутной сети городского пассажирского транспорта. / В.В. Яворский //Динамика неоднородных систем.: Сб. трудов. – М.: ВНИИСИ, 1986.-Вып.9.– С. 90-101.



7. Грановский Б.И., Рогова Г.Л. Модели выбора и определение пассажиропотоков. / Б.И. Грановский, Г.Л. Рогова //Итоги науки и техники. Автомобильный и городской транспорт.– М.: Наука, 1968, с.7-12

8. Питтель Б.Г., Федоров В.П. Математическая модель прогноза пассажиропотоков в городской транспортной сети. / Б.Г. Питтель, В.П. Федоров //Экономика и математические методы.– 1989.– Т.5.–С.744-754.

## СРЕДСТВА СОВЕРШЕНСТВОВАНИЯ БЮДЖЕТИРОВАНИЯ ГОРНОДОБЫВАЮЩЕГО ПРЕДПРИЯТИЯ

<sup>1</sup>Яворский В.В., <sup>2</sup>Проценко А.В., <sup>3</sup>Байдикова Н.В.

<sup>1</sup> Карагандинский государственный технический университет, Казахстан

<sup>2</sup> Национальный исследовательский технологический университет "МИСиС",  
Россия

<sup>3</sup> Карагандинский государственный индустриальный университет, Казахстан  
e-mail: yavorskiy-v-v@mail.ru

***Аннотация.** Рассмотрены основные аспекты внедрения систем бюджетирования для повышения эффективности работы предприятий на примере горнорудной промышленности. На горнорудном предприятии требуется внедрение разнообразных бизнес-приложений. В такой ситуации важно обеспечить согласованность максимального числа факторов, определяющих результативность реализации каждого проекта. Отдельным направлением в данном аспекте должно быть совершенствование системы бюджетирования предприятия. Этого можно достичь путем обеспечения гибкости и прозрачности процедур планирования, учета, контроля и анализа по всем уровням организационной структуры предприятия.*

На горнорудном предприятии важно обеспечить согласованность максимального числа факторов, определяющих результативность реализации каждого проекта. Для решения этого вопроса целесообразно:

–рассмотреть совокупность взаимосвязанных проектов по внедрению бизнес-приложений в качестве программы проектов;

–реализовать кросс-проектный механизм утверждения программ функционирования и развития предприятия.

Целесообразным является формирование у заказчика проектного офиса – подразделения, определяющего и обеспечивающего выполнение стандартов проектного управления в компании [1].

Возникает необходимость использования новой концепции построения информационной системы развития предприятия, сущностью которой является установление наличия возможностей в системах горнодобывающего предприятия для развития в текущем и долгосрочном периодах. Для реализации модели управления развитием требуется определенная организационная подготовка и перестройка организационной системы предприятия по следующим направлениям [2]:

1. Изменения в управленческом учете развитие аналитических блоков по мониторингу внешней среды и расчету финансовых последствий принимаемых решений;

2. Разработка стандартов эффективности работы производственных подразделений и служб предприятия;

3. Распределение прав и ответственности при принятии решений по уровням управления;

4. Внедрение системы планирования, основанной на определении вариантов развития и согласовании целевых функций и показателей деятельности структурных подразделений в условиях рыночных отношений;

5. Использование системы бизнес-планирования развития предприятия;

6. Квалификационная подготовка персонала в части освоения необходимых навыков, формирование компетентного персонала.

Последовательность разработки при функциональном подходе базируется на следующем:

– выбор автоматизируемых функций в соответствии с текущим этапом развития КИС (разработка функционального обеспечения);

– разбивка функций на ряд частных задач;

– обследование информационных потоков и определение набора и структуры данных, необходимых для решения поставленных задач и выполнения автоматизируемых функций.

На рисунке 1 представлена общая схема процесса планирования, бюджетирования и прогнозирования предприятия.

Главным в данном аспекте является совершенствование системы бюджетирования предприятия [3]. Это достигается за счет обеспечения гибкости и прозрачности процедур планирования, учета, контроля и анализа по всем уровням организационной структуры предприятия, интеграции со смежными информационными системами, увеличения точности, достоверности и аналитической ценности получаемой информации по показателям производственной деятельности. Обеспечить это можно на основе использования алгоритмов многомерного хранения и анализа данных (OLAP – многомерное моделирование).



Рисунок 1 – Общая схема процесса планирования, бюджетирования и прогнозирования предприятия

При внедрении КИС в горнодобывающих компаниях, использующих функциональный подход, отмечается ряд характерных особенностей.

Интеграция с бизнес приложениями [4]:

- отсутствие интеграции между системами ERP, MES, АСУТП – ведет к значительным трудозатратам при переносе и консолидации данных (повторный ввод данных), системы класса MES зачастую отсутствуют;

- внедрение систем приостанавливается или завершается без достижения поставленных целей;

- используемые аналитики дают противоречивые результаты.

Автоматизация процессов управления производством:

- не соответствуют современному уровню системы производственного учета (не обеспечивают сведение материально энергетических балансов);

- средний уровень автоматизации технологических процессов. Автоматизация носит «лоскутный» характер.

Стандартизация программного обеспечения и аппаратных платформ систем АСУ ТП;

- низкий уровень покрытия, основные технологические процессы покрыты системами АСУ ТП, но многие используемые системы требуют модернизации;

- множество различных типов систем (часть систем собственной разработки) - приводит к сложности и высокой стоимости их поддержки, зависимости от собственных специалистов-разработчиков, потере знаний о системах;

- используются платформы различных производителей, отсутствуют единые стандарты и унифицированные решения на автоматизацию технологических процессов;

- значительная часть документации на системы отсутствует, либо находится в неудовлетворительном состоянии – сложность поддержки и развития систем, высокие риски и эксплуатационные затраты.

Объектно-ориентированный подход, как альтернатива функциональному, обеспечивает более гибкое решение:

- позволяет последовательно организовывать сбор первичных данных и реализовывать на основе этих данных любые функции, решать любые задачи учета, планирования, контроля и управления.

- при изменении функциональных потребностей корректируется только программное обеспечение ИС, а структура баз данных и первичные информационные потоки остаются неизменными.

Однако его использование создает определенные сложности. Необходима организационная подготовка и модернизация организационной системы компании по следующим направлениям [5]:

- изменения в управленческом учете – развитие аналитических блоков по мониторингу внешней среды и расчету финансовых последствий принимаемых решений;

- разработка системы стандартов эффективности работы производственных подразделений и служб предприятия;

- распределение прав и ответственности при принятии решений по уровням управления;

- внедрение системы планирования, основанной на определении вариантов развития и согласовании целевых функций и показателей деятельности структурных подразделений в условиях рыночных отношений;

- использование системы бизнес-планирования развития предприятия;
- квалификационная подготовка персонала в части освоения необходимых навыков, формирование компетентного персонала.

Одним из подходов решения является внедрение хранилища данных, что часто связано с внедрением систем класса BI и систем планирования. Хранилища данных применяются для сбора, обработки, хранения и предоставления пользователям больших объемов аналитической информации, хранящейся первоначально в различных источниках.

Хранилище данных строится на основе реляционных СУБД, таких как Oracle Database, MS SQL Server, DB2 и подобных. Использование реляционных СУБД и многомерных моделей внутри хранилища данных позволяет, с одной стороны, обеспечить хранение детальных данных (вплоть до операций, транзакций), а с другой – обеспечить удобную обработку больших объемов таких данных.

Сохранение исторических данных в хранилище позволит ограничить период, за который данные хранятся в моделях IBM Cognos TM1, тем самым «разгрузив» модели и сняв с них задачу по накоплению больших объёмов плановых и фактических данных. На рисунке 2 представлена схема взаимодействия служб отделов при использовании моделей IBM Cognos.



**Рисунок 2 – Схема взаимодействия служб отделов при использовании моделей IBM Cognos**

Реализация новых решений будет способствовать повышению эффективности работы, как руководящего состава, так и исполнителей. На рисунке 3 представлены примеры визуализации данных для более наглядного представления.

Для руководителей система позволит [3,4]:

- повысить качество управленческой информации за счет получения более детализированных данных и расширенного перечня отчетов.

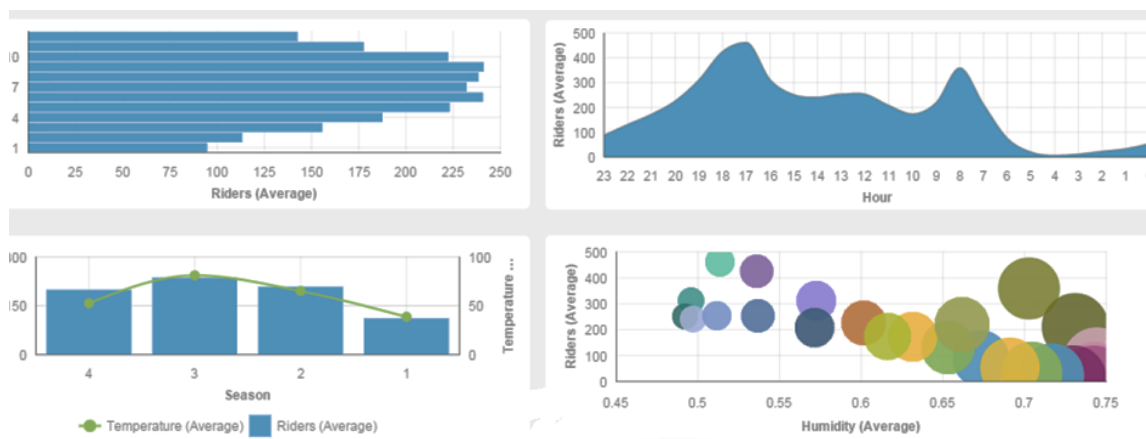
–повысить эффективность процесса формирования и согласования бюджетов: сокращение сроков и трудоемкости.

–предоставить возможность бюджетирования и мониторинга проектов в рамках всего жизненного цикла.

–перейти от дискретной модели планирования и контроля к скользящим бюджетам на весь срок реализации проектов.

–предоставить возможность эффективного управления оборотным капиталом.

–предоставить возможность задания стратегических целей на уровне всей компании, и отслеживание их достижения.



**Рисунок 3 - Пример визуализации данных**

Например: снижение себестоимости, или контроль издержек по "контролируемым расходам" - расходам, которые менеджмент может контролировать.

