

# **СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ**

**Научный журнал Российской академии наук  
(издается под руководством Отделения нанотехнологий  
и информационных технологий РАН)**

Издается с 1989 года  
Журнал выходит ежеквартально

## **Учредители:**

**Российская академия наук  
Федеральный исследовательский центр  
«Информатика и управление» Российской академии наук**

## **РЕДАКЦИОННЫЙ СОВЕТ**

академик РАН И. А. Соколов — председатель Редакционного совета  
академик РАН Г. И. Савин  
академик РАН А. Л. Стемпковский  
член-корреспондент РАН Ю. Б. Зубарев  
профессор Ш. Долев (S. Dolev, Beer-Sheva, Israel)  
профессор Ю. Кабанов (Yu. Kabanov, Besancon, France)  
профессор М. Никулин (M. Nikulin, Bordeaux, France)  
профессор В. Ротарь (V. Rotar, San-Diego, USA)  
профессор М. Финкельштейн (M. Finkelstein, Rostok, Germany)

## **РЕДАКЦИОННАЯ КОЛЛЕГИЯ**

академик РАН И. А. Соколов — главный редактор  
профессор, д.ф.-м.н. С. Я. Шоргин — заместитель главного редактора  
д.т.н. В. Н. Захаров проф., д.ф.-м.н. В. Ю. Королев  
проф., д.т.н. В. Д. Ильин проф., д.г.-м.н. Р. Б. Сейфуль-Мулюков  
проф., д.ф.-м.н. Л. А. Калининченко проф., д.т.н. И. Н. Синицын  
д.т.н. В. А. Козмидади к.т.н. А. В. Филин  
проф., д.т.н. К. К. Колин к.ф.-м.н. С. А. Христочевский

## **Редакция**

профессор, д.г.-м.н. Р. Б. Сейфуль-Мулюков  
к.ф.-м.н. Е. Н. Арутюнов  
С. Н. Стригина (ответственный секретарь)

© Федеральный исследовательский центр «Информатика  
и управление» Российской академии наук, 2015

Журнал входит в систему Российского индекса научного цитирования (РИНЦ):

**[http://elibrary.ru/title\\_about.asp?id=28980](http://elibrary.ru/title_about.asp?id=28980)**

Журнал включен в базу данных CrossRef (систему DOI — Digital Object Identifier),  
в базу данных Ulrich's periodicals directory  
и в информационную систему «Общероссийский математический портал Math-Net.Ru»

Журнал реферируется в «Реферативном журнале» ВИНТИ  
и в системе Google Scholar

Журнал «Системы и средства информатики»  
включен в формируемый Минобрнауки России Перечень рецензируемых научных изданий,  
в которых должны быть опубликованы основные научные результаты диссертаций  
на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук

**<http://www.ipiran.ru/journal/collected>**

# СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ

Том 25 № 4 Год 2015

## СОДЕРЖАНИЕ

Аналитическое моделирование нормальных процессов в стохастических системах со сложными дробно-рациональными нелинейностями <b>И. Н. Сеницын, В. И. Сеницын, Э. Р. Корепанов</b>	<b>4</b>
Анализ энергоэффективности вычислительного комплекса, моделируемого с помощью системы обслуживания с пороговым управлением и интенсивностями, зависящими от времени <b>Р. В. Разумчик, А. И. Зейфман, А. В. Коротышева, Я. А. Сатин</b>	<b>19</b>
Приближенная оптимизация стратегии распределения вычислительных ресурсов на примере системы андеррайтинга <b>М. Г. Коновалов, Р. В. Разумчик</b>	<b>31</b>
Метрическая классификация временных рядов с выравниванием относительно центроидов классов <b>А. В. Гончаров, М. С. Попова, В. В. Стрижов</b>	<b>52</b>
Архитектура стенда для экспериментального исследования моделей, алгоритмов и решений по обеспечению информационной безопасности в облачных вычислительных средах <b>А. А. Грушо, М. И. Забежайло, А. А. Зацаринный, А. В. Николаев, В. О. Писковский</b>	<b>65</b>
Аппаратно-программное моделирование и тестирование рекуррентного операционного устройства <b>Д. В. Хилько, Ю. А. Степченков, Ю. Г. Дьяченко, Ю. И. Шикунев, Н. В. Морозов</b>	<b>78</b>
Исследование особенностей сигналов миограммы <b>Т. В. Захарова, В. Ю. Королев, А. А. Щемирова</b>	<b>91</b>
Метод синхронизации сигналов магнитоэнцефалограмм и миограмм <b>Е. В. Щенявская, Т. В. Захарова</b>	<b>101</b>

# СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ

Том 25 № 4 Год 2015

## СОДЕРЖАНИЕ

Некоторые подходы к формированию обобщенной архитектуры информационных систем организаций — участников единого информационного пространства России <b>А. А. Зацаринный, Э. В. Киселев</b>	<b>114</b>
Основы применения онтологии и компьютерной лингвистики при проектировании перспективных автоматизированных информационных систем <b>И. И. Быстров, Б. В. Тарасов, А. А. Хорошилов, С. И. Радоманов</b>	<b>128</b>
Некоторые вопросы оценки надежности корпоративной информационной сети с учетом управления резервированием ее элементов <b>В. В. Бородин, Е. А. Растрелин</b>	<b>150</b>
Оценка стойкости кодового зашумления в задаче распределенного хранения данных <b>Ю. В. Косолапов, А. В. Поздняков</b>	<b>158</b>
Социальная информатика: российская научная школа и перспективные направления исследований <b>К. К. Колин</b>	<b>175</b>
Символьное моделирование и средства поддержки интеллектуальной деятельности <b>В. Д. Ильин</b>	<b>194</b>
Нормализация показателей в системе информационного мониторинга национальной безопасности <b>Г. В. Лукьянов, Д. А. Никишин, Г. Ф. Веревкин, В. В. Косарик</b>	<b>201</b>
Об авторах	<b>214</b>
Авторский указатель за 2015 г.	<b>218</b>
2015 Author Index	<b>225</b>
Правила подготовки рукописей статей	<b>231</b>
Requirements for manuscripts	<b>235</b>

## АНАЛИТИЧЕСКОЕ МОДЕЛИРОВАНИЕ НОРМАЛЬНЫХ ПРОЦЕССОВ В СТОХАСТИЧЕСКИХ СИСТЕМАХ СО СЛОЖНЫМИ ДРОБНО-РАЦИОНАЛЬНЫМИ НЕЛИНЕЙНОСТЯМИ\*

*И. Н. Сеницын<sup>1</sup>, В. И. Сеницын<sup>2</sup>, Э. Р. Корепанов<sup>3</sup>*

**Аннотация:** Рассматриваются методы аналитического моделирования нормальных стохастических процессов (СтП) в непрерывных стохастических системах (СтС) (в том числе на многообразиях — МСтС) с винеровскими и пуассоновскими шумами и со сложными дробно-рациональными нелинейностями (СДРН). Даны типовые представления скалярных и векторных СДРН. Получены уравнения методов нормальной аппроксимации (МНА) и статистической линеаризации (МСЛ). Представлено алгоритмическое обеспечение МНА (МСЛ) для СтС с СДРН. Приведены тестовые примеры. Рассмотрены возможные обобщения полученных результатов.

**Ключевые слова:** аналитическое моделирование; метод нормальной аппроксимации (МНА); метод статистической линеаризации (МСЛ); многочлены Эрмита; нормальный стохастический процесс; сложные дробно-рациональные нелинейности (СДРН); стохастическая система (СтС)

**DOI:** 10.14357/08696527150401

### 1 Введение

Рассмотрим методы аналитического моделирования нормальных СтП, приведенные в [1–5] для непрерывных СтС, в том числе и на многообразиях, с СДРН. В разд. 2 даны типовые представления дробно-рациональных функций (ДРФ) и соответствующих СДРН. Раздел 3 посвящен уравнениям МНА и МСЛ. Алгоритмическое обеспечение МНА (МСЛ) для СтС с СДРН представлено в разд. 4. Тестовые примеры приведены в разд. 5. В заключении даны выводы и рассмотрены некоторые обобщения полученных результатов.

---

\*Работа выполнена при поддержке РФФИ (проект 15-07-02244).

<sup>1</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, sinitsin@dol.ru

<sup>2</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, vsinitsin@ipiran.ru

<sup>3</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, ekorpanov@ipiran.ru

## 2 Дробно-рациональные функции и нелинейности

Как известно [6, 7], скалярной ДРФ над числовым полем  $\mathbb{P}$  называется любая функция  $R^{\text{ДРФ}}(y)$ , представимая в виде отношения двух полиномов  $P^+(y)$  и  $P^-(y)$ :

$$R^{\text{ДРФ}}(y) = \frac{P^-(y)}{P^+(y)} \quad (1)$$

с коэффициентами из поля  $\mathbb{P}$ . Для правильной ДРФ степень числителя  $P^-(y)$  не больше степени знаменателя  $P^+(y)$ .

В общем случае скалярную ДРФ можно однозначно представить в виде полинома  $P(y)$  и правильной ДРФ:

$$\frac{P^-(y)}{P^+(y)} = P(y) + \frac{P_1^-(y)}{P^+(y)}. \quad (2)$$

Следуя [6, 7], приведем основные свойства скалярных правильных ДРФ.

<sup>1</sup>0 Правильная ДРФ, знаменатель  $P^+(y)$  которой представим в виде произведения попарно взаимно простых множителей

$$P^+(y) = P_1^+(y)P_2^+(y) \cdots, \quad (3)$$

однозначно представима в виде суммы правильных ДРФ

$$\frac{P_1^-(y)}{P^+(y)} = \frac{P_{11}^-(y)}{P_1^+(y)} + \frac{P_{12}^-(y)}{P_2^+(y)} + \cdots. \quad (4)$$

<sup>2</sup>0 Пусть  $P^-(y)/P^+(y)$  — правильная ДРФ с коэффициентами из поля комплексных чисел  $\mathbb{C}$ , а знаменатель  $P^+(y)$  имеет вид:

$$P^+(y) = a_0(y - y_1)^{k_1}(y - y_2)^{k_2} \cdots (y - y_m)^{k_m}. \quad (5)$$

Тогда имеет место следующее единственное представление:

$$\frac{P_1^-(y)}{P^+(y)} = \sum_{\nu=1}^m \left[ \frac{A_{\nu 0}}{(y - y_\nu)^{k_\nu}} + \frac{A_{\nu 1}}{(y - y_\nu)^{k_\nu-1}} + \cdots + \frac{A_{\nu, k_\nu-1}}{y - y_\nu} \right], \quad (6)$$

где  $A_{\nu, \mu}$  — комплексные коэффициенты.

<sup>3</sup>0 Пусть  $P_1^-(y)/P^+(y)$  является правильной ДРФ с коэффициентами из поля рациональных чисел  $\mathbb{R}$ , а разложение знаменателя  $P^+(y)$  на неприводимые над полем  $\mathbb{R}$  множители имеет вид:

$$P^+(y) = a_0(y - y_1)^{k_1} \cdots (y - y_r)^{k_r} (y^2 + p_1y + q_1)^{l_1} \cdots (y^2 + p_1y + q_3)^{l_s}.$$

Тогда имеет место следующее единственное представление правильной ДРФ:

$$\frac{P_1^-(y)}{P^+(y)} = \sum_{\nu=1}^h \sum_{\mu=1}^{k_\nu} \frac{A_{\nu\mu}}{(y - y_\nu)^\mu} + \sum_{\nu=1}^s \sum_{\mu=1}^{l_\nu} \frac{B_{\nu\mu}y + C_{\nu\mu}}{(y^2 + p_\nu y + q_\nu)^\mu},$$

где  $A_{\nu\mu}$ ,  $B_{\nu\mu}$  и  $C_{\nu\mu}$  — вещественные коэффициенты.

4<sup>0</sup> Дифференцирование ДРФ (1) проводится по формуле:

$$\frac{dR^{\text{ДРФ}}(y)}{dy} = \frac{(P^-(y))'P^+(y) - (P^+(y))'P^-(y)}{(P^+(y))^2}.$$

5<sup>0</sup> Неопределенный интеграл от ДРФ всегда выражается через элементарные функции, так что

$$\begin{aligned} \int \frac{P^-(y)}{P^+(y)} dy &= \int P(y) dy + \int \frac{P_1^-(y)}{P^+(y)} dy = \\ &= \int P(y) dy + \sum_{\nu=1}^h \sum_{\mu=1}^{k_\nu} A_{\nu\mu} \int \frac{dy}{(y - y_\nu)^\mu} + \sum_{\nu=1}^s \sum_{\mu=1}^{l_\nu} \int \frac{B_{\nu\mu}y + C_{\nu\mu}}{(y^2 + p_\nu y + q_\nu)^\mu} dy. \end{aligned}$$

При этом интеграл от полинома  $P(y)$  всегда выражается в конечном виде. Интегралы первой суммы равны

$$\int \frac{dy}{(y - y_\nu)^\mu} = \frac{1}{1 - \mu} (y - y_\nu)^{1-\mu} + C \quad (\mu > 1).$$

Далее

$$\int \frac{By + C}{(y^2 + py + q)^\mu} = \frac{B}{2(1 - \mu)(y^2 + py + q)^{\mu-1}} + \frac{2C - Bp}{2} I_\mu(\xi).$$

Здесь  $I_\mu(\xi)$  вычисляется по рекуррентной формуле:

$$I_\mu = \frac{\xi(\xi^2 + a^2)^{1-\mu}}{2a^2(\mu - 1)} + \frac{2\mu - 3}{a^2(2\mu - 2)} I_{\mu-1}(\xi),$$

где  $\xi = y + p/2$ ;  $a^2 = q - p^2/4$ .

6<sup>0</sup> Имеет место следующее правило Остроградского для выделения рациональной части  $R_1(y)$  интеграла

$$\int \frac{P^-(y)}{P^+(y)} dy = R_1(y) + R_2(y),$$

где  $R_2(y)$  — трансцендентная часть. Если предположить, что знаменатель  $P^+(y)$  допускает представление вида

$$P^+(y) = L(y)K(y),$$

где

$$L(y) = a_0(y - y_1)^{k_1-1} \cdots (y - y_r)^{k_r-1} (y^2 + p_1y + q_1)^{l_1-1} \cdots (y^2 + p_sy + q_s)^{l_s-1};$$

$$K(y) = (y - y_1) \cdots (y - y_r) (y^2 + py + q) \cdots (y^2 + p_sy + q_s),$$

то имеет место единственное представление

$$R_1(y) = \frac{M(y)}{L(y)}; \quad R_2(y) = \int \frac{N(y) dy}{K(y)}.$$

Коэффициенты полиномов  $M(y)$  и  $N(y)$  находятся методом неопределенных коэффициентов, поскольку  $M(y)/L(y)$ ,  $N(y)/K(y)$  и  $L'(y)K(y)/L(y)$  являются правильными ДРФ.

Примерами СДРН, получаемых посредством отрезков сумм типовых ДРФ, могут служить следующие:

$$\varphi^{\text{СДРН}}(Y, t) = \sum_{r=1}^n l_{rt} \varphi_r^{\text{ДРН}}(Y); \quad (7)$$

$$\varphi^{\text{СДРН}}(Y, t) = \sum_{r=1}^n l_{rt} \varphi_r^{\text{ДРН}}(Y) \varphi_r(Y); \quad (8)$$

$$\varphi^{\text{СДРН}}(Y, t) = \frac{\sum_{r=1}^{n'} l'_{rt} \varphi_r^{\text{ДРН}}(Y)}{\sum_{r=1}^{n''} l''_{rt} \varphi_r^{\text{ДРН}}(Y)}; \quad (9)$$

$$\varphi^{\text{СДРН}}(Y, t) = \frac{\sum_{r=1}^{n'} l'_{rt} \varphi_r^{\text{ДРН}}(Y) \varphi'_r(Y)}{\sum_{r=1}^{n''} l''_{rt} \varphi_r^{\text{ДРН}}(Y) \varphi''_r(Y)}, \quad (10)$$

где  $\varphi_r^{\text{ДРН}}(Y)$ ,  $\varphi_r^{\text{ДРН}}(Y)$  и  $\varphi_r^{\text{ДРН}}(Y)$  — типовые ДРН;  $l_{rt}$ ,  $l'_{rt}$  и  $l''_{rt}$  — коэффициенты, зависящие от времени  $t$ ;  $\varphi_r(Y)$ ,  $\varphi'_r(Y)$  и  $\varphi''_r(Y)$  — известные функции.

Другими примерами СДРН являются нелинейности, получаемые путем соответствующего преобразования аргумента:

$$\varphi^{\text{СДРН}}(Y, t) = \varphi(\psi^{\text{ДРН}}(Y, t), t); \quad (11)$$

$$\varphi^{\text{СДРН}}(Y, t) = \varphi^{\text{ДРН}}(\psi(Y, t), t). \quad (12)$$

Операция интегрирования ДРФ может выводить из класса типовых алгебраических и трансцендентных ДРФ, т. е. интеграл от элементарной ДРФ не всегда может быть выражен через элементарные функции (алгебраические или трансцендентные). В результате приходится обращаться к специальным функциям [6–9]. Простейшими примерами специальных функций могут служить интегралы от типовых ДРФ, а также функций, получающихся из них с помощью конечного числа вычислительных операций и операций дифференцирования [6]. Сюда относятся и функции, обратные указанным.

Дробно-рациональная функция векторного аргумента определяется как отношение конечных сумм [7]:

$$R^{\text{ДРФ}}(y_1, \dots, y_n) = \frac{\sum_{k_1, \dots, k_n} a_{k_1, \dots, k_n} (y_1)^{k_1} \dots (y_n)^{k_n}}{\sum_{i_1, \dots, i_n} b_{i_1, \dots, i_n} (y_1)^{i_1} \dots (y_n)^{i_n}}. \quad (13)$$

В качестве примеров скалярных СДРН векторного аргумента  $Y = [Y_1 \dots Y_p]^T$  рассмотрим следующие:

$$\varphi^{\text{СДРН}}(Y, t) = \sum_{r=1}^n \prod_{h=1}^H l_{rh,t} \varphi_{rh}^{\text{ДРН}}(Y_h); \quad (14)$$

$$\varphi^{\text{СДРН}}(Y, t) = \frac{\sum_{r=1}^{n'} \prod_{h=1}^{H'} l'_{rh,t} \varphi'_{rh}{}^{\text{ДРН}}(Y_h)}{\sum_{r=1}^{n''} \prod_{h=1}^{H''} l''_{rh,t} \varphi''_{rh}{}^{\text{ДРН}}(Y_h)}. \quad (15)$$

В случае векторных и матричных СДРН формулы (7)–(15) имеют место как соответствующие формулы для компонент.

### 3 Уравнения методов нормальной аппроксимации и статистической линеаризации

Следуя [5], уравнения конечномерных непрерывных нелинейных систем со стохастическими возмущениями путем расширения вектора состояния СтС могут быть записаны в виде следующего векторного стохастического дифференциального уравнения Ито:

$$dY_t = a(Y_t, t) dt + b(Y_t, t) dW_0 + \int_{R_0} c(Y_t, t, v) P^0(dt, dv), \quad Y(t_0) = Y_0. \quad (16)$$



Здесь  $Y_t$  —  $(p \times 1)$ -мерный вектор состояния,  $Y_t \in \Delta_y$  ( $\Delta_y$  — многообразие состояний);  $a = a(Y_t, t)$  и  $b = b(y_t, t)$  — известные  $(p \times 1)$ -мерная и  $(p \times m)$ -мерная функции  $Y_t$  и  $t$ ;  $W_0 = W_0(t)$  —  $(r \times 1)$ -мерный винеровский СтП интенсивности  $\nu_0 = \nu_0(t)$ ;  $c(Y_t, t, v)$  —  $(p \times 1)$ -мерная функция  $Y_t, t$  и вспомогательного  $(q \times 1)$ -мерного параметра  $v$ ;  $\int_{\Delta} dP^0(t, A)$  — центрированная пуассоновская мера, определяемая соотношением

$$\int_{\Delta} dP^0(t, A) = \int_{\Delta} dP(t, A) = \int_{\Delta} \nu_P(t, A) dt.$$

В (16) принято:  $\int_{\Delta}$  — число скачков пуассоновского СтП в интервале времени  $\Delta = (t_1, t_2]$ ;  $\nu_P(t, A)$  — интенсивность пуассоновского СтП  $P(t, A)$ ;  $A$  — некоторое борелевское множество пространства  $R_0^q$  с выколотым началом. Начальное значение  $Y_0$  представляет собой случайную величину (с.в.), не зависящую от приращений  $W_0(t)$  и  $P(t, A)$  на интервалах времени, следующих за  $t_0$ ,  $t_0 \leq t_1 \leq t_2$ , для любого множества  $A$ .

В случае аддитивных нормальных (гауссовских) и обобщенных пуассоновских возмущений уравнение (16) имеет вид:

$$\dot{Y} = a(Y_t, t) + b_0(t)V, \quad V = \dot{W}; \quad Y(t_0) = Y_0. \quad (17)$$

Здесь  $W$  — СтП с независимыми приращениями, представляющий собой смесь нормального и обобщенного пуассоновского СтП.

Для компонент  $\varphi(Y_t, t) = \{a_h, b_{kj}, c_h\}$  функций  $a, b$  и  $c$ , являющихся СДРН, примем представления (1)–(6). Если предположить существование конечных вероятностных моментов второго порядка для моментов времени  $t_1$  и  $t_2$ , то уравнения МНА примут следующий вид [10, 11]:

– для характеристических функций:

$$\left. \begin{aligned} g_1^N(\lambda; t) &= \exp \left[ i\lambda^T m_t - \frac{1}{2} \lambda^T K_t \lambda \right]; \\ g_{t_1, t_2}^N(\lambda_1, \lambda_2; t_1, t_2) &= \exp \left[ i\bar{\lambda}^T \bar{m}_2 - \frac{1}{2} \bar{\lambda}^T \bar{K}_2 \bar{\lambda} \right], \end{aligned} \right\} \quad (18)$$

где

$$\bar{\lambda} = \left[ \lambda_1^T \lambda_2^T \right]^T; \quad \bar{m}_2 = \left[ m_{t_1}^T m_{t_2}^T \right]^T; \quad \bar{K}_2 = \begin{bmatrix} K(t_1, t_1) & K(t_1, t_2) \\ K(t_2, t_1) & K(t_2, t_2) \end{bmatrix};$$

– для математических ожиданий  $m_t$ , ковариационной матрицы  $K_t$  и матрицы ковариационных функций  $K(t_1, t_2)$ :

$$\dot{m}_t = a_1(m_t, K_t, t), \quad m_0 = m(t_0); \quad (19)$$

$$\dot{K}_t = a_2(m_t, K_t, t), \quad K_0 = K(t_0); \quad (20)$$

$$\frac{\partial K(t_1, t_2)}{\partial t_2} = K(t_1, t_2) a_{21}(m_{t_2}, K_{t_2}, t_2)^T, \quad K(t_1, t_1) = K_{t_1}. \quad (21)$$

Здесь приняты следующие обозначения:

$$m_t = M_{\Delta_y}^N Y_t; \quad Y_t^0 = Y_t - m_t; \quad K_t = M_{\Delta_y}^N Y_t^0 Y_t^{0T}; \quad K(t_1, t_2) = M_{\Delta_y}^N Y_{t_1}^0 Y_{t_2}^{0T};$$

$$a_1 = a_1(m_t, K_t, t) = M_{\Delta_y}^N a(Y_t, t);$$

$$a_2 = a_2(m_t, K_t, t) = a_{21}(m_t, K_t, t) + a_{21}(m_t, K_t, t)^T + a_{22}(m_t, K_t, t);$$

$$a_{21} = a_{21}(m_t, K_t, t) = M_{\Delta_y}^N a(Y_t, t) Y_t^{0T};$$

$$a_{22} = a_{22}(m_t, K_t, t) = M_{\Delta_y}^N \sigma(Y_t, t);$$

$$\sigma(Y_t, t) = b(Y_t, t) \nu_0(t) b(Y_t, t)^T + \int_{R_0^q} c(Y_t, t, v) c(Y_t, t, v)^T \nu_P(t, dv),$$

где  $M_{\Delta_y}^N$  — символ вычисления математического ожидания для нормальных распределений (18) на многообразии  $\Delta_y$ .

Для стационарных СтС нормальные стационарные СтП — если они существуют, то  $m_t = m^*$ ,  $K_t = K^*$ ,  $K(t_1, t_2) = k(\tau)$  ( $\tau = t_1 - t_2$ ), — определяются уравнениями [9–11]:

$$\left. \begin{aligned} a_1(m^*, K^*) &= 0; \quad a_2(m^*, K^*) = 0; \\ \dot{k}_\tau(\tau) &= a_{21}(m^*, K^*) K^{*-1} k(\tau); \quad k(0) = \bar{K} \quad (\forall \tau > 0); \\ k(\tau) &= k(-\tau)^T \quad (\forall \tau < 0). \end{aligned} \right\} \quad (22)$$

При этом необходимо, чтобы матрица  $a_{21}(m^*, K^*) = a_{21}^*$  была асимптотически устойчивой.

Уравнения МНА в случае СтС (17) переходят в уравнения МСЛ [10, 11]:

$$\left. \begin{aligned} \dot{m}_t &= a_1(m_t, K_t, t), \quad m_0 = m(t_0); \\ \dot{K}_t &= k_1^a(m_t, K_t, t) K_t + K_t k_1^a(m_t, K_t, t)^T + \sigma_0(t), \quad K_0 = K(t_0); \\ \frac{\partial K(t_1, t_2)}{\partial t_2} &= K(t_1, t_2) K_{t_2} k_1^a(m_{t_2}, K_{t_2}, t_2)^T, \quad K(t_1, t_2) = K_{t_1}, \end{aligned} \right\} \quad (23)$$

если принять

$$\left. \begin{aligned} a(Y_t, t) &= a_1(m_t, K_t) + k_1^a(m_t, K_t)Y_t^0; \\ b(Y_t, t) &= b_0(t); \quad \sigma(Y_t, t) = b_0(t)\nu(t)b_0(t)^T = \sigma_0(t); \\ k_1^a(m_t, K_t, t) &= \left[ \left( \frac{\partial}{\partial m_t} \right) a_0(m_t, K_t, t)^T \right]^T. \end{aligned} \right\} \quad (24)$$

Для стационарных СтС (17) при условии асимптотической устойчивости матрицы  $k_1^a(m^*, K^*)$  в основе МСЛ лежат уравнения (22), записанные в виде:

$$\left. \begin{aligned} a_1(m^*, K^*) &= 0; \quad k_1^a(m^*, K^*)K^* + K^*k_1^a(m^*, K^*)^T + \sigma_0^* = 0; \\ \dot{k}_\tau(\tau) &= k_1^a(m^*, K^*)k(\tau); \quad k(0) = K^* \quad (\forall \tau > 0); \\ k(\tau) &= k(-\tau)^T \quad (\forall \tau < 0). \end{aligned} \right\} \quad (25)$$

Уравнения (18)–(21) лежат в основе МНА для МСтС (16), а уравнения (23) и (24) — в основе МСЛ для СтС (17). Для определения стационарных СтП согласно МНА служат соотношения (22), а МСЛ — (25).

#### 4 Алгоритмическое обеспечение аналитического моделирования

Для МНА необходимо уметь вычислять следующие интегралы:

$$\begin{aligned} I_0^a &= I_0^a(m_t, K_t, t) = a_1(m_t, K_t, t) = M_{\Delta_y}^N a(Y_t, t); \\ I_1^a &= I_1^a(m_t, K_t, t) = a_{21}(m_t, K_t, t) = M_{\Delta_y}^N a(Y_t, t)Y_t^{0T}; \\ I_0^{\bar{\sigma}} &= I_0^{\bar{\sigma}}(m_t, K_t, t) = a_{22}(m_t, K_t, t) = M_N \bar{\sigma}(Y_t, t). \end{aligned} \quad (26)$$

Для МСЛ достаточно вычислить интеграл (26), причем интеграл  $I_1^a$  вычисляется по формуле [10, 11]

$$k_1^a = k_1^a(m_t, K_t, t) = \left[ \left( \frac{\partial}{\partial m_t} \right) I_0^a(m_t, K_t, t)^T \right]^T.$$

Как следует из свойства ДРФ  $5^0$ , базовыми интегралами при вычислении (26) будут интегралы:

$$M_{\Delta_y}^N \left[ \frac{1}{(Y - Y_\nu)^\mu} \right]; \quad M_{\Delta_y}^N \left[ \frac{1}{(Y^2 + p_\nu Y + q_\nu)^\mu} \right]; \quad M_{\Delta_y}^N \left[ \frac{Y}{(Y^2 + p_\nu Y + q_\nu)^\mu} \right]. \quad (27)$$

Подробно различные численные методы вычисления интегралов (27) описаны в [9, 12]. В случае  $\Delta_y = (-\infty, \infty)$  можно рекомендовать квадратурные формулы

на основе ортогональных полиномов Эрмита  $H_n(x)$ . В этом случае используется следующий алгоритм [9, 12].

$$\int_{-\infty}^{\infty} e^{-x^2} \varphi(x, \pi) dx = \sum_{i=1}^n w_i \varphi(x_i, \pi) + \rho_n. \quad (28)$$

Здесь  $\pi$  — вектор параметров;  $x_i$  — нуль  $H_n(x)$ ;  $w_i$  — весовой коэффициент:

$$w_i = \frac{2^{n-1} n! \sqrt{\pi}}{n^2 [H_{n-1}(x_i)]^2}; \quad (29)$$

$\rho_n$  — остаточный член:

$$\rho_n = \frac{n! \sqrt{\pi}}{2^n (2n)!} \varphi^{(2n)}(\xi) \quad (-\infty < \xi < \infty). \quad (30)$$

Входящий в (28) вектор параметров обычно включает в себя дисперсию  $D$  и отношение сигнал/шум  $\zeta_y = m_y / \sqrt{D_y}$ .

Уравнения МНА (МСЛ) содержат интегралы  $I_0^a$ ,  $I_1^a$  и  $I_0^g$  в виде соответствующих коэффициентов, поэтому процедура вычисления интегралов должна быть согласована с методом численного решения обыкновенных дифференциальных уравнений для  $m_t$ ,  $K_t$  и  $K(t_1, t_2)$ . Эти коэффициенты допускают дифференцирование по  $m_t$  и  $K_t$ , так как под интегралом стоит сглаживающая нормальная плотность.

В [13] изложены алгоритмы аналитического и статистического моделирования распределений (в том числе нормальных) в нелинейных МСТС. Алгоритмы аналитического и статистического моделирования для МСТС с СДРН, а также смешанные алгоритмы различной степени точности относительно шага интегрирования также представлены в [13].

## 5 Тестовые примеры

**5.1.** Для типовой ДРН вида  $R^{\text{ДРН}}(y) = 1/(y^2 + b^2)$ , учитывая соотношение [9]

$$\int_0^{\infty} \frac{e^{-h^2 \eta^2}}{\eta^2 + b^2} d\eta = \frac{\pi}{2b} e^{h^2 b^2} \operatorname{erfc}(hb) \quad (h^2 > 0, \quad b^2 > 0),$$

где

$$\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2} dt = 1 - \operatorname{erf} x, \quad \operatorname{erf} x = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt,$$

имеем следующее выражение для интеграла (26) при  $h^2 = 1/(2D_y)$ :

$$\begin{aligned} I_0^{\text{ДРН}} &= M^N \left[ \frac{1}{Y^2 + b^2} \right] = \frac{1}{\sqrt{2\pi D_y}} \int_{-\infty}^{\infty} \frac{e^{-\eta^2/(2D_y)} d\eta}{\eta^2 + b^2} = \\ &= \frac{\pi}{b} e^{b^2/(2D_y)} \operatorname{erfc} \left( \frac{b}{\sqrt{2D_y}} \right). \end{aligned}$$

**5.2.** Для СДРН вида  $R^{\text{СДРН}}(y) = (y+b)^{-1} \mathbf{1}(y)$  при  $h^2 > 0$ ,  $b > 0$ , учитывая соотношение [9]

$$\int_0^{\infty} \frac{e^{-h^2 \eta^2}}{(\eta + b)} d\eta = e^{-h^2 b^2} \left[ \sqrt{\pi} \int_0^{h\sqrt{b}} e^{\eta^2} d\eta - \frac{1}{2} \operatorname{Ei}(h^2 b^2) \right]$$

(Ei — интегральная показательная функция [6]), имеем следующее выражение для интеграла (26) при  $h^2 = 1/(2D_y)$ :

$$\begin{aligned} I_0^{\text{ДРН}} &= M^N \left[ \frac{\mathbf{1}(Y)}{Y + b} \right] = \frac{1}{\sqrt{2\pi D_y}} \int_{-\infty}^{\infty} \frac{\mathbf{1}(\eta) e^{-\eta^2/(2D_y)}}{\eta + b} d\eta = \\ &= \frac{1}{\sqrt{2\pi D_y}} e^{-b^2/(2D_y)} \left[ \sqrt{\pi} \int_0^{\sqrt{b/(2D_y)}} e^{\eta^2} d\eta - \frac{1}{2} \operatorname{Ei} \left( \frac{b^2}{2D_y} \right) \right]. \end{aligned}$$

**5.3.** Рассмотрим динамическую систему с дробно-рациональными нелинейностями в линейной стохастической гауссовской среде

$$\ddot{X}_t + \omega_0^2 X_t + \frac{c_0}{X_t^2 + c^2} = a - 2\varepsilon\omega_0 \dot{X}_t + bV. \quad (31)$$

Здесь  $\varepsilon$ ,  $\omega_0$ ,  $c_0$ ,  $c$ ,  $a$  и  $b$  — постоянные параметры;  $V$  — гауссовский белый шум интенсивности  $\nu$ .

Полагая  $X_t = Y_1$  и  $\dot{Y}_1 = Y_2$  и учитывая  $\dot{Y}_2 = Y_2(dY_2/dY_1)$ , получим следующие дифференциальные соотношения для фазового портрета рассматриваемой динамической системы при  $a = 0$ ,  $b = 0$  и  $\varepsilon = 0$ :

$$\frac{dY_2}{dY_1} = -\frac{\omega_0^2 Y_1(Y_1^2 + c^2) + c_0}{Y_2(Y_1^2 + c^2)}; \quad \frac{Y_2^2}{2} + \Pi(Y_1) = \frac{Y_{20}^2}{2} + \Pi(Y_{10}),$$

где

$$\Pi(Y_1) = - \int \left( \frac{1}{2} \omega_0^2 Y_1^2 + \frac{c_0}{Y_1^2 + c^2} \right) dY_1 = -\frac{1}{2} \omega_0^2 Y_1^2 - \frac{c_0}{c} \operatorname{arctg} \frac{Y_1}{c} + \operatorname{const}. \quad (32)$$

Уравнения (31) при  $a = 0$ ,  $b = 1$  и  $c_0 > 0$  допускают режим стационарных стохастических колебаний с одномерной плотностью, вычисляемой по формуле Гиббса [1]:

$$f_1^*(y_1, y_2) = \operatorname{const} \cdot e^{-\alpha \mathcal{H}(y_1, y_2)}.$$

Здесь  $\mathcal{H} = \mathcal{H}(y_1, y_2)$  — функция Гамильтона  $\mathcal{H} = Y_2^2/2 + \Pi(Y_1)$ , где  $\Pi(Y_1)$  определена в (32). Отсюда видно, что распределение  $Y_2$  — гауссовское, а распределение  $Y_1$  — негауссовское.

Уравнения МСЛ для определения вероятностных моментов первого и второго порядка имеют следующий вид:

$$\dot{m}_1 = m_2, \quad \dot{m}_2 = -\omega_0^2 m_1 - 2\varepsilon \omega_0 m_2 + R_0(m_1, D_1); \quad (33)$$

$$\dot{D}_1 = 2K_{12}; \quad \dot{D}_2 = \nu - 2\lambda K_{12} - 4\varepsilon \omega_0 D_2; \quad \dot{K}_{12} = D_2 - \lambda D_1 - 2\varepsilon \omega_0 K_{12}. \quad (34)$$

Здесь

$$\begin{aligned} \lambda &= \lambda(m_1, D_1) = \omega_0^2 \left[ 1 - \frac{R_1(m_1, D_1)}{\omega_0^2} \right]; \\ R_0(m_1, D_1) &= \frac{1}{\sqrt{2\pi D_1}} \int_{-\infty}^{\infty} \frac{c_0}{\eta^2 + c^2} e^{-(\eta - m_1)^2 / (2D_1)} d\eta = \\ &= \frac{2c_0}{\sqrt{2\pi D_1}} e^{-m_1^2 / (2D_1)} \int_0^{\infty} e^{-(\eta^2 / (2D_1)) + \eta m_1 / D_1} \frac{d\eta}{\eta^2 + c^2}; \quad (35) \\ R_1(m_1, D_1) &= \frac{\partial R_0(m_1, D_1)}{\partial m_1}. \end{aligned}$$

Интеграл (35) вычисляется численно согласно (28)–(30) при

$$\varphi(\eta) = \frac{1}{\eta^2 + c^2} e^{-\eta m_1 / D_1}.$$

Из (33) и (34) при  $\varepsilon > 0$  для стационарного режима имеем соотношения:

$$\begin{aligned} m_2^* &= 0; \quad m_1^* = R_0(m_1^*, D_1^*) \omega_0^{-2}; \\ D_2^* &= \alpha = \frac{\nu^*}{4\varepsilon \omega_0}; \quad K_{12}^* = 0; \quad D_1^* \lambda(m_1^*, D_1^*) = \alpha. \end{aligned}$$

Таким образом, аналитическое моделирование нормальных процессов в СтС с СДРН на основе МСЛ обеспечивает достаточную для применений точность. Так, при малых отношениях сигнал/шум  $\zeta = m_1^*/\sqrt{D_1^*}$  ошибка метода аналитического моделирования на основе МСЛ составляет не более 20%, а при больших — не более 1%.