Например, контролируемые расходы - это постоянные издержки, например, аренда или SG&A. А неконтролируемые - это переменные, зависящие от выработки, например, топливо.

Сравнение финансовых потоков - целевых и фактических, глобально, в целом по компании, и выявление нестыковок.

Например, прибыль есть (на бумаге), а денег в кассе нет (фактически) - за счет манипуляций со складом и т.п.

### **Литература**

1. MRP ERP системы: бизнес-приложения, охватывающие все сферы деятельности компании: <http://asapcg.com/press-center/articles/mrp-erp-sistemy/> (дата обращения: 29.05.2018).

2. Бородулин А. Н. Программные средства бизнес-аналитики в системе управления современным предприятием [Текст] / А. Н. Бородулин // Экономическая наука сегодня: теория и практика : материалы III Междунар. науч.-практ. конф. (Чебоксары, 26 дек. 2015 г.) / редкол.: О. Н. Широков [и др.]. — Чебоксары: ЦНС «Интерактив плюс», 2015. — С. 286–289.

3. Карпов А. Автоматизация бюджетирования и управленческого учета (новый подход). -: [http://bud-tech.ru/budgeting\\_it.html](http://bud-tech.ru/budgeting_it.html). (дата обращения: 30.05.2018).

4. Нуруллина Л.Ш., Тихонова А.В., Соколова И.Н. Методология автоматизации процесса бюджетирования на предприятиях // Проблемы экономики и

менеджмента. 2015. №12 (52). - <http://cyberleninka.ru/article/n/metodologiya-avtomatizatsii-protssessa-byudzhetrovaniya-na-predpriyatiyah> (дата обращения: 30.05.2018).

5. Алиева З.Г., Ибрагимова А.Х. Значение бюджетирования для управления предприятием. / З.Г. Алиева, А.Х. Ибрагимова. // Символ науки. - 2016. - №11-1. с.17-18.

## РАЗРАБОТКА МЕТОДА ПРЕДСТАВЛЕНИЯ ТРАНСПОРТНОЙ СИСТЕМЫ НА ОСНОВЕ ТЕОРИИ ГИПЕРСЕТЕЙ ДЛЯ ПРОЕКТИРОВАНИЯ И УПРАВЛЕНИЯ

<sup>1</sup>Яворский В.В., <sup>1</sup>Утепбергенов И.Т., <sup>2</sup>Байдикова Н.В., <sup>2</sup>Чванова А.О.

<sup>1</sup> Институт информационных и вычислительных технологий КН МОН РК

<sup>2</sup> Карагандинский государственный индустриальный университет  
e-mail: yavorskiy-v-v@mail.ru

*Аннотация.* Рассмотрен математический аппарат формализованного описания транспортной системы. Транспортная сеть, маршрутная сеть, корреспонденции между пунктами на транспортной сети, потоки на маршрутной сети, пути передвижения между пунктами транспортной системы предлагается представлять гиперсетью, дуги которой представляют собой пути передвижения. Введено точное понятие пути передвижения между корреспондирующими пунктами на городском общественном транспорте. Описана рекуррентная схема получения путей передвижения между вершинами гиперсети любого порядка.

Актуальность разработки формализованного описания транспортной системы заключается в том, что эффективные модели и алгоритмы, разработанные применительно к одним транспортным системам, после дополнительного учета ряда специфических особенностей, могут быть использованы и для других систем. Это связано с тем, что в рассматриваемых транспортных системах имеется ряд общих основных элементов: транспортная сеть, маршрутная сеть, корреспонденции между пунктами на транспортной сети (ТС), потоки на маршрутной сети (МС), пути передвижения между пунктами ТС, способы обслуживания потоков на МС и т.п. Очевидно также, что взаимосвязь элементов и процессы движения в транспортной системе можно описать, используя графовое представление системы [1-3] в виде взаимосвязанных сетей, на которых формируются потоки. Интегрированное описание транспортной системы представляет собой гиперсеть [4].

Пусть имеется некоторая транспортная система, основу которой составляет один из видов городского общественного транспорта (ГОТ). Рассмотрим подробно каждый из основных элементов транспортной системы.

Под транспортной сетью понимается совокупность пунктов, характеризующихся определенным местоположением и взаимными связями между собой. ТС может быть представлена в виде графа  $S(Z, W)$ , где  $Z$  – конечное непустое множество пунктов  $i \in Z$ ,  $|Z|=n$ ;  $W$  – множество связей (дуг или ребер) –  $(i, j) \in W$ , длина которых

соответствует расстоянию между пунктами  $i$  и  $j$ ,  $|W|=d$ ;  $i \neq j$ ;  $i, j = \overline{1, n}$ , здесь  $n$ ,  $d$  – соответственно число пунктов и связей на графе  $S(Z, W)$ .

МС задается совокупностью транспортных маршрутов (ТМ) –  $\{M^k\}$ , где  $k \in I^M$ ;  $|I^M| = m$ ;  $I^M$  – множество индексов всех ТМ  $k$ , функционирующих на МС;  $m$  – число ТМ на МС.

Каждый ТМ  $k \in I^M$  описывается последовательностью прохождения пунктов  $i^k_\xi$ ,  $\xi = \overline{1, n_k}$ , составляющих траекторию его движения:

$$M^k = \{i^k_1, i^k_2, \dots, i^k_{n_k}\}, \quad (1)$$

где  $n_k$  – число пунктов в ТМ  $k$ ; отметим, что на рассматриваемых ТМ  $k$  начальный и конечный пункты в траектории  $M^k$  совпадают, т.е.  $i^k_1 = i^k_{n_k}$ .

Связь  $(i^k_\xi, i^k_\eta)$  является маршрутной связью  $k$ -го ТМ. Множество всех маршрутных связей  $(i^k_\xi, i^k_\eta)$ ,  $k \in I^M$ ,  $k = \overline{1, m}$ ,  $\xi \neq \eta$ ;  $\xi, \eta \in \overline{1, n_k}$  на маршрутной сети можно представить в виде ориентированного гиперграфа  $M(Z_l, W_l)$ .  $Z_l$  – множество пунктов  $i \in Z_l$ ,  $|Z_l| = n_l$ ,  $Z_l \subseteq Z$ ;  $W_l$  – множество маршрутных связей  $(i^k_\xi, i^k_\eta) \in W_l$ , в качестве весов которым можно присвоить затраты времени при следовании из пункта  $i^k_\xi$  в пункт  $i^k_\eta$  на ТМ  $k$ , или же длина между  $i^k_\xi$  и  $i^k_\eta$  на графе  $S(Z, W)$  и т.д. Количество маршрутных связей  $|W_l|=d_l$  таково, что их число  $d_l$ , как правило, значительно больше  $d$ , кроме того, очевидно  $n_l \leq n$ .

По транспортным маршрутам  $k \in I^M$  в соответствии с его траекторией движения  $M^k$  перемещаются транспортные единицы, которые обеспечивают производство перевозок на ТМ.

Функционирование каждого транспортного маршрута  $k$  характеризуется набором параметров  $\{\Psi^k\} = \{\psi^k_1, \psi^k_2, \dots, \psi^k_l\}$ , где  $\psi^k_i$  –  $i$ -й параметр ТМ  $k \in I^M$ ;  $i = \overline{1, l}$ ,  $l$  – число параметров  $\psi^k_i \in \Psi^k$ .

Отметим, что тип транспортного маршрута  $k$  определяется типом транспортных единиц (ТЕ)  $\varphi_k$ , функционирующих на данном ТМ.

Между пунктами  $i$  и  $j$ ,  $i \neq j$ ;  $i, j \in Z$  существуют определенные потоки передвижений. Если эти потоки можно считать стационарными, то потенциальные потоки между пунктами  $i$  и  $j$  определяются корреспонденциями, которые можно представить в виде матрицы

$$\Lambda = \|\lambda_{ij}\|, \quad (2)$$

где  $\lambda_{ij}$  – интенсивность потока из  $i$  в  $j$ ,  $i \neq j$ ;  $i, j \in Z$ ,  $i, j \in \overline{1, n}$ .

Корреспонденции  $\lambda_{ij}$  распределяются по маршрутной сети в виде потоков на маршрутах [5-7], которые можно представить с помощью матриц потоков:

$$X^k = \|x^k_{\xi\eta}\|, \quad (3)$$

где  $x^k_{\xi\eta}$  – интенсивность потока из пункта  $\xi$  в пункт  $\eta$  на транспортном маршруте  $k \in I^M$ ,  $k = \overline{1, m}$ ;  $\xi, \eta = \overline{1, n_k}$ ;  $n_k$  – число пунктов на ТМ  $k$ .

Корреспонденция  $\lambda_{ij}$  состоит из отдельных потребностей в передвижении, которые будем называть элементами корреспонденции или просто элементами. Реализация элементов корреспонденции может происходить разными способами.

Например, возможно индивидуальное передвижение каждого элемента, при этом ограничения на такое передвижение накладываются в виде его продолжительности по времени. Наряду с индивидуальным передвижением элементы корреспонденций могут совершать передвижения с использованием одного или нескольких транспортных маршрутов типа  $\varphi_k, k \in I^M$ .

В процессе передвижения с использованием ТМ  $k$  возможна пересадка с одного типа ТЕ  $\varphi_s$  на другой тип ТЕ  $\varphi_b, s \neq b$ , или с одной ТЕ на другую ТЕ одного и того же типа  $\varphi_k$ , которая осуществляется с использованием индивидуального передвижения. Необходимость выполнения пересадки обуславливается чаще всего тем, что маршрутные связи  $(i^k_\xi, i^k_\eta)$  существуют не между всеми парами пунктов  $i$  и  $j$  на ТС. Кроме того, для различных связей отдельные виды транспорта могут иметь различные приоритеты.

Реализация передвижений элементами корреспонденции, как правило, осуществляется по нескольким конкурирующим (альтернативным) путям передвижений. Наиболее предпочтительным при этом является кратчайший путь и близкие к нему. Пути передвижения состоят из последовательности чередующихся индивидуальных и маршрутных связей, причем путь передвижения должен начинаться и кончаться индивидуальной связью. Введем более точное понятие пути передвижения между корреспондирующими пунктами на транспортной сети.

Путь  $h^{p_{kij}}$  является  $k$ -м путем передвижения между пунктами  $i$  и  $j$  на МС порядка  $p$ , если число маршрутных связей  $(i^l_\xi, i^l_\eta)$  на данном пути равно  $p$ , а количество пересадок на единицу меньше. Здесь  $l$  – номер маршрута:  $l \in I^M$ . Число маршрутных связей в маршруте (число пересадок) можно ограничить  $p = \overline{1, S}$ ,  $S$  – заданное допустимое число маршрутных связей  $(i^l_\xi, i^l_\eta)$  в пути  $h^{p_{kij}}$ . Так как отклонение от кратчайшего пути также можно рассматривать лишь в допустимых пределах, то число путей между пунктами  $i$  и  $j; i \neq j; i, j \in Z$ , также ограничено:  $k = \overline{1, N}$ .

Путь  $h^{p_{kij}}$  между пунктами  $i$  и  $j$  на МС можно описать следующим образом:

$$h^{p_{kij}} = \{i, j; t(h^{p_{kij}}); p; k_1, \xi_1, \eta_1; k_2, \xi_2, \eta_2; \dots; k_l, \xi_l, \eta_l, \dots, k_p, \xi_p, \eta_p\}, \quad (4)$$

где  $t(h^{p_{kij}})$  – затраты времени на передвижение по пути  $h^{p_{kij}}$ ;  $p$  – порядок пути;  $k_l$  – индекс ТМ, осуществляющего  $l$ -ю маршрутную связь  $(\xi_l, \eta_l)$ ;  $\xi_l$  и  $\eta_l$  – начальный и конечный пункты в маршрутной связи  $(\xi_l, \eta_l)$ ,  $l = \overline{1, p}$ .

В дальнейшем для сокращения обозначений вместо  $h^{p_{kij}}$  будем писать  $h_{kij}$ .

Обычно при определении разумных и целесообразных ограничений на число конкурирующих путей  $h_{kij}$  между пунктами  $i$  и  $j$  на МС можно руководствоваться определенной стратегией поведения. Такая стратегия задается вероятностной функцией предпочтения  $p(h_{kij})$  путей передвижения  $h_{kij}$ , связывающих корреспондирующие пункты  $i$  и  $j$ :

$$p(h_{kij}) = \mathcal{F}(v^k_1, v^k_2, \dots, v^k_l), \quad (5)$$

где  $v^k_\rho$  – величина  $\rho$ -го параметра, характеризующего путь  $h_{kij}$ ;  $\rho = \overline{1, l}$ ,  $l$  – число параметров.

Вероятностная функция предпочтения (5) позволяет произвести распределение корреспонденций  $\lambda_{ij}$  по путям  $h_{kij}$  пропорционально вероятностям:

$$x(h_{kij}) = \lambda_{ij} \cdot p(h_{kij}), \quad (6)$$



где  $x(h_{kij})$  – поток по пути  $h_{kij}$ ,  $k = \overline{1, N}$ ,  $N$  – число путей между пунктами  $i$  и  $j$ .

Поскольку распределение корреспонденций  $\lambda_{ij}$  может быть выполнено только по путям  $h_{kij}$ , связывающим пункты  $i$  и  $j$ , то, естественно, выполняются следующие равенства:

$$\sum_{h_{kij} \in H_{ij}} p(h_{kij}) = 1; \quad (7)$$

$$\sum_{h_{kij} \in H_{ij}} x(h_{kij}) = \lambda_{ij}, \quad (8)$$

где  $H_{ij}$  – множество всех путей, связывающих пункты  $i$  и  $j$ .

Приведем интерпретацию предложенного выше формализованного описания транспортной системы для системы ГОТ.

Транспортной сетью для ГОТ является улично-дорожная сеть (УДС) города, которая представляется графом УДС –  $S(Z, W)$ , где  $Z$  – конечное непустое множество зон города  $i \in Z$ ,  $|Z| = n$ ;  $W$  – множество ребер  $(i, j) \in W$ , длина которых соответствует расстоянию между зонами  $i$  и  $j$  города на графе УДС;  $|W| = d$ ,  $i, j = \overline{1, n}$ ,  $i \neq j$ ;  $n, d$  – соответственно число зон и ребер на графе УДС  $S(Z, W)$ .

Маршрутная сеть ГОТ состоит из множества маршрутов. Маршруты в городе бывают разного типа. Тип маршрута определяется, как правило, типом транспортных единиц, осуществляющих перевозки на маршруте.

Каждый  $k$ -й маршрут ГОТ описывается последовательностью прохождения зон  $i^k_{\xi}$ ,  $\xi = \overline{1, n_k}$  в прямом направлении до возвращения в пункт отстоя:

$$M^k = \{i_1^k, i_2^k, \dots, i_{n_k}^k\} \quad (9)$$

где  $n_k$  – число зон, через которые проходит маршрут  $k$ .

Отметим, что  $i_1^k = i_{n_k}^k$ , то есть начальная и конечная зоны на маршруте  $k$  совпадают, так как траектория любого маршрута  $M^k$  представляет собой кругорейс маршрута. Номера маршрутов образуют множество:  $k \in I^M$ ,  $m$  – число маршрутов в МС ГОТ.

Маршрутные связи ГОТ обозначим  $h^M_{k\xi\eta} = (i^k_{\xi}, i^k_{\eta})$ ,  $k \in I^M$ ,  $\xi \neq \eta$ ;  $\xi, \eta = \overline{1, n_k}$ . Они образуют гиперсеть беспересадочных связей  $M(Z_1, W_1)$ , где  $Z_1$  – конечное непустое множество зон  $i \in Z_1$ ,  $|Z_1| = n_1$ ,  $Z_1 \subseteq Z$ ;  $W_1$  – множество беспересадочных маршрутных связей  $(i^k_{\xi}, i^k_{\eta}) \in W_1$ , весами которых является время проезда по ним,  $|W_1| = d_1$ .

Поскольку в общем случае маршруты могут проходить не через все зоны города  $i \in Z$ , то  $n_1 \leq n$ , а множество беспересадочных связей в достаточно развитых сетях ГОТ на несколько порядков превышает множество связей на УДС, то есть  $d_1 > d$ .

Каждый маршрут ГОТ  $k \in I^M$  характеризуется следующими параметрами: числом ТЕ, осуществляющих перевозки,  $a_k$ , скоростью движения ТЕ  $V^k$ , интервалом движения ТЕ  $\tau_k$ , протяженностью маршрута  $L_k$ , типом ТЕ  $\varphi_k$ , вместимостью ТЕ  $\aleph_k$ .

Потенциальными потоками между зонами  $i$  и  $j$  города являются потенциальные корреспонденции, которые можно представить в виде матрицы:

$$A = \|\lambda_{ij}\|, \quad (10)$$

где  $\lambda_{ij}$  – интенсивность передвижения из зоны  $i$  в зону  $j$ ,  $i, j = \overline{1, n}$ ,  $i, j \in Z$ .

Матрицы потоков на ТМ – это матрицы реальных пассажиропотоков на маршрутах ГОТ  $k$ :

$$X^k = \parallel x^k_{\xi\eta} \parallel, \quad (11)$$

где  $x^k_{\xi\eta}$  – интенсивность пассажиропотока из зоны  $i$  в зону  $j$  на маршруте  $k$ ;  $\xi \neq \eta$ ;  $\xi, \eta = \overline{1, n_k}$ ;  $n_k$  – число зон в траектории маршрута  $k \in I^M$ .

Элементами корреспонденций в условиях ГОТ являются отдельные передвижения. Пассажиры при передвижениях между зонами города используют пешие (индивидуальные) передвижения, а также передвижения с использованием маршрутов ГОТ.

Пассажир в процессе своего передвижения также совершает возможные пересадки, что связано с отсутствием беспересадочных связей между всеми парами зон  $i$  и  $j$  города на УДС.

Передвижение пассажиров между зонами города осуществляется по конкурирующим путям передвижения в соответствии со стратегией поведения пассажиров [8-10]. В качестве основных параметров, являющихся характеристиками путей, выступают: время передвижения, включающее в себя время на пеший подход и отход от зоны начала и конца передвижения; время на пересадку, время проезда на ТЕ, время ожидания обслуживания; количество пересадок, комфортность проезда.

Процедура свободного распределения корреспонденций  $\lambda_{ij}$  на МС ГОТ производится в соответствии с (6).

Рассмотрим более подробно представление элементов, из которых составляются пути передвижения  $h_{kij}$ ,  $k = \overline{1, N}$ ,  $N$  – число путей между зонами ТС  $i$  и  $j$ . Приведем определение базовых элементов пути передвижения.

Индивидуальная связь между пунктами  $i$  и  $j$  на ТС отражает возможность пешего передвижения при реализации элементов корреспонденции  $\lambda_{ij}$ :

$$h^M_{ij} = \{i, j; t^M_{ij}\}, \quad (12)$$

где  $t^M_{ij}$  – время индивидуального передвижения из пункта  $i$  в пункт  $j$ ,  $t^M_{ij} = \infty$ , если между пунктами  $i$  и  $j$  нет индивидуального сообщения на графе ТС;  $i, j = \overline{1, n}$ ,  $i, j \in Z$ .

Между зонами  $i$  и  $j$  на ТС в общем случае может быть несколько индивидуальных связей, поэтому будем рассматривать множество индивидуальных связей:

$$H^M_{ij}: h^M_{ij} \in H^M_{ij}. \quad (13)$$

Индивидуальная связь  $h^M_{ij} = \{i, i; t^M_{ii}\}$  определяет возможности подхода к остановкам маршрутов в зоне  $i$ ;  $t^M_{ii}$  – среднее время подхода к остановке ТМ в зоне  $i$ ,  $i = \overline{1, n}$ ,  $i \in Z$ .

Маршрутная связь  $h^M_{k\xi\eta} = (i^k_{\xi}, i^k_{\eta})$  соответствует поездке, совершаемой элементами корреспонденций на некотором транспортном маршруте  $k$  без пересадки:

$$h^M_{k\xi\eta} = \{k, \xi, \eta; t^M_{k\xi\eta}\}, \quad (14)$$

где  $\xi$  и  $\eta$  – номера начальной и конечной зоны  $k$ -го маршрута;  $t_{k\xi\eta}^M$  – время передвижения по связи  $(\xi, \eta)_k$ .

Обозначим  $H_{\xi\eta}^M$  – множество всех маршрутных связей между  $\xi$  и  $\eta$  с использованием всех возможных маршрутов:  $h_{k\xi\eta}^M \in H_{\xi\eta}^M$ . Маршрутная связь  $h_{kii}^M$  является недопустимой, поэтому время передвижения по ней выбирается  $t_{kii}^M = \infty$ .