## 6 Заключение

Рассматриваются дифференциальные СтС (в том числе и МСтС) с винеровскими и пуассоновскими шумами и с СДРН. Такие модели описывают поведение ряда современных нано- и квантовооптических технических средств информатики. Приводятся уравнения МНА и МСЛ для моделирования нестационарных и стационарных нормальных процессов. Рассматриваются методы вычисления типовых интегралов для одно- и многомерных СДРН скалярного и векторного аргумента, получающихся из суперпозиции типовых ДРН. Обсуждается алгоритмическое и программное обеспечение аналитического моделирования. Приводятся тестовые примеры.

Результаты допускают обобщение на случай дискретных СтС, а также интегродифференциальных и операторных СтС с СДРН, в том числе с автокоррелированными шумами.

## Литература

1. *Синицын И. Н., Синицын В. И.* Аналитическое моделирование нормальных процессов в стохастических системах со сложными нелинейностями // Информатика и её применения, 2014. Т. 8. Вып. 3. С. 2–4.
2. *Синицын И. Н., Синицын В. И., Сергеев И. В., Белоусов В. В., Шоргин В. С.* Математическое обеспечение аналитического моделирования стохастических систем со сложными нелинейностями // Системы и средства информатики, 2014. Т. 24. № 3. С. 4–29.
3. *Синицын И. Н., Синицын В. И., Корепанов Э. Р.* Моделирование нормальных процессов в стохастических системах со сложными иррациональными нелинейностями // Информатика и её применения, 2015. Т. 9. Вып. 1. С. 2–8.
4. *Синицын И. Н., Синицын В. И., Сергеев И. В., Корепанов Э. Р., Белоусов В. В., Шоргин В. С.* Математическое обеспечение моделирования нормальных процессов в стохастических системах со сложными иррациональными нелинейностями // Системы и средства информатики, 2015. Т. 25. № 2. С. 3–19.
5. *Синицын И. Н., Синицын В. И., Корепанов Э. Р.* Моделирование нормальных процессов в стохастических системах со сложными трансцендентными нелинейностями // Информатика и её применения, 2015. Т. 9. Вып. 2. С. 23–29.
6. *Корн Г., Корн Т.* Справочник по математике / Пер. с англ. — М.: Наука, 1984. 831 с. (*Korn G. A., Korn T. M.* Mathematical handbook. — New York, NY, USA: McGraw-Hill Book Co., 1968. 943 p.)

7. *Зверович Э. И.* Вещественный и комплексный анализ. Кн. 2. Ч. 2. Интегральное исчисление функций скалярного аргумента. — Минск: Высшая школа, 2008. 319 с.
8. *Градштейн И. С., Рыжик И. М.* Таблицы интегралов, сумм, рядов и произведений. — М.: ГИФМЛ, 1963. 1100 с.
9. Справочник по специальным функциям / Под ред. М. Абрамовича, И. Стигана. — М.: Наука, 1979. 832 с.
10. *Пугачёв В. С., Синицын И. Н.* Теория стохастических систем. — М.: Логос, 2000; 2004. 1000 с. (Англ. пер. *Pugachev V. S., Sinitsyn I. N.* Stochastic systems. Theory and applications. — Singapore: World Scientific, 2001. 908 p.)
11. *Синицын И. Н., Синицын В. И.* Лекции по нормальной и эллипсоидальной аппроксимации распределений в стохастических системах. — М.: ТОРУС ПРЕСС, 2013. 488 с.
12. *Попов Б. А., Теслер Г. С.* Вычисление функций на ЭВМ: Справочник. — Киев: Наукова Думка, 1984. 599 с.
13. *Синицын И. Н.* Параметрическое статистическое и аналитическое моделирование распределений в нелинейных стохастических системах на многообразиях // Информатика и её применения, 2013. Т. 7. Вып. 2. С. 4–16.

*Поступила в редакцию 18.08.15*

---

---

## ANALYTICAL MODELING OF NORMAL PROCESSES IN STOCHASTIC SYSTEMS WITH COMPLEX FRACTION-RATIONAL NONLINEARITIES

*I. N. Sinitsyn, V. I. Sinitsyn, and E. R. Korepanov*

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** Methods of analytical modeling for continuous stochastic systems (StS) (including on manifolds) with Wiener and Poisson noises and with complex fraction-rational nonlinearities (FRN) are given. Typical representations and vector FRN are considered. Equations for the normal approximation method (NAM) and the method of statistical linearization (MSL) are deduced. The NAM and MSL for StS with FRN algorithms are given. Test examples are presented. Some generalizations are given.

**Keywords:** analytical modeling; complex fraction-rational nonlinearities (FRN); Hermite polynomials; method of statistical linearization (MSL); normal approximation method (NAM); stochastic systems (StS)

**DOI:** 10.14357/08696527150401



## Acknowledgments

The work was supported by the Russian Foundation for Basic Research (project 15-07-02244).

## References

1. Sinitsyn, I. N., and V. I. Sinitsyn. 2014. Analiticheskoe modelirovanie normal'nykh protsessov v stokhasticheskikh sistemakh so slozhnymi nelineynostyami [Analytical modeling of normal processes in stochastic systems with complex nonlinearities]. *Informatika i ee Primeneniya — Inform. Appl.* 8(3):2–4.
2. Sinitsyn, I. N., V. I. Sinitsyn, I. V. Sergeev, V. V. Belousov, and V. S. Shorgin. 2014. Matematicheskoe obespechenie analiticheskogo modelirovaniya stokhasticheskikh sistem so slozhnymi nelineynostyami [Mathematical software for analytical modeling of stochastic systems with complex nonlinearities]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 24(3):4–29.
3. Sinitsyn, I. N., V. I. Sinitsyn, and E. R. Korepanov. 2015. Modelirovanie normal'nykh protsessov v stokhasticheskikh sistemakh so slozhnymi irratsional'nymi nelineynostyami [Modeling of normal processes in stochastic systems with complex irrational nonlinearities]. *Informatika i ee Primeneniya — Inform. Appl.* 9(1):2–8.
4. Sinitsyn, I. N., V. I. Sinitsyn, I. V. Sergeev, E. R. Korepanov, V. V. Belousov, and V. S. Shorgin. 2015. Matematicheskoe obespechenie modelirovaniya normal'nykh protsessov v stokhasticheskikh sistemakh so slozhnymi irratsional'nymi nelineynostyami [Mathematical software for modeling of normal processes in stochastic systems with complex irrational nonlinearities]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 25(2): 3–19.
5. Sinitsyn, I. N., V. I. Sinitsyn, and E. R. Korepanov. 2015. Modelirovanie normal'nykh protsessov v stokhasticheskikh sistemakh so slozhnymi transtsendentnymi nelineynostyami [Modeling of normal processes in stochastic systems with complex transcendental nonlinearities]. *Informatika i ee Primeneniya — Inform. Appl.* 9(2):23–29.
6. Korn, G., and T. Korn. 1984. 1968. *Mathematical handbook*. New York, NY: McGraw-Hill Book Co. 943 p.
7. Zverovich, E. I. 2008. *Veshchestvennyy i kompleksnyy analiz*. Kn. 2. Ch. 2. Integral'noe ischislenie funktsiy skalyarnogo argumenta [Real complex analysis. Bk. 2. Pt. 2. Integral calculus of scalar argument]. Minsk: Vysshaya Shkola. 319 p.
8. Gradshteyn, I. S., and I. M. Ryzhik. 1963. *Tablitsy integralov, summ, ryadov i proizvedeniy*. Moscow: GIFML. 1100 p.
9. Abramovich, M., and I. Stigan, eds. 1979. *Spravochnik po spetsial'nym funktsiyam*. Moscow: Nauka. 832 p.
10. Pugachev, V. S., and I. N. Sinitsyn. 2001. *Stochastic systems. Theory and applications*. Singapore: World Scientific. 908 p.
11. Sinitsyn, I. N., and V. I. Sinitsyn. 2013. *Lektsii po normal'noy i ellipsoidal'noy approksimatsii raspredeleniy v stokhasticheskikh sistemakh* [Lectures on normal and ellipsoidal approximation of distributions in stochastic systems]. Moscow: TORUS PRESS. 488 p.

12. Popov, B. A., and G. S. Tesler. 1984. *Vychislenie funktsiy na EVM: Spravochnik* [Calculations of functions on computers. Reference book]. Kiev: Naukova Dumka. 599 p.
13. Sinitsyn, I. N. 2013. Parametricheskoe statisticheskoe i analiticheskoe modelirovanie raspredeleniy v nelineynykh stokhasticheskikh sistemakh na mnogoobraznykh [Parametric statistical and analytical modeling of distributions in stochastic systems on manifolds]. *Informatika i ee Primeneniya — Inform. Appl.* 7(2):4–16.

*Received August 18, 2015*

## **Contributors**

**Sinitsyn Igor N.** (b. 1940) — Doctor of Science in technology, professor, Honored scientist of RF, Head of Department, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; sinitsin@dol.ru

**Sinitsyn Vladimir I.** (b. 1968) — Doctor of Science in physics and mathematics, associate professor, Head of Department, Institute of Informatics Problems, Federal Research Center “Computer Science and Control,” Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; vSinitsyn@ipiran.ru

**Korepanov Eduard R.** (b. 1966) — Candidate of Science (PhD) in technology, Head of Laboratory, Institute of Informatics Problems, Federal Research Center “Computer Science and Control,” Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; ekorepanov@ipiran.ru

## АНАЛИЗ ЭНЕРГОЭФФЕКТИВНОСТИ ВЫЧИСЛИТЕЛЬНОГО КОМПЛЕКСА, МОДЕЛИРУЕМОГО С ПОМОЩЬЮ СИСТЕМЫ ОБСЛУЖИВАНИЯ С ПОРОГОВЫМ УПРАВЛЕНИЕМ И ИНТЕНСИВНОСТЯМИ, ЗАВИСЯЩИМИ ОТ ВРЕМЕНИ\*

*Р. В. Разумчик<sup>1</sup>, А. И. Зейфман<sup>2</sup>, А. В. Коротышева<sup>3</sup>, Я. А. Сатин<sup>4</sup>*

**Аннотация:** Рассматривается задача анализа энергоэффективности компонент вычислительного комплекса (серверов), в которых может быть реализована пороговая стратегия управления питанием (изменением напряжения и частоты работы процессоров). В ряде работ для решения этой задачи используются марковские модели массового обслуживания с пороговыми стратегиями управления интенсивностями обслуживания. Подобный подход позволяет вычислять основные показатели энергоэффективности серверов и находить оптимальные (в классе пороговых) стратегии управления. В большинстве работ одним из основных предположений является то, что интенсивности переходов процессов, описывающих функционирование системы, не зависят от времени. В данной работе рассматривается некоторое обобщение ряда известных результатов на случай, когда интенсивности переходов являются неслучайными функциями времени (в частности, периодическими). Рассматривается система  $M(t)/M(t)/1/\infty$  с однопороговой стратегией управления интенсивностью обслуживания. Показано, как, используя ряд общих результатов для неоднородных процессов рождения и гибели, можно получить сравнительно простые (приближенные) расчетные алгоритмы для основных характеристик системы. Представлены некоторые результаты численных расчетов значений параметров энергоэффективности серверов, проведенные на основе полученных формул.

**Ключевые слова:** система массового обслуживания; неоднородный процесс рождения и гибели; энергоэффективность

**DOI:** 10.14357/08696527150402

---

\* Работа выполнена при поддержке Российского научного фонда (проект 14-11-00397).

<sup>1</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук; Российский университет дружбы народов, grazumchik@ipiran.ru

<sup>2</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук; Вологодский государственный университет; Институт социально-экономического развития территорий Российской академии наук, a\_zeifman@mail.ru

<sup>3</sup>Вологодский государственный университет; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, a\_korotysheva@mail.ru

<sup>4</sup>Вологодский государственный университет; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, yacovi@mail.ru

## 1 Введение

В настоящее время задачам повышения энергоэффективности инфотелекоммуникационных систем, вычислительных комплексов (центров обработки данных) и других систем, потребляющих значительные объемы электрической энергии, уделяется большое внимание. Для снижения энергопотребления серверов уже разработано несколько подходов и сопутствующих технологий, каждая из которых обладает своими достоинствами и недостатками. Не вдаваясь в подробности технологий (см., например, [1]), отметим, что среди наиболее популярных, согласно [2], можно выделить технологию on-off, когда полностью отключаются простаивающие компоненты системы, и технологию DVFS (Dynamic Voltage Frequency Scaling), позволяющую уменьшать в том числе и энергопотребление путем снижения тактовой частоты процессора и напряжения его питания. Для крупномасштабных гетерогенных вычислительных комплексов существуют также подходы, не требующие динамического изменения состояния компонентов серверов и/или их характеристик, а предполагающие использование специальных дисциплин распределения задач, которые при планировании вычислений используют информацию о текущей загрузке и энергопотреблении каждого сервера. Примеры алгоритмов, использующих данный подход, можно найти в [3, 4].

Одним из направлений в анализе различных схем контроля расхода электроэнергии является моделирование компонентов вычислительных комплексов с помощью систем массового обслуживания (СМО) с пороговым управлением и нахождение подходящей схемы управления компонентами комплекса путем анализа тех или иных характеристик систем обслуживания при различных начальных условиях. Среди многочисленных научных и прикладных работ по данной тематике стоит отметить работы [5–7] (системы с  $N$ -политикой), [8, 9] (системы с  $D$ -политикой), [10, 11] (системы с  $T$ -политикой). Несмотря на то что за последние 30 лет в этом направлении получено большое число результатов, активные исследования продолжают, что мотивируется практическими задачами и подтверждается большим числом работ, публикующихся каждый год.

В недавней работе [1] рассмотрена задача повышения энергоэффективности одной из компонент вычислительного комплекса (сервера) за счет реализации в нем механизма порогового управления тактовой частотой процессора. В качестве модели сервера была выбрана СМО  $M/M/1$  с  $T$ -политикой. Согласно этой политике, когда общее число заявок в системе меньше  $T$ , интенсивность обслуживания заявок равна  $\mu_0$ ; когда же общее число заявок больше или равно  $T$ , интенсивность обслуживания становится равной (мгновенно)  $\mu > \mu_0$ . Фактически рассмотренная авторами  $T$ -политика совпадает с обычной пороговой стратегией (с порогом  $T$ ) управления интенсивностью обслуживания, неоднократно описанной в литературе. Не останавливаясь на новых математических результатах, полученных авторами [1] и касающихся свойств (монотонности, выпуклости) основных стационарных характеристик системы в зависимости от  $T$ , сразу перейдем к ее практической части. В работе [1] показано, каким образом

можно перейти от математической модели к характеристикам энергоэффективности сервера, и затем приводятся результаты численных расчетов (зависимости среднего энергопотребления в стационарном режиме от порогового значения  $T$ , тактовой частоты и др.). Все изложенные в [1] результаты получены в предположении того, что интенсивности поступления заявок в систему (или заданий в вычислительный комплекс), а также интенсивности обслуживания не зависят от времени. Несомненный интерес представляет изучение вопросов, поднятых в [1], в случае, когда все интенсивности являются неслучайными функциями времени, т. е. процесс, описывающий функционирование сервера, по-прежнему является марковским (процессом рождения и гибели), но неоднородным. В работе [12] для таких процессов были получены результаты, касающиеся равномерных по времени оценок погрешности аппроксимации с помощью процессов с меньшим числом состояний, что позволяет находить с заданной точностью значения вероятностных характеристик моделируемых систем. В данной работе будет показано, как данные результаты можно применить и получить оценки энергоэффективности, аналогичные тем, что получены в [1], но при условии наличия временных зависимостей.

Статья организована следующим образом. В разд. 2 приводится описание марковской СМО, аналогичной той, что рассматривается в [1] в качестве модели компонентов вычислительного комплекса, за исключением того, что в ней допускается зависимость интенсивности переходов от времени. В разд. 3 показано, каким образом, используя общие результаты для неоднородных процессов рождения и гибели, можно рассчитывать вероятностные характеристики этой системы. Заключительный раздел посвящен изложению результатов численных экспериментов, в которых исследовались в новых предположениях показатели энергоэффективности, предложенные в [1].

## 2 Описание системы

Рассмотрим систему обслуживания  $M(t)/M(t)/1/\infty$  с однопороговой стратегией управления интенсивностью обслуживания, которая определяется следующим образом. Пусть задано некоторое натуральное число  $T \geq 2$ , которое будем называть порогом. В течение времени, пока общее число заявок в системе меньше  $T$ , прибор обслуживает заявки по экспоненциальному закону с параметром  $\mu_0(t)$ . Однако, когда общее число заявок в системе превышает или равно  $T$ , интенсивность обслуживания прибором заявок равна  $\mu(t) > \mu_0(t) \forall t \geq 0$ . Функции  $\mu_0(t)$  и  $\mu(t)$  будем дальше называть до- и послепороговыми интенсивностями обслуживания соответственно. Переключение интенсивности обслуживания заявок прибором происходит мгновенно, а заявки из очереди обслуживаются в порядке поступления.

Функционирование рассмотренной системы может быть описано неоднородным (по времени) марковским процессом  $\{X(t), t \geq 0\}$  с непрерывным временем и дискретным множеством состояний  $\mathcal{X} = \{0, 1, 2, \dots\}$ . Состояние  $i \in \mathcal{X}$  озна-

чает, что в системе всего  $i$  заявок. Из состояния  $i$ ,  $i \geq 1$ , возможны переходы только в соседние состояния  $(i - 1)$  или  $(i + 1)$ . Предполагается, что интенсивности переходов являются (неслучайными) функциями времени. Обозначим через  $p_{ij}(s, t) = \Pr \{X(t) = j | X(s) = i\}$ ,  $0 \leq s \leq t$ , переходные вероятности процесса  $X(t)$ , а через  $p_i(t) = \Pr \{X(t) = i\}$ ,  $i \geq 0$ , вероятности того, что процесс  $X(t)$  находится в состоянии  $i$  в момент времени  $t$ . Введем вектор

$$\mathbf{p}(t) = (p_0(t), p_1(t), p_2(t), \dots)^T.$$

Из описания системы и процесса  $X(t)$  следует, что за малое время  $h > 0$  вероятности переходов имеют вид:

$$\begin{aligned} \Pr (X(t+h) = j | X(t) = i) &= \\ &= \begin{cases} \lambda(t)h + \alpha_{ij}(t, h), & \text{если } j = i + 1, i \geq 0; \\ \mu_0(t)h + \alpha_{ij}(t, h), & \text{если } j = i - 1, i = \overline{1, T-1}; \\ \mu(t)h + \alpha_{ij}(t, h), & \text{если } j = i - 1, i \geq T; \\ \alpha_{ij}(t, h) & \text{в остальных случаях,} \end{cases} \quad (1) \end{aligned}$$

где  $\sup_i |\alpha_{ij}(t, h)| = o(h)$ . Для сокращения записи положим  $\gamma_0(t) = \lambda(t) + \mu_0(t)$ ;  $\gamma(t) = \lambda(t) + \mu(t)$ . Нетрудно видеть, что матрица интенсивностей переходов процесса  $X(t)$  является трехдиагональной и имеет вид:

$$Q(t) = \begin{pmatrix} -\lambda(t) & \lambda(t) & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots \\ \mu_0(t) & -\gamma_0(t) & \lambda(t) & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots \\ 0 & \mu_0(t) & -\gamma_0(t) & \lambda(t) & \ddots & 0 & 0 & 0 & 0 & \cdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ddots & \mu_0(t) & -\gamma_0(t) & \lambda(t) & 0 & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & \mu(t) & -\gamma(t) & \lambda(t) & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & \mu(t) & -\gamma(t) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix},$$

а вероятности  $p_i(t)$  удовлетворяют системе дифференциальных уравнений Колмогорова:

$$\frac{d\mathbf{p}(t)}{dt} = A(t)\mathbf{p}(t), \quad (2)$$

где  $A(t) = Q^T(t)$ .

Напомним, что неоднородная марковская цепь  $X(t)$  называется слабо эргодичной, если  $\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\| \rightarrow 0$  при  $t \rightarrow \infty$  для любых начальных условий

$\mathbf{p}^*(0)$  и  $\mathbf{p}^{**}(0)$ . Здесь  $\mathbf{p}^*(t)$  и  $\mathbf{p}^{**}(t)$  — соответствующие решения (2), а  $\|\cdot\|$  —  $l_1$ -норма.

Обозначим через  $E_k(t) = E\{X(t) | X(0) = k\}$  среднее число требований (математическое ожидание процесса  $X(t)$ ) с начальным условием  $X(0) = \mathbf{e}_k$ .

Говорят, что функция  $\varphi(t)$  называется предельным средним для  $X(t)$ , если  $\lim_{t \rightarrow \infty} (\varphi(t) - E_k(t)) = 0$  при любом  $k$ .

Далее для простоты формулировок будем предполагать, что интенсивности поступления и обслуживания заявок периодичны с некоторым общим периодом  $\omega$  и, кроме того, что до- и послепороговая интенсивности обслуживания пропорциональны, т. е. что  $\mu_0(t) = k\mu(t)$  при некотором  $k < 1$ .

### 3 Вероятности состояний

Введем обозначения для усредненных интенсивностей поступления и обслуживания требований:

$$\mu_* = \omega^{-1} \int_0^\omega \mu(t) dt; \quad \mu_{0*} = \omega^{-1} \int_0^\omega \mu_0(t) dt; \quad \lambda_* = \omega^{-1} \int_0^\omega \lambda(t) dt.$$

Тогда, воспользовавшись теоремой 1 и следствием 1 из [12], можно получить следующие утверждения.

**Утверждение 1.** Пусть выполнено условие  $\lambda_* < \mu_*$ . Тогда процесс  $X(t)$  слабо эргодичен и имеет предельное среднее.

**Утверждение 2.** Пусть, кроме того, выполнено условие  $\lambda_* < \mu_{0*}$ . Тогда справедливы следующие оценки скорости сходимости: при всех  $t \geq 0$  и любых начальных условиях  $\mathbf{p}^*(0)$  и  $\mathbf{p}^{**}(0)$

$$\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\| \leq 4Re^{-\gamma_* t} \sum_{i \geq 1} g_i |p_i^*(0) - p_i^{**}(0)|, \quad (3)$$

где

$$\beta = \sqrt{\frac{\mu_{0*}}{\lambda_*}}; \quad R = e^{\int_0^\omega \mu_0(t) dt}; \quad g_i = \sum_{n=1}^i d_n; \quad d_n = \beta^{n-1}; \quad \gamma_* = \left(\sqrt{\mu_{0*}} - \sqrt{\lambda_*}\right)^2,$$

и при любом  $k$

$$|\phi(t) - E_k(t)| \leq \frac{4R}{W} e^{-\gamma_* t} \|\mathbf{p}(0) - \mathbf{e}_k\|_{1D}, \quad (4)$$

где

$$W = \inf_{n \geq 1} \frac{\beta^{n-1}}{n}.$$

Оценки погрешности, получаемой при решении прямой системы Колмогорова для «усеченного» процесса с теми же интенсивностями и множеством состояний  $\{0, 1, \dots, N\}$ , можно получить, воспользовавшись теоремой 6 из [12], что дает следующее утверждение.

**Утверждение 3.** Для вероятностей  $\mathbf{p}(t)$  и  $\mathbf{p}_N(t)$  исходного и «усеченного» процесса справедлива оценка

$$\|\mathbf{p}(t) - \mathbf{p}_N(t)\| \leq \frac{8LR^{3/2}(L+M)}{\gamma_*^{3/2} \beta^{N-1}} \quad (5)$$

при  $\mathbf{p}(0) = \mathbf{p}_N(0) = \mathbf{e}_0$ , а для соответствующих предельных средних значений — оценка

$$|E_0(t) - E_{0,N}(t)| \leq \frac{4LR^{3/2}(L+M)}{W\gamma_*^{3/2} \beta^{N-1}}, \quad (6)$$

где  $\lambda(t) \leq L$ ,  $\mu(t) \leq M$ ,  $\mu_0(t) \leq M$  при всех  $t$ .

#### 4 Численный эксперимент

В работе [1] авторы моделируют часть вычислительного комплекса (сервер) с помощью СМО  $M/M/1/\infty$  с однопороговой стратегией и далее, используя известные модели энергоэффективности, показывают, как результаты для математической модели можно пересчитать в показатели энергоэффективности работы сервера. Известно (см., например, [13]), что энергопотребление (или потребляемая мощность)  $P$  вычислительной системы может быть рассчитано по формуле:

$$P = CfV^2 + P_s, \quad (7)$$

где  $C$  — емкость затворов транзисторов;  $V$  — напряжение питания;  $f$  — тактовая частота процессора;  $P_s$  — постоянная потребляемая мощность. Значения  $C$ ,  $V$  и  $P_s$  являются константами для конкретной конфигурации вычислительной системы. Скорость обработки процессором заданий изменяется путем изменения тактовой частоты  $f$  процессора. Пусть, когда число заданий меньше  $T$ , тактовая частота равна  $f_0$ , а когда число заданий больше или равно  $T$ , она равна  $f > f_0$ . Таким образом, введение порога  $T$  позволяет путем изменения его значения регулировать потребляемую мощность  $P$ .



Будем считать, что интенсивность обслуживания заявок есть линейная функция тактовой частоты, т. е.

$$\mu_0(t) = \alpha f_0 a(t); \quad \mu(t) = \alpha f a(t), \quad (8)$$

где  $a(t)$  — неслучайная периодическая функция времени, а коэффициент  $\alpha$  однозначно определяется для каждой модели процессора (например, с помощью процедур, разрабатываемых SPEC [14]). В работе [1] предполагается, что  $a(t) = 1$ . С учетом введенных предположений потребляемая мощность системы  $M(t)/M(t)/1/\infty$  с порогом  $T$  теперь зависит от времени  $t$  и может быть определена следующим образом:

$$P(t) = P(t, f_0, f, T) = C \left[ f_0 \sum_{i=0}^{T-1} p_i(t) + f \left( 1 - \sum_{i=0}^{T-1} p_i(t) \right) \right] V^2 + P_s. \quad (9)$$

Поскольку в (9) имеется зависимость от  $t$ , то представляет интерес рассмотрение усредненного значения  $P(t, f_0, f, T)$ . Предположим, что интересуют характеристики сервера в период  $[t, t + t^*]$  при достаточно большом  $t$ . На практике этот период может быть равен суткам, неделе, году и т. д. после того, как система проработала уже достаточно долгое время. Тогда значение

$$P^*(f_0, f, T) = \frac{1}{t^*} \int_0^{t^*} P(t, f_0, f, T) dt \quad (10)$$

можно интерпретировать как среднее энергопотребление сервера (средняя потребляемая мощность) за период  $[t, t + t^*]$ .

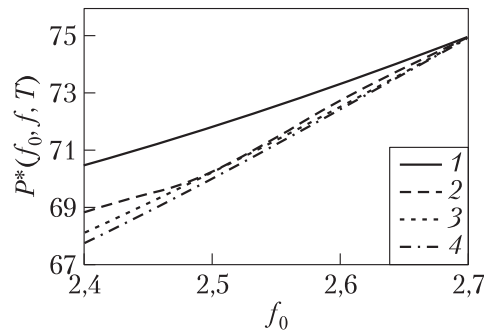
Построим кривые зависимости среднего энергопотребления  $P^*(f_0, f, T)$  от  $f_0$ ,  $f$  и  $T$  (рис. 1–3), а также кривые зависимости энергопотребления  $P(t, f_0, f, T)$  от времени  $t$  и порога  $T$  при  $f_0 = 2,6$  и  $f = 2,7$  (рис. 4). Для этого определим функции  $\lambda(t)$  и  $a(t)$  следующим образом:

$$\lambda(t) = \frac{4}{5} + \frac{1}{100} \sin(2\pi t); \quad a(t) = 1 + \frac{1}{30} \sin(0,04\pi t).$$

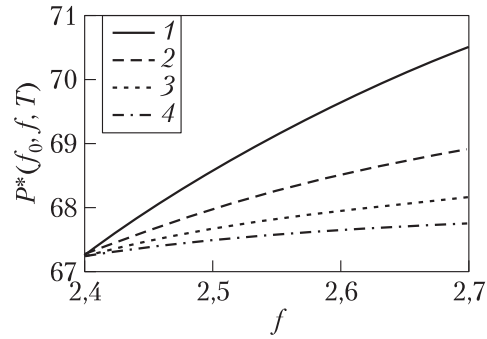
Остальные неизвестные константы во всех случаях примем равными  $\alpha = 0,37$ ,  $V = 1,35$ ,  $P_s = 5$  и  $C = 14,23$ , как это принято и в работе [1]. Таким образом, до- и послепороговая интенсивности обслуживания равны

$$\mu(t) = 0,37f \left( 1 + \frac{1}{30} \sin(0,04\pi t) \right); \quad \mu_0(t) = 0,37f_0 \left( 1 + \frac{1}{30} \sin(0,04\pi t) \right).$$

Как и ожидалось, результаты расчетов среднего энергопотребления  $P^*(f_0, f, T)$  очень близки к тем, что получены в работе [1]. Отметим еще

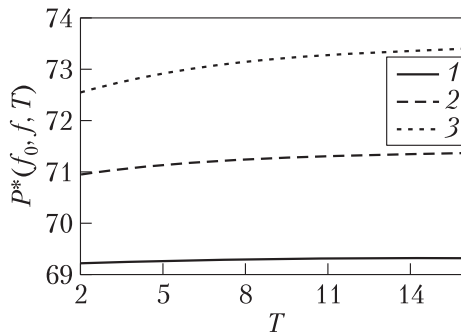


**Рис. 1** Зависимость значения  $P^*(f_0, f, T)$  от  $f_0$  при  $f = 2,7$  и  $T = 5$  (1), 10 (2), 15 (3) и 20 (4)

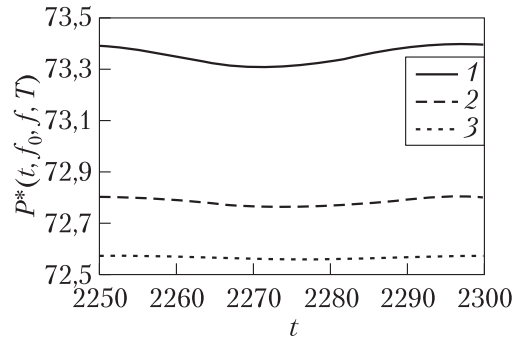


**Рис. 2** Зависимость значения  $P^*(f_0, f, T)$  от  $f$  при  $f_0 = 2,6$  и  $T = 5$  (1), 10 (2), 15 (3) и 20 (4)

раз, как это сделано в [1], некоторые характерные особенности поведения  $P^*(f_0, f, T)$ . Как видно из рис. 1, с ростом  $f_0$  среднее энергопотребление  $P^*(f_0, f, T)$  монотонно возрастает (и, следовательно, среднее время ожидания начала обслуживания, наоборот, убывает). Из рис. 2 следует, что увеличение значения  $f$  вне зависимости от значений  $T$  влечет и рост среднего энергопотребления при фиксированном  $f_0$ . При рассмотрении же  $P^*(f_0, f, T)$  как функции от  $T$ , из рис. 3 видно, что большему значению потребляемой мощности  $f$  соответствует и большее энергопотребление при фиксированном  $T$ . Примечательно, что при уменьшении значения  $f$  (до критического значения, гарантирующего еще существование стационарного режима) среднее энергопотребление стремится к пределу, не зависящему от значения порога  $T$ . Примечательно также, что при отсутствии усреднения, т. е. при рассмотрении  $P(t, f_0, f, T)$  как функции



**Рис. 3** Зависимость значения  $P^*(f_0, f, T)$  от  $T$  при  $f_0 = 2,4$  и  $f = 2,5$  (1); 2,6 (2) и 2,7 (4)



**Рис. 4** Зависимость значения  $P(t, f_0, f, T)$  от  $t \in [2250, 2300]$  при  $f_0 = 2,4$ ,  $f = 2,7$  и  $T = 5$  (1), 10 (2) и 15 (3)

от времени  $t$  и порога  $T$ , наблюдается похожая закономерность:  $P(t, f_0, f, T)$  стремится к пределу с ростом  $T$ . Кроме того, необходимо отметить, что периодичность интенсивностей входящего потока и обслуживания с ростом  $T$  все меньше оказывает влияние на значение энергопотребления.

## Литература

1. Zhang X., Wang J., Do T. V. Threshold properties of the  $M/M/1$  queue under T-policy with applications // *Appl. Math. Comput.*, 2015. Vol. 261. P. 284–301.
2. Benini L., Bogliolo A., De Micheli G. A survey of design techniques for system-level dynamic power management // *IEEE Trans. VLSI Syst.*, 2000. Vol. 8. Iss. 3. P. 299–316.
3. Zikos S., Karatza H. D. Performance and energy aware cluster-level scheduling of compute-intensive jobs with unknown service times // *Simul. Model. Pract. Th.*, 2011. Vol. 19. Iss. 1. P. 239–250.
4. Do T. V., Vu B. T., Tran X. T., Nguyen A. P. A generalized model for investigating scheduling schemes in computational clusters // *Simul. Model. Pract. Th.*, 2013. Vol. 37. P. 30–42.
5. Yadin M., Naor P. Queueing system with a removable service station // *Oper. Res. Quart.*, 1963. Vol. 14. P. 393–405.
6. Lee D. H., Yang W. S. The N-policy of a discrete time Geo/G/1 queue with disasters and its application to wireless sensor networks // *Appl. Math. Model.*, 2013. Vol. 37. Iss. 23. P. 9722–9731.
7. Jayachitra P., Albert A. J. Recent developments in queueing models under N-policy: A short survey // *Int. J. Math. Arch.*, 2014. Vol. 5. Iss. 3. P. 227–233.
8. Balachandran K. R. Control policies for a single server system // *Manage. Sci.*, 1973. Vol. 19. Iss. 9. P. 1013–1018.
9. Artalejo J. R. On the  $M/G/1$  queue with D-policy // *Appl. Math. Model.*, 2001. Vol. 25. Iss. 12. P. 1055–1069.
10. Heyman D. P. The T-policy for the  $M/G/1$  queue // *Manage. Sci.*, 1977. Vol. 23. Iss. 7. P. 775–778.
11. Wang T. Y., Wang K. H., Pearn W. L. Optimization of the T policy  $M/G/1$  queue with server breakdowns and general startup times // *J. Comput. Appl. Math.*, 2009. Vol. 228. Iss. 1. P. 270–278.
12. Zeifman A. I., Satin Ya. A., Korolev V. Yu., Shorgin S. Ya. On truncations for weakly ergodic inhomogeneous birth and death processes // *Int. J. Appl. Math. Comp. Sci.*, 2014. Vol. 24. Iss. 3. P. 503–518.
13. Sueur E. L., Heiser G. Dynamic voltage and frequency scaling: The laws of diminishing returns // *HotPower'10: 2010 Conference (International) on Power Aware Computing and Systems Proceedings*. — Vancouver: USENIX Association, 2010. P. 1–8.
14. Standard Performance Evaluation Corporation. <http://www.spec.org>.

Поступила в редакцию 16.09.15

## **ENERGY EFFICIENCY ESTIMATION OF THE COMPUTATIONAL CLUSTER USING A QUEUEING SYSTEM WITH THRESHOLD CONTROL AND TIME-DEPENDENT SERVICE AND ARRIVAL RATES**

*R. V. Razumchik*<sup>1,2</sup>, *A. I. Zeifman*<sup>1,3,4</sup>, *A. V. Korotysheva*<sup>1,3</sup>, and *Ya. A. Satin*<sup>1,3</sup>

<sup>1</sup>Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

<sup>2</sup>Peoples’ Friendship University, 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation

<sup>3</sup>Vologda State University, 15 Lenin Str., Vologda 160000, Russian Federation

<sup>4</sup>ISED T RAS, 56-A Gorky Str., Vologda 16001, Russian Federation

**Abstract:** Consideration is given to the problem of energy efficiency of the computational cluster (servers), in which energy consumption threshold control policy (by adjusting operating frequency of processors) can be implemented. There are the papers in which Markov queueing systems with (queue-size) thresholds are used to solve such problems. Such approach allows one to compute main stationary performance characteristics of the computational cluster and obtain optimal energy-saving strategies (within the class of threshold strategies). It is usually assumed that service and arrival rates of the underlying processes are stationary. The present authors obtain a generalization of the known results for the case when service and arrival rates are the nonrandom functions of time (in particular, periodic). The authors consider  $M(t)/M(t)/1/\infty$  queue with single-threshold service rate control policy. It is shown how, using general results for inhomogeneous birth-and-death processes, one can obtain relatively simple approximate algorithms for computation of the main (stationary) performance characteristics of the system. The numerical section illustrates the application of the obtained results to estimation of power consumption efficiency of the computational server.