Маршрутно-индивидуальная связь между пунктами  $i$  и  $j$  – это связь, представляющая собой сочетание некоторой маршрутной связи  $h_{k\xi\eta}^M$  и затем индивидуальной связи  $h_{\eta j}^I$ :

$$h_{k\xi j}^{MI} = \{ \xi, j; t_{k\xi j}^{MI}; k, \xi, \eta \}, \quad (15)$$

где  $t_{k\xi j}^{MI}$  – время передвижения по маршрутно-индивидуальной связи  $h_{k\xi j}^{MI}$ ;  $\xi$  и  $\eta$  – начальный и конечный индексы пунктов беспересадочной маршрутной связи  $(\xi, \eta)_k$ ;  $k$  – номер ТМ, осуществляющего связь  $(\xi, \eta)_k$ .

Индивидуальное передвижение осуществляется между пунктами  $\eta$  и  $j$ , информация о возможности такого передвижения имеется в  $h_{\eta j}^I$  – индивидуальной связи между пунктами  $\eta$  и  $j$ . Обозначим  $H_{ij}^{MI}$  – множество всех маршрутно-индивидуальных связей, связывающих пункты  $i$  и  $j$  на ТС,  $i, j = \overline{1, n}$ ,  $i \neq j$ ;  $i, j \in Z$ .

В общем случае между заданными пунктами  $i$  и  $j$  на ТС существует несколько индивидуальных, маршрутных и маршрутно-индивидуальных связей.

Обозначим множество индивидуальных связей между всеми парами пунктов  $i$  и  $j$

$$H^I = \bigcup_{i,j} H_{ij}^I; \quad (16)$$

множество маршрутных (беспересадочных) связей между всеми парами пунктов  $\xi$  и  $\eta$

$$H^M = \bigcup_{\xi\eta} H_{\xi\eta}^M; \quad (17)$$

множество маршрутно-индивидуальных связей между всеми парами пунктов  $\xi$  и  $j$

$$H^{MI} = \bigcup_{\xi,j} H_{\xi j}^{MI}. \quad (18)$$

Множество допустимых путей передвижения между пунктами  $i$  и  $j$  порядка  $p$  обозначим  $H_{ij}^p$ . Для определения таких путей введем специальную операцию склеивания путей –  $\otimes$ . Беспересадочные пути передвижения между зонами  $i$  и  $j$   $H_{ij}^1$  будем определять, производя склеивание индивидуальных и маршрутно-индивидуальных связей:

$$H_{ij}^1 = H_{i\xi}^I \otimes H_{\xi j}^{MI}. \quad (19)$$

В случае, если мы сформировали пути передвижения  $s$ -го порядка между всеми парами вершин  $i, \xi \in Z - H_{i\xi}^s$ , можно получить пути передвижения  $(s+1)$ -го порядка по схеме

$$H_{ij}^{s+1} = H_{i\xi}^s \otimes H_{\xi j}^{MI}. \quad (20)$$

Таким образом, соотношения (19) и (20) задают рекуррентную схему получения путей передвижения между вершинами  $i, j \in Z$  любого порядка. Для реальных транспортных систем допустимые пути передвижения имеют ограниченный порядок. Например, для системы ГОТ практически не используются пути передвижения выше 3-го порядка (имеющие три пересадки). Обозначим высший порядок рассматриваемых путей  $q^M$ , тогда все множество путей передвижения, связывающих вершины  $i$  и  $j$ , определяется как

$$H_{ij} = H_{ij}^M \cup \left[ \bigcup_{S=1}^{q^M} H_{ij}^S \right], \quad (21)$$

а все множество путей передвижения

$$H = \bigcup_{i,j} H_{ij}. \quad (22)$$

Определим теперь более детально операцию склеивания  $\otimes$ , используемую для моделирования системы городского пассажирского транспорта. Для формирования базы данных путей передвижения (2.83) используется граф улично-дорожной сети и информация о маршрутах ГОТ. На подготовительном этапе производится формирование гиперсетей пешеходных и маршрутно-пешеходных связей в городе. Для дальнейшего формирования базы данных используется гиперсеть маршрутно-пешеходных связей, дуги которой "склеиваются" с пешеходными путями, а затем путями первого и любого другого порядка. Таким образом формируются дуги гиперсети путей передвижения между всеми парами вершин, соответствующих транспортным зонам города. Множество этих дуг можно существенно ограничить, учитывая реальное поведение населения города при передвижении. Прежде всего, путь передвижения можно считать допустимым и вводить соответствующую дугу в гиперсеть, если число совершаемых на нём пересадок не больше некоторой максимальной величины  $q^M$ . В современных городских транспортных системах  $q^M$  можно принимать равным двум для городов до миллиона жителей и равным трём для больших городов.

Вполне обоснованным является требование, чтобы время передвижения  $t_{kij}$  отличалось от минимально возможного времени передвижения между зонами  $i$  и  $j$   $t_{ij}^M$  не более чем в  $\alpha$  раз:  $t_{kij} \leq \alpha t_{ij}^M$ . Необходимо также, чтобы последовательность используемых в пути передвижения маршрутов была элементарной (номера маршрутов ГОТ не должны повторяться). Элементарным должен быть и путь  $i, \xi_1, \eta_1, \dots, \xi_q, \eta_q, j$ , при этом, однако, допустимы пешие переходы внутри одной зоны. Это означает, что допустимы равенства:  $i = \xi_1, \eta_l = \xi_{l+1}, l = \overline{1, q-1}, \eta_q = j$ . Очевидно, что для последовательности  $i, \xi_1, \eta_1, \dots, \xi_q, \eta_q, j$  можно сформулировать ряд других естественных ограничений, описывающих процессы передвижения в

конкретном городе и являющихся фактически константами самоорганизации транспортной системы города [11].

Для получения совокупности дуг  $\{H_{ij}^{q+1}\}$  берутся совокупности  $\{H_{ij}^q\}$  и к ним присоединяются маршрутно-пешеходные связи из совокупности  $\{H_{ij}^{МП}\}$ .

Операцию склеивания  $\otimes$  можно определить следующим образом:

$$(i, \xi, t_s; q, k_1, \dots, k_q; \xi_1, \eta_1, \dots, \xi_q, \eta_q) \otimes (\xi, j; t_V; k_V; \eta_V) = \begin{cases} (i, j; t_s + t_V; q+1; k_1, \dots, k_q, k_V; \xi_1, \eta_1, \dots, \xi_q, \eta_q, \xi, \eta_V), \text{ если} \\ q+1 \leq q^M \wedge t_s + t_V \leq \alpha_{ij}^M \wedge (k_1, \dots, k_q, k_V) \text{ — элементарная после} \\ \text{довательность} \wedge (i, \xi_1, \eta_1, \dots, \xi_q, \eta_q, \xi, \eta_V, j) \text{ — элементарный путь} \\ \emptyset \text{ — в противном случае.} \end{cases} \quad (23)$$

Заключительным этапом формирования базы данных путей передвижения является объединение данных о путях, содержащихся в массивах пешеходных, беспересадочных и прочих путей:  $H_{ij}^n, H_{ij}^1, H_{ij}^2, H_{ij}^3, H_{ij}^4$ , и создание массива  $H = \{H_{ij}\}$ , записи которого описывают дуги гиперсети связей между транспортными зонами города.

### Литература

1. Кристофидес Н. Теория графов. –М.:Мир, 1978.-432с.
2. Хейт Ф. Математическая теория транспортных потоков./ Ф. Хейт –М.:Мир, 1966.-187с.
3. Дрю Д. Теория транспортных потоков и управление ими / Д.Дрю. –М.:Транспорт, 1972.-424с.
4. Попков В.К. Математические модели связности. / В.К. Попков – Новосибирск: Изд. ИВМиМГ СО РАН, 2006. - 409 с.
5. Вильсон А.Д. Энтропийные методы моделирования сложных систем./А.Д. Вильсон –М.: Наука, 1978. – 248с.
6. Попков Ю.С. Системный анализ и проблемы развития городов. / Ю.С. Попков –М.: Наука, 1983.-512с.
7. Математические методы в управлении городскими транспортными системами.– Л.:Наука,1979.-151с.
8. Яворский В.В., Утепбергенов И.Т. Структурные методы совершенствования управления транспортными системами городов./ В.В. Яворский, И.Т. Утепбергенов - Караганда: КарГТУ, 2006. – 227с.
9. Яворский В.В. Моделирование распределения пассажиропотоков на маршрутной сети городского пассажирского транспорта. / В.В. Яворский //Динамика неоднородных систем.: Сб. трудов. –М.: ВНИИСИ, 1986.-Вып.9.– С. 90-101.
10. Яворский В.В. Моделирование процессов формирования и обслуживания транспортных потребностей населения города. / В.В. Яворский //Прикладные проблемы управления макросистемами: Тр. Первой всесоюзной школы семинара. – М.: ВНИИСИ, 1986. -С. 53-59.
11. Гольц Г.А. Транспорт и расселение. / Г.А. Гольц –М.:Наука, 1983.-248с.

# МОДЕЛИРОВАНИЕ ПРОЦЕССОВ ДЛЯ ИНТЕЛЛЕКТУАЛЬНОЙ ТРАНСПОРТНОЙ СИСТЕМЫ ГОРОДА

<sup>1</sup>Яворский В.В., <sup>2</sup>Чванова А.О., <sup>2</sup>Байдикова Н.В.

<sup>1</sup> Карагандинский государственный технический университет, Казахстан

<sup>2</sup> Карагандинский государственный индустриальный университет, Казахстан  
e-mail: yavorskiy-v-v@mail.ru

***Аннотация.** В работе рассмотрены основные задачи моделирования процессов передвижения для интеллектуальной транспортной системы города, которая содержит приложения по управлению работой городского общественного транспорта, информированию пассажиров общественного транспорта, модули интегрированной дорожно-транспортной среды, а также приложения по интермодальной маршрутизации, аварийной связи, и экологическому управлению дорожным движением. Для внедрения этих средств необходимо исследовать текущую дорожную ситуацию и создать цифровое хранилище данных о дорожной ситуации. Решение этих задач невозможно без создания информационной цифровой транспортной модели ИТС.*

Совершенствования планирования и управления транспортными системами города на современном уровне предполагает реализацию цифровых технологий управления. В настоящее время решение этих задач реализуется на уровне SMART-систем.

Несомненно, что важнейшей составной частью интеллектуальных транспортных систем (ИТС) является внедрение электронных систем управления дорожным движением, управления паркингами и информирования участников движения с целью решения задач приоритизации ГПТ (с использованием систем “park&ride”). Необходимо стимулировать смещение пикового периода на время слабой загрузки УДС, обеспечить участников движения информацией о текущей транспортной ситуации в городе и наличии свободных мест для парковки автомобилей для уменьшения интенсивности движения, связанной с поиском места для парковки. Необходимо повышать эффективность работы динамичных транспортных маршрутов (микроавтобусов, такси). Телематические технологии применяются для отслеживания местоположения транспортного средства и обмена информацией о местонахождении и текущем графике движения посредством бортовых электронных систем.

Обязательным при внедрении этих передовых технологий являются меры нормативно-правового регулирования. Повышение пропускной способности за счет развития улично-дорожной сети, организация парковочного пространства. Все это также предполагает использование моделирования процессов передвижения. И закрепления полученных результатов в нормативно-правовых актах.

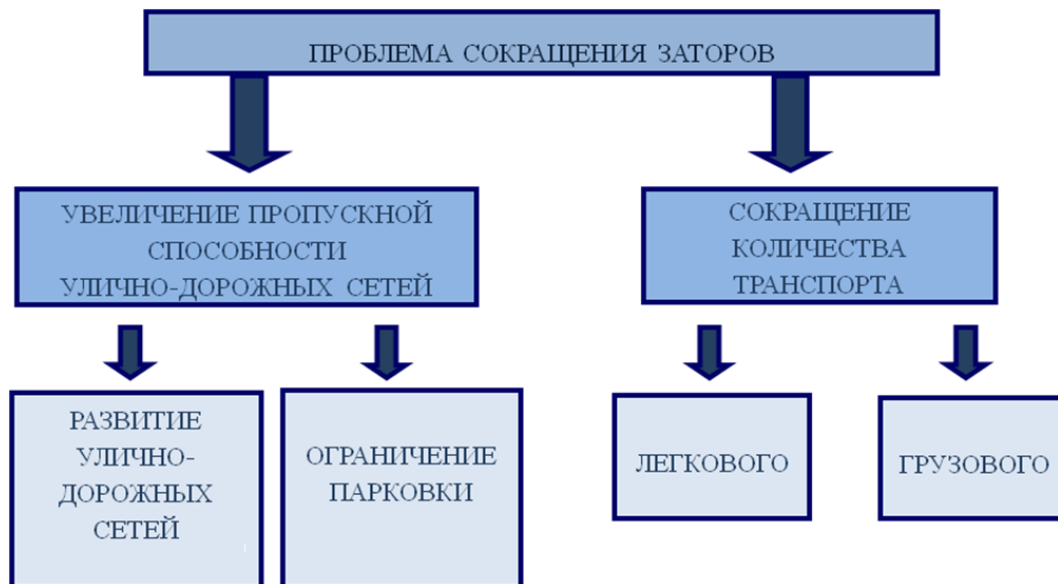
В состав внедряемых в рамках ИТС подсистем в разных сочетаниях входят приложения по управлению работой городского общественного транспорта (ГОТ), информированию пассажиров общественного транспорта, модули интегрированной дорожно-транспортной среды, а также приложения по интермодальной маршрутизации, аварийной связи, и экологическому управлению дорожным движением. Для внедрения этих средств необходимо исследовать текущую дорожную ситуацию и создать цифровое хранилище данных о дорожной ситуации.

Одной из основных проблемы современных городов является то, что загруженность дорог в крупных городах растет. Актуальная проблема сокращения заторов на дорогах может быть решена двумя путями: увеличение пропускной способности улично-дорожных сетей и сокращение количества как грузового, так и легкового транспорта. Второй путь реализуется за счет двух основных мероприятий: организации движения грузового транспорта и ограничение присутствия легкового.

Уменьшение присутствия легкового транспорта на перегруженной дорожной сети ограничивается посредством перераспределения с легкового на городской общественный транспорт (ГОТ) за счет повышения качества обслуживания и привлекательности общественного транспорта. Важно совершенствовать материально-техническую базу управления ГОТ, создания цифрового хранилища данных о передвижениях [1].

Движение грузового транспорта необходимо регулировать посредством интеллектуальной транспортной системы (отслеживать местоположение, маршруты).

Решение этих задач невозможно без создания информационной цифровой транспортной модели ИТС. В мировой практике накоплен опыт использования транспортных моделей разного уровня и степени детализации, ориентированных на решение разных классов задач.



**Рисунок 1 - Решение проблемы сокращения заторов**

В целом, транспортные модели [2] могут быть разной степени детализации: локальный уровень, уровень города и межрегиональный и региональный уровень. Каждый из этих уровней решает определенные задачи.

На локальном уровне строятся имитационные модели объектов транспортной инфраструктуры, на уровне города – моделируются пассажирские потоки, спрос на них, а на уровень региона – прогнозируются состояние и развитие автомобильно-дорожной сети, а также строится модель распределения транспортных и пассажирских потоков на мультимодальной транспортной сети. Основные элементы цифровой транспортной модели представлены на рисунке 2.

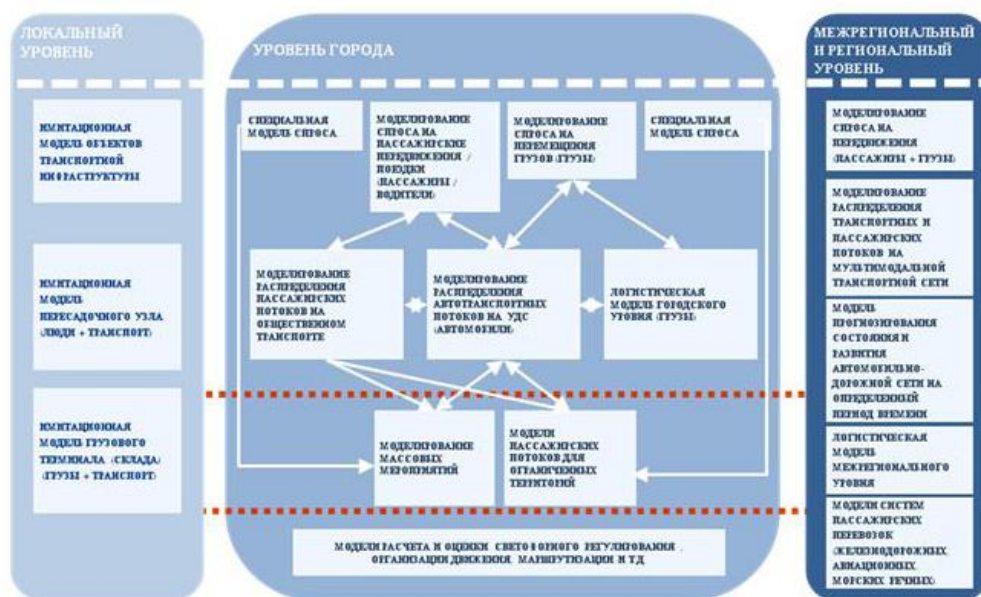


Рисунок 2 – Обобщенная классификация транспортных моделей

Исследования по исследованию возможностей применения smart-технологий в сфере городского транспорта выполняются в рамках проекта AP05133699 Комитета науки МОН РК.

### Литература

1. Калимолдаев М.Н., Утепбергенов И.Т., Ахмедиярова А.Т. Об одной задаче маршрутизации транспорта в мегаполисе. / М.Н. Камимолбаев, И.Т. Утепбергенов, А.Т. Ахмедиярова. // Вестник КазНТУ имени К.Сатпаева. - 2016. - №1. - С.409-414.
2. Яворский В.В., Утепбергенов И.Т. Структурные методы совершенствования управления транспортными системами городов. / В.В. Яворский, И.Т. Утепбергенов. – Караганда: Изд-во КарГТУ, 2006. – 272с.

УДК 004.056

## ГУМАНИТАРНЫЕ АСПЕКТЫ РЕШЕНИЯ ПРОБЛЕМ КИБЕРБЕЗОПАСНОСТИ РЕСПУБЛИКИ КАЗАХСТАН

Покусов Виктор Владимирович

*Казахстанская ассоциация информационной безопасности*

### Введение

Динамика развития информационных технологий в социально-экономической и культурной жизни общества и государства предъявляет повышенные требования к обеспечению информационной безопасности (ИБ). Для этого государству необходимо использование комплексного подхода, включающего организационные, технические,



программные, социальные и другие механизмы, способные реализовать конституционные права и свободу человека и общества на информацию. Это необходимо как для защиты конституционного строя, суверенитета и территориальной целостности Республики Казахстан, так и для политической, экономической и социальной стабильности, законности и правопорядка, а также безусловно развития взаимовыгодного международного сотрудничества в области ИБ.

Происходящие в настоящее время процессы преобразования в политической жизни и экономике оказывают непосредственное влияние на состояние их ИБ. При этом постоянно возникают новые факторы, которые необходимо учитывать при оценке реального состояния дел, определении ключевых проблем и выборе направлений для их разрешения.

### **Проблемы влияния Интернета на человека и общество**

Сегодняшние условия политического и социально-экономического развития вызывают обострение противоречий между потребностями общества в расширении свободного обмена информацией и необходимостью ограничения ее бесконтрольного распространения. Так, с возрастанием роли Интернета, предоставляющего свободный доступ к множеству научно-познавательных и технических ресурсов, в информационном пространстве возникает необходимость защиты прав и свобод человека и общества от информации, пропагандирующей насилие и жестокость, навязывающей недостоверные представления о ценностях, целенаправленно формирующей негативное мировоззрение молодого поколения и т. п. Все это говорит о существовании проблемы влияния Интернета на человека и общество, необходимость решения которой не вызывает сомнений. При этом с каждым годом наблюдается ее усиление.

Экспоненциальное увеличение числа пользователей Интернета повышает критичность и делает еще более ощутимыми последствия в случае отказов техники. Количество вредоносных программ для мобильных устройств растет вместе с числом их владельцев. При этом подавляющее большинство пользователей не использует специализированное программное обеспечение для защиты своих смартфонов, планшетов и других персональных устройств с выходом в Интернет. Это успешно эксплуатируются хакерами, каждый день приводя к увеличению количества успешных атак, нацеленных на абонентские устройства [1].