**Keywords:** queueing system; special functions; two-dimensional Markov chain; joint distribution

**DOI:** 10.14357/08696527150402

### **Acknowledgments**

This work was financially supported by the Russian Science Foundation (grant No. 14-11-00397).

## References

1. Zhang, X., J. Wang, and V. T. Do. 2015. Threshold properties of the  $M/M/1$  queue under T-policy with applications. *Appl. Math. Comput.* 261:284–301.
2. Benini, L., A. Bogliolo, and G. De Micheli. 2000. A survey of design techniques for system-level dynamic power management. *IEEE Trans. VLSI Syst.* 8(3):299–316.
3. Zikos, S., and H. D. Karatza. 2011. Performance and energy aware cluster-level scheduling of compute-intensive jobs with unknown service times. *Simulation Modelling Practice and Theory* 19(1):239–250.
4. Do, T. V., B. T. Vu, X. T. Tran, and A. P. Nguyen. 2013. A generalized model for investigating scheduling schemes in computational clusters. *Simul. Model. Pract. Th.* 37:30–42.
5. Yadin, M., and P. Naor. 1963. Queueing system with a removable service station. *Oper. Res. Quart.* 14:393–405.
6. Lee, D. H., and W. S. Yang. 2013. The N-policy of a discrete time  $Geo/G/1$  queue with disasters and its application to wireless sensor networks. *Appl. Math. Model.* 37(23):9722–9731.
7. Jayachitra, P., and A. J. Albert. 2014. Recent developments in queueing models under N-policy: A short survey. *Int. J. Math. Arch.* 5(3):227–233.
8. Balachandran, K. R. 1973. Control policies for a single server system. *Manage. Sci.* 19(9):1013–1018.
9. Artalejo, J. R. 2001. On the  $M/G/1$  queue with D-policy. *Appl. Math. Model.* 25(12):1055–1069.
10. Heyman, D. P. 1977. The T-policy for the  $M/G/1$  queue. *Manage. Sci.* 23(7):775–778.
11. Wang, T. Y., K. H. Wang, and W. L. Pearn. 2009. Optimization of the T policy  $M/G/1$  queue with server breakdowns and general startup times. *J. Comput. Appl. Math.* 228(1):270–278.
12. Zeifman, A. I., Ya. A. Satin, V. Yu. Korolev, and S. Ya. Shorgin. 2014. On truncations for weakly ergodic inhomogeneous birth and death processes. *Int. J. Appl. Math. Comp. Sci.* 24(3):503–518.
13. Sueur, E. L., and G. Heiser. 2010. Dynamic voltage and frequency scaling: The laws of diminishing returns. *HotPower'10: 2010 Conference (International) on Power Aware Computing and Systems Proceedings*. Vancouver: USENIX Association. 1–8.
14. Standard Performance Evaluation Corporation. Available at: <http://www.spec.org> (accessed September 16, 2015).

*Received September 16, 2015*

## Contributors

**Razumchik Rostislav V.** (b. 1984) — Candidate of Science (PhD) in physics and mathematics, senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; associate professor, Peoples’

Friendship University of Russia, 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; rrazumchik@ipiran.ru

**Zeifman Alexander I.** (b. 1954) — Doctor of Science in physics and mathematics, professor, Head of Department, Vologda State University, 15 Lenin Str., Vologda 160000, Russian Federation; senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; principal scientist, ISEDT RAS, 56-A Gorky Str., Vologda 160001, Russian Federation; a\_zeifman@mail.ru

**Korotysheva Anna V.** (b. 1988) — Candidate of Science (PhD) in physics and mathematics, associate professor, Vologda State University, 15 Lenin Str., Vologda 160000, Russia; Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; a\_korotysheva@mail.ru

**Satin Yakov A.** (b. 1978) — Candidate of Science (PhD) in physics and mathematics, associate professor, Vologda State University, 15 Lenin Str., Vologda 160000, Russian Federation; Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; yacovi@mail.ru

## ПРИБЛИЖЕННАЯ ОПТИМИЗАЦИЯ СТРАТЕГИИ РАСПРЕДЕЛЕНИЯ ВЫЧИСЛИТЕЛЬНЫХ РЕСУРСОВ НА ПРИМЕРЕ СИСТЕМЫ АНДЕРРАЙТИНГА\*

*М. Г. Коновалов<sup>1</sup>, Р. В. Разумчик<sup>2</sup>*

**Аннотация:** Рассмотрена проблема оптимизации процесса оценки риска при заключении договоров (андеррайтинг). Процесс моделируется как управляемая система массового обслуживания, обладающая специфическими особенностями: наличием нескольких потоков заданий, имеющих дедлайн; сложной структурой заданий; двухфазным процессом обработки задач; непостоянной и неполной доступностью ресурсов. Ставится задача приближенной максимизации доли заданий, выполненных в срок. Для решения задачи используется методика, которая предполагает создание имитационной модели и последующее применение адаптивных оптимизационных алгоритмов на имитируемых траекториях. На численных примерах показано, что предлагаемый подход позволяет находить наиболее эффективный алгоритм из заданного набора эвристических стратегий. Методика применима в задачах распределения вычислительных ресурсов, для которых трудно получить точное математическое решение и которые возникают в связи с необходимостью многоэтапной обработки информации, неопределенностью сроков выполнения заданий и человеческим фактором.

**Ключевые слова:** управляемые системы массового обслуживания; имитационные модели; адаптивные алгоритмы

**DOI:** 10.14357/08696527150403

### 1 Введение

Системы обработки, хранения и передачи больших объемов информации в реальном или отложенном времени (системы распределенных и облачных вычислений, грид-системы, системы обработки транзакций) играют большую роль в различных областях деятельности, таких как электронная коммерция, банковское обслуживание, авиа- и железнодорожное сообщение. Организационная и экономическая выгода от применения подобных систем во многом зависит от того, обеспечивают ли они необходимый уровень качества обслуживания. По-

\* Работа выполнена при частичной поддержке РФФИ (проект № 15-07-03406).

<sup>1</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, mkonvalov@ipiran.ru

<sup>2</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук; Российский университет дружбы народов, grazumchik@ipiran.ru

скольку ресурсы системы обычно ограничены, а потребности в них изменяются во времени, на практике невозможно обеспечить одновременно оптимальное качество выполнения запросов каждого пользователя системы. Часто пользователи объединены в группы, и тогда естественной целью оптимальной организации работы системы является обеспечение наилучшего качества для групп пользователей в порядке ранжирования по значимости. Оптимальная работа системы зависит не столько от аппаратного обеспечения, подобрать которое обычно не составляет труда, сколько от того, какая логическая структура и какие правила обработки поступающих заданий (алгоритмы управления) лежат в ее основе. Последнее составляет суть задачи оптимального распределения заданий в подобных системах и является предметом настоящей статьи. Общая правильная организация работы системы позволяет ее владельцу предлагать своим пользователям соглашения о качестве обслуживания (Service Level Agreements, SLA), в которых среди прочего фиксируются требования к качественным и количественным характеристикам обслуживания. Требования могут быть как простыми (ограничение на среднее время выполнения запроса по каждой группе), так и сложными (ограничение на отклонение времени выполнения 90% запросов по каждой группе от среднего). Усложнение требований зачастую приводит к усложнению как логической схемы работы системы, так и алгоритмов управления. Можно выделить два подхода, которые используются для нахождения алгоритмов управления, соответствующих заданным требованиям владельца системы. Первый подход связан с использованием экспертных мнений лучших практиков в рассматриваемой области. Не останавливаясь на очевидных достоинствах и недостатках этого подхода, рассмотрим подробнее второй, который заключается в использовании средств математического моделирования.

Системы, о которых шла речь выше, допускают достаточно удобную и гибкую формализацию в рамках широкого класса разработанных к настоящему времени математических моделей. С математической точки зрения описанную выше проблематику можно отнести к области управления потоками заданий в системах вычислительных ресурсов. Суть возникающих задач можно кратко охарактеризовать следующим образом. Требуется определить алгоритмы, согласно которым следует распоряжаться заданиями (запросами от пользователей), поступающими в систему вычислительных ресурсов.

Несмотря на то что задачи распределения заданий обладают многими общими чертами, они отличаются многообразием постановок, вызванным особенностями конкретных систем и ситуаций. За несколько прошедших десятилетий, в течение которых подобным задачам уделялось значительное внимание, было выработано несколько подходов к их решению. Эти подходы используют в различных комбинациях методы теории массового обслуживания, теории расписаний и потокового программирования наряду с общими разделами теории вероятностей и методами оптимизации. Дадим краткую характеристику самым основным из них.



Первый подход заключается в том, что по экспертной оценке выбирается наиболее привлекательная стратегия и делается попытка проверить, является ли она оптимальной. Так, например, обстоит дело со стратегией «назначать задание всегда элементу с наикратчайшей очередью», которая является, в частности, оптимальной в случае одинакового размера всех заданий, пуассоновского входящего потока, экспоненциальных времен обслуживания (с одинаковыми параметрами) на всех элементах с дисциплинами FIFO (first in, first out).

Второй подход заключается в том, что выбирается стратегия, зависящая от параметра, далее находится выражение для целевого функционала и значение параметра, при котором значение функционала оптимально. Примером подобной стратегии является популярная в последнее время стратегия SITA (и ее всевозможные разновидности), согласно которой решение о назначении задания тому или иному элементу принимается только на основе его размера (который считается известным в момент поступления задания в систему). Не останавливаясь на описании этой стратегии [1, 2], отметим, что SITA является оптимальной в случае пуассоновского входящего потока с точки зрения среднего времени ожидания и пребывания в системе.

Третий подход основывается на идеях из области марковских процессов принятия решений. Задается функция стоимости (функция доходов или потерь) и с помощью специальных итерационных процедур ищется наилучшая стратегия. Несмотря на то что данный подход рассматривался в огромном числе работ (см., например, [3–12] и ссылки в них), его основным недостатком остается так называемое «проклятие размерности» (curse of dimensionality), которое удается обойти лишь в некоторых очень частных случаях. В связи с этим в рамках данного подхода уже на протяжении многих лет разрабатываются методы нахождения приближенных решений [12].

Четвертый подход характеризуется привлечением методов статистического (имитационного) моделирования для воспроизведения процесса обслуживания заданий и использования имитационной модели для выбора приближенно оптимального при заданных параметрах и ограничениях алгоритма. Этот, пожалуй, наиболее гибкий из всех описанных подходов используется в настоящее время для разработки различных методов и алгоритмов оптимального управления, о чем свидетельствует большое число работ по данной тематике, появляющихся в периодической печати каждый год [13, 14].

В случаях, когда рассматриваемая система вычислительных ресурсов является сложной и практика диктует необходимость учета многих технических подробностей ее функционирования, подход, связанный с имитационным моделированием, в настоящее время является наиболее продуктивным. В настоящей работе представлен пример использования данного подхода для решения новой масштабной задачи оптимального распределения заданий в системе вычислительных ресурсов. Изучаемая система обладает рядом особенностей:

- наличием нескольких потоков заданий;
- наличием внутренней структуры заданий, состоящих из нескольких задач;
- двухфазным процессом обработки задач;
- специализацией ресурсов на выполнении лишь определенного типа заданий;
- непостоянной во времени доступностью ресурсов.

В совокупности эти особенности делают невозможным получение точного математического решения задачи. Для приближенной оптимизации используется методика, реализующая единый подход к решению обширного круга задач. Она предполагает создание имитационной модели системы вычислительных ресурсов и последующее использование адаптивных оптимизационных алгоритмов на имитируемых траекториях.

Статья организована следующим образом. В разд. 2 приводится краткое описание реальной системы, которая лежит в основе рассматриваемой задачи распределения ресурсов. В разд. 3 описана модель управляемой системы обслуживания и изложена методика нахождения приближенно оптимальной стратегии управления потоками заданий. В разд. 4 содержится численный пример и результаты компьютерного моделирования. В заключении сделаны некоторые выводы.

## 2 Андеррайтинг

Процесс андеррайтинга характерен для любого крупного кредитного учреждения. Андеррайтинг определяется как процесс оценки риска заключения любого договора с физическим или юридическим лицом. Существует два типа андеррайтинга: автоматический и индивидуальный. Ввиду особенностей автоматического андеррайтинга он не представляет, на взгляд авторов, большого интереса с точки зрения моделирования и решения на его основе каких-либо оптимизационных задач. Напротив, организация индивидуального андеррайтинга, применяемого при кредитовании на крупные денежные суммы и требующего взаимодействия нескольких внутренних и внешних служб, оставляет простор для выбора последовательности действий, временных рамок и др. Осуществлять этот процесс исключительно из экспертных соображений сложно и, как правило, неоптимально.

Учитывая, что в настоящее время на процедуры андеррайтинга вышестоящими регулирующими организациями устанавливаются требования по соблюдению так называемых показателей качества, которые фиксируются в SLA, их также необходимо принимать во внимание при принятии решений на любой стадии андеррайтинга. Далее пойдет речь только об индивидуальном андеррайтинге.

Процесс андеррайтинга протекает в рамках единой системы, которая представляет собой совокупность типовых объектов: договоров, работников, внутренних и внешних служб. Суть процесса андеррайтинга состоит в передаче

договоров между работниками<sup>1</sup> и службами. Важно отметить, что процесс андеррайтинга «запускается» по каждому договору в отдельности; он имеет начало, соответствующее моменту поступления договора в кредитное учреждение, и имеет конец, соответствующий моменту принятия решения о заключении или невозможности заключения договора. Показатели качества андеррайтинга устанавливаются самим учреждением (например, минимальное время принятия решения по договору) или вышестоящими регулирующими организациями (например, минимальное время рассмотрения договора о реструктуризации кредита). Для соблюдения этих требований необходимо отыскать соответствующие способы (алгоритмы) управления объектами системы или их действиями.

В рассматриваемой системе не всем можно управлять. Например, нельзя управлять моментами поступления договоров в кредитное учреждение, однако можно управлять правилами выбора очередного договора для рассмотрения. Нельзя управлять временем рассмотрения договора внешней службой, но можно управлять временем рассмотрения договора внутренней службой или сотрудником, а также управлять правилом выбора сотрудника для рассмотрения договора (если по каждому сотруднику известно, как он работает с каждым типом договоров).

Некоторые свойства объектов системы невозможно полностью определить. К примеру, работник учреждения может заболеть, и поэтому время рассмотрения договора увеличится на неопределенную величину. Внутренняя или внешняя служба может задержать рассмотрение договора, а поскольку на это обстоятельство работник имеет ограниченное влияние, время рассмотрения договора опять же возрастет. Таким образом, информация о рассматриваемой системе страдает неполнотой. Некоторые параметры могут быть определены только вероятностным образом. Например, время рассмотрения договора в Федеральной налоговой службе есть случайная величина с неизвестным распределением, которое оценивается исходя из статистических данных кредитного учреждения об истории взаимодействия с Федеральной налоговой службой. Однако структура системы (число объектов, правила рассмотрения договоров) известна заранее и не меняется в процессе андеррайтинга.

Легко усмотреть аналогию между системой андеррайтинга и системой распределения вычислительных ресурсов. Действительно, с вычислительным ресурсом можно ассоциировать работников и службы. Каждому заданию, поступающему в систему и состоящему из задач, ставится в соответствие договор. Задача соответствует процедуре рассмотрения договора исполнителем или какой-либо службой, а сама процедура рассмотрения есть обработка задачи на ресурсе системы. После завершения всех обязательных для договора процедур (задач), решение по нему (задание) считается принятым.

---

<sup>1</sup>Работник — инспектор банка — должен определить вероятность исполнения клиентом условий договора (например, погашения заемщиком выдаваемого кредита) полностью и в срок исходя из информации, содержащейся в заявлении клиента.

В следующем разделе показано, как проблема управления договорами в системе андеррайтинга формулируется и решается в терминах математической модели и задачи оптимального управления.

### 3 Описание модели управляемой системы обслуживания

#### 3.1 Основные параметры модели

Общей содержательной интерпретацией модели является компьютерная или человеко-компьютерная система, в которую поступают случайные потоки рабочей нагрузки. Эта нагрузка должна быть распределена между имеющимися обслуживающими ресурсами и выполнена на этих ресурсах. Под обслуживающим ресурсом подразумевается человек или устройство, способное выполнять определенную работу, в частности компьютер (процессор). Механизм распределения нагрузки между ресурсами является предметом оптимизации. Далее в этом разделе приводится описание составных элементов модели и взаимодействия между ними.

Основными категориями объектов являются «потоки» и «ресурсы».

Поток, или поток заданий, представляет собой рекуррентный случайный поток событий, причем каждое событие означает появление ровно одного нового задания. Каждое задание сопровождается указанием срока (дедлайна), в течение которого его следует выполнить. Задание состоит из упорядоченного набора задач. Выполнение задания заключается в последовательном решении всех составляющих его задач.

Задача является минимальной единицей рабочей нагрузки. Выполнение задачи состоит из двух фаз:

- фазы подготовки ( $I$ ), которая не требует занятия обслуживающего ресурса;
- фазы решения ( $II$ ).

Задача характеризуется двумя параметрами: временем на подготовку и объемом работы, определяющим (вместе с производительностью ресурса, на котором выполняется задача) время решения. Соответственно, тип задачи задается парой вероятностных распределений, устанавливающих в модели значения этих параметров для каждой задачи:

$$t = (D^I, D^{II}) .$$

Задания, поступающие из одного потока, являются однотипными. Тип задания определяется набором

$$j = (d, t_1, \dots, t_k) ,$$

в котором  $d > 0$  означает дедлайн, а  $t_i$  — типы последовательных задач, общее количество которых  $k$  является фиксированным для конкретного типа заданий.

Тип потока  $f$  характеризуется функцией распределения времени между поступлениями заданий  $D$ , а также типом поступающих заданий  $j$ ,  $f = (D, j)$ . Всего в модели имеется  $M$  потоков заданий, типы которых составляют множество  $\mathbb{F} = \{f_1, \dots, f_M\}$ .

Ресурс, или сервер, предназначен для выполнения задач в фазе решения. Задачи для решения выбираются из тех заданий, которые закреплены за сервером. На сервере одновременно может выполняться ограниченное число заданий. Это ограничение называется емкостью ресурса, обозначается буквой  $c$  и, вообще говоря, различно для разных ресурсов. Подготовленная к решению задача занимает ресурс на время, необходимое для выполнения связанного с ней объема работы. В каждый момент времени ресурс может быть занят решением только одной задачи. Время непосредственного решения задач определяется его производительностью, которая обозначается  $r > 0$ . Для задачи с объемом работы  $v$  это время равно  $v : r$ . Производительность ресурсов в общем случае различна. Предполагается, что ресурс периодически и самопроизвольно переходит из работоспособного состояния в состояние бездействия и обратно. Период работоспособности обозначается  $T_0$ , а период бездействия —  $T_1$ . Тип ресурса  $\tau$  характеризуется набором параметров

$$\tau = (c, r, T_0, T_1) .$$

Всего в модели имеется  $N$  ресурсов, типы которых составляют множество  $\mathbb{R} = \{\tau, \dots, \tau_N\}$ .

Совокупность  $\mathbb{M} = (\mathbb{F}, \mathbb{R})$  относится к набору входных параметров модели.

### 3.2 Управление в модели

Динамика модели определяется стратегией распределения заданий. Эта стратегия понимается как совокупность правил и алгоритмов, согласно которым производятся определенные управленческие действия по размещению заданий на ресурсах и установлению очередности их выполнения. Такие действия, как правило, вариативны, что и вызывает потребность в оптимизации стратегии. В качестве цели, которую должна обеспечить стратегия управления, выберем максимизацию среднего числа выполненных в срок заданий.

Исходя из описания прикладного прообраза модели, которое было сделано в разд. 2, предполагается, что обслуживание заданий подчинено следующим правилам.

- П1. Все задания, поступающие в систему, принимаются к выполнению. Ограничение доступа в данной модели не рассматривается.
- П2. Каждый ресурс, вообще говоря, может выполнять задания только определенных типов. (В этом смысле будем говорить о ресурсах, «подходящих» для данного задания.) Перечень подходящих ресурсов для каждого потока дополняет список параметров модели.

П3. Каждое задание «закрепляется» за каким-либо подходящим ресурсом. Перемещение и обмен заданий между ресурсами не допускаются.

П4. Фаза подготовки задачи, так же как и фаза непосредственного решения, может быть начата только в работоспособном состоянии ресурса.

П5. Фаза подготовки задачи, начавшись, не прерывается, не зависит от состояния ресурса и не влияет на него. Подготовленная задача в любой момент готова к непосредственному решению, но может также ожидать неопределенное время.

П6. Фаза непосредственного решения задачи прерывается, если начался период бездействия сервера. По окончании этого периода решение задачи возобновляется. Рабочая нагрузка, обрабатываемая в ходе непосредственного решения задачи, предполагается делимой. Это означает, что объем работы в момент возобновления решения меньше первоначального на величину, равную части, выполненной до прерывания.

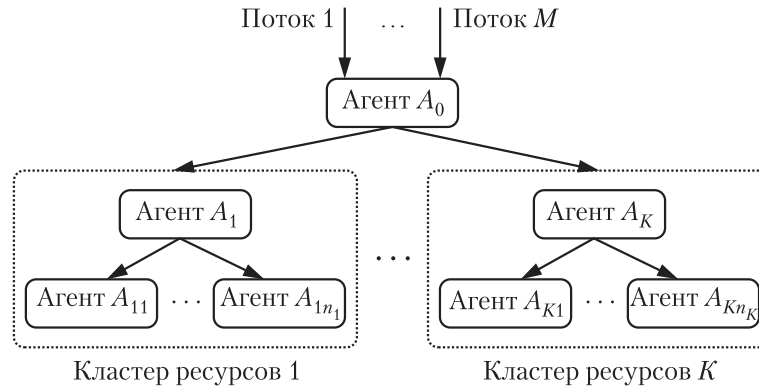
Поскольку все без исключения задания допускаются в систему, в процессе обслуживания заданий остаются две основные ситуации, когда требуется принять решение по выбору того или иного распределяющего действия. Это

- (1) назначение ресурса для выполнения задания;
- (2) выбор очередной задачи для непосредственного решения после освобождения сервера от предыдущей задачи.

Априори в обеих ситуациях представляется достаточно широкий простор для действий. Например, можно по-разному выбирать момент прикрепления задания к ресурсу: назначать ресурс одновременно с поступлением задания или же откладывать решение на некоторый срок. Сам ресурс может выбираться с помощью различных алгоритмов, основанных, как подсказывает интуиция, на учете производительности и текущей загруженности серверов. Аналогичная ситуация с выбором очередной задачи для загрузки ресурса. Хотя момент принятия решения в данном случае очевиден — в силу неэффективности простоя сервера это надо делать немедленно, как только сервер освободился, алгоритм выбора неоднозначен. В общем случае выбор должен зависеть от имеющегося списка подготовленных к решению задач и таких показателей, как объем работы и время до наступления дедлайна.

Чтобы несколько убавить разнообразие и, соответственно, неопределенность в возможностях распределения заданий между ресурсами, примем еще одно условие в дополнение к правилам П1–П6.

П7. Все ресурсы системы разбиты на  $K$  непересекающихся множеств, которые будем называть кластерами. Количество ресурсов  $n_k$ ,  $k = 1, \dots, K$ , в кластерах может быть неодинаковым, а состав ресурсов — неоднородным. Потребуем, чтобы каждое поступающее задание было немедленно распределено в один из кластеров. При этом если в кластере есть ресурсы, которые



**Рис. 1** Трехуровневая структура стратегии распределения заданий

могут принять задание, то задание обязательно закрепляется за одним из серверов. В ином случае задание остается в накопителе выбранного кластера и в дальнейшем может быть выполнено только на ресурсах этого кластера.

В рамках описанной модели и с учетом предположений П1–П7 общая стратегия  $A$  распределения заданий между ресурсами естественным образом представляется как агрегированная из отдельных компонент, которые также будем обозначать буквой  $A$  с соответствующими индексами и для краткости называть агентами. Структура общей стратегии представлена на рис. 1.

Опишем кратко роль отдельных агентов — компонент стратегии — в процессе управления.

Агент  $A_0$  распределяет каждое вновь поступившее задание в один из кластеров. Распределение происходит немедленно в момент поступления задания.

Агенты  $A_k$ ,  $k = 1, \dots, K$ , распределяют задания среди ресурсов внутри кластеров. Каждый такой агент, в свою очередь, содержит две компоненты:  $A_k = (A'_k; A''_k)$ . Агент  $A'_k$  отвечает за выбор ресурса в момент поступления задания в кластер (наступающий по решению агента  $A_0$ ). Согласно П7, если все ресурсы заняты до предела своих емкостей  $c$ , то задание помещается в накопитель данного кластера. Агент  $A''_k$  принимает решения в ситуации, когда накопитель не пуст, а на каком-либо сервере появилось свободное место.

Агент  $A_{ki}$  влияет исключительно на поведение ресурса  $i$  из кластера  $k$ . Его назначение — выбор очередной задачи из числа задач, готовых к решению (если таковые имеются), в моменты, когда вычислительные мощности сервера освобождаются от предыдущей задачи.

Стратегия распределения ресурсов содержит  $1 + K + N$  агентов и символически записывается в виде:

$$A = (A_0, A', A'', \tilde{A}), \quad (1)$$



где  $A' = (A'_1, \dots, A'_K)$ ;  $A'' = (A''_1, \dots, A''_K)$ ;  $\tilde{A} = (A_{ki}, i = 1, \dots, n_K; k = 1, \dots, K)$ . Такая запись позволяет говорить о четырех составляющих стратегии  $A$ .

Реализация перечисленных составляющих и компонент стратегии описана в следующем подразделе.

### 3.3 Компоненты стратегии распределения заданий

Поскольку для изложенной модели нет теоретических результатов, которые можно было бы применить для построения алгоритмов распределения потоков заданий, то остается путь использования эвристических конструкций и статистического моделирования для нахождения наилучшей стратегии. Перечислим некоторые простые эвристики, на базе которых будут строиться алгоритмы для агентов.

**Равновероятный выбор.** Этот алгоритм является тривиальным и рассматривается в качестве «точки отсчета», с которой можно сравнивать другие, более эффективные алгоритмы. Его применение подразумевает наличие конечного множества, из которого осуществляется равновероятный выбор элемента (действия). Для обозначения алгоритма используется символ  $\alpha$ .

**Случайный выбор** элемента из множества осуществляется с помощью вероятностного распределения, заданного на этом множестве, причем это распределение не обязательно равномерное. Этот способ обозначается символом  $\beta$ . Заметим, что на самом деле речь идет о семействе алгоритмов, параметризованном указанным распределением и, возможно, способом трансформации этого распределения, если таковое предусмотрено.

**Выбор по наименьшей нагрузке.** Предполагается, что выбор осуществляется из некоторого набора ресурсов или из набора групп ресурсов. Под нагрузкой понимается некоторая числовая характеристика ресурса или группы ресурсов, минимальное значение которой является основанием для выбора соответствующего элемента. Поскольку обычно нагрузку можно определять разными способами, то этот метод, по существу, представляет собой семейство алгоритмов, параметризованное типом назначенной характеристики нагрузки. Эвристика обозначается символом  $\gamma$ .

**Выбор по наименьшей нагрузке с поправками.** Идея этого алгоритма основана на следующем замечании, которое справедливо для некоторых задач распределения нагрузки и в ряде случаев имеет теоретическое обоснование [14]. Оказывается, что среди ресурсов с одинаковой нагрузкой, как бы она ни определялась, выгоднее выбирать более производительный ресурс. Частое использование этого обстоятельства сводится к тому, что к вычисляемой фактической нагрузке добавляется фиктивная константа, которая тем больше, чем меньше производительность ресурса. Такие алгоритмы обычно называют пороговыми. В данной работе используются имеющие аналогичный смысл мультипликативные поправки. Это означает, что выбор ресурса или группы ресурсов осуществляется



на основе массива нагрузок, умноженных на некоторые числовые коэффициенты. В данном случае опять имеем дело с семейством алгоритмов, параметризованных набором поправочных коэффициентов. Обозначение для эвристики —  $\delta$ .

**Выбор в зависимости от запаса времени.** В этом случае речь идет о выборе задания или задачи из числа находящихся в системе. Запас времени — это время до наступления дедлайна минус оценка оставшегося времени выполнения задания. Если выбирать, например, задание с бóльшим запасом времени, или задачу (относящуюся к такому заданию), то можно рассчитывать на бóльшую вероятность того, что задание будет выполнено в срок. Противоположный подход, т. е. выбор заданий с наименьшим запасом времени, дает преимущество заданиям, находящимся в критическом по срокам положении. Эвристика обозначается символом  $\varepsilon$ .

Далее речь будет о выборе той или иной эвристики для компонент стратегии, названных агентами. Для обозначения алгоритма, выбранного для агента, используются те же греческие буквы  $\alpha, \beta, \dots$ , но с индексами, повторяющими либо индексы агента, который работает в соответствии с данным алгоритмом, либо индексы одной из четырех составляющих в (1). Например, запись  $A' = \alpha'$  означает, что все агенты  $A'_1, \dots, A'_K$  реализованы в виде равновероятного выбора.

### 3.4 Методика оптимизации

Оптимизация проводится на основе компьютерного моделирования, поэтому первоначальным шагом является создание формальной, а затем программной имитационной модели. Математический аппарат для имитационного моделирования обсуждался в [15, 16]. Формальная имитационная модель строится на основе определенного в подразд. 3.1 множества параметров  $M$  и перечисленных в подразд. 3.2 правил П1–П7, которые регулируют взаимодействие заданий и ресурсов. Построение имитационной модели предполагает, в частности, написание специфических уравнений, задающих все потенциально возможные траектории событий, которые целесообразно рассматривать в модели с точки зрения изучаемой проблемы, такие как поступление заданий, занятие и освобождение серверов, потеря и восстановление работоспособности ресурсов и пр. Отметим, что программная реализация полностью воспроизводит описанный в этом разделе механизм взаимодействия заданий и ресурсов.

Компьютерная имитационная модель позволяет получать для любого фиксированного алгоритма распределения ресурсов  $A = A$  выборочные траектории<sup>1</sup> и строить оценку эффективности стратегии  $w(A)$ , определяемую как долю заданий, выполненных до наступления дедлайна. Это дает возможность проводить

<sup>1</sup>Для приводимого далее в разд. 4 примера, содержащего 5 потоков и 55 ресурсов, практически достаточная для статистических оценок траектория длиной в  $10^7$  событий моделируется на обычном персональном компьютере за время порядка 1,5 мин.

оптимизацию, последовательно проводя эксперименты с разными алгоритмами для  $A$  и выбирая среди них алгоритм с наилучшей оценкой. Есть, однако, причины, по которым такой простой и надежный способ (назовем его «простым перебором») не всегда пригоден и, в частности, неудовлетворителен в рассматриваемой ситуации.

Оценка эффективности стратегии, в принципе, может быть сделана сколь угодно точной, однако точность зависит от длины наблюдаемой траектории и определяется скоростью сходимости в законе больших чисел. Может оказаться, что для получения одной оценки требуется много времени. Тогда, особенно если число вариантов велико, простой перебор будет неприемлем. Также затруднение с простым перебором возникает, когда множество испытываемых алгоритмов для стратегии  $A$  бесконечно. Например, рассмотренные в предыдущем подразделе алгоритмы  $\beta$  и  $\delta$  зависят от многомерного непрерывного параметра, для которого выбор значения совсем не очевиден.

Существует подход, который позволяет заменить простой перебор, использующий для оценки одного варианта стратегии отдельную траекторию, некоторым алгоритмом, действующим на единственной траектории и достигающим той же цели выбора. Суть подхода (назовем его «адаптивным перебором») в общем виде и без строгих формулировок заключается в следующем.

Пусть имеется некоторое множество  $A = \{\alpha\}$  алгоритмов для реализации стратегии  $A$ , согласно которой должна управляться некоторая система, и пусть  $\sigma_A$  — вероятностное распределение на множестве  $A$ .

Рассмотрим реализацию стратегии  $A$ , определяемую следующими условиями ( $\sigma_A$ -алгоритм):

- (1) каждый раз в момент принятия решения разыгрывается распределение  $\sigma_A$  и выбирается элемент  $\alpha$  из множества  $A$ ;
- (2) решение в этот момент принимается с помощью алгоритма  $\alpha$ , т. е. так, как если бы выполнялось условие  $A = \alpha$ .

Адаптивный перебор — это  $\sigma_A$ -алгоритм, в котором распределение  $\sigma_A$  трансформируется в процессе управления на основе наблюдений за траекторией, причем таким образом, что вероятность «наиболее эффективных» с точки зрения выбранной цели элементов из  $A$  увеличивается. В [17] содержатся элементы теории адаптивного перебора, а также некоторые приемы конструктивного решения проблемы трансформации управляющего распределения  $\sigma_A$ .

Применительно к рассматриваемой модели адаптивный перебор использован в двух ситуациях:

- (1) для сокращения перебора при нахождении наилучшей комбинации значений компонент стратегии;
- (2) для оптимизации параметров в алгоритмах типа  $\beta$  и  $\delta$ .

## 4 Численный пример

### 4.1 Параметры модели

Рассмотрим пример приближенной оптимизации стратегии распределения заданий с точки зрения числа выполненных в срок заданий. В систему, содержащую ресурсы пяти типов ( $N = 5$ ), поступают потоки заданий пяти типов  $M = 5$ . Потоки заданий — пуассоновские. Все задания состоят из трех задач. Во всех задачах независимо от типа заданий:

- распределение времени на подготовку — Парето с параметрами  $x_{\min}$  и  $a$ ,  $D^I = P(x_{\min}, a)$ ;
- распределение времени непосредственного решения — нормальное с параметрами  $m$  и  $s$ ,  $D^{II} = N(m, s)$ .

Всего в системе 5 потоков заданий (все разного типа) и 55 ресурсов. Входные параметры модели перечислены в табл. 1–4 (единица измерения времени — минута).

**Таблица 1** Типы задач

Тип задачи	Время на подготовку, $D^I = P(x_{\min}, a)$	Время решения, $D^{II} = N(m, s)$
$t_1$	$x_{\min} = 800, a = 6$ , (среднее 960)	$m = 960, s = 20$
$t_2$	$x_{\min} = 480, a = 3$ , (среднее 720)	$m = 480, s = 20$
$t_3$	$x_{\min} = 320, a = 3$ , (среднее 480)	$m = 240, s = 10$
$t_4$	$x_{\min} = 160, a = 3$ , (среднее 240)	$m = 480, s = 1$
$t_5$	$x_{\min} = 960, a = 3$ , (среднее 120)	$m = 120, s = 1$

**Таблица 2** Типы заданий

Тип задания	Типы задач в задании	Дедлайн, $d$
$j_1$	$t_1, t_2, t_3$	8640
$j_2$	$t_1, t_2, t_5$	7200
$j_3$	$t_1, t_3, t_4$	5760
$j_4$	$t_2, t_3, t_4$	5760
$j_5$	$t_2, t_3, t_5$	4320

**Таблица 3** Типы потоков

Тип потока	Интенсивность	Тип заданий	Подходящие типы ресурсов
$f_1$	0,0138 ( $\approx 20$ в сутки)	$j_1$	$r_2; r_4$
$f_2$	0,000694 ( $\approx 1$ в сутки)	$j_2$	$r_1$
$f_3$	0,0208 ( $\approx 30$ в сутки)	$j_3$	$r_1; r_3$
$f_4$	0,00694 ( $\approx 10$ в сутки)	$j_4$	$r_4$
$f_5$	0,027 ( $\approx 40$ в сутки)	$j_5$	$r_2; r_5$

**Таблица 4** Типы ресурсов

Тип ресурса	Емкость, $c$	Производительность, $r$	Расписание (работа, $T_0$ /простой, $T_1$ )	Количество ресурсов данного типа
$\tau_1$	4	3,0	720/720	5
$\tau_2$	3	2,5	600/840	10
$\tau_3$	3	2,0	600/840	10
$\tau_4$	2	1,5	600/840	10
$\tau_5$	2	1,0	480/960	20

## 4.2 Кластеры из однородных ресурсов

Пусть ресурсы разбиты на 5 кластеров ( $K = 5$ ), причем кластер  $i$  содержит все ресурсы, имеющие тип  $\tau_i$ , так что  $n_1 = 5$ ,  $n_2 = n_3 = n_4 = 10$ ,  $n_5 = 20$ .