Пренебрежение соображениями безопасности при использовании Интернет-ресурсов и социальных сетей ведет к повышенному риску для неприкосновенности частной жизни, модификации или уничтожению общедоступных персональных данных, разглашению персональных данных ограниченного доступа, а также их экстерриториальной доступности для преступных сообществ или разведывательных служб других государств.

В то время, как количество онлайн-устройств увеличивается, а большинство пользователей продолжают игнорировать меры «цифровой гигиены» в отношении себя и принадлежащих им устройств, внедрение концепции «Интернета вещей» только усиливает необходимость решения задач кибербезопасности. Так, если традиционные электронные устройства, такие как персональные компьютеры и ноутбуки, имеют возможность установки и обновления антивирусного программного обеспечения, а также различных авторских систем защиты информации [2], то пользователи «Интернета вещей» часто даже не знают, как обезопасить их функционирование, поскольку такие устройства пока создаются без учета технологических рисков. При этом, пользователи могут невольно становиться участниками масштабных

целенаправленных атак, построенных на целых сетях зараженных устройств (так называемых «ботнет сетях»), хотя непосредственно данным пользователей ничего не угрожает. При этом, слепое применение комплексных систем защиты информации не является *панацеей* от всех угроз [3], поскольку ни одна из реально применяемых систем защиты не обладает 100%-ной эффективностью [4].

Как результат, низкий уровень компьютерной грамотности конечных пользователей при отсутствии базовых знаний об основных методах компьютерных атак (таких, как фишинг, поддельные Интернет-магазины и т. д.), приводит к тысячам случаев, когда граждане становятся как жертвами, так и орудиями противоправного использования информационных компьютерных технологий (ИКТ). Одна из причин этого, связанная в том числе с нехваткой специалистов ИБ, в значительной мере вызвана непропорциональным распределением образовательных грантов по ИБ для системы подготовки специалистов в области ИКТ, а также их низкой востребованностью со стороны работодателей.

Таким образом можно утверждать, что недостаточное количество специалистов ИБ оказывает существенное негативное влияние на организацию ИБ.

### **Проблема образовательной национальной кибербезопасности Республики Казахстан**

Казахстан, как страна, которая пока в значительной мере заимствует передовые информационные технологии, и что важно – в том числе технологии обеспечение кибербезопасности, в любой момент может столкнуться с ситуацией, когда окажется в качестве объекта экспериментов или реальных атак на критически важные объекты информационно-коммуникационной инфраструктуры. При этом вектор атаки может исходить как из самой страны, так и от других стран, а в качестве нарушителей будут выступать отдельные лица или же целые преступные организации. Необходимо уточнить, что под кибербезопасностью в целом понимается обеспечение состояния защищенности средств телекоммуникаций (средства связи), электронных информационных ресурсов, информационных систем и информационно-коммуникационной инфраструктуры от внешних и внутренних угроз.

Однако существующая казахстанская модель высшего и послевузовского образования в области ИКТ, включающая ИБ, не в полной мере соответствует современным потребностям и тенденциям развития общества. Она не дает достаточных знаний для студентов как бакалавриата, так и магистратуры, а также ограничена по числу дисциплин и специальностей ИБ; при этом некоторые из них попросту отсутствуют в образовательных и профессиональных стандартах. Кроме того, не всегда прослеживается необходимая взаимосвязь профессиональных стандартов с типовыми и рабочими учебными планами высших и средне-специальных учебных заведений. Все это подчеркивает размах проблемы национальной кибербезопасности Республики Казахстан, а точнее отсутствие возможностей для реализации такой безопасности в требуемом объеме.

Тем не менее в последние годы наблюдаются положительные тенденции развития страны в сторону популяризации кибербезопасности: стали продвигаться продукты обеспечения корпоративной ИБ, резко выросло количество специалистов по ИБ (хотя имеющих и не вполне достаточный уровень), почти каждая IT-компания имеет предложения на рынке услуги по внедрению собственных и сторонних продуктов ИБ. Изучение данной ситуации показало, что модный тренд и популяризация ИБ вызвана проведенными реформами и наросшей необходимостью обеспечения ИБ, так как государственные и частные компании столкнулись с

реальными угрозами и существенными ущербами от их реализации. Естественно это дает некоторые плюсы для национальной кибербезопасности, поскольку руководители государственных предприятий и собственники бизнеса стали не без оснований чувствовать себя более защищенными.

Впрочем, что крайне важно, отсутствие системного подхода в этом вопросе на национальном уровне не только существенно снижает итоговую эффективность, но и может даже принести больше отрицательных, чем положительные, последствия. Так, например, одномоментное появление достаточного количества высококвалифицированных специалистов по ИБ представляется крайне сомнительным, что в конечном итоге приведет к существенному снижению эффективности любых высокотехнологичных и дорогостоящих средств обеспечения ИБ, поскольку их единовременное внедрение и настройка будет осуществляться не профессионалами. Также ускоренный рост числа псевдо-специалистов, создаваемых для нужд каждой отдельно взятой организации лишь при возникновении инцидентов ИБ, в условиях отсутствия единой системы обучения специалистов, приведет к различающемуся пониманию основ ИБ, не возможности обмена опытом, их раскоординированной работе, торможению развития научных направлений области и другим негативным эффектам.

### **Гуманитарные аспекты решения проблем**

Решением проблемы национальной кибербезопасности с одной стороны может стать введение организационно-правовых мер на государственном уровне, а с другой стороны правильное выстраивание целостной системы подготовки специалистов. Для этого на данный момент в Казахстане принимается концепция кибербезопасности (называемая «Кибершит Казахстана» [5]). В ней, в том числе, уделяется особое внимание подготовке работников, ответственных за состояние ИБ в государственных органах, а также отводится существенная роль высшим учебным заведениям Казахстана для реализации образовательных и исследовательских задач в сфере ИБ. Также, для наращивания казахстанского потенциала в сфере научной, научно-технической и образовательной деятельности подчеркивается необходимость сосредоточивания на создании научно-исследовательских и опытно-конструкторских лабораторий, обеспечения тесной связи учебного процесса с производственной деятельностью предприятий электронной промышленности, приведения учебных программ в соответствие с отраслевыми профессиональными стандартами и современным уровнем развития технологий. Так же отмечается, что приоритет должен отдаваться исследованиям и развитию собственной школы прикладной математики, криптологии, разработкам по программируемым логическим интегральным схемам (ПЛИС) и созданию защищенных систем передачи, обработки и хранения информации [6].

Уже сейчас идет создание центров подготовки кадров, открытие в ВУЗах специальностей из области ИБ; появляются компании с курсами по международным стандартам ИБ и много другое. Тем не менее, несмотря на положительные стороны процесса, все это открывает ряд новых растущих проблем. Так, например, каждое учебное заведение старается сделать собственную «уникальную» программу обучения, используя для этого разные подходы и отличные терминологии, что в итоге выливается в противоборство ведомств на предмет того, чьи специалисты *лучше*. Как результат, все это негативно отражается на развитии национальной кибербезопасности.

Организации и общество в целом нуждаются в квалифицированных кадрах, которые будут выстраивать процессы национальной кибербезопасности, в том числе связанные с подготовкой самих кадров и развитием научной базы. Для осуществления этого необходимо уже на начальном этапе создавать центры подготовки специалистов и обмениваться опытом между образовательными учреждениями, подготавливать общие стандарты и методические пособия, а также в дальнейшем совершенствовать научно-практическую базу.

### **Выводы**

Анализ состояния и тенденций развития проблем кибербезопасности Республики Казахстан позволяет говорить о неудовлетворительности текущих процессов их разрешения и требуют привлечения мировой общественности к проблемам педагогики в сфере ИБ. Существенное улучшение ситуации может быть достигнуто следующим образом. Во-первых, системный и целостный подход к формированию учебно-практических научных центров даст возможность готовить специалистов по ИБ, обладающих при этом высоким уровнем универсальности. Во-вторых, использование единых международных стандартов с применением зарубежного опыта позволит осуществлять слаженные с другими странами мероприятия в области кибербезопасности. И, в-третьих, исследование и разработка собственных актуальных направлений ИБ учтет специфику Республики Казахстан, ее положительные и отрицательные стороны. Все это безусловно переведет национальную кибербезопасность на качественно новый уровень.

### **Литература**

1. Буйневич М.В. Проблемные вопросы и тенденции обеспечения ИБ в сфере телекоммуникаций // Защита информации. Инсайд. 2017. № 1(73). С. 49-55.
2. Покусов В.В. Система анализа действий на электронно-вычислительных устройствах и защиты информации. Патент Республики Казахстан на полезную модель № 1409 от 15.03.2016.
3. Израйлов К.Е., Покусов В.В. Актуальные вопросы взаимодействия элементов комплексных систем защиты информации // Актуальные проблемы инфотелекоммуникаций в науке и образовании (АПИНО 2017): сборник научных статей VI Международной научно-технической и научно-методической конференции. 2017. С. 255-260.
4. Буйневич М.В., Покусов В.В., Ярошенко А.Ю., Хорошенко С.В. Категориальный подход в приложении к синтезу архитектуры интегрированной системы обеспечения безопасности информации // Проблемы управления рисками в техносфере. 2017. № 4 (44). С. 95-102
5. Концепция кибербезопасности ("Кибершит Казахстана") // Постановление Правительства Республики Казахстан от 30 июня 2017 года № 407
6. Буйневич М.В., Васильева И.Н., Воробьев Т.М., Гниденко И.Г., Егорова И.В., Еникеева Л.А., Зельман С.Г., Израйлов К.Е., Ишанханов С.Р., Куватов В.И., Локтионов О.В., Малахова П.А., Мердина О.Д., Полегенько А.М., Попов М.А., Примакин А.И., Семенова Т.Г., Семенова С.О., Соловьев А.И., Соловьев С.А., Солодянников А.В., Соколов Р.В., Соколовская С.А., Стельмашонок В.Л., Стельмашонок Е.В., Торосян Е.К., Федоров Д.Ю., Чернокнижный Г.М. Защита информации в компьютерных системах: монография / СПб.: СПГУЭУ, 2017. 163 с.