Каждая стратегия распределения заданий является многоуровневой (в соответствии с рис. 1) и описывается четырьмя типами агентов в соответствии с (1). Алгоритмы для агентов многовариантны в соответствии с эвристиками, описанными в подразд. 3.3.

При реализации эвристик  $\gamma$  и  $\delta$  нагрузка на ресурс определяется либо как число заданий, либо как оценка «остаточного времени», т.е. среднего времени, оставшегося до полного выполнения рабочей нагрузки. При вычислении оценки остаточного времени делается предположение о том, что первые моменты распределений  $D^I$  и  $D^{II}$  известны. Нагрузка на кластер определяется как сумма нагрузки по всем ресурсам, входящим в кластер.

В рамках принятой структуры общей стратегии и выбранного набора эвристических алгоритмов проблема оптимизации сводится к двум подзадачам: необходимо выбрать лучшее семейство алгоритмов для каждого агента и в случае необходимости найти значения многомерных параметров, выделяющих из выбранного семейства конкретную лучшую модификацию. Обе подзадачи решаются с существенным использованием упомянутого в подразд. 3.3 адаптивного перебора.

Точкой отсчета для сравнения стратегий служит стратегия  $A = A_0 = \{\alpha_0, \alpha', \alpha'', \tilde{\alpha}\}$ , в которой все решения принимаются случайно и равновероятно. Для нее получено значение целевой функции  $w(A_0) = 58,50$ . (Здесь и далее значения эффективности стратегий, под которой понимается доля выполненных в срок заданий, приведены в процентах.)

Наилучшая стратегия, которую удалось получить, имеет символическое обозначение  $A_{\text{опт}} = \{\delta_0, \gamma', \varepsilon''_{\min}, \tilde{\varepsilon}_{\min}\}$ . Ее компоненты означают следующее.

Выбор кластера для каждого нового задания (функция агента  $A_0$ ) осуществляется по минимальному значению нагрузки на кластер, умноженному на поправочный коэффициент. При этом нагрузка определяется как остаточное

время. Поправочные коэффициенты достаточно было определить только для потоков 0, 2 и 4, поскольку потоки 1 и 3 направляются однозначно в кластеры 0 и 3 соответственно. Поток 0 может быть направлен либо в кластер 1 (коэффициент принят за 1), либо в кластер 3 (найденное оптимальное значение коэффициента — 1,40). Поток 2 может быть направлен либо в кластер 0 (коэффициент принят за 1), либо в кластер 2 (найденное оптимальное значение коэффициента — 1,06). Поток 4 может быть направлен либо в кластер 1 (найденное оптимальное значение коэффициента — 1,72), либо в кластер 3 (коэффициент принят за 1).

Выбор ресурса для нового задания, уже распределенного в один из кластеров  $k$  (функция агентов  $A'_k$ ), также осуществляется по минимальному значению остаточного времени, но без поправочных коэффициентов. Поправки не нужны, поскольку ресурсный состав кластеров в данном случае однороден.

Выбор задания из накопителя при появлении свободного места в кластере  $k$  (функция агентов  $A''_k$ ) осуществляется по минимальному значению запаса времени.

Выбор задачи внутри ресурса  $i$  из кластера  $k$  в момент освобождения ресурса (функция агентов  $A_{ki}$ ) осуществляется по минимальному значению запаса времени у задания, к которому относится задача.

Для стратегии  $A_{\text{опт}}$  получено значение целевой функции  $w(A_{\text{опт}}) = 82,17$ , что на 40,5% лучше равновероятного выбора.

Интересно сравнить эффективность «крайних» стратегий с показателями некоторых «промежуточных» стратегий. Их список содержится в табл. 5, а сравнительная эффективность показана на рис. 2.

Стратегии  $A_1$ – $A_5$  структурно близки к наилучшей стратегии и каждая из них отличается от  $A_{\text{опт}}$  всего одной компонентой. Например, отличие  $A_3$  от  $A_{\text{опт}}$  в том, что алгоритм выбора кластера не использует поправочных коэффициентов. Остальные четыре стратегии получаются из  $A_{\text{опт}}$ , если поочередно ровно одну из компонент заменить на равновероятный выбор. Сравнивая стратегии, можно сделать следующие выводы.

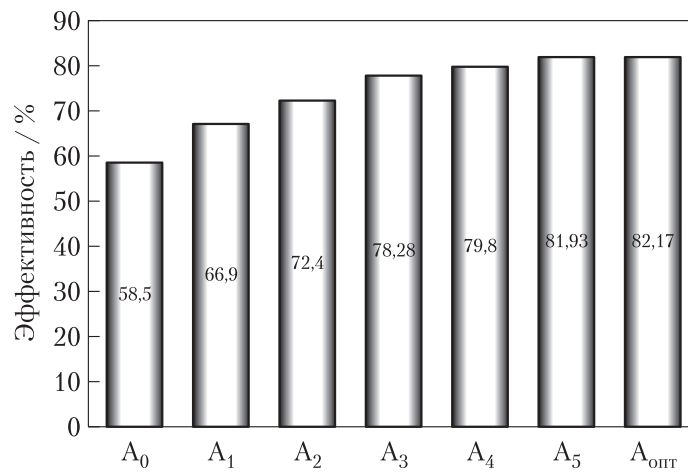
Наиболее существенный вклад в оптимизацию вносят первые две составляющие стратегии (1),  $A_0$  и  $A'$ , которые определяют выбор ресурса для вновь поступившего задания. Их замена на равновероятный выбор уменьшает целевую функцию соответственно на 18,6% и 11,9%.

Влияние на долю выполненных в срок заданий двух других компонент, отвечающих за загрузку сервера в моменты освобождения от предыдущего задания или от предыдущей задачи, значительно меньше: соответственно около 2,8% и 0,3%. Этого, однако, достаточно, чтобы говорить об управляемости системы по всем указанным в (1) составляющим.

Стратегия  $A_3$ , действующая без поправочных коэффициентов и, соответственно, построенная без применения достаточно изощренного метода адаптивного выбора, «всего» на 4,7% менее эффективна, чем наилучшая стратегия.

**Таблица 5** Описание вариантов стратегий

Условное обозначение	$A_0$	Все $A'_k$	Все $A''_k$	Все $A_{ki}$
$A_1 = \{\alpha_0, \gamma', \varepsilon'', \tilde{\varepsilon}\}$	Равновероятно	По минимуму остаточного времени	По минимальному запасу времени	По минимальному запасу времени
$A_2 = \{\delta_0, \alpha', \varepsilon'', \tilde{\varepsilon}\}$	По минимуму остаточного времени с адаптивным выбором поправочных коэффициентов	Равновероятно		
$A_3 = \{\gamma_0, \gamma', \varepsilon'', \tilde{\varepsilon}\}$	По минимуму остаточного времени			Равновероятно
$A_4 = \{\delta_0, \gamma', \varepsilon'', \tilde{\alpha}\}$	По минимуму остаточного времени с адаптивным выбором поправочных коэффициентов			
$A_5 = \{\delta_0, \gamma', \alpha'', \tilde{\varepsilon}\}$		Равновероятно		По минимальному запасу времени



**Рис. 2** Эффективность различных стратегий

### 4.3 Кластеры из неоднородных ресурсов

В табл. 6 указан состав пяти кластеров, на которые разбито множество ресурсов, описанное в подразд. 4.1.

Таким образом,  $n_1 = 11$ ,  $n_2 = 8$ ,  $n_3 = 13$ ,  $n_4 = 9$ ,  $n_5 = 14$ .

Для стратегии равновероятного выбора  $A_0 = \{\alpha_0, \alpha', \alpha'', \tilde{\alpha}\}$  получена оценка эффективности  $w(A_0) = 65,46$ , что на 11,9% больше, чем в предыдущем случае.

Лучшая стратегия, которую удалось найти, имеет структуру, отличающуюся первыми тремя компонентами от предыдущего случая. Символическое представление ее выглядит как  $A_{\text{опт}} = \{\gamma_0, \delta', \varepsilon'', \tilde{\varepsilon}_{\min}\}$ , а содержательно эта запись имеет следующую интерпретацию.

Выбор кластера (первая составляющая) происходит согласно алгоритму минимальной нагрузки. При этом оказалось, что, во-первых, эффективнее оценивать нагрузку по количеству заданий, чем по остаточному времени, и, во-вторых, что поправочные коэффициенты в этом случае не дают дополнительного выигрыша.

Выбор ресурса внутри кластера (вторая составляющая) осуществляется по минимуму остаточного времени, но, в отличие от подразд. 4.2, поправочные коэффициенты дали заметный выигрыш. В данном случае этих коэффициентов больше и они зависят не только от типа потока, но и от кластера.

Выбор нового задания в случае появления свободного места на сервере происходит по-разному в разных кластерах. В 1-м и 5-м кластерах оказалось выгоднее выбирать задание с наибольшим запасом времени, а во 2–4 кластерах — с наименьшим.

Эффективность лучшей стратегии оценена как  $w(A_{\text{опт}}) = 83,36$ , что на 27,3% больше, чем эффективность равновероятного выбора. Это также на 6,5% больше, чем для лучшей из найденных стратегий, не использующих поправочные коэффициенты.

Обращает внимание, что в одной и той же по составу ресурсов системе лучшие стратегии, которые удалось получить при различных вариантах разбиения на кластеры, имеют разную эффективность. Это означает, что разбиение может служить одним из инструментов оптимизации стратегии, если оно не является жестко предписанным.

**Таблица 6** Неоднородный состав кластеров

Кластер	Тип ресурсов				
	1	2	3	4	5
1	1	2	2	2	4
2	1	3	4	0	0
3	1	3	2	3	4
4	1	1	2	3	2
5	1	1	0	2	10

## 5 Заключение

В статье изложена методика нахождения приближенной стратегии распределения ресурсов в сложной системе обслуживания на примере андеррайтинга — процесса оценки риска заключения договора с физическим или юридическим лицом. Проблему эффективности обслуживания предлагается решать путем создания компьютерной модели и последующей численной оптимизации на имитируемых траекториях системы.

Рассмотренная система обладает большим числом специфических особенностей, связанных с гетерогенностью ресурсов, периодичностью их работы, сложной структурой выполняемых заданий, жесткими правилами обслуживания, большой размерностью. Из-за отсутствия теоретических результатов в основу алгоритмов управления кладутся эвристические соображения. При этом возникает задача выбора наилучшей эвристической стратегии и оптимизации ее параметров. Такая задача решается путем использования оригинальных приемов адаптивного выбора вариантов.

Результаты численных экспериментов еще раз, наряду с другими работами в этой области, показывают, что статистическое моделирование в сочетании с адаптивными алгоритмами позволяет решать некоторые значимые с практической точки зрения задачи оптимизации и особенно полезно, когда получение точного математического решения невозможно.

Практическая реализация предполагает дальнейшее усложнение модели. В частности, применительно к андеррайтингу потребуется учесть дополнительные факторы описания объектов, увеличить размерность (до порядка тысячи ресурсов) и рассмотреть более сложные критерии эффективности. Описанный подход сохраняет все свои возможности и в этих условиях и, более того, представляется почти единственным способом численной оптимизации.

## Литература

1. *Crovella M. E., Harchol-Balter M., Murta C. D.* Task assignment in a distributed system: Improving performance by unbalancing load // SIGMETRICS'98 Proceedings. — Madison, WI, USA, 1998. P. 268–269.
2. *Harchol-Balter M., Crovella M. E., Murta C. D.* On choosing a task assignment policy for a distributed server system // J. Parallel Distr. Com., 1999. Vol. 59. P. 204–228.
3. *Krishnan K. R., Ott T. J.* State-dependent routing for telephone traffic: Theory and results // IEEE Conference on Decision and Control, 1986. Vol. 25. P. 2124–2128.
4. *Krishnan K. R.* Markov decision algorithms for dynamic routing // IEEE Commun. Mag., 1990. Vol. 28. Iss. 10. P. 66–69.
5. *Krishnan K. R.* Joining the right queue: A state-dependent decision rule // IEEE Trans. Automat. Contr., 1990. Vol. 35. Iss. 1. P. 104–108.
6. *Sassen S. A. E., Tijms H. C., Nobel R. D.* A heuristic rule for routing customers to parallel servers // Stat. Neerl., 1997. Vol. 51. Iss. 1. P. 107–121.



7. Van Leeuwen J., Aalto S., Virtamo J. Load balancing in cellular networks using first policy iteration. — Helsinki: Helsinki University of Technology, Networking Laboratory, 2001. Technical Report.
8. Bhulai S. On the value function of the  $M/\text{Cox}(r)/1$  queue // J. Appl. Probab., 2006. Vol. 43. Iss. 2. P. 363–376.
9. Hyytia E., Virtamo J., Aalto S., Penttinen A.  $M/M/1$ -PS queue and size-aware task assignment // Perform. Evaluation, 2011. Vol. 68. Iss. 11. P. 1136–1148.
10. Hyytia E., Penttinen A., Aalto S. Size- and state-aware dispatching problem with queue-specific job sizes // Eur. J. Oper. Res., 2012. Vol. 217. Iss. 2. P. 357–370.
11. Hyytia E., Aalto S., Penttinen A. Minimizing slowdown in heterogeneous size-aware dispatching systems // ACM SIGMETRICS Perform. Eval. Rev., 2012. Vol. 40. P. 29–40.
12. Hyytia E., Spyropoulos T., Ott J. Offload (only) the right jobs: Robust offloading using the Markov decision processes // IEEE Symposium (International) on a World of Wireless, Mobile and Multimedia Networks (IEEE WoWMoM 2015). — Boston, MA, USA, 2015. P. 1–9.
13. Hyytia E. Lookahead actions in dispatching to parallel queues // Perform. Eval., 2013. Vol. 70. Iss. 10. P. 859–872.
14. Hyytia E. Optimal routing of fixed size jobs to two parallel servers // INFOR: Inform. Syst. Oper. Res., 2013. Vol. 51. Iss. 4. P. 215–224.
15. Konovalov M., Razumchik R. Simulation of task distribution in parallel processing systems // ICUMT-2014: 6th Congress (International) on Ultra Modern Telecommunications and Control Systems Proceedings. — St. Petersburg: IEEE, 2014. P. 657–663.
16. Коновалов М. Г. Построение имитационной модели для решения задач планирования вычислительных ресурсов // Системы и средства информатики, 2014. Т. 24. № 4. С. 45–62.
17. Коновалов М. Г. Методы адаптивной обработки информации и их приложения. — М.: ИПИ РАН, 2007. 212 с.

*Поступила в редакцию 14.09.15*

---

## APPROXIMATE OPTIMIZATION OF RESOURCE ALLOCATION STRATEGY: THE CASE OF BANK UNDERWRITING SYSTEM

*M. G. Konovalov<sup>1</sup> and R. V. Razumchik<sup>1,2</sup>*

<sup>1</sup>Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

<sup>2</sup>Peoples’ Friendship University of Russia, 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation

**Abstract:** Consideration is given to the problem of risk assessment process optimization in commercial underwriting. The process is modeled by a dispatching

queueing system with multiple single-server queues, several input regenerative flows of jobs with deadlines, complex job structure, two-stage job service process and partial availability of resources. One is interested in maximization of percent of jobs served in time (i. e., within deadline). The paper proposes the optimization method which requires development of the simulation model and further application of adaptive optimization algorithms on simulated paths. Based on numerical examples, it is shown that the approach allows one to find the most efficient strategy in a given set of heuristic strategies. The methodology is also applicable in general cases of resource allocation problems, which are not analytically tractable and which arise frequently when one needs to take into consideration multistage job service process, unpredictability of load and service times, and human factors.

**Keywords:** controlled queueing system; simulation model; adaptive strategy

**DOI:** 10.14357/08696527150403

## Acknowledgments

The work was partly supported by the Russian Foundation for Basic Research (project No. 15-07-03406).

## References

1. Crovella, M. E., M. Harchol-Balter, and C. D. Murta. 1998. Task assignment in a distributed system: Improving performance by unbalancing load. *SIGMETRICS'98 Proceedings*. Madison, WI. 268–269.
2. Harchol-Balter, M., M. E. Crovella, and C. D. Murta. 1999. On choosing a task assignment policy for a distributed server system. *J. Parallel Distr. Com.* 59:204–228.
3. Krishnan, K. R., and T. J. Ott. 1986. State-dependent routing for telephone traffic: Theory and results. *IEEE Conference on Decision and Control* 25:2124–2128.
4. Krishnan, K. R. 1990. Markov decision algorithms for dynamic routing. *IEEE Commun. Mag.* 28(10):66–69.
5. Krishnan, K. R. 1990. Joining the right queue: A state-dependent decision rule. *IEEE Trans. Automat. Contr.* 35(1):104–108.
6. Sassen, S. A. E., H. C. Tijms, and R. D. Nobel. 1997. A heuristic rule for routing customers to parallel servers. *Stat. Neerl.* 51(1):107–121.
7. Van Leeuwen, J., S. Aalto, and J. Virtamo. 2001. Load balancing in cellular networks using first policy iteration. Helsinki: Helsinki University of Technology, Networking Laboratory. Technical Report.
8. Bhulai, S. 2006. On the value function of the  $M/\text{Cox}(r)/1$  queue. *J. Appl. Probab.* 43(2):363–376.
9. Hyttia, E., J. Virtamo, S. Aalto, and A. Penttinen. 2011.  $M/M/1$ -PS queue and size-aware task assignment. *Perform. Evaluation* 68(11):1136–1148.
10. Hyttia, E., A. Penttinen, and S. Aalto. 2012. Size- and state-aware dispatching problem with queue-specific job sizes. *Eur. J. Oper. Res.* 217(2):357–370.

11. Hyytia, E., S. Aalto, and A. Penttinen. 2012. Minimizing slowdown in heterogeneous size-aware dispatching systems. *ACM SIGMETRICS Perform. Eval. Rev.* 40:29–40.
12. Hyytia, E., T. Spyropoulos, and J. Ott. 2015. Offload (only) the right jobs: Robust offloading using the Markov decision processes. *IEEE Symposium (International) on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)*. Boston, MA. 1–9.
13. Hyytia, E. 2013. Lookahead actions in dispatching to parallel queues. *Perform. Eval.* 70(10):859–872.
14. Hyytia, E. 2013. Optimal routing of fixed size jobs to two parallel Servers. *INFOR: Inform. Syst. Oper. Res.* 51(4):215–224.
15. Konovalov, M., and R. Razumchik. 2014. Simulation of task distribution in parallel processing systems. *ICUMT-2014: 6th Congress (International) on Ultra Modern Telecommunications and Control Systems Proceedings*. St. Petersburg. 657–663.
16. Konovalov, M. G. 2014. Postroenie imitatsionnoy modeli dlya resheniya zadach planirovaniya vychislitel'nykh resursov [Building a simulation model for solving scheduling problems of computing resources]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 24(4):45–62.
17. Konovalov, M. G. 2007. *Metody adaptivnoy obrabotki informatsii i ikh prilozheniya* [Methods of adaptive information processing and their applications]. Moscow: IPI RAN. 212 p.

*Received September 14, 2015*

## Contributors

**Konovalov Mikhail G.** (b. 1950) — Doctor of Science in technology, Head of Laboratory, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; mkonovalov@ipiran.ru

**Razumchik Rostislav V.** (b. 1984) — Candidate of Science (PhD) in physics and mathematics, senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; Peoples’ Friendship University of Russia, 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; rrazumchik@ipiran.ru

## МЕТРИЧЕСКАЯ КЛАССИФИКАЦИЯ ВРЕМЕННЫХ РЯДОВ С ВЫРАВНИВАНИЕМ ОТНОСИТЕЛЬНО ЦЕНТРОИДОВ КЛАССОВ\*

*А. В. Гончаров<sup>1</sup>, М. С. Попова<sup>2</sup>, В. В. Стрижов<sup>3</sup>*

**Аннотация:** Рассматривается задача многоклассовой классификации временных рядов. Классификация производится с помощью метрических методов, использующих матрицу попарных расстояний между временными рядами. Вычисление такой матрицы является трудоемким, так как ее размерность равна числу объектов выборки. С целью снижения размерности предлагается предварительно выделять эталонные объекты, а именно: центроиды каждого класса, и строить матрицу попарных расстояний между объектами выборки и эталонными объектами классов. Для вычисления попарных расстояний между объектами предлагается использовать метод динамического выравнивания временных рядов. В качестве прикладной задачи рассматривается задача распознавания типа движения по данным акселерометра мобильного телефона. Метрический алгоритм классификации, исследованный в этой работе, сравнивается по точности с алгоритмом разделяющей классификации.

**Ключевые слова:** метрическая классификация; динамическое выравнивание; классификация временных рядов; центроид; функция расстояния

**DOI:** 10.14357/08696527150404

### 1 Введение

В работе решается задача идентификации движений человека по временным рядам акселерометра [1]. Данные [2] представляют собой измерения акселерометра некоторого носимого устройства, например мобильного телефона, находящегося в кармане человека, и используются для идентификации действий человека в конкретный момент времени [3]. Метками классов движений служат: ходьба, бег, подъем по лестнице, спуск по лестнице, сидение, лежание.

В работах [4, 5] приведены методы, решающие задачу классификации временных рядов акселерометра. В [4] исследованы различные методы построения признакового описания временных рядов, в частности метод экспертного

---

\* Работа выполнена при финансовой поддержке РФФИ (проект 13-07-00678) и в рамках инициативы «МИТ-Сколтех».

<sup>1</sup>Московский физико-технический институт, alex.goncharov@phystech.edu

<sup>2</sup>Московский физико-технический институт, maria\_popova@phystech.edu

<sup>3</sup>Вычислительный центр им. А. А. Дородницына Федерального исследовательского центра «Информатика и управление» Российской академии наук, strijov@ccas.ru

построения признаков и метод построения признакового описания на основе гипотезы порождения данных. Результаты показывают, что построенное признаковое пространство достаточным образом описывает зависимую переменную и приводит к высокой точности классификации. В [5] для решения задачи использованы нейронные сети с небольшим числом связей между нейронами, которые достаточно точно решают задачу классификации и обладают свойством устойчивости к возмущениям данных.

Особенность этой работы заключается в использовании метрических способов классификации с предварительным отбором эталонных объектов. Предлагается ввести функции расстояния между объектами выборки, определить признаковое пространство при помощи этой функции и провести дальнейшую классификацию объектов методом  $k$  ближайших соседей.

Функция расстояния между временными рядами может быть задана различными способами: евклидово расстояние [6], метод динамического выравнивания временных рядов [7, 8], метод, основанный на нахождение наибольшей общей последовательности [9], Edit Distance with Real Penalty [10], Edit Distance on Real sequence [11], DISSIM [12], Sequence Weighted Alignment model [13], Spatial Assembling Distance [14] и др.

В [15] показано, что разность между значениями временного ряда, соответствующими различным временным отсчетам, не может рассматриваться в качестве описания расстояния между двумя объектами: эта мера расстояния чувствительна к шуму и локальным временным сдвигам. Для решения задачи предлагается использовать метод динамического выравнивания временных рядов (*англ.* Dynamic Time Warping — DTW) [16]. Как показано в [17], этот метод находит наилучшее соответствие между двумя временными рядами, если они нелинейно деформированы друг относительно друга — растянуты, сжаты или смещены вдоль оси времени. Метод DTW используется для определения сходства между ними и введения расстояния между двумя объектами.

Для снижения размерности признакового пространства предлагается искать расстояния от некоторого объекта до эталонных объектов каждого класса. Этими эталонными объектами предлагается назначать центроиды классов. Для определения центроида требуется одновременно выравнивать все временные ряды, принадлежащие данному классу. Эта задача обладает степенной сложностью  $m^2n^2$ , где  $m$  — число объектов класса;  $n$  — длина каждого временного ряда.

Процедура классификации выполняется в три шага. Первый — отбор эталонных объектов каждого класса. Второй — построение матрицы попарных расстояний сниженной размерности между временными рядами и эталонными объектами каждого класса. Третий — классификация временных рядов методом  $k$  ближайших соседей с помощью матрицы попарных расстояний.

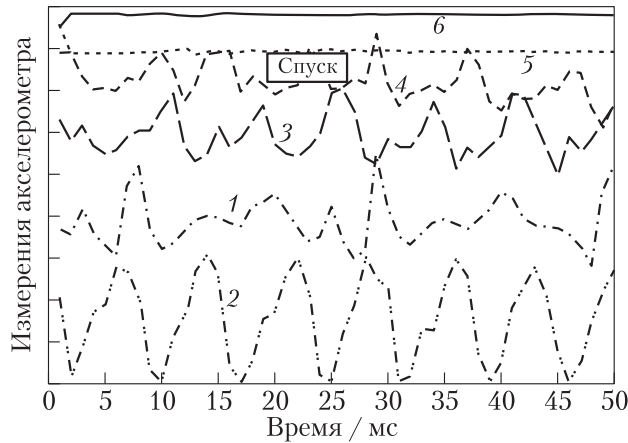
С целью сравнения результатов классификации методом, предложенным в этой работе, и результатов, полученных в [5], проведен вычислительный эксперимент на исходных данных [2].

## 2 Постановка задачи

Задана выборка  $\mathcal{D} = \{(\mathbf{s}_i, y_i)\}_{i=1}^m$ , состоящая из пар объект–ответ. Объектами служат временные ряды  $\mathbf{s}_i \in \mathbb{R}^n$ , а ответами являются метки класса —  $y_i \in Y = \{1, \dots, E\}$ , где  $E \ll m$ . Примеры временных рядов показаны на рис. 1. Выборка разбита на обучающую  $\mathcal{D}_l$  и контрольную  $\mathcal{D}_t$ .

Принята функция ошибки  $S$ . Требуется построить модель классификации  $f: \mathbb{R}^n \rightarrow Y$ , минимизирующую функцию ошибки  $S$  на контрольной выборке:

$$f = \operatorname{argmin}_f (S(f(\mathbf{s}_i), y_i | \mathcal{D}_t)) = \operatorname{argmin}_f \left( \frac{1}{|\mathcal{D}_t|} \sum_{i=1}^{|\mathcal{D}_t|} [f(\mathbf{s}_i) \neq y_i] \right). \quad (1)$$



**Рис. 1** Примеры временных рядов измерений акселерометра для разных видов физической активности: 1 — бег; 2 — ходьба; 3 — подъем; 4 — спуск; 5 — сидение; 6 — стояние

## 3 Описание алгоритма метрической классификации

Модель метрической классификации состоит из трех функций. Первая функция выделяет центроиды классов, вторая функция строит матрицу попарных расстояний между центроидами и объектами, а третья функция классифицирует объекты по полученной матрице.

### 3.1 Отбор эталонных объектов

В общем случае матрица попарных расстояний  $\mathbf{A}$  имеет размерность  $m \times m$ , где  $m$  — число объектов выборки  $\mathcal{D}$ . Для снижения размерности матрицы попарных

расстояний предлагается строить ее не для всех пар элементов выборки  $\mathcal{D}$ , а для пар  $(\mathbf{s}_i, \mathbf{z}_e)$ , где  $\mathbf{s}_i$  — элемент выборки, а  $\mathbf{z}_e$  — эталонный объект класса  $e$  из  $Y$ . Предлагается в качестве эталонного объекта использовать *центроид* этого класса. Это решение даст значительное снижение размерности матрицы попарных расстояний до  $m \times E$ , где  $E$  — число меток классов.

**Определение 1.** Пусть  $\mathcal{D}_e$  — множество элементов из  $\mathcal{D}$ , принадлежащих одному классу  $e$  из  $Y$ . Центроидом множества векторов  $\mathcal{D}_e = \{\mathbf{s}_i | y_i = e\}_{i=1}^m$  по расстоянию  $\rho$  назовем вектор  $\mathbf{z}_e \in \mathbb{R}^n$  такой, что

$$\mathbf{z}_e = \operatorname{argmin}_{\mathbf{z} \in \mathbb{R}^n} \sum_{\mathbf{s}_i \in \mathcal{D}_e} \rho(\mathbf{s}_i, \mathbf{z}).$$

Опишем метод оценки *центроида*  $\mathbf{z}_e$  из множества векторов  $\mathcal{D}_e \subset \mathbb{R}^n$ . Строим матрицу попарных расстояний  $\mathbf{A}$  между всеми элементами множества  $\mathcal{D}_e$ . С помощью метода главных компонент снижаем размерность пространства столбцов матрицы  $\mathbf{A}$ . Построим такие проекции исходных объектов на новую систему координат, называемых главными компонентами, чтобы сумма квадратов расстояний от объектов до их проекций на первые  $h$  ( $h < n$ ) главных компонент была минимальной. Опишем этот метод более подробно. Рассмотрим такую ортогональную матрицу  $\mathbf{W}$  в линейной комбинации  $\mathbf{B} = \mathbf{W}\mathbf{A}^T$  векторов-строк матрицы  $\mathbf{A}$ , что векторы-столбцы  $\mathbf{b}_1, \dots, \mathbf{b}_n$  матрицы  $\mathbf{B}$  будут иметь максимальную сумму их дисперсий:

$$\sum_{j=1}^n \sigma^2(\mathbf{b}_j) \rightarrow \max,$$

где

$$\sigma^2(\mathbf{b}) = \frac{1}{m} \sum_{i=1}^m (\mathbf{b}_i - \bar{\mathbf{b}})^2; \quad \bar{\mathbf{b}} = \frac{1}{m} \sum_{i=1}^m \mathbf{b}_i.$$

Матрица попарных расстояний между объектами в пространстве меньшей размерности  $\mathbf{A}^T = \mathbf{W}'\mathbf{A}^T$  есть проекция векторов-строк матрицы  $\mathbf{A}$  на  $h$  первых главных компонент.

Центроидом множества векторов  $\mathcal{D}_e$  по расстоянию  $\rho$  будем приближенно считать вектор  $\mathbf{z}_e \in \mathcal{D}_e$  такой, что в пространстве первых  $h$  главных компонент  $\mathcal{D}_e$  он оказывается ближайшим к остальным объектам, т. е.

$$\mathbf{z}_e = \operatorname{argmin}_{\mathbf{z} \in \mathcal{D}_e} \sum_{\mathbf{s}_i \in \mathcal{D}_e} \rho(\mathbf{W}'^T \mathbf{s}_i, \mathbf{W}'^T \mathbf{z}). \quad (2)$$

Рассматривая подмножества объектов каждого класса по отдельности, вводим функцию  $g_\rho$ , которая каждому классу из  $Y$  ставит в соответствие центроид данного класса по расстоянию  $\rho$  по формуле (2):

$$g_\rho : Y \rightarrow \mathcal{D}',$$

где  $\mathcal{D}' \subset \mathcal{D}_l$  — множество центроидов каждого класса.

### 3.2 Введение функции расстояния между объектами

В данной работе в качестве метрического расстояния между объектами предлагается использовать стоимость *пути наименьшей стоимости* между этими объектами.

Дадим определение пути наименьшей стоимости. Пусть  $\Omega^{n \times n}$  — это матрица, такая что ее элемент  $\Omega_{ij}$  равен разности между  $i$ -м и  $j$ -м элементами последовательностей  $\mathbf{s}_1$  и  $\mathbf{s}_2$ :

$$\Omega_{ij} = |\mathbf{s}_{1i} - \mathbf{s}_{2j}|.$$

**Определение 2.** Путем  $\pi$  между последовательностями  $\mathbf{s}_1$  и  $\mathbf{s}_2$  назовем упорядоченное множество пар индексов элементов матрицы  $\Omega$ :

$$\pi = \{\pi_r\} = \{(i_r, j_r)\}, \quad r = 1, \dots, R, \quad i, j \in \{1, \dots, n\},$$

где  $R$  — длина пути, зависящая от выбора пути. Он должен удовлетворять следующим условиям:

**Граничные условия:**  $\pi_1 = (1, 1)$  и  $\pi_R = (n, n)$ , т. е. начало и конец  $\pi$  находятся на диагонали в противоположных углах  $\Omega$ .

**Непрерывность.** Пусть  $\pi_r = (p_1, p_2)$  и  $\pi_{r-1} = (q_1, q_2)$ ,  $r = 2, \dots, R$ . Тогда

$$p_1 - q_1 \leq 1; \quad p_2 - q_2 \leq 1.$$

Это ограничение нужно, чтобы в шаге пути  $\pi$  участвовали только соседние элементы матрицы (включая соседние по диагонали).

**Монотонность.** Пусть  $\pi_r = (p_1, p_2)$  и  $\pi_{r-1} = (q_1, q_2)$ ,  $r = 2, \dots, R$ . Тогда выполняется хотя бы одно из условий

$$p_1 - q_1 \geq 1; \quad p_2 - q_2 \geq 1.$$

Это ограничение обусловлено природой рассматриваемых последовательностей и предназначено для монотонности функции выравнивания времени.

**Определение 3.** Стоимостью  $\text{Cost}(\mathbf{s}_1, \mathbf{s}_2, \pi)$  пути  $\pi$  между последовательностями  $\mathbf{s}_1$  и  $\mathbf{s}_2$  назовем

$$\text{Cost}(\mathbf{s}_1, \mathbf{s}_2, \pi) = \sum_{(i,j) \in \pi} \Omega_{ij}.$$

**Определение 4.** Путем наименьшей стоимости (выравнивающим путем)  $\hat{\pi}$  между последовательностями  $\mathbf{s}_1$  и  $\mathbf{s}_2$  назовем путь, который имеет наименьшую стоимость среди всех возможных путей между последовательностями  $\mathbf{s}_1$  и  $\mathbf{s}_2$ :

$$\hat{\pi} = \underset{\pi}{\operatorname{argmin}} \text{Cost}(\mathbf{s}_1, \mathbf{s}_2, \pi).$$



Обозначим стоимость выравнивающего пути между последовательностями  $\mathbf{s}_1$  и  $\mathbf{s}_2$  через  $\delta(\mathbf{s}_1, \mathbf{s}_2) = \text{Cost}(\mathbf{s}_1, \mathbf{s}_2, \hat{\boldsymbol{\pi}})$ .

Для вычисления стоимости выравнивающего пути в данной работе используется метод DTW. Согласно этому методу необходимо построить новую матрицу  $\gamma$ , элементы которой определяются следующим образом:

$$\begin{aligned} \gamma_{1j} &= \Omega_{1j}, \quad \gamma_{i1} = \Omega_{i1}, \quad i, j = 1, \dots, n; \\ \gamma_{ij} &= \Omega_{ij} + \min(\gamma_{i,j-1}, \gamma_{i-1,j}, \gamma_{i-1,j-1}), \quad i, j = 2, \dots, n. \end{aligned}$$

Элемент  $\gamma_{ij}$  матрицы  $\gamma$  равен стоимости выравнивающего пути между последовательностями  $\{s_{1a}\}_{a=1}^i$  и  $\{s_{2a}\}_{a=1}^j$ .

В качестве значения функции расстояния между двумя объектами выберем стоимость выравнивающего пути между ними из определения 3:

$$\rho(\mathbf{s}_1, \mathbf{s}_2) = \delta(\mathbf{s}_1, \mathbf{s}_2) = \gamma_{nn}.$$

Иначе говоря, для множеств декартова произведения  $\mathfrak{D} \times \mathfrak{D}'$

$$\rho : \mathfrak{D} \times \mathfrak{D}' \rightarrow \mathbf{X},$$

где  $\mathbf{X}$  — матрица попарных расстояний между объектами и центроидами классов. Ее элементы равны  $X_{ij} = \rho(\mathbf{s}_i, \mathbf{d}_j)$ ,  $\mathbf{s}_i \in \mathfrak{D}$ ,  $\mathbf{d}_j \in \mathfrak{D}'$ .

### 3.3 Описание алгоритма классификации

Считаем матрицу  $\mathbf{X}$  множеством векторов-строк, служащих описанием объектов  $\{\mathbf{x}_i\}_{i=1}^m$ . Матрица  $\mathbf{X}$  используется как матрица объект-признак в алгоритме классификации. Применяется метод  $k$  ближайших соседей. При этом  $\mathbf{X}^l = \rho(\mathfrak{D}^l \times \mathfrak{D}')$  — обучающая выборка из множества  $\mathbf{X}$ , а  $\mathbf{X}^t = \rho((\mathfrak{D} \setminus \mathfrak{D}^l) \times \mathfrak{D}')$  — тестовая. Каждому объекту обучающей выборки  $\mathbf{x}_i \in \mathbf{X}^l$ , соответствующему исходному объекту  $\mathbf{s}_i \in \mathfrak{D}$ , поставлена в соответствие метка класса  $y_i$ .

Для произвольного объекта  $\mathbf{u} \in \mathbf{X}$  расположим элементы обучающей выборки  $\mathbf{x}_1, \dots, \mathbf{x}_{|\mathbf{X}^l|} = \mathbf{X}^l$  в порядке возрастания расстояний до  $\mathbf{u}$ :

$$\eta(\mathbf{u}, \mathbf{x}_{\mathbf{u}}^{(1)}) \leq \eta(\mathbf{u}, \mathbf{x}_{\mathbf{u}}^{(2)}) \leq \dots \leq \eta(\mathbf{u}, \mathbf{x}_{\mathbf{u}}^{(|\mathbf{X}^l|)}),$$

где  $\eta(\mathbf{u}, \mathbf{x}_{\mathbf{u}}^{(i)})$  — евклидово расстояние между объектами  $\mathbf{u}$  и  $\mathbf{x}_{\mathbf{u}}^{(i)}$ , вычисленное по матрице объект-признак  $\mathbf{X}$ . Будем относить объект  $\mathbf{u}$  к тому классу, элементов которого окажется больше среди  $k$  его ближайших соседей  $\mathbf{x}_{\mathbf{u}}^{(i)}$ ,  $i = 1, \dots, k$ :

$$h_{\mathbf{u}}(\mathbf{u}, \mathbf{X}^l, k) = \operatorname{argmax}_{y \in Y} \sum_{i=1}^k \left[ y_{\mathbf{u}}^{(i)} = y \right].$$

Оптимальное значение параметра  $k$  определяется по критерию скользящего контроля с исключением объектов группами: производится  $T$  исключений из обучающей выборки произвольных групп  $\mathbf{X}^{l_t}$ ,  $t = 1, \dots, T$ , размер которых составляет 30% обучающей выборки. Измеряется величина ошибки классификации на этих группах при фиксированном числе  $k$  ближайших соседей. Эта ошибка суммируется для всех таких групп и минимизируется по числу ближайших соседей:

$$S(k, \mathbf{X}^l) = \sum_{t=1}^T \sum_{i=1}^{l_t} \left[ h_{\mathbf{u}_i}(\mathbf{u}_i, \mathbf{X}^l \setminus \{\mathbf{X}^{l_t}\}, k) \neq y_i \right] \rightarrow \min_{k \in \mathbb{N}}. \quad (3)$$

Таким образом, предлагаемая модель метрической классификации  $f$  может быть представлена в виде композиции нескольких функций:

$$f(\mathcal{D}, k) = h \left( \rho(\mathcal{D} \times g_\rho(Y)), \rho(\mathcal{D}^l \times g_\rho(Y)), k \right).$$

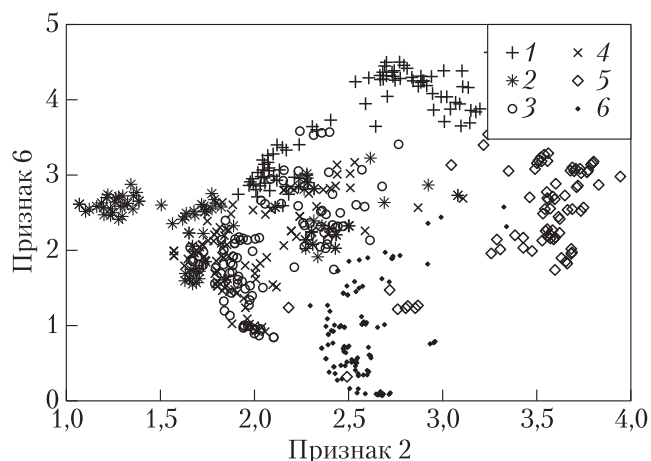
При этом решаемая задача (1) выбора структуры модели классификации имеет вид:

$$f(\mathcal{D}, k) = \arg \min_{k \in \mathbb{N}} (S(f(\mathcal{D}, k), Y | \mathcal{D}_t)).$$

#### 4 Вычислительный эксперимент

С целью получения оценки ошибки предложенной модели и ее сравнения с моделью классификации из [5] был проведен вычислительный эксперимент. Использовались данные акселерометра мобильного телефона. Данные представляли собой 1260 временных рядов длиной 600 точек, каждый из которых последовательно объединял три временных ряда: временной ряд ускорения по оси  $X$  (200 измерений), оси  $Y$  (200 измерений) и оси  $Z$  (200 измерений). Выделено шесть типов физической активности: ходьба, бег, сидение, стояние, подъем, спуск. Временные ряды записывались акселерометром, который находился в кармане у человека, выполняющего один из типов физической активности, после чего разделялись на 10-секундные сегменты.

Для предварительной оценки качества классификации описанного алгоритма были построены диаграммы рассеяния для множества признаков  $\mathbf{X}$ . Пример диаграммы рассеяния для признаков 6 и 2 приведен на рис. 2 для всех объектов обучающей выборки. По оси  $X$  отложены значения признака два, т. е. расстояние до центроида второго класса, а по оси  $Y$  — значения шестого признака, т. е. расстояние до центроида шестого класса. Из-за плохой визуальной разделимости на диаграммах рассеяния выдвигается предположение о низком качестве классификации объектов, принадлежащих классам «спуск — ходьба вниз по лестнице» и «подъем — ходьба вверх по лестнице». Остальные же классы визуально



**Рис. 2** Диаграмма рассеяния для признаков 6 и 2, по которой выдвигаются предположения о качестве классификации: 1 — бег; 2 — ходьба; 3 — подъем; 4 — спуск; 5 — сидение; 6 — стояние

**Таблица 1** Средние расстояния между объектами различных классов

Класс	Бег	Ходьба	Вверх	Вниз	Сидение	Стояние
Бег	<b>1,7221</b>	2,7483	3,3432	3,2631	3,4754	3,8243
Ходьба	2,7483	<b>1,0665</b>	1,4622	1,3133	2,9301	2,1336
Подъем	3,3432	1,4622	<b>1,4405</b>	1,3760	3,1055	1,8948
Спуск	3,2631	1,3133	1,3760	<b>1,2323</b>	3,1764	1,9166
Сидение	3,4754	2,9301	3,1055	3,1764	<b>1,5114</b>	3,2813
Стояние	3,8243	2,1336	1,8948	1,9166	3,2813	<b>1,2692</b>

разделимы на этих диаграммах. Предполагается, что на них величина ошибки будет значительно ниже.

В табл. 1 показаны средние расстояние между объектами различных классов. Для большинства классов средние расстояния до объектов других классов больше, чем среднее расстояние между объектами самого класса. Выдвигается предположение о низкой величине ошибки на таких классах.

Все вышеизложенные предположения подтверждаются последующими результатами эксперимента.

При выполнении классификации методом  $k$  ближайших соседей при помощи метода скользящего контроля (3) было найдено наилучшее  $k$ . На рис. 3 показано среднее значение функции ошибки  $S$  из выражения (1) и ее стандартное отклонение в зависимости от величины  $k$ . Имеется минимум при  $k = 1$ .

В вычислительном эксперименте исследовалась точность модели метрической классификации, описанной в разд. 3. Оценивалась точность модели по критерию

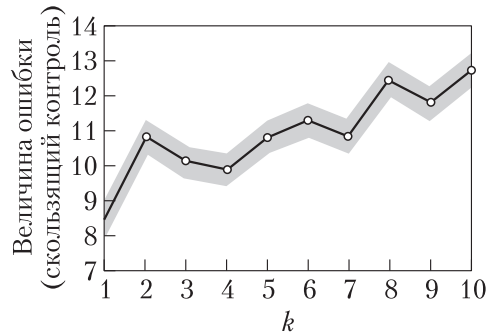


Рис. 3 Зависимость ошибки от величины  $k$

Таблица 2 Сравнение эффективности предложенного метрического алгоритма классификации и алгоритма разделяющей классификации на данных [2]

Классификация	Бег	Ходьба	Подъем	Спуск	Сидение	Стояние	Общее
Метрическая классификация. Точность по критерию скользящего контроля, %	100	83	80	87	99	100	91
Разделяющая классификация [5]. Относительная точность на тестовых данных, %	98	96	75	72	100	89	88

скользящего контроля (3). В табл. 2 приводятся результаты работы модели, основанной на методе DTW, и модели, описанной в [5]. Качество классификации в классе «ходьба» ухудшилось, в то время как в классах «подъем» и «спуск» увеличилось. У построенной модели качество классификации выше, чем у модели разделяющей классификации.

### Заключение

В работе представлена модель метрической классификации. В рамках модели предложено использование расстояния, вычисленного методом динамического выравнивания временных рядов, использование центроидов для снижения размерности задачи и последующей классификации внутри построенного таким образом признакового пространства. С целью сравнения качества классификации предложенной модели и модели разделяющей классификации был проведен вычислительный эксперимент, демонстрирующий высокое качество классификации модели метрической классификации. Программная реализация алгоритма метрической классификации в среде разработки MatLab находится в свободном доступе [18].

**Литература**

1. *Kwapisz J. R., Weiss G. M., Moore S.* Activity recognition using cell phone accelerometers // SIGKDD Explorations, 2010. Vol. 12. No. 2. P. 74–82.
2. *Kwapisz J. R.* Data from accelerometer // Sourceforge.net: веб-сайт для разработчиков открытого программного обеспечения, 2010. <http://sourceforge.net/p/mlalgorithms/TSLearning/data/preprocessedJarge.csv>.
3. *Hoseini-Tabatabaei S. A., Gluhak A., Tafazolli R.* A survey on smartphone-based systems for opportunistic user context recognition // ACM Comput. Surv., 2013. Vol. 45. No. 3. P. 27.
4. *Кузнецов М. П.* Признаковая классификация временных рядов // Sourceforge.net: веб-сайт для разработчиков открытого программного обеспечения, 2014. <https://sourceforge.net/p/mlalgorithms/code/HEAD/tree/TSLearning/doc/TSClassification/TSClassification.pdf>.
5. *Попова М. С., Стрижов В. В.* Выбор оптимальной модели классификации физической активности по измерениям акселерометра // Информатика и её применения, 2015. Т. 9. Вып. 1. С. 79–89.
6. *Faloutsos C., Ranganathan M., Manolopoulos Y.* Fast subsequence matching in time-series databases // ACM Conference (International) on Management of Data (SIGMOD), 1994. Minneapolis, MN, USA. P. 419–429.
7. *Berndt D. J., Clifford J.* Using dynamic time warping to find patterns in time series // Workshop on Knowledge Discovery in Databases (KDD Workshop), 1994. Seattle, WA, USA. P. 359–370.
8. *Keogh E. J., Ratanamahatana C. A.* Exact indexing of dynamic time warping // Knowl. Inf. Syst., 2005. Vol. 7. No. 3. P. 358–386.
9. *Vlachos M., Gunopulos D., Kollios G.* Discovering similar multidimensional trajectories // IEEE Conference (International) on Data Engineering (ICDE), 2002. San Jose, CA, USA. P. 673–684.
10. *Chen L., Ng R. T.* On the marriage of lp-norms and edit distance // Very Large Data Bases (VLDB) Conference, 2004. Toronto, Ontario, Canada. P. 792–803.
11. *Chen L., Özsu M. T., Oria V.* Robust and fast similarity search for moving object trajectories // ACM Conference (International) on Management of Data (SIGMOD), 2005. Baltimore, MD, USA. P. 491–502.
12. *Frentzos E., Gratsias K., Theodoridis Y.* Index-based most similar trajectory search // IEEE Conference (International) on Data Engineering (ICDE), 2007. Istanbul, Turkey. P. 816–825.
13. *Morse M. D., Patel J. M.* An efficient and accurate method for evaluating time series similarity // ACM Conference (International) on Management of Data (SIGMOD), 2007. Beijing, China. P. 569–580.
14. *Chen Y., Nascimento M. A., Ooi B. C., Tung A. K. H.* SpADe: On shape-based pattern detection in streaming time series // IEEE Conference (International) on Data Engineering (ICDE), 2007. Istanbul, Turkey. P. 786–795.
15. *Scheuermann P., Wang X., Ding H., Trajcevski G., Keogh E.* Querying and mining of time series data: Experimental comparison of representations and distance measures // Very Large Data Bases (VLDB) Conference, 2008. Auckland, New Zealand. P. 1542–1552.

16. *Keogh E. J., Pazzani M. J.* Scaling up dynamic time warping to massive datasets // Principles of data mining and knowledge discovery / Eds. J. M. Zytkow, J. Rauch. — Lecture notes in artificial intelligence ser. — Berlin–Heidelberg: Springer-Verlag, 1999. Vol. 1704. P. 1–11.
17. *Salvador S., Chan P.* Fastdtw: Toward accurate dynamic time warping in linear time and space // Workshop on Mining Temporal and Sequential Data, 2004. Seattle, WA, USA. P. 11.
18. *Гончаров А. В.* Реализация метрической классификации временных рядов // Sourceforge.net, 2015. <http://sourceforge.net/p/mlalgorithms/code/HEAD/tree/Group274/Goncharov2015MetricClassification/code/main.m>

*Поступила в редакцию 16.07.15*

---

---

## METRIC TIME SERIES CLASSIFICATION USING DYNAMIC WARPING RELATIVE TO CENTROIDS OF CLASSES

*A. V. Goncharov<sup>1</sup>, M. S. Popova<sup>1</sup>, and V. V. Strijov<sup>2</sup>*

<sup>1</sup>Moscow Institute of Physics and Technology, 9 Institutskiy Per., Dolgoprudny, Moscow Region 141700, Russian Federation

<sup>2</sup>Dorodnicyn Computing Center, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 40 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** The paper discusses the problem of time series classification in the case of several classes. The proposed classification model uses the matrix of distance between time series. This distance measure is defined by the dynamic time warping method. The dimension of the distance matrix is very high. The paper introduces centroids of each class as reference objects used to decrease this dimension. The distance matrix with lower dimension describes the distance between all objects and reference objects. This method is used for human activity recognition. The quality of classification on data from a mobile accelerometer is investigated. This metric algorithm of classification is compared with the separating classification algorithm.

**Keywords:** metric classification; dynamic time warping; time series classification; centroid; distance function

**DOI:** 10.14357/08696527150404

### Acknowledgments

The work was financially supported by the Russian Foundation for Basic Research (project 13-07-00678) and under the initiative “MIT-Skoltech.”

## References

1. Kwapisz, J. R., G. M. Weiss, and S. Moore. 2010. Activity recognition using cell phone accelerometers. *SIGKDD Explorations* 12(2):74–82.
2. Kwapisz, J. R. 2010. Data from accelerometer. Available at: <http://sourceforge.net/p/mlalgorithms/TSLearning/data/preprocessedLARGE.csv> (accessed April 21, 2014).
3. Hoseini-Tabatabaei, S. A., A. Gluhak, and R. Tafazolli. 2013. A survey on smartphone-based systems for opportunistic user context recognition. *ACM Comput. Surv.* 45(3):27.
4. Kuznecov, M. P. Priznakovaya klassifikatsiya vremennykh ryadov [Time series feature-based classification]. Available at: <https://sourceforge.net/p/mlalgorithms/code/HEAD/tree/TSLearning/doc/TSClassification/TSClassification.pdf> (accessed June 18, 2015).
5. Popova, M. S., and V. V. Strijov. 2015. Vybór optimal'noy modeli klassifikatsii fizicheskoy aktivnosti po izmereniyam akselerometra [Selection of optimal physical activity classification model using measurements of accelerometer]. *Informatika i ee Primeneniya — Inform. Appl.* 9(1):79–89.
6. Faloutsos, C., M. Ranganathan, and Y. Manolopoulos. 1994. Fast subsequence matching in time-series database. *ACM Conference (International) on Management of Data (SIGMOD)*. Minneapolis, MN. 419–429.
7. Berndt, D. J., and J. Clifford. 1994. Using dynamic time warping to find patterns in time series. *Workshop on Knowledge Discovery in Databases (KDD Workshop)*. Seattle, WA. 359–370.
8. Keogh, E. J., and C. A. Ratanamahatana. 2005. Exact indexing of dynamic time warping. *Knowl. Inf. Syst.* 7(3):358–386.
9. Vlachos, M., D. Gunopulos, and G. Kollios. 2002. Discovering similar multidimensional trajectories. *IEEE Conference (International) on Data Engineering (ICDE)*. San Jose, CA. 673–684.
10. Chen, L., and R. T. Ng. 2004. On the marriage of lp-norms and edit distance. *Very Large Data Bases (VLDB) Conference*. Toronto, Ontario, Canada. 792–803.
11. Chen, L., M. T. Özsu, and V. Oria. 2005. Robust and fast similarity search for moving object trajectories. *ACM Conference (International) on Management of Data (SIGMOD)*. Baltimore, MD. 491–502.
12. Frentzos, E., K. Gratsias, and Y. Theodoridis. 2007. Index-based most similar trajectory search. *IEEE Conference (International) on Data Engineering (ICDE)*. Istanbul, Turkey. 816–825.
13. Morse, M. D., and J. M. Patel. 2007. An efficient and accurate method for evaluating time series similarity. *ACM Conference (International) on Management of Data (SIGMOD)*. Beijing, China. 569–580.
14. Chen, Y., M. A. Nascimento, B. C. Ooi, and A. K. H. Tung. 2007. SpADe: On shape-based pattern detection in streaming time series. *IEEE Conference (International) on Data Engineering (ICDE)*. Istanbul, Turkey. 786–795.
15. Scheuermann, P., X. Wang, H. Ding, G. Trajcevski, and E. Keogh. 2008. Querying and mining of time series data: Experimental comparison of representations and distance measures. *Very Large Data Bases (VLDB) Conference*. Auckland, New Zealand. 1542–1552.

16. Keogh, E. J., and M. J. Pazzani. 1999. Scaling up dynamic time warping to massive datasets. *Principles of data mining and knowledge discovery*. Eds. J. M. Zytlow and J. Rauch. Lecture notes in artificial intelligence ser. Berlin–Heidelberg: Springer-Verlag. 1704:1–11.
17. Salvador, S., and P. Chan. 2004. Fastdtw: Toward accurate dynamic time warping in linear time and space. *3rd SIGKDD Workshop on Mining Temporal and Sequential Data (KDD/TDM)*. Seattle, WA. 11.
18. Goncharov, A. V. Programnaya realizatsiya [Program realization]. Available at: <http://sourceforge.net/p/mlalgorithms/code/HEAD/tree/Group274/Goncharov2015MetricClassification/code/main.m> (accessed April 20, 2015).

*Received July 16, 2015*

## **Contributors**

**Goncharov Alexey V.** (b. 1995) — student, Moscow Institute of Physics and Technology, 9 Institutskiy Per., Dolgoprudny, Moscow Region 141700, Russian Federation; alex.goncharov@phystech.edu

**Popova Maria S.** (b. 1994) — student, Moscow Institute of Physics and Technology, 9 Institutskiy Per., Dolgoprudny, Moscow Region 141700, Russian Federation; mariewelt@gmail.com

**Strijov Vadim V.** (b. 1967) — Doctor of Science in physics and mathematics, leading scientist, Dorodnicyn Computing Center, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 40 Vavilov Str., Moscow 119333, Russian Federation; strijov@ccas.ru



## АРХИТЕКТУРА СТЕНДА ДЛЯ ЭКСПЕРИМЕНТАЛЬНОГО ИССЛЕДОВАНИЯ МОДЕЛЕЙ, АЛГОРИТМОВ И РЕШЕНИЙ ПО ОБЕСПЕЧЕНИЮ ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТИ В ОБЛАЧНЫХ ВЫЧИСЛИТЕЛЬНЫХ СРЕДАХ\*

*А. А. Грушо<sup>1</sup>, М. И. Забейжайло<sup>2</sup>, А. А. Зацаринный<sup>3</sup>, А. В. Николаев<sup>4</sup>,  
В. О. Писковский<sup>5</sup>*

**Аннотация:** Рассматривается задача разработки архитектуры стенда для моделирования виртуальных сетей и экспериментальных исследований проверки методов обеспечения информационной безопасности (ИБ), применяемых в облачных вычислительных средах. Предложены подходы к моделированию сетей, применяемые с этой целью платформы и системы. Приведена диаграмма развертывания стенда. Указаны требования к аппаратному и математическому обеспечению. Свойства стенда проанализированы с точки зрения применимости к исследованиям механизмов защиты в облачных вычислительных средах.

**Ключевые слова:** информационная безопасность; облачные вычисления; программно-конфигурируемые сети; OpenStack

**DOI:** 10.14357/08696527150405

### 1 Введение

В настоящее время волна централизации вычислений, вызванная возможностью более эффективного использования вычислительных ресурсов и экономии средств на реализации и поддержке сетевых услуг, привела к появлению центров обработки данных (ЦОД). В ЦОД [1] сосредоточены вычислительные мощности, сервисы надежного хранения данных и сетевой конфигурации, обеспечивающие работу автоматизированных информационных систем в соответствии с требованиями ИБ.

---

\* Работа выполнена при поддержке РФФИ (проект 15-29-07981 офи-м).

<sup>1</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, grusho@yandex.ru

<sup>2</sup>Всероссийский институт научной и технической информации Российской академии наук (ВИНИТИ), zmizanv@ya.ru

<sup>3</sup>Федеральный исследовательский центр «Информатика и управление» Российской академии наук, azatsarginny@ipiran.ru

<sup>4</sup>Институт химической физики Российской академии наук, gentoorion@mail.ru

<sup>5</sup>НП «Центр прикладных исследований компьютерных сетей», vrvp80@yandex.ru

Средства виртуализации вычислительных мощностей и ресурсов хранения дали новый импульс к образованию ЦОД и облачных вычислительных систем (ОВС) [2], предоставляющих свои услуги в аренду. Необходимость соблюдения законодательных и нормативных требований стимулирует развитие автоматизированных решений для поддержки политик ИБ в ОВС [3, 4]. В условиях непрерывно усложняющейся топологии сетевых соединений и предъявляемых к ней требований централизации вычислений и хранения данных [5, 6] такое развитие представляется единственно возможным.

Целью данной работы является разработка требований и подходов к построению архитектуры стенда, моделирующего работу сложных IP (Internet Protocol) сетей или ОВС, с эмуляцией в рамках стенда типовой инфраструктуры тенанта, предназначенной для экспериментального исследования механизмов, алгоритмов и решений по обеспечению ИБ в IP-сетях или ОВС.

## **2 Перечень задач и предъявляемые к стенду требования**

Сложные IP-сети и ОВС представляют собой множество программно-аппаратных сущностей с независимой управляемой логикой. Среди средств и методов моделирования выделяют аналитический подход и построение моделей, имитирующих работу при помощи средств компьютерного моделирования.

Для аналитического подхода характерно представление IP-сетей, в частности ОВС, их топологии, способов и характера использования компонентов при помощи введения абстрактных понятий, отражающих как состояние в целом, так и отдельных элементов, а также формулирование и доказательство утверждений, непротиворечивых в рамках принятых в аналитической модели операций. При этом корректность модели определяется согласованностью вводимых понятий и операций, а ее полезность — соответствием теоретических выводов результатам, наблюдаемым на практике.

Второй подход, заключающийся в моделировании работы IP-сетей или ОВС на компьютерном оборудовании, возможно, и не обладает той общностью рассуждений, которую предоставляет аналитический подход. Однако такой подход позволяет проверить на практике в условиях, максимально приближенных к реальным, эффективность методов и алгоритмов решения задач, возникающих при эксплуатации сложных систем. В частности, рассматриваются задачи обеспечения политик ИБ в IP-сетях или ОВС.

Перечислим задачи и эксперименты, для которых собирается стенд.

1. Интеграция статистических и детерминистских методов выявления, а также анализа «аномалий» (в том числе формирования мер по противодействию или минимизации последствий таких воздействий) в штатном режиме функционирования сложных IP-сетей, в частности облачных сред [7].
2. Разработка и экспериментальное исследование проблемно-ориентированных методов, моделей и алгоритмов обучения на прецедентах [8] — иденти-

фицированных примерах и контрпримерах «аномалий» в штатном режиме функционирования сложных IP-сетей или облачных сред.

3. Оценка вычислительной сложности основных задач обучения на прецедентах [9].
4. Разработка моделей и алгоритмов оптимизации вычислительной сложности комбинаторного перебора вариантов при поиске решений задач обучения на прецедентах [10].
5. Разработка методов формализации политик безопасности, учитывающих возможности интеллектуального анализа данных (ИАД) (применение математических моделей, методов и алгоритмов искусственного интеллекта в компьютерном анализе данных), а также SDN (Software Defined Networks) и NFV (Network Function Virtualization) подходов. Разработка моделей и методов мониторинга исполнения политик безопасности, определяющих штатный режим функционирования сложных IP-сетей или облачных сред, и соблюдения соответствующих требований [11].
6. Разработка принципов формирования и архитектурных решений для тестовых сред, обеспечивающих экспериментальную проверку создаваемых моделей и алгоритмов выявления и анализа «аномалий» в штатном режиме функционирования сложных IP-сетей или облачных сред [12].
7. Исследование возможностей архитектурных решений на базе SDN/NFV-подходов для оптимизации архитектуры комплексных программно-аппаратных решений по идентификации и анализу «аномалий» в штатном режиме функционирования сложных IP-сетей или облачных сред. Разработка рекомендаций по перспективным архитектурам подобных программно-аппаратных комплексов [4, 12].

Для построения стенда также учтены следующие положения и требования [13, 14]:

1. Принцип максимального исключения влияния инструментов мониторинга и измерения на наблюдаемые показатели для объективного анализа измеримых показателей работы IP-сетей или ОВС, обнаружения аномалий, в том числе вызванных несанкционированными действиями.
2. Эмуляция детального поведения отдельных узлов и сегментов сетей.
3. Проведение исследований облачных решений по обеспечению ИБ на базе SDN/NFV-технологий.
4. Совместное использование статистических и детерминистских методов анализа «аномалий» в штатном функционировании IP-сетей или облачных сред.
5. Исследование возможностей повышения точности результатов статистического анализа, идентификации «аномалий» в контролируемой среде за счет

ее интеграции с проблемно-ориентированными методами ИАД, в частности путем учета контекстных зависимостей, порождаемых в процессе обучения на локальных коллекциях релевантных прецедентов.

6. Эмуляция работы нового типа сети — программно-конфигурируемой сети (ПКС) и динамический контроль исполнения требований политик ИБ в SDN-сетях.
7. Использование серийных недорогих широко распространенных серверов, собранных из типового оборудования на типовых чипсетах.
8. Аппаратное масштабирование.
9. Визуализация результатов анализа и мониторинга.
10. Автоматическое управление вычислительными ресурсами.

### 3 Архитектура и концепция реализации и свойства стенда

Для выполнения перечисленных выше требований предлагается использовать программно-аппаратный комплекс под управлением облачной платформы OpenStack, с помощью которой реализовать комбинированный подход к моделированию, а именно: построить детальные фрагменты сложных IP-сетей и ОВС на уровне отдельных виртуальных машин (ВМ), виртуальных сетевых коммутаторов [15] и построить сегменты IP-сетей или ОВС при помощи системы имитационного моделирования, установленной на одной или нескольких ВМ. Указанную комбинацию предлагается расширить установкой средств проведения испытаний, также размещаемых на одной или нескольких ВМ.

Платформу OpenStack [16, 17] предлагается установить на четырех физических серверах [18], как указано на диаграмме развертывания, представленной на рисунке.

Узел «Управление» содержит систему управления базами данных (СУБД) MySQL и центральные компоненты управления платформы OpenStack:

- управление виртуальными машинами (компонент nova: api, scheduler, cert, consoleauth, novncproxy, novaclient);
- управление сетевыми блочными устройствами (компонент cinder: api, volume, scheduler);
- управление сетью (компонент neutron: server, l3-agent, dhcp-agent, m12-plugin-agent);
- репозиторий образов виртуальных машин (компонент glance);
- систему аутентификации (компонент keystone-all);
- очередь сообщений (компонент rabbitMQ).

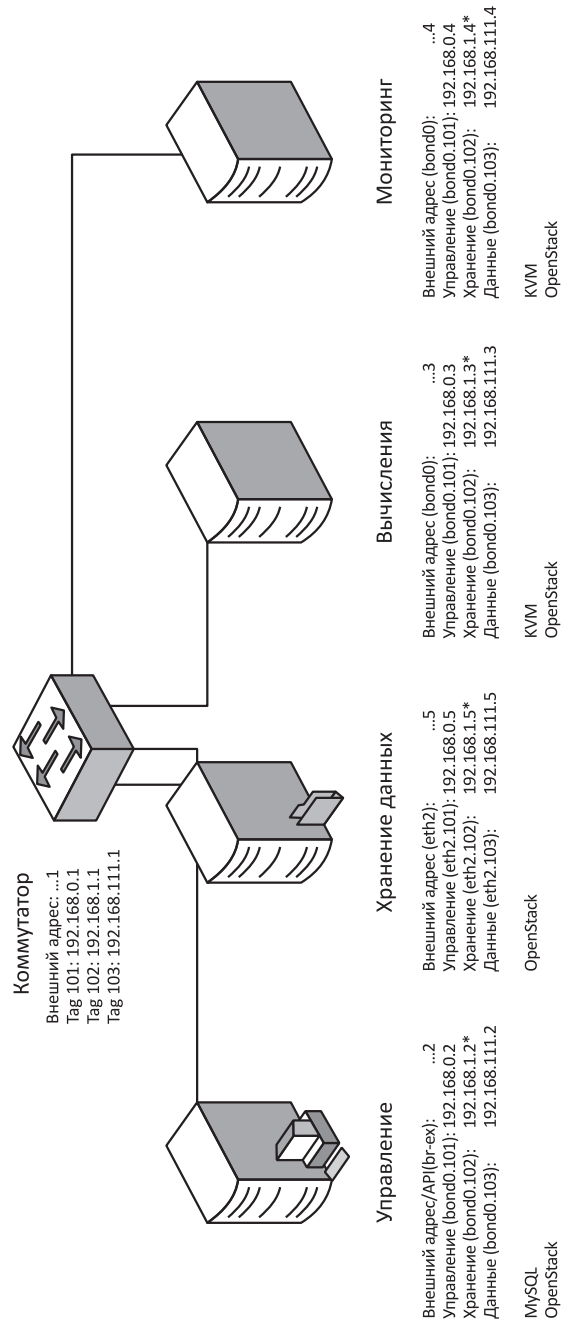


Диаграмма развертывания

На узле «Хранение данных» установлен cinder-volume, агент OpenStack для управления экспортируемыми по iSCSI-протоколу блочными устройствами. Узел работает с использованием Logical Volume Manager (LVM) для поддержки образов экспортируемых блочных устройств.

Узел «Вычисления» и узел «Мониторинг» работают под управлением системы компонента nova (compute и api-metadata) и используются для запуска и управления ВМ под гипервизором KVM (Kernel-based Virtual Machine) [19]. На этих узлах также установлены агенты Neutron (m12-plugin-agent и openvswitch-agent), которые служат для создания изолированных виртуальных L2-сетей с помощью GRE (Generic Routing Encapsulation) или VXLAN (Virtual eXtensible Local Area Network) [3] протокола.

Узел «Вычисления» планируется реализовать в одном или нескольких экземплярах в зависимости от исследуемых моделей и их сложности. В случае использования нескольких вычислительных узлов возникает задача размещения частей моделируемой IP-сети на различных вычислительных узлах. При этом несколько виртуальных каналов моделируемой IP-сети будут вложены в реальный канал, соединяющий вычислительные узлы стенда. Поскольку латентности моделируемых сетей практически невозможно сохранить на уровне физической сети, на первом этапе предлагается уделить внимание сохранению (или пропорциональному урезанию) емкости моделируемых соединений. Исходя из того, что в большинстве современных сетей используется технология 1000BASE-T (1 Gbit) и ее модификации, для моделирования таких сетей будет адекватно использовать межзловую сеть стандарта 10GBASE, возможно кратную, в режиме LACP (Link Aggregation Control Protocol) bonding. Тогда задача укладки частей модели на разные серверы сведется к проверке того, что емкость виртуальных каналов, вкладываемых в физическую межзловую сеть, не превосходит производительность этой сети с небольшой поправкой, значительно меньшей 1.

Узел «Мониторинг» предназначен для использования инструментами мониторинга и анализа данных экспериментов, например KALI Linux [20], ИАД с применением ДСМ-методов [8–11]. С целью исключения влияния инструментов мониторинга и измерения на наблюдаемые показатели для объективного анализа измеримых показателей работы IP-сетей или ОВС указанные инструменты используют этот узел в монопольном режиме. Узел используется для работы систем ИАД, сбора статистики, визуализации результатов, размещения инструментария для проведения тестов на безопасность и подготовки отчетности.

Вычислительный узел используется для размещения ВМ, на которых планируется как моделирование детального поведения отдельных узлов, так и сегментов сетей. Для моделирования детального поведения отдельных узлов используется полная виртуализация гостевых машин. Для моделирования сегментов сети на виртуальной гостевой машине устанавливается система имитационного моделирования, например Mininet [21], DevStack [17], NS-3 [22] или OMNeT++ [23] в зависимости от типа и целей проводимых численных экспериментов. Система имитационного моделирования позволяет средствами виртуализации имитиро-

вать работу компьютерной сети из нескольких десятков узлов или ОВС. При этом каждый узел моделируется процессом на уровне операционной системы с применением так называемых контейнеров, что представляет собой легкую виртуализацию узлов с собственным пространством процессов и сетевым стеком, работающих на одной ВМ и совместно использующих общие ресурсы в рамках операционной системы. В результате такая легкая виртуализация позволяет имитировать работу сети при относительно скромных требованиях к аппаратному обеспечению.

Требования к аппаратной части комплекса:

- серверы (4 шт.) с характеристиками: 16 ядер, частота не менее 2,4 ГГц, 64 ГБ оперативной памяти, аппаратная виртуализация, два порта по 10 ГБ, два порта по 1 ГБ, карточки 82599, жесткий диск 1 ТБ;
- коммутатор: 24 порта, 1 ГБ + 10 ГБ uplink, с поддержкой OpenFlow.

Сформированный таким образом стенд обладает следующими свойствами:

- в силу того, что инструментарий мониторинга и измерения расположен на монопольно используемом узле «Мониторинг», можно считать, что влияние инструментария на временные и иные моделируемые показатели минимально, что соответствует первому требованию из приведенного выше списка. Ограничением может стать производительность сетевого оборудования, так как в рамках физически одного сетевого интерфейса моделируется работа нескольких виртуальных сетевых каналов. Число таких каналов не может быть большим, порядка 10 на один сетевой интерфейс производительностью 1 ГБ/с;
- вычислительный узел допускает размещение и работу отдельных моделируемых узлов в режиме полной виртуализации, моделей сегментов сетей, работающих в рамках отдельных ВМ, работу облачных решений на базе SDN/NFV-технологий [24]. Таким образом, выполнены требования 2 и 3 из представленного выше списка;
- узел «Мониторинг» применяется для работы ВМ с установленными на них средствами визуального представления результатов, статистического и динамического анализа измеряемых показателей. Узел «Мониторинг» также дает исходную информацию системам искусственного интеллекта, в которых используются элементы каузального анализа, ДСМ-методов и др. Это открывает возможность применения статистических и детерминистских методов для анализа работы компонентов сложных IP-сетей или ОВС, что отвечает требованиям 4, 5 и 9 из вышеприведенного списка;
- размещение средств управления стендом и инструментария интеллектуального анализа на отдельных серверах, возможность применения OpenFlow-протокола при организации работы комплекса дают возможность эмуляции работы ПКС-сети в рамках стенда и динамического контроля требований политик ИБ в таких сетях, что соответствует требованию 6;

- применение облачной платформы и среды виртуализации KVM позволяет масштабировать комплекс, добавляя к нему вычислительные мощности на базе недорогих серийных серверов, единственное требование к которым — наличие средств аппаратной виртуализации, что соответствует требованиям 7 и 8 из списка.

Кроме того, стенд реализует ОВС и обеспечивает необходимые механизмы информационной безопасности:

- идентификации/аутентификации;
- разграничения доступа;
- управления криптографическими ключами средствами используемого в рамках OpenStack компонента KeyStone.

Также стенд допускает применение самых разнообразных средств соблюдения политик безопасности, управления информационными потоками, журналирования действий, анализа поведения и проактивного управления компонентами сложных IP-сетей или ОВС.

Таким образом, стенд располагает возможностями интеллектуального анализа работы компонентов сложных IP-сетей или ОВС, исследования эффективности и оценки вычислительной сложности разрабатываемых методов и алгоритмов ИБ, включая способы мониторинга соблюдения политик безопасности, применяемых для различных сегментов сложных IP-сетей или ОВС. Использование стенда позволит экспериментально изучать возможности архитектурных решений сложных IP-сетей или ОВС с целью выработки направлений их дальнейшего развития.

## 4 Заключение

В результате проведенных исследований предложена архитектура стенда для экспериментального исследования моделей, алгоритмов и решений по обеспечению ИБ в сложных IP-сетях или облачных вычислительных средах, рассмотрены и обоснованы характеристики стенда с точки зрения выполнения требований, предъявляемых к моделям и решаемым при помощи стенда задачам.