## СОДЕРЖАНИЕ

Arslanov M.Z., Mustafin S.A. Naizabayeva L.K.	CREDIT RISK ASSESSMENT AND MODELING METHODS USING RECOGNITION ALGORITHMS BASED ON CALCULATION OF ESTIMATES	4
Kudaykulov A.K., Tashev A.A., Arshidinova M.T., Begaliyeva K.B.	ENERGY METHOD RESEARCH THERMO-STRESS- DEFORMATION STATE IN A ROD OF VARIABLE CROSS SECTION	8
Амирханова Г.А., Дуйсенбаева А.Ж.	ЖАЗЫҚ КРИСТАЛДЫҢ ЕКІӨЛШЕМДІ МОДЕЛІ	14
Байбеков С.Н., Алтынбек С.А., Елеусинова А.У., Тургинбаева А.С.	ҰЛЫ ДАЛАНЫҢ БАЙЫРҒЫ КӨНЕ ЕСЕПТЕРІ – ҚАЗАҚ ХАЛҚЫНЫҢ ҚАСИЕТТІ РУХАНИ МҰРАСЫ	17
Капалова Н., Хаумен А., Абишева А.	КРИПТОГРАФИЯЛЫҚ ЖҮЙЕЛЕРДЕ КІЛТТЕРДІ БАСҚАРУ	20
Қожахмет Б., Калижанова А.У.	ЖҮЙКЕ ЖҮЙЕСІНІҢ ҚҰРЫЛЫМЫ МЕН ҚАСИЕТТЕРІН ИМИТАЦИЯЛАЙТЫН НЕЙРОНДЫ ЖЕЛІЛЕР	24
Мамырбаев О.Ж., Мекебаев Н.О., Тұрдалыұлы М.	СӨЙЛЕУДІ ТАНУДА MFCC ЖӘНЕ DTW АЛГОРИТМДЕРІНІҢ ЕРЕКШЕЛІГІ	28
Черикбаева Л.Ш., Калыбек уулы Б.	АРАҚАШЫҚТЫҚ МЕТРИКАСЫ АУЫСПАЛЫ БОЛЫП КЕЛЕТІН АЛГОРИТМДЕР	34
Хомпыш А.	МОДУЛЬ БОЙЫНША ДӘРЕЖЕГЕ ШЫҒАРУ ОПЕРАЦИЯСЫН ҚОЛДАНУ АРҚЫЛЫ АҚПАРАТТЫ ШИФРЛАУ ӘДІСІ	38
Абдилдаева А.А., Галиева Ф.М., Базарбекова М.О., Даулетбек Е.Т.	ОБ ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЕ И МАТЕМАТИЧЕСКОЙ МОДЕЛИ ЭЛЕКТРОЭНЕРГЕТИЧЕСКИХ СИСТЕМ	41
Абдилдаева А.А., Дрозденко А.А., Коплык И.В., Маринич Т.А.	АНАЛИЗ СОВРЕМЕННЫХ МЕТОДОВ МОДЕЛИРОВАНИЯ ПОТРЕБЛЕНИЯ ЭЛЕКТРИЧЕСКОЙ ЭНЕРГИИ	47

Абешев К.Ш., Бадаев С.А., Кабылжанова Д.К., Калмурзаев Б.С., Серикжан Р., Сыдыкбек А.Ж.	АЛГЕБРАИЧЕСКИЕ СВОЙСТВА СТЕПЕНЕЙ ПОЗИТИВНЫХ ПРЕДПОРЯДКОВ	50
Амиргалиев Е.Н., Иржанова А.А.	ПОТЕНЦИАЛ ВОЗОБНОВЛЯЕМОЙ ЭНЕРГЕТИКИ И ОБЗОР ИМЕЮЩИХСЯ ЭНЕРГО И ТЕПЛО СБЕРЕГАЮЩИХ ТЕХНОЛОГИЙ В КАЗАХСТАНЕ	54
Асанова А.Т., Бакирова Э.Б., Темешева С.М.	ОБ ОДНОМ ПОДХОДЕ К РЕШЕНИЮ ДВУХТОЧЕЧНОЙ ЗАДАЧИ С ПАРАМЕТРОМ ДЛЯ СИСТЕМЫ ИНТЕГРО-ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ	58
Байшемиров Ж.Д. , Жанбырбаев А.Б. , Бердышева Д.	ПРОБЛЕМЫ ПРОБЛЕМ УСРЕДНЕНИЯ НЕРАВНОВЕСНОГО ТЕЧЕНИЯ В ПОРИСТЫХ СРЕДАХ	68
Бегимбаева Е.Е.	КОНФЛИКТНЫЕ СИТУАЦИИ В СИСТЕМЕ ЗАЩИЩЕННОГО ТРАНСГРАНИЧНОГО ОБМЕНА	75
Бердышев А.С., Блиева Д.Н.	ПРИМЕНЕНИЕ ПРЕОБРАЗОВАНИЯ ЛАПЛАСА ДЛЯ СВЕДЕНИЯ УРАВНЕНИЙ ПОРОУПРУГОСТИ К ЭЛИПТИЧЕСКОЙ СИСТЕМЕ УРАВНЕНИЙ	80
Блейхер Г.А.,, Кривобоков В.П., Байдикова Н.В.	ИССЛЕДОВАНИЕ ПРОЦЕССОВ ТЕПЛОМАССОПЕРЕНОСА ПРИ ЭЛЕКТРОННО- ЛУЧЕВОЙ СВАРКЕ	83
Варенников А.В.	ФОРМИРОВАНИЕ СИСТЕМЫ РАБОЧИХ ОСНОВАНИЙ, ДЛЯ АЛГОРИТМА ШИФРОВАНИЯ НА БАЗЕ НЕПОЗИЦИОННЫХ ПОЛИНОМИАЛЬНЫХ СИСТЕМ СЧИСЛЕНИЯ	88
Вуйчик В., Калижанова А.У., Кисала П., Кашаганова Г.Б., Амиргалиева С.Н., Картбаев Т.С., Оразбеков Ж.	СПЕКТРАЛЬНЫЕ ХАРАКТЕРИСТИКИ ВОЛОКОННЫХ РЕШЕТОК БРЭГГА	93
Есеркепова И.Б., Юничева Н.Р., Нурсеитов Д.Б., Бостанбеков К., Гаврук С.В.	К ВОПРОСУ МОДЕЛИРОВАНИЯ БУДУЩЕГО КЛИМАТА ЦЕНТРАЛЬНОЙ АЗИИ В РАМКАХ МЕЖДУНАРОДНОГО ПРОЕКТА CORDEX (COORDINATED REGIONAL CLIMATE DOWNSCALING EXPERIMENT)	98

Жомартова Л.М., Рахимова Д.Р.	ИССЛЕДОВАНИЕ РЕКУРЕНТНЫХ НЕЙРОННЫХ СЕТЕЙ ДЛЯ МОДЕЛИРОВАНИЯ ЕСТЕСТВЕННЫХ ЯЗЫКОВ ФЛЕКТИВНЫХ КЛАССОВ	103
Искакова М.Т., Калижанова А.У. Айткулов Ж.С., Тогжанова Г.О.	РЕАЛИЗАЦИЯ КАРКАСНОЙ ТЕХНОЛОГИИ В AUTODESK INVENTOR	107
Калимолдаев М.Н., Бияшев Р.Г., Рог О.А.	МНОГОУРОВНЕВОЕ ОПРЕДЕЛЕНИЕ МОДЕЛИ ТИПИЗИРОВАННОГО АТТРИБУТНОГО РАЗГРАНИЧЕНИЯ ДОСТУПА	111
Калимолдаев М.Н., Магзом М.М., Косынбай Е.Б.	ИНСТРУМЕНТЫ ДЛЯ ОБРАБОТКИ ПОТОКОВЫХ ДАННЫХ	120
Калимолдаев М.Н., Тынымбаев С.Т., Магзом М.М.	О ПРЕДВАРИТЕЛЬНОМ РАСЧЕТЕ ПАРАМЕТРОВ НЕПОЗИЦИОННОЙ СИСТЕМЫ СЧИСЛЕНИЯ	128
Картбаев А.Ж.	АНАЛИЗ МЕТОДОВ АНГЛО-КАЗАХСКОГО ВЫРАВНИВАНИЯ СЛОВ	133
Кубеков Б.С., Утегенова А.У., Науменко В.В., Аленова Р.А.	ОНТОЛОГИЧЕСКИЙ ПОДХОД К СЕМАНТИЧЕСКОМУ МОДЕЛИРОВАНИЮ ОБРАЗОВАТЕЛЬНЫХ ПРОГРАММ В ВЫСШЕМ ОБРАЗОВАНИИ	136
Литвиненко Н.Г., Литвиненко А.Г., Шаяхметова А.С., Мамырбаев О.Ж.	К ВОПРОСУ О КЛАССИФИКАЦИИ ТИПОВ СВИДЕТЕЛЬСТВ В БАЙЕСОВСКИХ СЕТЯХ	143
Мазаков Т.Ж., Айпанов Ш.А., Тусупова С.А., Байрбекова Г.С., Зиятбекова Г.З., Мазакова А.Т.	БИОМЕТРИЧЕСКАЯ ИДЕНТИФИКАЦИЯ ЧЕЛОВЕКА ПО ЛИЦУ	152
Мазаков Т.Ж., Исимов Н.Т., Зиятбекова Г.З., Жолмагаметова Б.Р., Джомартова Д.Т., Ыдырышбаева М.Б.	ОБРАБОТКА СИГНАЛОВ ЭКГ НА МИКРОПРОЦЕССОРАХ	159