## Литература

1. Open Data Center Alliance Master Usage Model: Commercial Framework Rev 1.0. [http://www.opendatacenteralliance.org/docs/ODCA\\_Commercial\\_Framework\\_MasterUM\\_v1.0\\_Nov2012.pdf](http://www.opendatacenteralliance.org/docs/ODCA_Commercial_Framework_MasterUM_v1.0_Nov2012.pdf).
2. Amies A., Sluiman H., Tong Q. G., Liu G. N. Developing and hosting applications on the cloud. — IBM Press, 2012. 367 p.
3. IETF Virtual eXtensible Local Area Network (VXLAN): A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks. <https://tools.ietf.org/html/rfc7348>.



4. *Grusho A., Grusho N., Piskovski V., Timonina E.* Five SDN-oriented directions in information security // SDN&NFV: The next generation of computational infrastructure: 2014 Science and Technology Conference (International) — Modern Networking Technologies (MoNeTec) Proceedings. — Moscow: MAKS Press, 2014. P. 68–71.
5. Network virtualization. <http://networkheresy.com/2012/05/31/network-virtualization>.
6. Saha R. SDN, OpenFlow and Network Virtualization. — IBM. [http://www-304.ibm.com/connections/blogs/sdn/entry/did\\_you\\_learn\\_what\\_sdn\\_software\\_defined\\_networking\\_is\\_at\\_ons\\_2013?lang=en\\_us](http://www-304.ibm.com/connections/blogs/sdn/entry/did_you_learn_what_sdn_software_defined_networking_is_at_ons_2013?lang=en_us).
7. *Grusho A. A., Grusho N. A., Timonina E. E.* Statistical decision functions based on bans // AIP Conference Proceedings, 2015. Vol. 1648. P. 250002-1–250002-4.
8. *Забейайло М. И.* К вопросу о достаточности оснований для принятия результатов интеллектуального анализа данных средствами ДСМ-метода // Научно-техническая информация, 2015. Сер. 2. № 1. С. 1–9.
9. *Забейайло М. И.* О некоторых оценках сложности вычислений в ДСМ-рассуждениях // Искусственный интеллект и принятие решений, 2015. № 1. С. 3–17 (Ч. I); № 2. С. 3–17 (Ч. II).
10. *Забейайло М. И.* О некоторых возможностях управления перебором в ДСМ-методе // Искусственный интеллект и принятие решений, 2014. № 1. С. 95–110 (Ч. I); № 3. С. 3–21 (Ч. II).
11. *Забейайло М. И.* К проблеме быстрого анализа заголовков при адресации пакетов данных в компьютерных сетях // Научно-техническая информация, 2012. Сер. 2. № 12. С. 8–15.
12. *Грушо А., Грушо Н., Тимонина Е., Шоргин С.* Безопасные архитектуры распределенных систем // Системы и средства информатики, 2014. Т. 24. № 3. С. 18–31.
13. ENISA (European Network and Information Security Agency) Procure Secure: A guide to monitoring security service levels in the cloud environment. [http://www.enisa.europa.eu/activities/Resilience-and-CIIP/cloud-computing/procure-secure-a-guide-to-monitoring-of-securityservice-levels-in-cloud-contracts/at\\_download/fullReport](http://www.enisa.europa.eu/activities/Resilience-and-CIIP/cloud-computing/procure-secure-a-guide-to-monitoring-of-securityservice-levels-in-cloud-contracts/at_download/fullReport).
14. *Зацаринный А. А., Козлов С. В., Гаранин А. И.* Стенд главного конструктора — организационно-техническая основа разработки крупномасштабных информационно-телекоммуникационных систем // Системы и средства информатики, 2010. Т. 20. № 3. С. 174–190.
15. OpenVSwitch documentation. <http://openvswitch.org/support>.
16. *Pepple K.* Deploying OpenStack. — O'Reilly Media, 2011. 70 p.
17. Документация по OpenStack. <http://docs.openstack.org>.
18. *Kostenko V., Plakunov A., Nikolaev A., Tabolin V., Smeliansky R., Shakhova M.* Self-organizing cloud platform // SDN&NFV: The Next Generation of Computational Infrastructure: 2014 Science and Technology Conference (International) — Modern Networking Technologies (MoNeTec) Proceedings. — Moscow: MAKS Press, 2014. P. 77–82.
19. KVM. [http://www.linux-kvm.org/page/Main\\_Page](http://www.linux-kvm.org/page/Main_Page).
20. Kali Linux. <https://www.kali.org>.
21. Introduction to Mininet. <https://github.com/mininet/mininet/wiki/Introduction-to-Mininet>.

22. Network Simulator 3. <http://www.nsnam.org>.
23. OMNeT++ Network Simulation Framework Homepage. <http://www.omnetpp.org>.
24. OpenFlow-enabled SDN and Network Functions Virtualization ONF Solution Brief. February 17, 2014. <https://www.opennetworking.org/images/stories/downloads/sdn-resources/solution-briefs/sb-sdn-nvf-solution.pdf>.

*Поступила в редакцию 11.09.15*

---

---

## **ARCHITECTURE OF THE STAND FOR THE PILOT STUDY OF MODELS, ALGORITHMS, AND SOLUTIONS OF INFORMATION SECURITY IN CLOUD COMPUTING ENVIRONMENTS**

*A. A. Grusho<sup>1</sup>, M. I. Zabezhailo<sup>2</sup>, A. A. Zatsarinny<sup>3</sup>, A. V. Nikolaev<sup>4</sup>,  
and V. O. Piskovski<sup>5</sup>*

<sup>1</sup>Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

<sup>2</sup>All-Russian Institute for Scientific and Technical Information of Russian Academy of Sciences, 20 Usievicha Str., Moscow 125190, Russian Federation

<sup>3</sup>Federal Research Center “Computer Sciences and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

<sup>4</sup>Semenov Institute of Chemical Physics of the Russian Academy of Sciences, 4 Kosygina Str., Moscow 119991, Russian Federation

<sup>5</sup>Non-Profit Partnership “Applied Research Center for Computer Networks” (ARCCN), Bld. 1, suite 407B, Moscow Region, Leninsky District, Business Park “Rumyantsevo”

**Abstract:** The paper contains a concept to simulate virtual networks for experimental research in the field of information security methods in cloud computing environment. The paper proposes different approaches and systems to simulate networks and network segments. The paper presents a deployment diagram with technical requirements to hardware and software for workbench environment. Workbench properties cover prerequisites to research secure cloud computing solutions.

**Keywords:** information security; cloud computing; Software Defined Networks; OpenStack

**DOI:** 10.14357/08696527150405

## Acknowledgments

The work was supported by the Russian Foundation for Basic Research (project 15-29-07981 ofi-m).

## References

1. Open Data Center Alliance Master Usage Model: Commercial Framework Rev 1.0. Available at: [http://www.opendatacenteralliance.org/docs/ODCA\\_Commercial\\_Framework\\_MasterUM\\_v1.0\\_Nov2012.pdf](http://www.opendatacenteralliance.org/docs/ODCA_Commercial_Framework_MasterUM_v1.0_Nov2012.pdf) (accessed September 11, 2015).
2. Amies, A., H. Sluiman, Q. G. Tong, and G.N. Liu. 2012. Developing and hosting applications on the cloud. IBM Press. 367 p.
3. IETF Virtual eXtensible Local Area Network (VXLAN): A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks. Available at: <https://tools.ietf.org/html/rfc7348> (accessed September 11, 2015).
4. Grusho, A., N. Grusho, V. Piskovski, and E. Timonina. 2014. Five SDN-oriented directions in information security. *SDN&NFV: The Next Generation of Computational Infrastructure: 2014 Science and Technology Conference (International) — Modern Networking Technologies (MoNeTec) Proceedings*. Moscow: MAKS Press. 68–71.
5. Network virtualization. Available at: <http://networkheresy.com/2012/05/31/network-virtualization> (accessed September 11, 2015).
6. Saha, R. SDN, OpenFlow and Network Virtualization. IBM. Available at: [http://www-304.ibm.com/connections/blogs/sdn/entry/did\\_you\\_learn\\_what\\_sdn\\_software\\_defined\\_networking\\_is\\_at\\_ons\\_2013?lang=en\\_us](http://www-304.ibm.com/connections/blogs/sdn/entry/did_you_learn_what_sdn_software_defined_networking_is_at_ons_2013?lang=en_us) (accessed September 11, 2015).
7. Grusho, A. A., N. A. Grusho, and E. E. Timonina. 2015. Statistical decision functions based on bans. *AIP Conference Proceedings*. 1648:250002-1–250002-4.
8. Zabezhaylo, M. I. 2015. K voprosu o dostatochnosti osnovaniy dlya prinyatiya rezul'tatov intellektual'nogo analiza dannykh sredstvami DSM-metoda [To the question on the sufficiency of reasons set to admit results of intellectual analysis via DSM-methods]. *Nauchno-Tekhnicheskaya Informatsiya* [Scientific and Technical Information] 2(1):1–9.
9. Zabezhaylo, M. I. 2015. O nekotorykh otsenkakh slozhnosti vychisleniy v DSM-rassuzhdeniyakh [About some computational complexity estimations in DSM-considerations]. *Iskusstvennyy Intellekt i Prinyatie Resheniy* [Artificial Intelligence and Decision Making] 1:3–17 (Pt. I); 2:3–17 (Pt. II).
10. Zabezhaylo, M. I. 2014. O nekotorykh vozmozhnostyakh upravleniya pereborom v DSM-metode [About some prospects to control the enumeration of possibilities in DSM method]. *Iskusstvennyy Intellekt i Prinyatie Resheniy* [Artificial Intelligence and Decision Making] 1:95–110 (Pt. I); 3:3–21 (Pt. II).
11. Zabezhaylo, M. I. 2012. K probleme bystrogo analiza zagolovkov pri adresatsii paketov dannykh v komp'yuternykh setyakh [To the problem of fast analyzing data packet headings in computer networks]. *Nauchno-Tekhnicheskaya Informatsiya* [Scientific and Technical Information] 2(12):8–15.
12. Grusho, A., N. Grusho, E. Timonina, and S. Shorgin. 2014. Bezopasnye arkhitektury raspredelennykh sistem [Safety architectures of distributed systems]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 24(3):18–31.

13. ENISA (European Network and Information Security Agency) Procure Secure: A guide to monitoring security service levels in the cloud environment. Available at: [http://www.enisa.europa.eu/activities/Resilience-and-CIIP/cloud-computing/procure-secure-a-guide-to-monitoring-of-securityservice-levels-in-cloud-contracts/at\\_download/fullReport](http://www.enisa.europa.eu/activities/Resilience-and-CIIP/cloud-computing/procure-secure-a-guide-to-monitoring-of-securityservice-levels-in-cloud-contracts/at_download/fullReport) (accessed September 11, 2015).
14. Zatsarinnyy, A. A., S. V. Kozlov, and A. I. Garanin. 2010. Stend glavnogo konstruktora — organizatsionno-tekhnicheskaya osnova razrabotki krupnomasshtabnykh informatsionno-telekommunikatsionnykh system [The test-site of a chief designer is a base to develop large-scaled ICT systems]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 20(3):174–190.
15. OpenVSwitch documentation. Available at: <http://openvswitch.org/support> (accessed September 11, 2015).
16. Pepple, K. 2011. *Deploying OpenStack*. O'Reilly Media. 70 p.
17. Documentation on OpenStack. Available at: <http://docs.openstack.org> (accessed September 11, 2015).
18. Kostenko, V., A. Plakunov, A. Nikolaev, V. Tabolin, P. Smeliansky, and M. Shakhova. 2014. Self-organizing cloud platform. *SDN&NFV: The Next Generation of Computational Infrastructure: 2014 Science and Technology Conference (International) — Modern Networking Technologies (MoNeTec) Proceedings*. Moscow: MAKS Press. 77–82.
19. KVM. Available at: [http://www.linux-kvm.org/page/Main\\_Page](http://www.linux-kvm.org/page/Main_Page) (accessed September 11, 2015).
20. Kali Linux. Available at: <https://www.kali.org> (accessed September 11, 2015).
21. Introduction to Mininet. Available at: <https://github.com/mininet/mininet/wiki/Introduction-to-Mininet> (accessed September 11, 2015).
22. Network Simulator 3. Available at: <http://www.nsnam.org> (accessed September 11, 2015).
23. OMNeT++ Network Simulation Framework Homepage. Available at: <http://www.omnetpp.org> (accessed September 11, 2015).
24. OpenFlow-enabled SDN and Network Functions Virtualization ONF Solution Brief. February 17, 2014. Available at: <https://www.opennetworking.org/images/stories/downloads/sdn-resources/solution-briefs/sb-sdn-nvf-solution.pdf> (accessed September 11, 2015).

*Received September 11, 2015*

## Contributors

**Grusho Alexander A.** (b. 1946) — Doctor of Science in physics and mathematics, professor, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; [grusho@yandex.ru](mailto:grusho@yandex.ru)

**Zabezhailo Michael I.** (b. 1956) — Candidate of Science (PhD) in physics and mathematics, senior scientist, All-Russian Institute for Scientific and Technical

Information of Russian Academy of Sciences, 20 Usievicha Str., Moscow 125190, Russian Federation; zmizanv@ya.ru

**Zatsarinny Alexander A.** (b. 1951) — Doctor of Science in technology, professor, Deputy Director, Federal Research Center “Computer Sciences and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; azatsarinny@ipiran.ru

**Nikolaev Andrey V.** (b. 1973) — Candidate of Science (PhD) in physics and mathematics, scientist, Semenov Institute of Chemical Physics of the Russian Academy of Sciences, 4 Kosygina Str., Moscow 119991, Russian Federation; gentoorion@mail.ru

**Piskovski Viktor O.** (b. 1963) — Candidate of Science (PhD) in physics and mathematics, leading programmer, Non-Profit Partnership “Applied Research Center for Computer Networks” (ARCCN), Bld. 1, suite 407B, Moscow Region, Leninsky District, Business Park “Rumyantsevo;” vpv80@yandex.ru

## АППАРАТНО-ПРОГРАММНОЕ МОДЕЛИРОВАНИЕ И ТЕСТИРОВАНИЕ РЕКУРРЕНТНОГО ОПЕРАЦИОННОГО УСТРОЙСТВА\*

*Д. В. Хилько<sup>1</sup>, Ю. А. Степченков<sup>2</sup>, Ю. Г. Дьяченко<sup>3</sup>,  
Ю. И. Шикунов<sup>4</sup>, Н. В. Морозов<sup>5</sup>*

**Аннотация:** Рассматриваются основные проблемы моделирования и отладки операционного устройства, основанного на принципах нетрадиционной потоковой рекуррентной архитектуры. Приводится краткое описание методологии аппаратно-программного моделирования Hardware in the Loop, также рассматривается возможность ее применения для разработки рекуррентного операционного устройства (РОУ). Предложен подход и методика аппаратно-программного моделирования и тестирования устройства на основе новой архитектуры. Описываются разработанные программная и аппаратная модели устройства, средства моделирования и отладки, а также результат их интеграции в единую среду. Вводится понятие целевой платформы моделирования, названной GAROS IDE. Приводится описание ее основных компонент, функциональных возможностей и перспектив развития. Представлены результаты апробации платформы на одной из подзадач распознавания изолированных слов — алгоритме Rasta-фильтрации.

**Ключевые слова:** моделирование; тестирование; рекурсивность; рекуррентность; потоковая архитектура; методология программирования

**DOI:** 10.14357/08696527150406

### 1 Введение

Коллективом отдела «Архитектуры перспективных компьютерных систем» ИПИ РАН ведется разработка концепций принципиально новой архитектуры,

---

\*Исследование выполнено при частичной финансовой поддержке РФФИ (проект 13-07-12068).

<sup>1</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, dhilko@yandex.ru

<sup>2</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, ystepchenkov@ipiran.ru

<sup>3</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, diaura@mail.ru

<sup>4</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, yshikunov@yandex.ru

<sup>5</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, nmorozov@ipiran.ru

а также вычислительного устройства на ее основе. Предлагаемая нетрадиционная архитектура предназначена для реализации параллельных вычислений ограниченной размерности в области цифровой обработки сигналов (ЦОС).

В настоящее время разрабатывается опытный образец, названный рекуррентным обработчиком сигналов (РОС), исполняемый в гибридном, двухуровневом варианте с ведущим фон-неймановским процессором на управляющем (верхнем) уровне (УУ) и рядом потоковых процессоров на нижнем уровне — РОУ [1]. Данная реализация архитектуры была названа «Гибридная архитектура рекуррентного обработчика сигналов» (ГАРОС). В качестве демонстрационной задачи для ГАРОС была выбрана задача распознавания изолированных слов ввиду высокой степени готовности ее математического обеспечения и ее алгоритмов, воспроизводящих широкий спектр типовых задач ЦОС.

Предварительным этапом разработки любой технически сложной системы является создание ряда моделей: математических, программных и аппаратных, средствами которых осуществляется отладка основных механизмов функционирования системы. В ходе проектирования указанных моделей был выявлен ряд проблем, результат решения которых был оформлен в виде *методики (подхода) аппаратно-программного моделирования рекуррентного обработчика сигналов*. Настоящая статья посвящена описанию выявленных проблем, разработанного подхода и результатов его апробации.

## **2 Проблемы моделирования ГАРОС. Методика аппаратно-программного моделирования и отладки**

Существует несколько основных подходов к разработке систем ЦОС. В работе [2] автор приводит достаточно подробное описание данных подходов, а также приводит ключевые аспекты метода аппаратно-программного моделирования Hardware in the Loop (HIL). «Методология HIL обеспечивает промежуточный уровень контроля разрабатываемых систем — между программным обеспечением моделирования и тестированием оборудования проекта», — пишет автор. Согласно методологии HIL все основные этапы разработки систем ЦОС интегрируются в единый итерационный цикл проектирования. Другими словами, программные и аппаратные модели системы функционируют и отлаживаются одновременно.

В рамках HIL используются такие программные средства, как MATLAB/Simulink, Altera DSP Builder и др. С точки зрения разработчиков РОС приведенная выше методология HIL оказалась привлекательной для внедрения. Однако был обнаружен целый ряд проблем, препятствующих приложению HIL в полном объеме. Среди этих проблем можно выделить следующие:

- (1) разработчикам не удалось реализовать некоторые положения, лежащие в основе ГАРОС, в частности рекуррентность средствами среды MATLAB/Simulink;



- (2) построение VHDL (VHSIC (Very High Speed Integrated Circuits) Hardware Description Language) модели на основе математической трудно или невозможно завершить ввиду проблемы 1;
- (3) уникальность представления данных в ГАРОС в виде «элемента самоопределяющихся данных» сильно затрудняет или делает невозможным трансляцию в язык C++, используемый в системе MATLAB в качестве основного;
- (4) результат трансляции математической модели в VHDL и его последующий перенос в программируемую логическую интегральную схему (ПЛИС) не удовлетворяют физическим ограничениям микросхемы;
- (5) отсутствие средств разработки и отладки программного обеспечения для РОС привело к необходимости ручного проектирования и тестирования, что сильно затрудняло автоматизацию процесса моделирования.

Были выявлены и другие проблемы моделирования, но указанные являются основными. В процессе поиска решения данных проблем был выработан подход к организации процесса моделирования и отладки ГАРОС, определены компоненты аппаратно-программных средств, которые необходимо создать, осуществлена их разработка. Предлагаемый подход заключается в *разработке целевой платформы* для проверки работоспособности, оценки производительности оборудования системы ГАРОС в реальных условиях, отладки специализированного программного обеспечения (ПО) для ГАРОС.

Указанная целевая платформа должна включать в себя: средства математического моделирования ГАРОС (имитационную модель, написанную на языке высокого уровня); средства аппаратного моделирования (VHDL-модель, исполняемую в среде ModelSim); опытный образец на ПЛИС; средства разработки ПО, предназначенного для функционирования в среде ГАРОС.

С учетом выбранного подхода по аналогии с методологией НИЛ была разработана *методика аппаратно-программного моделирования и отладки ГАРОС*:

- (1) разработка с нуля либо усовершенствование математической (имитационной) модели, в качестве платформы разработки был выбран язык программирования C#;
- (2) разработка с нуля либо усовершенствование VHDL-модели средствами программных продуктов Altera;
- (3) разработка тестов проверки соответствия имитационной и VHDL-модели и их последующее моделирование;
- (4) реализация проекта на базе ПЛИС компании Altera;
- (5) отладка полученной системы;
- (6) интеграция этапов 1–5 в единый итерационный цикл проектирования;
- (7) создание средств разработки ПО, а также единой среды моделирования, интегрирующей в себе вышеуказанные модели и тесты;



- (8) итерационная разработка и отладка ГАРОС средствами единой среды моделирования.

Таким образом, была заложена основа интегрированной среды разработки для ГАРОС, названная авторами в работе [3] GAROS IDE.

### 3 Описание среды GAROS IDE и ее ключевых компонентов

В процессе поиска решения проблем программируемости ГАРОС была разработана методология программирования, названная авторами в работе [4] *рекуррентно-поточковой методологией программирования*. В рамках данной методологии удалось формализовать процесс разработки программного обеспечения, определить его ключевые этапы, а также подходы к организации отладки разрабатываемых программ. Формализованное представление процесса программирования ГАРОС послужило фундаментом для разработки прототипа интегрированной среды программирования GAROS IDE, функциональные возможности компонентов которого охватывают все введенные в рамках методологии этапы.

В настоящее время ведутся работы по развитию и усовершенствованию интегрированной среды разработки GAROS IDE, назначение которой эволюционировало от простой среды программирования до полноценного комплекса моделирования, программирования и отладки. В работе [5] приводится общая структура прототипа этой среды, ее архитектура и описание компонентов. На рис. 1 представлена структура GAROS IDE.

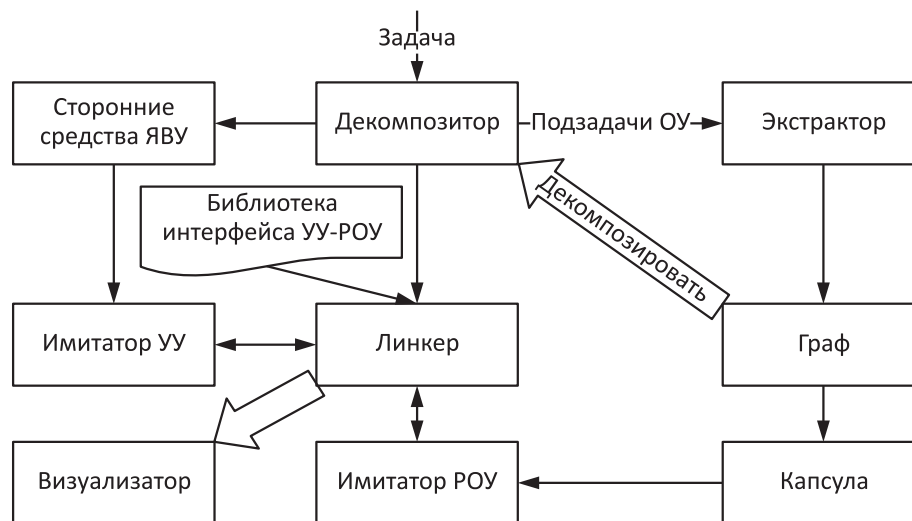


Рис. 1 Обобщенная структура GAROS IDE

*Компонент «Декомпозитор»* предназначен для декомпозиции решаемой задачи.

*Сторонние средства ЯВУ* (языков высокого уровня) — набор программных средств для разработки традиционного ПО: для реализации программы УУ.

*Компонент «Экстрактор»* предназначен для извлечения графа вычислительного процесса (ВП) из описания подзадачи на языке высокого уровня.

*Компонент «Граф»* — многофункциональный компонент, предоставляющий широкий спектр возможностей для работы с визуальными графами. В случае необходимости (в соответствии с рекуррентно-поточковой методологией) подзадача может быть передана в «Декомпозитор» для повторной декомпозиции.

*Компонент «Капсула»* предназначен для работы со специализированными программами, функционирующими в РОУ, — капсулами.

*Компонент «Имитатор УУ»* предназначен для интерпретирования программы УУ. Может работать в двух режимах: полной интерпретации программы и интерпретации при помощи программы ПРАПОР [6].

*Компонент «Имитатор РОУ»* предназначен для интерпретирования программы операционного уровня. Может работать в четырех различных режимах: интерпретация в режиме ПРАПОР, имитационное моделирование при помощи программы СИМПРА (программа реализует имитационную модель РОУ) [7], моделирование при помощи VHDL-модели, интерпретация на опытном образце.

*Компонент «Линкер»* предназначен для организации взаимодействия между имитаторами УУ и РОУ с учетом информации, хранящейся в картах памяти и данных. Осуществляет сборку и интерпретацию всей задачи в целом.

*Компонент «Визуализатор»* предназначен для отображения результатов моделирования и отладки ПО.

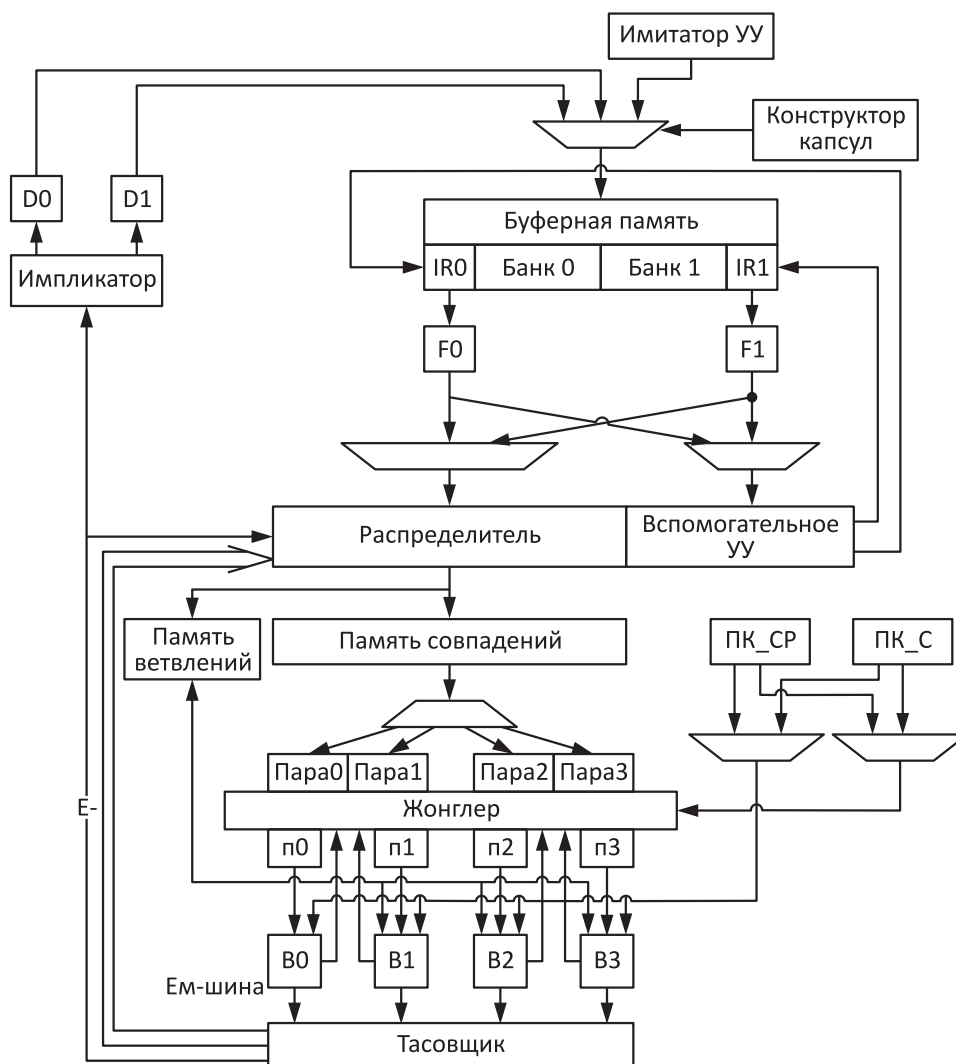
С точки зрения описываемого в статье подхода и методики аппаратно-программного моделирования и отладки ГАРОС наиболее важным компонентом является «Имитатор РОУ», включающий в себя как имитационную модель РОУ, так и аппаратную VHDL-модель РОУ. В работе [1] приводится подробное описание VHDL-модели РОУ, а в работе [8] — описание имитационной модели. Далее приводится их краткое описание.

#### **4 Имитационная модель рекуррентного операционного устройства**

Имитационная модель включает в себя следующие компоненты:

- (1) буферная память (БП) — компонент, логически разделенный на два блока (физически — единое целое); адресуется двумя индексными регистрами IR0 и IR1;
- (2) интерфейсные регистры IR0, IR1, F0, F1, D0, D1, предназначенные для моделирования взаимодействия между УУ, БП и РОУ, а также для поддержки библиотеки интерфейса;

- (3) вспомогательное устройство управления, организующее многократное исполнение капсулы и ее фрагментов;
- (4) распределитель — основной компонент, отвечающий за обработку потока операндов и их рассылку по секциям назначения;



**Рис. 2** Архитектура имитационной модели

- (5) импликатор — компонент обработки выходных результатов, отвечающий за своевременный отбор выходных данных потока промежуточных операндов и запись их в БП;
- (6) память констант (секционная и секционная-регистровая) — компонент, предназначенный для хранения постоянных данных;
- (7) память совпадений — упрощенный аналог ассоциативной;
- (8) память переходов, хранящая условия переходов и функциональные поля операнда — результата перехода, структурно аналогичная памяти совпадений;
- (9) жонглер, предназначенный для анализа пар операндов и принятия решения о характере протекания вычислений в вычислителе;
- (10) вычислитель (4 шт.) — совокупность вычислительных блоков (умножитель, АЛУ16 и АУ40), поддерживающая разные режимы суперскалярной работы;
- (11) шины E- и Em-магистралы пересылки промежуточных данных;
- (12) конструктор капсул.

Архитектура имитационной модели представлена на рис. 2.

### VHDL-модель рекуррентного операционного устройства

На рис. 3 приведена структура VHDL-модели ГАРОС. Ключевые компоненты соответствуют аналогичным в имитационной модели.

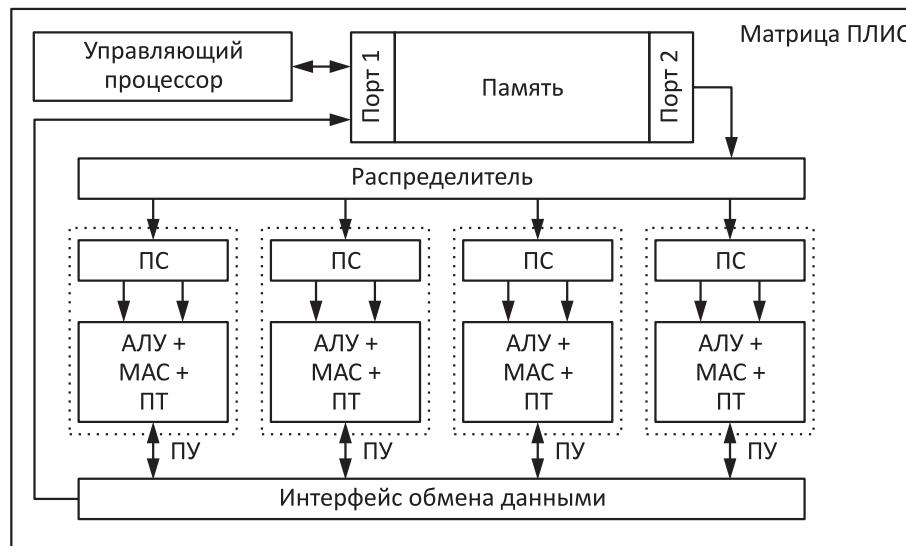


Рис. 3 Структура VHDL-модели ГАРОС

«В состав «Вычислителя» включено автономное устройство преобразования тегов (ПТ), которое обладает возможностью саморазвертки вычислительного процесса. Устройство ПТ представляет собой простую комбинационную схему, содержащую средства настройки на предметную область. В памяти есть только набор операндов и начальных значений их функциональных полей, которые динамически подвергаются рекуррентной саморазвертке устройством ПТ» [3].

## 5 Апробация аппаратно-программной среды моделирования

Среда разработки и моделирования GAROS IDE находится в активной стадии разработки и постоянно совершенствуется. В настоящее время она позволяет осуществлять разработку капсул, выполнять их моделирование и отладку как средствами имитационной модели, так и средствами аппаратной модели, причем выполнять оба этих моделирования можно одновременно с постоянным контролем результатов.

Для апробации возможностей созданной платформы был реализован и отлажен ряд подзадач распознавания изолированных слов. Каждая из подзадач была реализована в виде капсулы, были получены результаты их моделирования. Для демонстрации возможностей среды рассмотрим алгоритм RASTA-фильтрации. При помощи средств GAROS IDE полученная капсула, реализующая алгоритм RASTA-фильтрации, преобразуется во внутренний формат do-файла моделирования ModelSim. На рис. 4 представлен снимок экрана формирования до-капсулы.

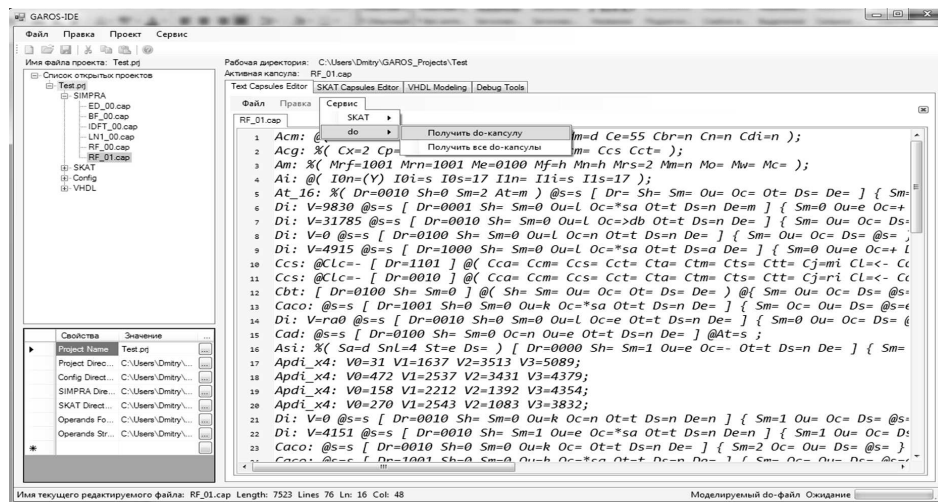


Рис. 4 Формирование капсулы и до-капсулы

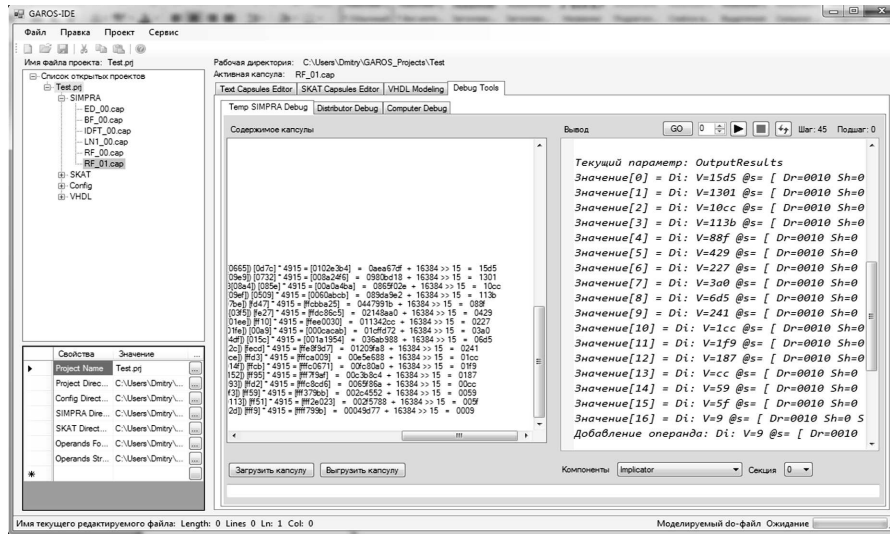


Рис. 5 Результаты имитационного моделирования

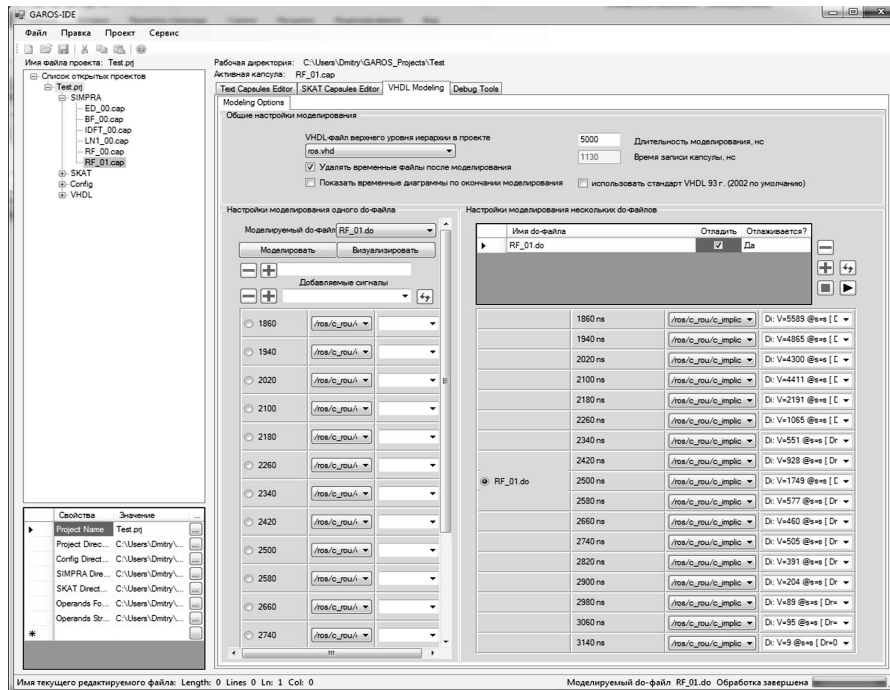


Рис. 6 Результаты VHDL-моделирования

Далее полученная капсула имплементируется в Си-программу распознавателя изолированных слов с целью извлечения всех промежуточных и выходных данных, необходимых для проверки корректности функционирования как капсулы, так и имитационной модели. Извлеченные данные могут быть вставлены в капсулу в виде комментариев для удобства. Следующим шагом является моделирование и отладка капсулы средствами имитационной модели РОУ. На рис. 5 представлен снимок экрана результатов моделирования, демонстрирующий совпадение всех выходных результатов с ожидаемыми.

Одновременно с имитационным испытанием можно запустить VHDL-модель средствами ModelSim, так как этот процесс в GAROS IDE выполняется в фоновом потоке. Результаты моделирования могут наблюдаться тремя способами: в виде временных диаграмм моделирования (наиболее детализированное представление); в виде SVG (Scalable Vector Graphics) диаграмм визуализации результатов (наиболее удобный для отладки); в виде набора наблюдаемых сигналов (для отлаженных капсул). На рис. 6 приводится снимок экрана результатов VHDL-моделирования капсулы RASTA-фильтрации. Представленные данные (Di) имеют десятичный формат. При переводе их в шестнадцатеричный формат нетрудно убедиться, что данные совпадают.

## 6 Заключение

В ходе работ по созданию вычислительного устройства, основанного на рекуррентно-поточковой модели вычислений, был разработан эффективный подход, а также методика аппаратно-программного моделирования и тестирования ГАРОС. Для реализации предложенного подхода были спроектированы аппаратно-программные средства, сформировавшие мощную платформу GAROS IDE для выполнения широкого спектра задач: от разработки ПО и моделирования до выработки способа расширения функциональных возможностей ГАРОС.

Таким образом, можно говорить, что платформа GAROS IDE, реализующая аппаратно-программный подход к моделированию и отладке ГАРОС, является эффективным средством поддержки разработки и поэтому требует дальнейшего развития и совершенствования функциональных возможностей.

## Литература

1. *Степченко Ю. А., Петрухин В. С.* Особенности гибридного варианта реализации на ПЛИС рекуррентного обработчика сигналов // Системы и средства информатики, 2008. Доп. вып. С. 118–129.
2. *Вычужанин В.* Методика аппаратно-программного моделирования и тестирования проектируемых систем // Современная электроника, 2015. № 2. С. 62–67.
3. *Хилько Д. В., Степченко Ю. А., Шикун Ю. И.* Инструментальная среда проектирования ПО для гибридной архитектуры рекуррентного обработчика сигналов

- (GAROS IDE). Свидетельство о государственной регистрации программы для ЭВМ № 2015614004.
4. Хилько Д. В., Степченков Ю. А. Теоретические аспекты разработки методологии программирования рекуррентной архитектуры // Системы и средства информатики, 2013. Т. 23. № 2. С. 133–156.
  5. Хилько Д. В., Шикунов Ю. И. Разработка инструментальной среды проектирования программного обеспечения для рекуррентно-поточковой модели вычислений // 4-я школа молодых ученых ИПИ РАН: Сб. докл. — М.: ИПИ РАН, 2013. С. 65–77.
  6. Хилько Д. В., Степченков Ю. А., Шнейдер А. Ю. Программа обработки результатов моделирования потоковой рекуррентной архитектуры (ПРАПОР). Свидетельство о государственной регистрации программы для ЭВМ № 2013610199.
  7. Хилько Д. В., Степченков Ю. А. Средства имитационного моделирования потоковой рекуррентной архитектуры (СИМПРА). Версия 2. Свидетельство о государственной регистрации программы для ЭВМ № 2014610123.
  8. Хилько Д. В., Степченков Ю. А., Шикунов Ю. И. Средства имитационного моделирования многоядерной потоковой рекуррентной архитектуры // Многоядерные процессоры, параллельное программирование, ПЛИС, системы обработки сигналов: Сб. научных статей II Всеросс. науч.-практич. конф. — Барнаул: АлтГУ, 2014. С. 58–69.

*Поступила в редакцию 30.09.15*

---

---

## **HARDWARE AND SOFTWARE MODELING AND TESTING OF THE RECURRENT OPERATIONAL DEVICE**

*D. V. Khilko, Yu. A. Stepchenkov, Yu. G. Diachenko, Yu. I. Shikunov,  
and N. V. Morozov*

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** The paper considers the main problems of modeling and debugging of the operational device based on the principles of the nonconventional data-flow recurrent architecture. The short description of the methodology of hardware and software modeling of Hardware in the Loop is provided; also, the possibility of its application for developing a recurrent operational device is considered. An approach and a technique of hardware and software modeling and testing of devices on the basis of the new architecture are suggested. The paper describes the developed program and models of the device’s hardware, simulating and debugging tools, and the result of their integration into the uniform environment. The concept of the target platform of modeling called GAROS IDE is introduced. The descriptions of its main components, functionality, and prospects of development are presented. The results of approbation of this platform on one of the subtasks of isolated word recognition (the algorithm of Rasta-filtration) are reported.



**Keywords:** modeling; testing; recursiveness; recurrence property; data-flow architecture; programming methodology

**DOI:** 10.14357/08696527150406

## Acknowledgments

The work was partially supported by the Russian Foundation for Basic Research (project No. 13-07-12068).

## References

1. Stepchenkov, Yu. A., and V. S. Petrukhin. 2008. Osobennosti gibridnogo varianta realizatsii na PLIS rekurrentnogo obrabotchika signalov [Features of the hybrid embodiment of the recurrent signal processor on FPGA]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics*. Add. iss.:118–129.
2. Vychuzhanin, V. 2015. Metodika apparatno-programmnogo modelirovaniya i testirovaniya proektiruemykh sistem [Methods of hardware and software simulation and testing the designed system]. *Sovremennaya Elektronika* [Modern Electronics] 2:62–67.
3. Khilko, D. V., Yu. A. Stepchenkov, and Yu. I. Shikunov. Instrumental'naya sreda proektirovaniya PO dlya gibridnoy arkhitektury rekurrentnogo obrabotchika signalov (GAROS IDE) [Instrumental development environment software for the hybrid architecture of recurrent signal processor (GAROS IDE)]. Certificate of state registration of computer program No. 2015614004.
4. Khilko, D. V., and Yu. A. Stepchenkov. 2013. Teoreticheskie aspekty razrabotki metodologii programmirovaniya rekurrentnoy arkhitektury [Theoretical aspects of programming methodology development of recurrent architecture]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 23(2):133–156.
5. Khilko, D. V., and Yu. I. Shikunov. 2013. Razrabotka instrumental'noy sredy proektirovaniya programmnogo obespecheniya dlya rekurrentno-potokovoy modeli vychisleniy [Development IDE software recursively-flow model calculations]. *4-ya shkola molodykh uchennykh IPI RAN* [4th School of Young Scientists of IPI RAN]. 65–77.
6. Khilko, D. V., Yu. A. Stepchenkov, and A. Yu. Shneyder. Programma obrabotki rezul'tatov modelirovaniya potokovoy rekurrentnoy arkhitektury (PRAPOR) [Processing simulation results program of data-flow recurrent architecture (PRAPOR)]. Certificate of state registration of computer program No. 2013610199.
7. Khilko, D. V., and Yu. A. Stepchenkov. Sredstva imitatsionnogo modelirovaniya potokovoy rekurrentnoy arkhitektury (SIMPRA). Versiya 2 [Simulation tools of data-flow recurrent architecture (SIMPRA). Version 2]. Certificate of state registration of computer program No. 2014610123.
8. Khilko, D. V., Yu. A. Stepchenkov, and Yu. I. Shikunov. 2014. Sredstva imitatsionnogo modelirovaniya mnogoyadernoy potokovoy rekurrentnoy arkhitektury [Simulation tools for multicore streaming recurrent architecture]. *Mnogoyadernye protsessy, parallel'noe programmirovaniye, PLIS, sistemy obrabotki signalov: Sb. nauchnykh statey Vserossiyskoy nauchno-praktich. konf.* [Multicore processes, parallel programming, FPGA, signal processing system: A collection of scientific papers of All-Russian Scientific-Practical Conference]. Barnaul. 58–69.

*Received September 30, 2015*

## **Contributors**

**Khilko Dmitri V.** (b. 1987) — scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; [dhilko@yandex.ru](mailto:dhilko@yandex.ru)

**Stepchenkov Yuri A.** (b. 1951) — Candidate of Science (PhD) in technology, Head of Department, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; [ystepchenkov@ipiran.ru](mailto:ystepchenkov@ipiran.ru)

**Diachenko Yuri G.** (b. 1958) — Candidate of Science (PhD) in technology, senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, Moscow 119333, Russian Federation; [diaura@mail.ru](mailto:diaura@mail.ru)

**Shikunov Yuri I.** (b. 1995) — apprentice researcher, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; [yishikunov@yandex.ru](mailto:yishikunov@yandex.ru)

**Morozov Nikolay V.** (b. 1956) — senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, Moscow 119333, Russian Federation; [nmorozov@ipiran.ru](mailto:nmorozov@ipiran.ru)

## ИССЛЕДОВАНИЕ ОСОБЕННОСТЕЙ СИГНАЛОВ МИОГРАММЫ\*

*Т. В. Захарова*<sup>1</sup>, *В. Ю. Королев*<sup>2</sup>, *А. А. Щемирова*<sup>3</sup>

**Аннотация:** Исследуется поведение сигналов мышечной активности человека — миограмм, регистрируемых с применением метода вызванных потенциалов. Авторами предлагаются новые методы обработки этих сигналов для определения так называемых опорных точек, которые соответствуют моментам начала движения в проводимом эксперименте. В дальнейшем локализованные опорные точки используются для построения ассоциативного фильтра обработки сверхслабых нестационарных сигналов магнитоэнцефалограмм. Новые методы сохраняют высокую точность ранее применяемых методов, но работают в несколько раз быстрее, что имеет важное значение при их использовании в режиме реального времени.

**Ключевые слова:** метод вызванных потенциалов; выборочная оконная дисперсия; миограмма; магнитоэнцефалограмма; проверка гипотез

**DOI:** 10.14357/08696527150407

### 1 Введение

Задача безопасной и безболезненной диагностики состояния пациента, особенно связанной с головным мозгом, который наиболее чувствителен к оперативному вмешательству, представляет большой практический интерес. Предоперационные методы локализации невосполнимых участков мозга играют исключительно важную роль для проведения успешной операции на мозге человека и для здоровья пациента. При определении таких участков предпочтительно, чтобы процедура носила неинвазивный характер [1]. Тем не менее неинвазивные методы пока не стали общепринятой практикой и их разработка и внедрение является важной и актуальной задачей.

---

\*Работа выполнена при частичной поддержке РФФИ (проект 15-07-02652).

<sup>1</sup>Московский государственный университет им. М. В. Ломоносова, факультет вычислительной математики и кибернетики; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, lsa@cs.msu.ru

<sup>2</sup>Московский государственный университет им. М. В. Ломоносова, факультет вычислительной математики и кибернетики; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, vko@lev@cs.msu.ru

<sup>3</sup>Московский государственный университет им. М. В. Ломоносова, факультет вычислительной математики и кибернетики, lsa@cs.msu.ru

Магнитоэнцефалограф (МЭГ) используется для считывания слабых магнитных сигналов с датчиков, установленных на голове человека. Обработкой сигналов магнитоэнцефалограмм можно определить границы функциональных зон мозга [2–5].

В 2008 г. в России появился первый МЭГ, и задача анализа сигналов МЭГ приобрела практический характер. Основной задачей исследования являлось выделение участка коры головного мозга, инициирующего движение. Требовалось определить датчик, регистрирующий МЭГ-сигнал с наиболее сильной компонентой движения. В ходе работы выяснилось, что стандартные методы не подходят для обработки таких сложных сигналов, как магнитоэнцефалограммы, по целому ряду причин: сверхслабое магнитное поле, низкое соотношение сигнал/шум, нестационарность, неоднородность и др. Моменты начала движения по МЭГ-сигналу определить было невозможно.

Свойства миограмм значительно лучше, так как энергия сигналов миограмм на несколько порядков выше сигналов МЭГ. Но миограмма является показателем мышечной активности, которая производится с некоторой задержкой относительно работы мозга, поэтому по миограммам определялись опорные точки — моменты начала мышечной активности, вызванной работой мозга. В дальнейшем, используя опорные точки, устанавливалась синхронизация МЭГ-сигналов и сигналов миограмм.

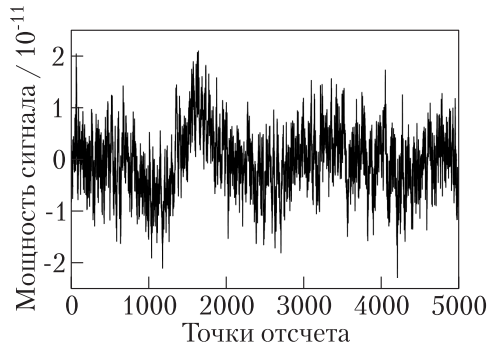
## **2 Схема проведения эксперимента**

В основе эксперимента лежит метод вызванных потенциалов, который заключается в периодическом повторении некоторого действия испытуемым. При этом записываются магнитоэнцефалограмма (сигналы активности головного мозга), миограмма движений руки испытуемого, сигналы с кнопки, а задача испытуемого состоит в периодическом нажатии на кнопку. Примеры получаемых сигналов показаны на рис. 1.

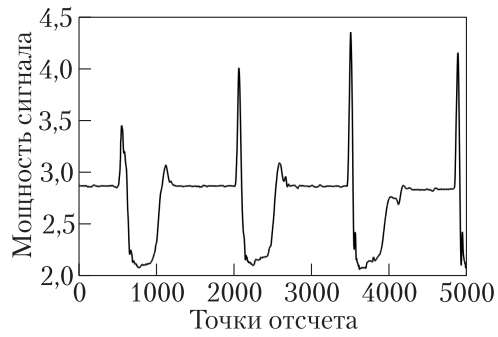
По рис. 2 хорошо видно, что сигнал с кнопки менее зашумлен, но он не отражает информацию о точном начале движения руки человека. Показания кнопки имеют значительное запаздывание относительно точек начала движения и большой разброс по времени, так как фиксируют показатели движения как результат работы мышцы. В связи с этим точность определения опорных точек по сигналам кнопки недостаточная. Но эти показания хороши для первичной, грубой обработки, так как по ним легко определять интервалы без движения (интервалы покоя) и тем самым разбивать сигнал на эпохи. Эпохой здесь называется интервал времени, содержащий одно движение.

Таким образом, наиболее ценным источником информации о начале движения является миограмма.

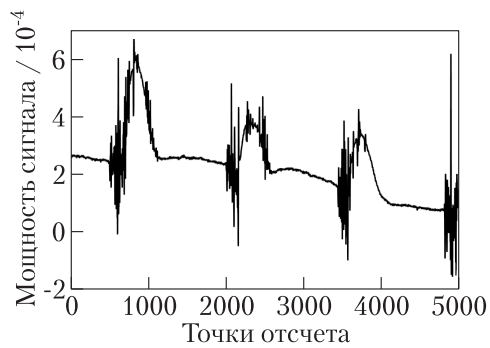
В данной статье рассматриваются методы нахождения опорных точек по миограмме (рис. 3). В схожих постановках задача исследовалась в работах [6–8].



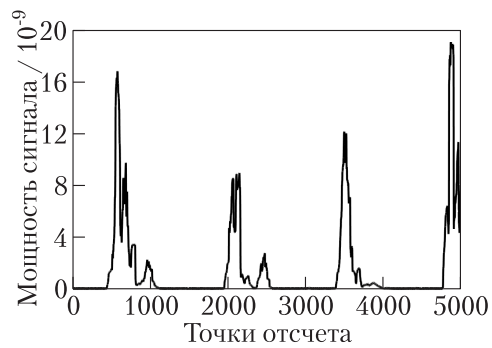
**Рис. 1** Магнитоэнцефалограмма



**Рис. 2** Сигнал с кнопки



**Рис. 3** Миограмма



**Рис. 4** Оконная дисперсия миограммы

Для улучшения выделения полезного сигнала (движения) по миограмме вычисляется оконная дисперсия [6, 9], определяемая формулой:

$$\frac{1}{N-1} \sum_{i=1}^N (X_i - \bar{X})^2,$$

где

$$\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i.$$

Далее везде рассматривается именно оконная дисперсия миограммы. Движением будем считать ту часть сигнала, на которой оконная выборочная дисперсия миограммы возрастает.

Оконная дисперсия миограммы показана на рис. 4. Ширина окна  $N$  была определена эмпирическим путем и равнялась 30 точкам отсчета.

### 3 Метод двухоконной прокрутки

Сигнал анализируется двумя смежными окнами размера  $N$ . Идея метода заключается в сравнении вероятностных характеристик сигнала в каждом из окон. Понятно, что на интервалах покоя в обоих окнах характеристики сигнала одинаковые и построенные на каждом окне эмпирические функции распределения практически не различимы. Различие же характеристик свидетельствует о переходе сигнала из состояния покоя в состояние движения. В работе были применены два метода для поиска участков покоя: метод Колмогорова–Смирнова (КС) и метод скользящего среднего.

Метод КС основан на предположении об одинаковой распределенности сигнала в соседних окнах на участках покоя.

В методе скользящего среднего используется принцип сравнения выборочных средних в двух соседних окнах. Рассматриваемые алгоритмы предназначены для локализации опорных точек на концах участков покоя.

#### 3.1 Метод Колмогорова–Смирнова

Метод КС предназначен для проверки гипотезы о принадлежности двух независимых выборок одному закону распределения. В данном случае в качестве выборки используется сигнал в отдельно взятом окне.

Эмпирическая функция распределения (э.ф.р.) рассчитывается по формуле

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I\{X_i \leq x\},$$

где  $I\{X_i \leq x\}$  — индикатор попадания  $X_i$  в область  $(-\infty, x]$ , определяемый как

$$I\{X_i \leq x\} = \begin{cases} 1, & X_i \leq x; \\ 0, & X_i > x. \end{cases}$$

Статистика сравнения эмпирической функции распределения первого окна  $F_n(x)$  с эмпирической функцией распределения второго окна  $F_m(x)$  задается формулой

$$D_{nm} = \sup_x |F_n(x) - F_m(x)|.$$

Пусть  $H_0$  — гипотеза о том, что выборки с э.ф.р.  $F_n$  и  $F_m$  одинаково распределены. Тогда по критерию КС гипотеза принимается, если  $\sqrt{nm/(n+m)} D_{nm}$  не превышает квантиль распределения Колмогорова  $K_\alpha$  для заданного уровня значимости  $\alpha$ .

В рассматриваемом случае  $n = m$  и ширина окна равнялась 60 точкам отсчета.

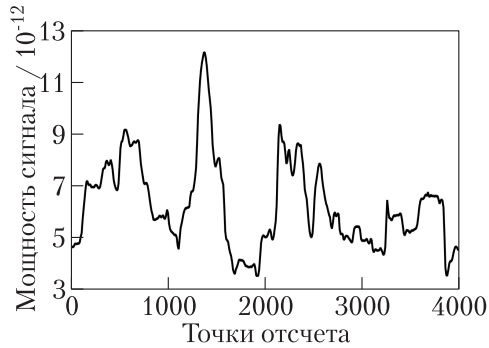


Рис. 5 Участок покоя при увеличении

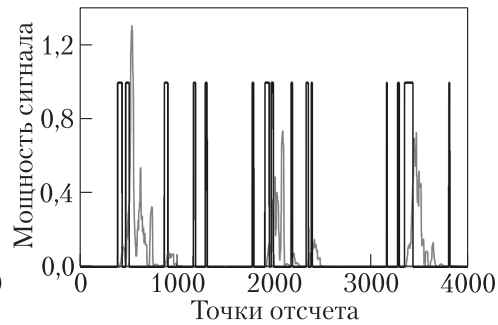


Рис. 6 Модифицированный КС-метод

В данной статье окном с номером  $n$  называется окно, которое начинается в точке  $n$  и состоит из  $N$  точек с номерами  $(n, n + 1, \dots, n + N - 1)$ .

Идея алгоритма состоит в последовательном применении КС-метода к каждой паре соседних окон с номерами  $n$  и  $n + N$ . Если выборки в них одинаково распределены, то первая точка окна  $(n + N)$  считается точкой покоя, иначе — точкой движения. Таким образом определяются участки покоя и участки движения по всему сигналу.

Следует отметить, что в оконной дисперсии миограммы значительно уменьшается влияние шума. Тем не менее на участках покоя сигнал не является константой, а колеблется в незначительных пределах (рис. 5).

Метод КС чрезвычайно чувствителен к шуму.

Для применения метода к сигналу требуется его модификация. Вместо квантили распределения Колмогорова возьмем эмпирически определяемую константу  $K = 0,82$ . Результаты работы метода представлены на рис. 6. Время работы метода на 5000 точках отсчета составляет примерно 10 с.

Вариацией метода является проход по сигналу тремя окнами размера  $N$ . Выберем три окна подряд с номерами  $n, n + N, n + 2N$ . Попарно сравним окна описанным выше КС-методом. Если все три сравнения привели к отклонению гипотезы об одинаковой распределенности выборок, то точка  $n + N$  будет считаться точкой движения. Результат работы алгоритма показан на рис. 7.

Данный метод с большой точностью определяет точку начала движения. На рис. 7 видно, что метод определил начало третьего движения до видимого изменения величины сигнала.

На рис. 8 увеличено изображение сигнала и рассчитанной точки отсчета третьего движения.

Трехоконная вариация метода дает выигрыш в точности, но время обработки значительно увеличивается.

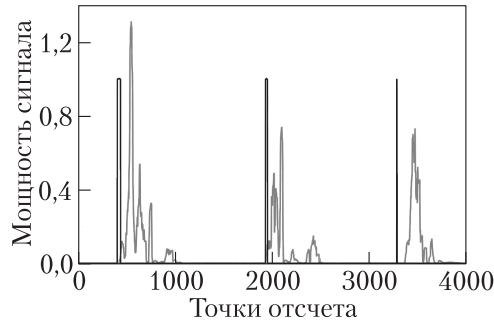


Рис. 7 Сигнал с опорными точками

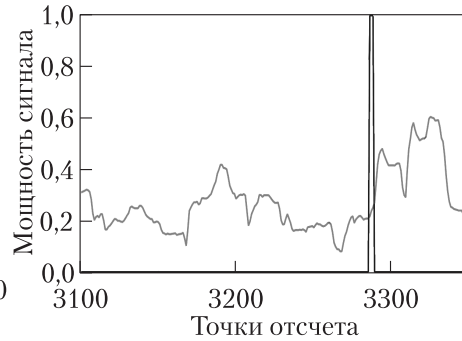


Рис. 8 Опорная точка третьего движения

### 3.2 Метод скользящего среднего

Метод скользящего среднего основан на сравнении математических ожиданий  $M_n$  и  $M_{n+N}$  в двух соседних окнах. Если модуль разности этих величин  $|M_n - M_{n+N}|$  выше заданного порога, то аналогично предыдущему алгоритму точка  $n + N$  является точкой движения, иначе она будет точкой покоя.

Порог будет определен через выборочную квантиль. Выборочной квантилью  $C_\alpha$  порядка  $\alpha$  выборки  $X = (X_1, \dots, X_n)$  называется порядковая статистика

$$C_\alpha = \begin{cases} X_{([n\alpha]+1)}, & n\alpha \in \mathbb{R} \setminus \mathbb{Z}; \\ X_{(n\alpha)}, & n\alpha \in \mathbb{Z}. \end{cases}$$

В данном методе порогом являлась выборочная квантиль порядка 0,9 величин  $|M_n - M_{n+N}|$  по всему сигналу. Результат работы алгоритма показан на рис. 9.

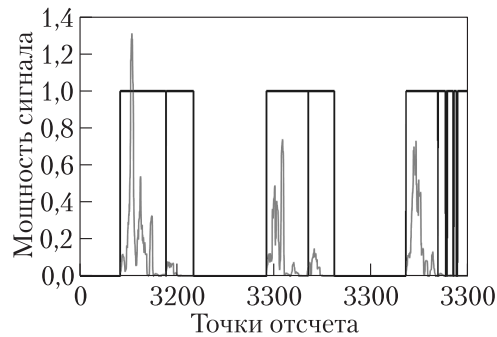


Рис. 9 Метод скользящего среднего

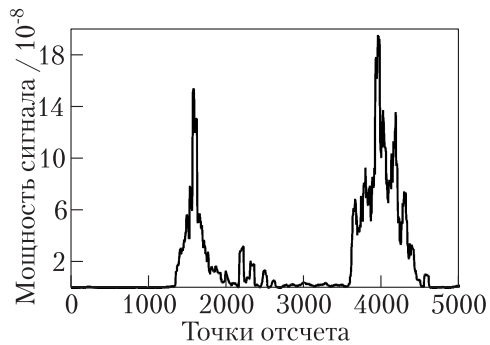


Рис. 10 Оконная дисперсия миограммы для левой руки

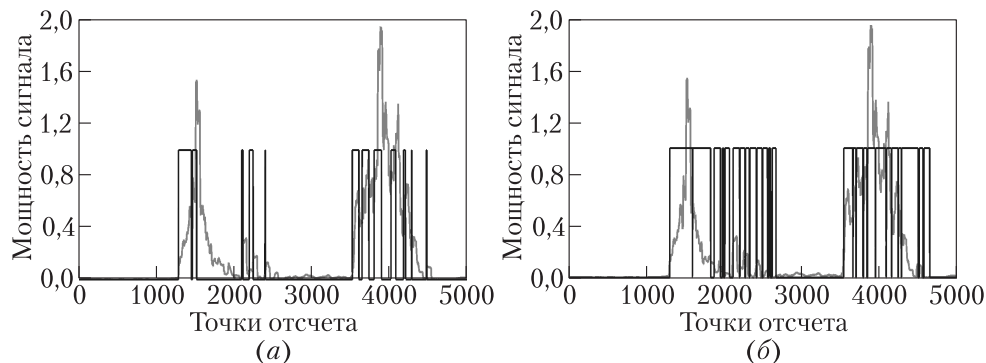


Преимущество данного метода в том, что он определяет не только начало, но и конец движения. Кроме того, этот метод дает значительный выигрыш в скорости. Время его работы на 5000 точках отсчета составляет 0,02 с.

#### 4 Результаты для левой руки

Интересные результаты были получены при применении тех же алгоритмов, но в эксперименте с левой рукой испытуемого (рис. 10). В силу физических особенностей сигнал миограммы получился более зашумленным.

Результаты работы методов представлены на рис. 11.



**Рис. 11** Метод КС (а) и метод скользящего окна (б)

За счет большого разброса значений сигнала на участках покоя в случае левой руки метод скользящего среднего в данном случае является менее точным.

#### 5 Заключение

Основной целью решения обратной задачи в магнитной энцефалографии является точное определение функциональных зон головного мозга. Данная работа посвящена точному определению местоположения моторных зон коры головного мозга человека.

Классический метод вызванных потенциалов предполагает усреднение мозговой активности (сигналов магнитоэнцефалограмм) относительно опорных точек по всему сигналу. Основная проблема метода усреднения МЭГ-сигналов относительно начала движения заключается в неопределенности точек синхронизации с миограммами.

В статье предлагаются усовершенствованные алгоритмы уже существующих методов, а также предложены новые подходы к определению опорных точек в сигнале, способствующие решению задачи по локализации областей мозга,

отвечающих за движения. Новые методы, сохраняя высокую точность ранее применяемых методов, работают на порядок быстрее, что имеет важное значение при их использовании в режиме реального времени.

Также в работе проведено сравнение миограмм для разных рук человека. Разницу в зашумленности миограмм для каждой руки также стоит учитывать при задании параметров алгоритмов обработки сигнала.

## Литература

1. *Baillet S., Mosher J. C., Leahy R. M.* Electromagnetic brain mapping // *IEEE Signal Proc. Mag.*, 2001. P. 14–30.
2. *Hämäläinen M., Hari R., Ilmoniemi R. J., Knuutila J., Lounasmaa O. V.* Magnetoencephalography — theory, instrumentation, and applications to noninvasive studies of the working human brain // *Rev. Mod. Phys.*, 1993. Vol. 65. P. 413–497.
3. *Friston K., Harrison L., Daunizeau J., Kiebel S., Phillips C., Trujillo-Barreto N., Henson R., Flandin G., Mattout J.* Multiple sparse priors for the M/EEG inverse problem // *NeuroImage*, 2008. Vol. 39. P. 1104–1120.
4. *Захарова Т. В., Гончаренко М. Б., Никифоров С. Ю.* Метод решения обратной задачи магнитоэнцефалографии, основанный на кластеризации поверхности мозга // *Статистические методы оценивания и проверки гипотез: Межвузовский сб. науч. тр.* — Пермь: ПГНИУ. 2013. Т. 25. С. 120–125.
5. *Бенинг В. Е., Драницына М. А., Захарова Т. В., Карпов П. И.* Решение обратной задачи в многодипольной модели источников магнитоэнцефалограмм методом независимых компонент // *Информатика и её применения*, 2014. Т. 8. Вып. 2. С. 79–87. doi: 10.14357/19922264140208.
6. *Захарова Т. В., Никифоров С. Ю., Гончаренко М. Б., Драницына М. А., Климов Г. А., Хазиахметов М. Ш., Чаянов Н. В.* Методы обработки сигналов для локализации невосполнимых областей головного мозга // *Системы и средства информатики*, 2012. Т. 22. Вып. 2. С. 157–176.
7. *Dranitsyna M., Zakharova T., Allakhverdiyeva V., Chshenyavskaya E.* Probability density function of myogram noise and its role in localization of brain activity // 32nd Seminar (International) on Stability Problems for Stochastic Models and 8th Workshop (International) “Applied Problems in Theory of Probabilities and Mathematical Statistics Related to Modeling of Information Systems.” Book of abstracts. — Moscow: IPI RAN, 2014. P. 25–26.
8. *Щенявская Е. В., Аллахвердиева В. М.* Модификация метода нахождения опорных точек и его применение для изучения свойств сигналов миограммы // *Задачи современной информатики: Тр. 1-й науч.-практич. конф. молодых ученых.* — М.: ИПИ РАН, 2014. С. 21–25.
9. *Хазиахметов М. Ш.* Свойства оконной дисперсии миограммы как случайного процесса // *Системы и средства информатики*, 2014. Т. 24. № 3. С. 110–120.

Поступила в редакцию 09.09.15

## SIGNAL MYOGRAM FEATURES RESEARCH

*T. V. Zakharova*<sup>1,2</sup>, *V. Yu. Korolev*<sup>1,2</sup>, and *A. A. Schemirova*<sup>1</sup>

<sup>1</sup>Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation

<sup>2</sup>Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** The paper investigates behavior of myograms, signals of muscle activity of a person. The paper proposes new ways of processing these signals to determine the reference points. The reference points are associated with the actual moment of the beginning of movement in the experiments. Further localized reference points are used to construct the associative filter processing ultraweak nonstationary signals of magnetoencephalogram. New methods retain high accuracy of previously used methods, but they are faster in several times.

**Keywords:** method of caused potentials; window sample variance; myogram; magnetoencephalogram; hypothesis testing

**DOI:** 10.14357/08696527150407

### Acknowledgments

The work was partly supported by the Russian Foundation for Basic Research (project 15-07-02652).

### References

1. Baillet, S., J. C. Mosher, and R. M. Leahy. 2001. Electromagnetic brain mapping. *IEEE Signal Proc. Mag.* 14–30.
2. Hämeäläinen, M., R. Hari, R. J. Ilmoniemi, J. Knuutila, and O. V. Lounasmaa. 1993. Magnetoencephalography — theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev. Mod. Phys.* 65:413–497.
3. Friston, K., L. Harrison, J. Daunizeau, S. Kiebel, C. Phillips, N. Trujillo-Barreto, R. Henson, G. Flandin, and J. Mattout. 2008. Multiple sparse priors for the M/EEG inverse problem. *NeuroImage* 39:1104–1120.
4. Zakharova, T. V., M. B. Goncharenko, and S. Yu. Nikiforov. 2013. Metod resheniya obratnoy zadachi magnitoentsefalografii, osnovanny na klasterizatsii poverkhnosti mozga [Method for solving the inverse problem of magnetoencephalography, based on brain surface clustering]. *Statisticheskie metody otsenivaniya i proverki gipotez: Mezhdvuzovskiy sb. nauch. tr.* [Statistical methods of estimation and hypothesis testing: Interuniversity collection of scientific papers]. Perm’: Perm’ State National Research University. 25:120–125.
5. Bening, V. E., M. A. Dranitsyna, T. V. Zakharova, and P. I. Karpov. 2014. Reshenie obratnoy zadachi v mnogodipol’noy modeli istochnikov magnitoentsefalogramm metodom

- nezavisimyykh komponent [Independent component analysis for the inverse problem in the multidipole model of magnetoencephalogram's sources]. *Informatika i ee Primeneniya — Inform. Appl.* 8(2):79–87. doi: 10.14357/19922264140208.
6. Zakharova, T. V., S. Yu. Nikiforov, M. B. Goncharenko, M. A. Dranitsyna, G. A. Klimov, M. Sh. Khaziakhmetov, and N. V. Chayanov. 2012. Metody obrabotki signalov dlya lokalizatsii nevospolnnykh oblastey golovnoy mozga [Signal processing methods for localization of nonrenewable brain regions]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 22(2):157–176.
  7. Dranitsyna, M., T. Zakharova, V. Allakhverdiyeva, and E. Chshenyavskaya. 2014. Probability density function of myogram noise and its role in localization of brain activity sources. *32nd Seminar (International) on Stability Problems for Stochastic Models (ISSPSM'2014) and 8th Workshop (International) "Applied Problems in Theory of Probabilities and Mathematical Statistics Related to Modeling of Information Systems" (APTP + MS'2014): Book of abstracts*. Moscow: IPI RAN. 25–26.
  8. Chshenyavskaya, E. V., and V. M. Allakhverdiyeva. 2014. Modifikatsiya metoda nakhozhdeniya opornykh tochek i ego primeneniya dlya izucheniya svoystv signalov miogrammy [Modification method of finding reference points and its application to study the properties of myogram signals]. *Zadachi sovremennoy informatiki: Tr. 1-y nauch.-praktich. konf. molodykh uchenykh* [Tasks of modern science: 1st Scientific Conference of Young Scientists Proceedings]. Moscow: IPI RAN. 21–25.
  9. Khaziakhmetov, M. Sh. 2014. Svoystva okonnoy dispersii miogrammy kak sluchaynogo protsesssa [Properties of window dispersion's increments of myogram as stochastic process]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 24(3):110–120.

Received September 9, 2015

## Contributors

**Zakharova Tatiana V.** (b. 1962) — Candidate of Science (PhD) in physics and mathematics; senior lecturer, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; lsa@cs.msu.ru

**Korolev Victor Yu.** (b. 1954) — Doctor of Science in physics and mathematics, professor, Head of the Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; leading scientist Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; vkorolev@cs.msu.ru

**Schemirova Anna A.** (b. 1995) — student, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; lsa@cs.msu.ru

## МЕТОД СИНХРОНИЗАЦИИ СИГНАЛОВ МАГНИТОЭНЦЕФАЛОГРАММ И МИОГРАММ\*

*Е. В. Щенявская<sup>1</sup>, Т. В. Захарова<sup>2</sup>*

**Аннотация:** Исследуется задача определения первично-моторной зоны головного мозга (зоны М1), которая относится к группе невосполнимых областей коры головного мозга человека. Для локализации зоны М1 был проведен эксперимент с использованием метода вызванных потенциалов, отвечающих движению указательного пальца руки. В ходе эксперимента одновременно регистрировались сигналы магнитоэнцефалограммы (МЭГ) и миограммы. Для обработки МЭГ был применен ассоциативный фильтр, использующий моменты начала движения (опорные точки) на сигнале миограммы. Представлен новый высокоточный метод определения опорных точек.