Макаренко Н.Г., Рыбинцев А.С., Князева И.С., Волобуев Д.М., Каримова Л.М., Пак И.Т.	МЕТРИКА ФИШЕРА – РАО НА СФЕРЕ ДЛЯ ТОПОЛОГИЧЕСКОГО РАСПОЗНАВАНИЯ ТЕКСТУР ЦИФРОВЫХ ИЗОБРАЖЕНИЙ	163
Мамырбаев О.Ж., Сундетов Т., Жакупбеков Т.Е.	СИСТЕМА УПРАВЛЕНИЯ БИОНИЧЕСКОЙ РУКОЙ С ПОМОЩЬЮ ЭМГ ДАТЧИКОВ	168
Марденов Е.М.	АНАЛИЗ ЗАЩИТЫ ДОСТУПА СВЯЗИ РОБОТОТЕХНИЧЕСКИХ СИСТЕМ	173
Мусабаев Р.Р., Красовицкий А.М., Уалиева И.М., Касымжанов Б.К., Ибраева В.М., Мейрамбеккызы Ж., Тасболатулы Н., Мукашев А.Ш.	СИСТЕМА МОНИТОРИНГА И ПРОВЕРКИ ДОСТОВЕРНОСТИ ПУБЛИКАЦИЙ В СМИ	178
Мусабаев Р.Р., Красовицкий А.М.	АЛГОРИТМ ИЕРАРХИЧЕСКОЙ КЛАСТЕРИЗАЦИИ НА ОСНОВЕ ЭНЕРГЕТИЧЕСКОГО КРИТЕРИЯ	185
Мусаев М.С., Рахимова Д.Р.	ОСОБЕННОСТИ ОБРАБОТКИ ТЕКСТОВ ЕСТЕСТВЕННОГО ЯЗЫКА В РАЗРАБОТКЕ ИНТЕЛЛЕКТУАЛЬНОЙ ПОИСКОВОЙ СИСТЕМЫ	189
Мустафин С.А. Сериков Б.Б.	МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ДИНАМИЧЕСКОЙ СИСТЕМЫ	193
Найзабаева Л., Нуржанов Ч., Pidlisnyuk V., Сатымбеков М.	МАТЕМАТИЧЕСКАЯ МОДЕЛЬ НАКОПЛЕНИЯ ТЯЖЕЛЫХ МЕТАЛЛОВ ИЗ ЗАГРЯЗНЕННОЙ ПОЧВЫ В ВЕГЕТАТИВНЫХ ОРГАНАХ РАСТЕНИЙ	196
Нуртазин А.Т., Хисамиев З.Г., Досанбай С.П.	ЛЮБОЙ ФОРСИНГ-ТИП РЕАЛИЗУЕТСЯ В НЕКОТОРОЙ ФОРСИНГ-МОДЕЛИ И СЕМАНТИЧЕСКИЙ КРИТЕРИЙ E-ФОРМУЛ	205
Нысанбаева С.Т., Капалова Н.А., Дюсенбаев Д.С., Алгазы К.Т.	ИССЛЕДОВАНИЯ СТАТИСТИЧЕСКИХ СВОЙСТВ РАЗРАБОТАННОГО ГЕНЕРАТОРА ПСЕВДОСЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ	210
Нысанбаева С.Е., Нюсупов А.Т.	ДОСТИЖЕНИЕ КОНСЕНСУСА В СОВРЕМЕННЫХ ДЕЦЕНТРАЛИЗОВАННЫХ СИСТЕМАХ	217
Нысанбаева С.Е., Усатова О.А.	КРИПТОГРАФИЧЕСКАЯ ЗАЩИТА В АВТОМАТИЗИРОВАННЫХ СИСТЕМАХ	220



Пак А.А., Джаксылыкова А.Б.	КОНЦЕПЦИЯ ЦИФРОВОЙ ИНТЕЛЛЕКТУАЛЬНОЙ МЕДИЦИНЫ	223
Посыпкин М.А., Ахметжанов М.А., Намазбаев Б.Д.	МЕТОДЫ АППРОКСИМАЦИИ РЕШЕНИЯ СИСТЕМ УРАВНЕНИЙ И НЕРАВЕНСТВ	228
Самбетбаева М.А., Тусупов Д.А., Федотов А.М., Батырханов А.Г.	МОДЕЛИ И МЕТОДЫ СОЗДАНИЯ ИНФОРМАЦИОННЫХ СИСТЕМ ПОДДЕРЖКИ НАУЧНО-ОБРАЗОВАТЕЛЬНОЙ ДЕЯТЕЛЬНОСТИ, ИНТЕГРИРОВАННЫХ В ОТКРЫТОЕ СЕМАНТИЧЕСКОЕ ПРОСТРАНСТВО	235
Самигулина Г.А., Лукманова Ж.С.	КОГНИТИВНАЯ SMART – ТЕХНОЛОГИЯ ДЛЯ ПЕРСОНАЛИЗИРОВАННОГО ДИСТАНЦИОННОГО ОБУЧЕНИЯ ЛЮДЕЙ С ОСЛАБЛЕННЫМ ЗРЕНИЕМ ИНЖЕНЕРНЫМ СПЕЦИАЛЬНОСТЯМ В НЕФТЕГАЗОВОЙ ОТРАСЛИ	244
Самигулина Г.А., Масимканова Ж.А.	ДИАГНОСТИКА ПРОМЫШЛЕННОГО ОБОРУДОВАНИЯ НА ОСНОВЕ МОДИФИЦИРОВАННОГО АЛГОРИТМА РОЯ ЧАСТИЦ И ПОДХОДА ИСКУССТВЕННЫХ ИММУННЫХ СИСТЕМ	248
Самигулина З.И., Алдибекова А.Н.	ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА УПРАВЛЕНИЯ СЛОЖНЫМИ ОБЪЕКТАМИ НА ОСНОВЕ АЛГОРИТМА КЛОНАЛЬНОГО ОТБОРА	253
Суйменбаев Б.Т., Трушляков В.И., Ермолдина Г.Т., Бапыев А.М.	К ВОПРОСУ СОЗДАНИЯ РН С УЛУЧШЕННЫМИ ЭКОЛОГИЧЕСКИМИ ХАРАКТЕРИСТИКАМИ ЗА СЧЕТ ИЗВЛЕЧЕНИЯ ЭНЕРГЕТИЧЕСКИХ РЕСУРСОВ, НАХОДЯЩИХСЯ В НЕВЫРАБАТЫВАЕМЫХ ОСТАТКАХ ТОКСИЧНОГО КОМПОНЕНТА ТОПЛИВА В БАКАХ И МАГИСТРАЛЯХ ОС, ЗА СЧЁТ ИХ ПРИНУДИТЕЛЬНОГО ИСПАРЕНИЯ	258
Сулейменов И.Э., Пак И.Т., Бакиров А.С.	ПРОЛЕГОМЕНЫ К ЧЕТВЕРТОМУ ЗАКОНУ ДИАЛЕКТИКИ	266
Терехов А.Г.	ОСНОВНЫЕ НАПРАВЛЕНИЯ МОНИТОРИНГА ТЕРРИТОРИИ СИНЬЦЗЯН-УЙГУРСКОГО АВТОНОМНОГО РАЙОНА КНР В ПРЕДЕЛАХ БАССЕЙНОВ ТРАНСГРАНИЧНЫХ РЕК И ЗАВИСИМЫХ ТЕРРИТОРИЙ НА БАЗЕ СПУТНИКОВОЙ ИНФОРМАЦИИ	276

Тусупова М.Д., Мурзин Ф.А., Еримбетова А.С.	ТЕХНОЛОГИИ ОПРЕДЕЛЕНИЯ ГРАММАТИЧЕСКИХ КАТЕГОРИЙ КАЗАХСКОГО ЯЗЫКА	286
Уалиева И.М., Красовицкий А.М., Мусабаев Р.Р., Мейрамбеккызы Ж.	ЧИСЛЕННАЯ ОЦЕНКА ГЕНЕРАЛИЗАЦИИ В ТЕКСТАХ СМИ НА РУССКОМ ЯЗЫКЕ С УЧЕТОМ ИСТОЧНИКОВ ТОНАЛЬНОСТИ: ЛЕКСИЧЕСКИЙ ПОДХОД	289
Утепбергенов И.Т., Тойбаева Ш.Д., Буранбаева А.И., Коржаспаев А.Е.	ПОДХОД К СТАТИСТИЧЕСКОМУ ИССЛЕДОВАНИЮ ПРОБЛЕМ ИНФОРМАЦИОННОГО ОБЕСПЕЧЕНИЯ ИННОВАЦИОННОЙ ДЕЯТЕЛЬНОСТИ ПРЕДПРИЯТИЙ В КАЗАХСТАНЕ	298
Хайрова Н.Ф., Мамырбаев О.Ж., Избасаров Е.Ж., Мухсина К.Ж.	ФОРМАЛЬНАЯ МОДЕЛЬ ОЦЕНИВАНИЯ КАЧЕСТВА ЭКСТРАКЦИИ И ИДЕНТИФИКАЦИИ ЗНАНИЙ ИЗ СЛАБОСТРУКТУРИРОВАННОЙ ТЕКСТОВОЙ ИНФОРМАЦИИ	306
Шокишалов Ж.М.	ПРИМЕНЕНИЕ ТЕХНОЛОГИИ ОТСЛЕЖИВАНИЯ ГЛАЗ В ИССЛЕДОВАТЕЛЬКИХ РАБОТАХ	310
Яворский В.В., Ахмедиярова А.Т., Байдикова Н.В., Чванова А.О.	МОДЕЛИРОВАНИЕ ТРАНСПОРТНОЙ ИНФРАСТРУКТУРЫ ГОРОДСКОГО ОБЩЕСТВЕННОГО ТРАНСПОРТА	314
Яворский В.В., Проценко А.В., Байдикова Н.В.	СРЕДСТВА СОВЕРШЕНСТВОВАНИЯ БЮДЖЕТИРОВАНИЯ ГОРНОДОБЫВАЮЩЕГО ПРЕДПРИЯТИЯ	321
Яворский В.В., Утепбергенов И.Т., Байдикова Н.В., Чванова А.О.	РАЗРАБОТКА МЕТОДА ПРЕДСТАВЛЕНИЯ ТРАНСПОРТНОЙ СИСТЕМЫ НА ОСНОВЕ ТЕОРИИ ГИПЕРСЕТЕЙ ДЛЯ ПРОЕКТИРОВАНИЯ И УПРАВЛЕНИЯ	326
Яворский В.В., Чванова А.О., Байдикова Н.В.	МОДЕЛИРОВАНИЕ ПРОЦЕССОВ ДЛЯ ИНТЕЛЛЕКТУАЛЬНОЙ ТРАНСПОРТНОЙ СИСТЕМЫ ГОРОДА	334
Покусов Виктор Владимирович	ГУМАНИТАРНЫЕ АСПЕКТЫ РЕШЕНИЯ ПРОБЛЕМ КИБЕРБЕЗОПАСНОСТИ РЕСПУБЛИКИ КАЗАХСТАН	336
<b>Содержание</b>		341

МАТЕРИАЛЫ  
научной конференции  
ИИВТ МОН РК  
«Современные проблемы информатики и  
вычислительных технологий»  
2 - 5 июля 2018 года  
  
(Сборник статей)

Подписано в печать 20.06.2018г. Формат А4.  
Печать цифровая. Бумага офсетная. Усл. печ. л. 18,32.  
Тираж 300 экз. Заказ № 01265.  
Отпечатано в типографии НЦ ГНТЭ.  
Алматы, ул. Богенбай батыра, 221.