**Ключевые слова:** магнитоэнцефалография; миограмма; метод вызванных потенциалов; опорные точки

**DOI:** 10.14357/08696527150408

### 1 Введение

Наиболее неизученным и сложным органом в живой природе является головной мозг человека, который очень чувствителен к оперативному вмешательству и требует вследствие этого безопасных методов исследования.

Магнитоэнцефалография — бесконтактный метод исследования функций мозга с регистрацией сверхслабых магнитных полей, которые возникают в результате протекания в головном мозге электрических токов. Магнитные поля, генерируемые нейронами головного мозга, имеют порядок  $10^{-15}$  Тл, их записи называются МЭГ. Возможность регистрации МЭГ появилась в связи с применением СКВИД (SQUID — Superconducting QUantum Interference Device) — специальных магнитометров, работающих в жидком гелии и обладающих достаточной чувствительностью для регистрации столь малых полей мозга. Подробнее о методах по обработке МЭГ-сигналов можно прочитать в [1, 2].

---

\* Работа выполнена при частичной поддержке РФФИ (проект 15-07-02652).

<sup>1</sup> Московский государственный университет имени М. В. Ломоносова, факультет вычислительной математики и кибернетики, elena24051993@mail.ru

<sup>2</sup> Московский государственный университет имени М. В. Ломоносова, факультет вычислительной математики и кибернетики; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, lsa@cs.msu.ru

Активность мозга всегда представлена синхронной активностью большого количества нервных клеток, которые сопровождаются слабыми электрическими токами, а те, в свою очередь, порождают магнитные поля. Особенностью магнитного поля является то, что череп и мозговые оболочки практически не оказывают влияния на его величину, они «прозрачны» для магнитных силовых линий.

Неинвазивность данного исследования, которая заключается в отсутствии необходимости операционного вмешательства, является одним из главных преимуществ МЭГ.

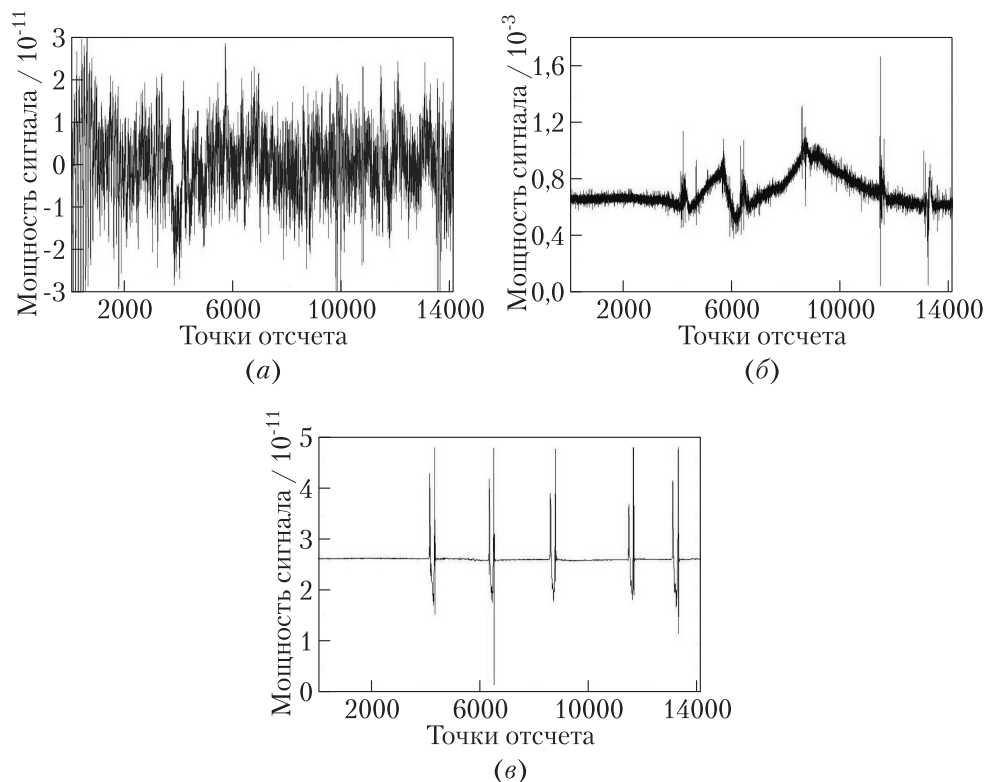
При анализе данных МЭГ важное значение имеет решение обратной задачи [3]. Обратная задача состоит в нахождении источника активности в коре головного мозга. Результаты данной работы могут применяться в исследованиях, использующих метод вызванных потенциалов при решении обратной задачи магнитоэнцефалографии.

## 2 Постановка задачи

Для точной локализации первичной моторной зоны (M1) общепринято применять методику вызванных потенциалов [4, 5]. Классический метод вызванных потенциалов подразумевает усреднение мозговой активности относительно точек привязки [5] по всему сигналу. В качестве «вызванных потенциалов» может выступать произвольная активность испытуемого (например, движение). Данная методика позволяет увеличить отношение сигнал/шум, что делает возможным выделить из сильно зашумленной активности мозга связанную с событием активность, несущую полезную информацию. Под шумом в данной задаче понимается суперпозиция физического шума (исходящего от оборудования) и физиологического (так как одновременно происходит множество мозговых процессов, не связанных с искомой областью коры мозга человека).

В рамках метода вызванных потенциалов был проведен следующий эксперимент. Испытуемый кладет ладонь на стол и многократно выполняет удары указательным пальцем по нему. При этом фиксируется его мозговая и мышечная активность. Дополнительно могут сниматься данные с акселерометра, закрепленного на пальце, и с кнопки, которую можно разместить на столе под пальцем испытуемого. Полученные сигналы проиллюстрированы на рис. 1.

Поставленная задача сводится к определению так называемых опорных точек. Под опорными точками в данном случае понимаются точки начала движения (ТНД). Их нахождение непосредственно по МЭГ-сигналу не представляется возможным ввиду малого отношения сигнал/шум (см. рис. 1, *a*), поэтому их определяют по миограмме. Казалось бы, на сигнале акселерометра (который не всегда сопровождается при съемке МЭГ-сигналов) значительно яснее просматриваются скачки активности, отвечающие внешнему проявлению движения, ввиду того что данный сигнал слабо зашумлен. Однако точки всплесков акселерометра не соответствуют именно началу движения в отличие от сигналов



**Рис. 1** Сигналы МЭГ (а), миограмма (б) и акселерометр (в)

миограммы. Свойства миограмм значительно лучше, так как миограмма является показателем именно мышечной активности, которая производится с небольшой задержкой относительно работы мозга по сравнению с внешним проявлением активности. С другой стороны, акселерометр удобно использовать в одном из методов определения интервалов покоя [6], что также играет значительную роль в данной задаче. В дальнейшем с помощью опорных точек устанавливалась синхронизация МЭГ-сигналов и сигналов миограмм.

Согласно методу вызванных потенциалов вокруг ТНД из МЭГ-сигнала вырезаются окрестности фиксированного радиуса, данные которых усредняются для получения искомого отклика. Затем определяется отклик с наибольшей амплитудой среди всех полученных откликов, соответствующих каждому сенсору. После определения максимального отклика возможно нахождение искомой области на коре головного мозга. Ввиду того что сигналы МЭГ и миограммы синхронизированы, было принято решение искать опорные точки по сигналу миограммы. Исходный алгоритм нахождения опорных точек использовал прос-



тое свойство миограммы — ее оконная дисперсия при осуществлении движения значительно больше по сравнению с состоянием, когда мышца находится в покое (ширина окна 30–50 мс). Подробное обоснование приведено в [2, 3].

При анализе данных МЭГ встает проблема решения так называемой обратной задачи, которая состоит в восстановлении распределения активности нейронных источников на поверхности коры головного мозга. Различные примеры решения обратной задачи рассматриваются в работах [3, 7].

### 3 Методы обработки магнитных энцефалограмм

Рассмотрим сигнал миограммы  $\xi(t), t \geq 0$ , как циклический случайный процесс

$$\xi(t) = \sum_i ((s_i(t) + \varepsilon_i(t)) \mathbb{I}\{t_i \leq t < t_{i+1}\}),$$

где  $s_i(t)$  — полезный сигнал, отвечающий движению;  $\varepsilon_i(t)$  — шум на интервале времени  $[t_i, t_{i+1})$ , который в нейрофизиологии именуется эпохой (отрезок времени, содержащий одно движение);  $i$  — номер эпохи.

Сигнал миограммы является циклическим, но не стационарным [2]. Применение Фурье-преобразования как стандартного, так и оконного, не позволило определить с нужной точностью даже разбивку сигнала на эпохи.

В конце 1970-х гг. появился новый метод вейвлет-обработки сигналов, который решал проблему анализа сигналов, характеризующихся высокочастотной компонентой в течение короткого промежутка времени и низкочастотными колебаниями при рассмотрении больших временных интервалов. Оконные преобразования позволяли проанализировать либо высокие частоты в коротком окне времени, либо низкочастотную компоненту, но не оба колебания одновременно. В результате был предложен подход, в котором для разных диапазонов использовались временные окна разной длительности. Оконные функции получались в результате растяжения-сжатия и смещения по времени гауссиана. В дальнейшем эти базисные функции были названы вейвлетами [8].

Казалось бы, вейвлет-преобразование по своему смыслу отлично подходит для решения поставленной задачи. Однако при помощи вейвлет-анализа хорошо восстанавливаются сигналы, у которых шум независим и нормально распределен [8].

В данном исследовании при проверке гипотезы о нормальном распределении шума миограммы гипотеза была отклонена. Проверка производилась по критерию согласия хи-квадрат К. Пирсона. Уровень значимости был выбран стандартный  $\alpha = 0,05$ , объем выборки 76 000. Значение критерия составило 35 855 при критической границе 14,1, что позволило сделать вывод о том, что гипотеза о нормальном распределении не согласуется с выбранным наблюдением [2]. Как следствие, вейвлет-обработка сигнала миограммы давала неустойчивые результа-



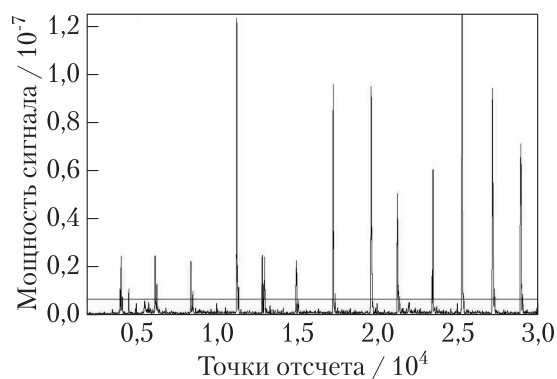
ты, иногда совершенно неверные, поэтому стал разрабатываться альтернативный подход для обработки таких сложных биосигналов.

### 3.1 Метод, основанный на порядковых статистиках

Оконная дисперсия служит хорошим инструментом для обработки слабых сигналов. В рассматриваемом случае она несет смысл «энергии», ведь именно характерные скачки оконной дисперсии миограммы свидетельствуют о наличии движения, на которое выделяется энергия. Предложенный в [2] метод нахождения опорных точек использовал простое свойство миограммы: ее оконная дисперсия при осуществлении движения значительно больше по сравнению с оконной дисперсией на интервале покоя (ширина окна 30–50 мс) [9, 10].

Основные этапы работы алгоритма:

1. Строится эмпирическая функция распределения оконной дисперсии миограммы (ОДМ) и по ней определяется квантиль высокого порядка.
2. Отмечаются участки пересечения квантили отрезками возрастания ОДМ.
3. К полученному массиву точек применяется фильтрация. Оставшиеся точки принимаются соответствующими началам движения (назовем их базовыми точками).
4. Оконная дисперсия миограммы при помощи базовых точек делится на интервалы с минимальной длиной не меньше заданной, на каждом из которых строится эмпирическая функция распределения, находится квантиль высокого порядка.
5. Для каждого интервала отдельно применяется шаг 2 с соответствующей квантилью. Для полученного общего массива точек применяется фильтрация. Точки полученного массива назовем предварительными точками.



**Рис. 2** Оконная дисперсия миограммы

6. Для каждой предварительной точки находится «интервал покоя» — отрезок между предшествующим и текущим движением, содержащий исключительно шумовую мышечную активность. Для него строится эмпирическая функция распределения и вычисляется квантиль высокого порядка. За точку начала движения берется последний момент времени перед предварительной точкой, когда значение ОДМ меньше этой квантили.

Предложенный метод позволил значительно повысить качество определения опорных точек, а именно существенно возросло количество определяемых опорных точек — более чем в 1,5 раза. На рис. 2 можно увидеть ярко выраженные скачки, соответствующие движению, а пересечение их с квантилью высокого порядка говорит о редкости данного события.

### 3.2 Метод двухоконной прокрутки

В дальнейшем потребуется разбивать сигнал на интервалы покоя и движения. Рассмотрим наиболее точный метод определения интервалов покоя — метод двухоконной прокрутки. Оставшиеся участки сигналы будем считать подозрительными на активность.

Дадим краткое описание этого метода.

**Определение.** Выборочной  $\alpha$ -квантилью  $C_\alpha$  выборки  $X = (X_1, \dots, X_n)$  называется порядковая статистика

$$C_\alpha = \begin{cases} X_{[n\alpha]+1}, & n\alpha \in \mathbb{R} \setminus \mathbb{Z}; \\ \frac{1}{2}(X_{n\alpha} + X_{n\alpha+1}), & n\alpha \in \mathbb{Z}. \end{cases}$$

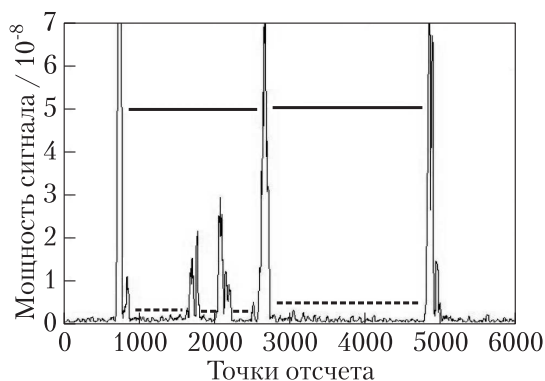
Пусть  $i$  — некоторая точка отсчета. Выберем окно шириной  $2N$ , левая граница которого находится в точке  $i$ . По значениям сигнала  $(X_i, \dots, X_{i+N-1})$  и  $(X_{i+N}, \dots, X_{i+2N-1})$  рассчитаем выборочные математические ожидания  $M_i$  и  $M_{i+N}$  соответственно. В данном случае можно говорить и о двух соседних окнах шириной  $N$ , по каждому из которых рассчитывается выборочное среднее. Далее считается разность  $M_i - M_{i+N}$  и берется ее абсолютное значение.

Пройдя двумя окнами по всему сигналу слева направо и проделав указанную выше операцию для каждой точки отсчета  $i$ , получим набор модулей разности выборочных средних, рассчитанных на соседних окнах. По полученному набору модулей разности определим порог, равный выборочной квантили  $C_\alpha$  высокого порядка  $\alpha$ . В данном случае было выбрано значение  $\alpha = 0,9$ , а также эмпирически подобрано значение параметра  $N$ , равное 100 точкам отсчета.

Далее для каждого значения  $i$  будет проверяться следующее условие:

$$|M_i - M_{i+N}| \geq C_\alpha.$$

Если разность превышает порог, то на правом окне фиксируется изменение состояния сигнала, т. е. можно говорить о том, что на двух окнах сигнал имеет



**Рис. 3** Работа алгоритма двухоконной прокрутки

различный характер поведения. Рассматриваемый сигнал может соответствовать только двум состояниям: покою и движению.

Далее всю вышеизложенную процедуру повторяем при обратной прокрутке двумя окнами, т. е. при движении по сигналу справа налево. В этом случае также появятся соседние интервалы (на каждом из двух окон) с различными состояниями. Теперь фиксируется изменение состояния сигнала на левом окне.

Те интервалы, которые оба раза отмечены с измененным состоянием как слева направо, так и справа налево, определяются как интервалы движения. Оставшиеся интервалы определяются как интервалы покоя.

Работа метода двухоконной прокрутки проиллюстрирована на рис. 3.

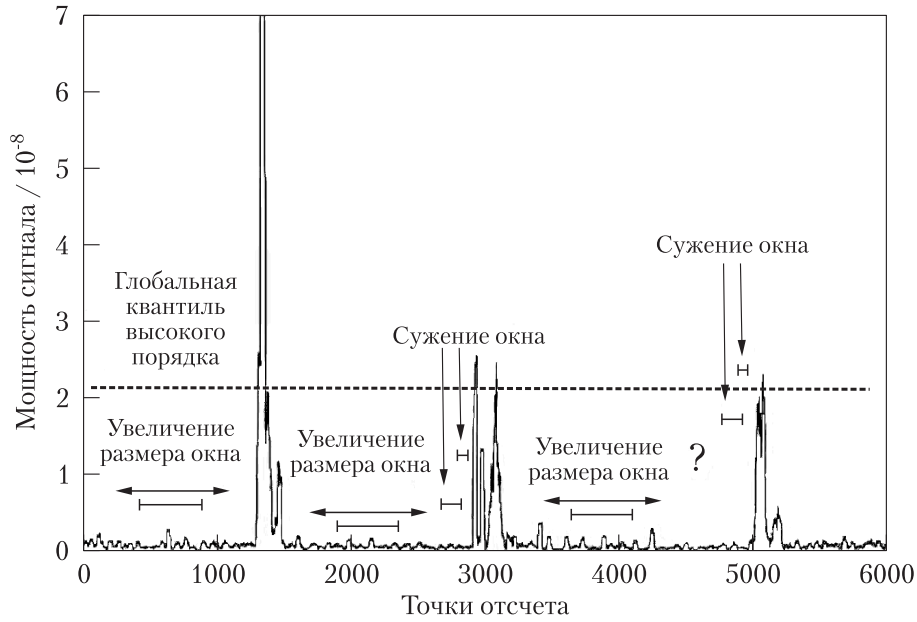
Стоит, однако, отметить, что регистрируемая активность необязательно отвечает полезному сигналу. Она может быть вызвана некоторым человеческим фактором, нарушающим чистоту эксперимента: испытуемый дернул или напряг сильно руку, повернулся и т. п.

В отличие от алгоритма на основе глобальной квантили, данный алгоритм является более чувствительным к изменениям в сигнале. На рис. 3 проиллюстрировано это свойство алгоритма, участки покоя здесь выделены горизонтальной линией, разными кривыми для каждого случая.

Таким образом, новым методом можно находить однородные участки сигнала и более точно определять участки покоя сигнала.

### 3.3 Модифицированный метод

Напомним, что поставленная задача локализации моторной зоны сводится к точному определению опорных точек. Именно для повышения этой точности было предложено объединить метод, основанный на порядковых статистиках, с идеей вейвлет-анализа, а именно: ранее применявшееся окно с постоянной



**Рис. 4** Определение размера окна

шириной заменить на динамичное окно, ширина которого будет меняться в зависимости от характера изменения сигнала [11].

Предлагаемый в этом разделе алгоритм преследует две цели: подавление излишнего шума на интервалах покоя с помощью большого окна оконной дисперсии и сохранение большей доли полезной информации на интервалах активности с использованием меньшего окна. Последнее позволит точнее зафиксировать момент начала движения для дальнейшего решения обратной задачи (рис. 4). Отделить участки покоя от участков с возможным движением можно с помощью метода двухоконной прокрутки, о котором говорилось ранее.

Основная проблема заключается в выборе ширины окна в каждой точке активности исходного сигнала в зависимости от изменчивости исходного сигнала миограммы в окрестности данной точки. Для решения этой задачи применялся следующий подход: в пределах каждой эпохи в моменты времени, подозрительные на движение, определяли минимальное и максимальное значения разностной аппроксимации производной сигнала миограммы с некоторым шагом. Затем интервал значений полученных производных сопоставлялся с интервалом возможных значений окон в диапазоне от фиксированного минимального до фиксированного максимального (минимальному значению производной сопоставлялся наибольший размер окна и наоборот):

$$\Delta(t) = (W_{\min} + k(W_{\max} - W_{\min})) \mathbb{I}\{t \in \text{Move}\};$$

$$k = \frac{f(t) - \min_{t \in [t_i, t_{i+1}]} f}{\max_{t \in [t_i, t_{i+1}]} f - \min_{t \in [t_i, t_{i+1}]} f},$$

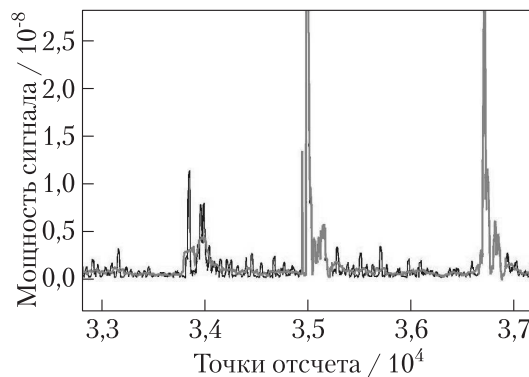
где  $\Delta(t)$  — значение ширины окна в текущий момент времени  $t$ ;  $W_{\min}$  и  $W_{\max}$  — фиксированные значения минимального и максимального окна дисперсии соответственно;  $\mathbb{I}\{t \in \text{Move}\}$  — индикатор движения в текущий момент времени  $t$ ;  $\max_{t \in [t_i, t_{i+1}]} f$  и  $\min_{t \in [t_i, t_{i+1}]} f$  — максимальное и минимальное значение всех разностных производных в точках, полученных разбиением сеткой с фиксированным шагом относительно каждой эпохи  $[t_i, t_{i+1}]$ ;  $f(t)$  — текущее значение данной разностной аппроксимации миограммы в момент времени  $t$ , вычисляемой по формуле:

$$f(t) = \frac{\xi(t+h) - \xi(t)}{h}. \quad (1)$$

Эмпирически было подобрано значение шага сетки, равное 5 точкам отсчета. С помощью сетки стало возможным повысить тенденцию роста энергии, несмотря на зашумление сигнала.

Далее к таким образом обработанному сигналу применяется алгоритм нахождения опорных точек, описанный в подразд. 3.1.

При обработке сигнала миограммы оконной дисперсией с динамическим окном моменты начала движения всегда определяются раньше, чем при стандартной, что иллюстрирует рис. 5. Новый метод является более чувствительным к изменениям именно полезного сигнала и тем самым повышает точность определения опорных точек. Более того, можно заметить, что различного рода выбросы (обусловленные человеческим фактором и нередко встречающиеся на реальных сигналах электромиограммы) новый метод стремится сгладить, что также является дополнительным преимуществом.



**Рис. 5** Сравнение сигнала с динамическим окном (серый) с сигналом, обработанным постоянным окном (черный)

## 4 Результаты и выводы

Был проведен анализ использовавшихся ранее методов обработки слабых нестационарных сигналов и предложен новый метод, использующий преимущества предшествующих методов, но работающий без жестких ограничений на характер шума сигнала. Описанный в данной работе модифицированный метод позволил повысить точность определения точек привязки на МЭГ-сигнале, а именно: новый метод позволил подавить излишний шум на интервалах покоя с помощью большего окна оконной дисперсии и сохранить больший процент полезного сигнала на интервалах активности с использованием меньшего окна. Последнее позволило точнее зафиксировать момент начала движения. В совокупности эти два качества способствовали повышению качества данных как при применении алгоритма нахождения опорных точек (описанного в подразд. 3.1), так и для дальнейшего решения обратной задачи.

Важной особенностью также является то, что модифицированный метод использует неотфильтрованный сигнал миограммы. Это дает возможность отразить больше полезной информации, несмотря на высокий шум, присутствующий в сигнале.

Результаты работы применялись для решения обратной задачи по локализации невосполнимых областей головного мозга и были проверены на экспериментальных данных, полученных Московским центром нейрокognитивных исследований. Решение обратной задачи имеет важное практическое значение, поскольку первичная моторная кора относится к зонам, не способным к самовосстановлению, и ее точная локализация критически важна для пациентов.

Результаты работы могут быть применены для предоперационного картирования мозговых функций, а также в определении патологической активности в мозге.

## Литература

1. *Hamalainen M., Hari R., Ilmoniemi R. J., Knuutila J., Lounasmaa O. V.* Magnetoencephalography — theory, instrumentation, and applications to noninvasive studies of the working human brain // *Rev. Mod. Phys.*, 1993. Vol. 65. P. 413–497.
2. *Захарова Т. В., Никифоров С. Ю., Гончаренко М. Б., Драницына М. А., Климов Г. А., Хазиахметов М. Ш., Чаянов Н. В.* Методы обработки сигналов для локализации невосполнимых областей головного мозга // *Системы и средства информатики*, 2012. Т. 22. № 2. С. 157–175.
3. *Бенинг В. Е., Драницына М. А., Захарова Т. В., Карнов П. И.* Решение обратной задачи в многодипольной модели источников магнитоэнцефалограмм методом независимых компонент // *Информатика и её применения*, 2014. Т. 8. Вып. 2. С. 79–87. doi: 10.14357/19922264140208.
4. *McGillem C. D., Aunon A. J.* Analysis of event-related potentials methods of analysis of brain electrical and magnetic signals: EEG handbook. — Amsterdam: Elsevier Science Publishers, 1987. 436 p.

5. *Fabiani M., Gratton G., Federmeier K. D.* Event-related brain potentials: Methods, theory, and applications // Handbook of Psychophysiology / Eds. J. T. Cacioppo, L. G. Tassinary, G. G. Berntson. — Cambridge: Cambridge University Press, 2007. P. 85–119.
6. *Dranitsyna M., Zakharova T., Allakhverdiyeva V., Chshenyavskaya E.* Probability density function of myogram noise and its role in localization of brain activity // 32nd Seminar (International) on Stability Problems for Stochastic Models and 8th Workshop (International) “Applied Problems in Theory of Probabilities and Mathematical Statistics Related to Modeling of Information Systems.” Book of abstracts. — Moscow: IPI RAN, 2014. P. 25–26.
7. *Захарова Т. В., Гончаренко М. Б., Никифоров С. Ю.* Метод решения обратной задачи магнитоэнцефалографии, основанный на кластеризации поверхности мозга // Статистические методы оценивания и проверки гипотез: Межвузовский сб. науч. тр. — Пермь: ПГНИУ, 2013. Т. 25. С. 120–125.
8. *Захарова Т. В., Шестаков О. В.* Вейвлет-анализ и его приложения. — 2-е изд. — М.: ИНФРА-М, 2012. 158 с.
9. *Хазиахметов М. Ш., Захарова Т. В.* Об алгоритмах нахождения опорных точек миограммы для использования в локализации невосполнимых областей головного мозга // Статистические методы оценивания и проверки гипотез: Межвузовский сб. науч. тр. — Пермь: ПГНИУ, 2013. Т. 25. С. 56–63.
10. *Захарова Т. В., Хазиахметов М. Ш.* Свойства оконной дисперсии миограммы как случайного процесса // Системы и средства информатики, 2014. Т. 24. № 3. С. 110–120.
11. *Щенявская Е. В., Алахвердиева В. М.* Модификация метода нахождения опорных точек и его применение для изучения свойств сигналов миограммы // Задачи современной информатики: Тр. 1-й науч.-практич. конф. молодых ученых. — М.: ИПИ РАН, 2014. С. 21–25.

*Поступила в редакцию 03.10.15*

---

---

## METHOD OF SYNCHRONIZATION OF MAGNETOENCEPHALOGRAMS AND MYOGRAMS SIGNALS

*E. V. Chshenyavskaya<sup>1</sup> and T. V. Zakharova<sup>1,2</sup>*

<sup>1</sup>Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation

<sup>2</sup>Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** The paper examines the problem of the primary motor cortex (or M1) determination, which is one of the principal brain areas involved in the motor function. The experiment, based on the method of caused potentials, was conducted to locate M1 area: caused potentials were obtained from the

nervous system in response to tactile stimulation of a forefinger. During the experiment, both myogram and magnetoencephalogram (MEG) signals were recorded simultaneously. The associative filter, utilizing moments of the beginning of the movement (reference points) of the myogram signal, was applied for processing the MEG signal. This article introduces a new high-precision method for reference points detection.

**Keywords:** magnetoencephalogram; myogram; method of caused potentials; reference points

**DOI:** 10.14357/08696527150408

## Acknowledgments

The work was partly supported by the Russian Foundation for Basic Research (project 15-07-02652).

## References

1. Hamalainen, M., R. Hari, R. J. Ilmoniemi, J. Knuutila, and O. V. Lounasmaa. 1993. Magnetoencephalography — theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev. Mod. Phys.* 65:413–497.
2. Zakharova, T. V., S. Yu. Nikiforov, M. B. Goncharenko, M. A. Dranitsyna, G. A. Klimov, M. Sh. Khaziakhmetov, and N. V. Chayanov. 2012. Metody obrabotki signalov dlya lokalizatsii nevospolnimykh oblastey golovnogo mozga [Signal processing methods for localization of nonrenewable brain regions]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 22(2):157–176.
3. Bening, V. E., M. A. Dranitsyna, T. V. Zakharova, and P. I. Karpov. 2014. Reshenie obratnoy zadachi v mnogodipol'noy modeli istochnikov magnitoentsefalogramm metodom nezavisimykh komponent [Independent component analysis for the inverse problem in the multidipole model of magnetoencephalogram's sources]. *Informatika i ee Primeneniya — Inform. Appl.* 8(2):79–87. doi: 10.14357/19922264140208.
4. McGillem, C. D., and A. J. Aunon. 1987. *Analysis of event-related potentials methods of analysis of brain electrical and magnetic signals: EEG handbook*. Elsevier Science Publishers. 436 p.
5. Fabiani, M., G. Gratton, and K. D. Federmeier. 2007. Event-related brain potentials: Methods, theory, and applications. *Handbook of psychophysiology*. Eds. J. T. Cacioppo, L. G. Tassinary, and G. G. Berntson. Cambridge: Cambridge University Press, 2007. 85–119.
6. Dranitsyna, M., T. Zakharova, V. Allakhverdiyeva, and E. Chshenyavskaya. 2014. Probability density function of myogram noise and its role in localization of brain activity. *32nd Seminar (International) on Stability Problems for Stochastic Models and 8th Workshop (International) "Applied Problems in Theory of Probabilities and Mathematical Statistics Related to Modeling of Information Systems:" Book of abstracts*. Moscow: IPI RAN, 2014. 25–26.
7. Zakharova, T. V., M. B. Goncharenko, and S. Yu. Nikiforov. 2013. Metod resheniya obratnoy zadachi magnitoentsefalografii, osnovanny na klasterizatsii poverkhnosti



- mozga [Method for solving the inverse problem of magnetoencephalography, based on brain surface clustering]. *Statisticheskie metody otsenivaniya i proverki gipotez: Mezhvuzovskiy sb. nauch. tr.* [Statistical methods of estimation and hypothesis testing: Interuniversity collection of scientific papers]. Perm: Perm State National Research University. 25:120–125.
8. Zakharova, T. V., and O. V. Shestakov. 2012. *Veyvlet-analiz i ego prilozheniya* [Wavelet analysis and applications]. 2nd ed. Moscow: INFRA-M. 158 p.
  9. Khaziakhmetov, M. Sh., and T. V. Zakharova. 2013. Ob algoritmakh nakhozhdeniya opornykh tochek miogrammy dlya ispol'zovaniya v lokalizatsii nevospolnimykh oblastey golovnoy mozga [Algorithms of finding reference's points of a myogram for localization of nonrenewable brain regions]. *Statisticheskie metody otsenivaniya i proverki gipotez: Mezhvuzovskiy sb. nauch. tr.* [Statistical methods of estimation and hypothesis testing: Interuniversity collection of scientific papers]. Perm: Perm State National Research University. 25:56–63.
  10. Khaziakhmetov, M. Sh. 2014. Svoystva okonnoy dispersii miogrammy kak sluchayno-go protsessa [Properties of window dispersion's increments of myogram as stochastic process]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 24(3):110–120.
  11. Chshenyavskaya, E. V., and V. M. Allakhverdiyeva. 2014. Modifikatsiya metoda nakhozhdeniya opornykh tochek i ego primeneniye dlya izucheniya svoystv signalov miogrammy [Modification method of finding reference points and its application to study the properties of signals myogram]. *Zadachi sovremennoy informatiki: Tr. 1-y nauch.-praktich. konf. molodykh uchenykh* [Tasks of Modern Science: 1st Scientific Conference of Young Scientists Proceedings]. Moscow: IPI RAN. 21–25.

*Received October 3, 2015*

## Contributors

**Shchenyavskaya Yelena V.** (b. 1993) — Master Degree student, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; elena24051993@mail.ru

**Zakharova Tatiana V.** (b. 1962) — Candidate of Science (PhD) in physics and mathematics; senior lecturer, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University, 1-52 Leninskiye Gory, GSP-1, Moscow 119991, Russian Federation; senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; lsa@cs.msu.ru

## НЕКОТОРЫЕ ПОДХОДЫ К ФОРМИРОВАНИЮ ОБОБЩЕННОЙ АРХИТЕКТУРЫ ИНФОРМАЦИОННЫХ СИСТЕМ ОРГАНИЗАЦИЙ — УЧАСТНИКОВ ЕДИНОГО ИНФОРМАЦИОННОГО ПРОСТРАНСТВА РОССИИ

*А. А. Зацаринный<sup>1</sup>, Э. В. Киселев<sup>2</sup>*

**Аннотация:** Статья продолжает серию публикаций, посвященных формированию нормативно-технической базы для создания единого информационного пространства Российской Федерации (ЕИП РФ). В основу статьи положены требования к архитектуре единых информационных систем (ЕИС). В статье излагаются общие положения системного подхода к архитектуре информационных систем организаций — участников ЕИП РФ (УЕИП), предложения по обобщенной архитектуре ЕИС УЕИП и ее возможной типизации. Показано, что обобщенная архитектура ЕИС должна ориентироваться на три взаимосвязанных сегмента — открытый, конфиденциальный и закрытый, каждый из которых рассматривается как самостоятельная автоматизированная система. Определены в общем виде группы пользователей УЕИП, а также обслуживающего персонала, обеспечивающего функционирование единой системы. Определяется предметное понятие архитектуры ЕИС с базированием на трех ключевых положениях: фундаментальный подход к системе, рациональное структурирование системы, учет факторов внешней среды. В рамках рационального структурирования дается общий подход к определению номенклатуры составных частей единой системы — сегментов и компонентов (базовых систем, подсистем, сетей, комплексов и др.) и их взаимодействию между собой на принципах консолидации, интеграции и централизации. Предлагается подход, предусматривающий разделение ЕИС УЕИП на две группы базовых технологических систем: функциональных и обеспечивающих. Положения данной статьи могут использоваться в качестве научно-методической основы при формировании нормативно-технической базы построения информационных систем УЕИП.

**Ключевые слова:** единое информационное пространство Российской Федерации; участник единого информационного пространства; системный подход; архитектура единой информационной системы УЕИП; защищенные информационные ресурсы; централизация; интеграция; информационная безопасность; телекоммуникационное обеспечение; группа базовых функциональных технологических систем; группа базовых обеспечивающих технологических систем

**DOI:** 10.14357/08696527150409

<sup>1</sup>Федеральный исследовательский центр «Информатика и управление» Российской академии наук, azatsarinny@ipiran.ru

<sup>2</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, ekiselev@ipiran.ru

## 1 Основные понятия и положения

1. В настоящей статье рассматриваются вопросы, положения и предложения по обобщенной архитектуре ЕИС УЕИП.

Переходя к детализации построения обобщенной архитектуры ЕИС УЕИП, определим понятия *многочисленных взаимодействий информационных инфраструктур и их пользователей в ЕИП РФ*.

Информационные инфраструктуры представлены в виде совокупности ЕИС и единых информационно-телекоммуникационных (ЕИТКС) систем многочисленных УЕИП, имеющих свои информационные ресурсы (ИР). Пользователями ИР в ЕИП РФ [1, 2] являются: функциональные системы своей организации и других УЕИП; свои ЕИС и ЕИТКС; иные потребители — отдельные люди и группы людей.

Выделим внутренние и внешние взаимодействия.

Под *внутренними* будем понимать взаимодействия, относящиеся исключительно к компетенции сферы деятельности внутри каждой УЕИП и состоящие в предоставлении необходимых ИР и услуг внутренним пользователям и потребителям из состава конкретного участника. В статье внутренние взаимодействия не детализируются, так как они определяются спецификой деятельности каждого участника.

*Внешние взаимодействия* относятся к компетенции сферы деятельности данного участника с другими УЕИП и внешними физическими лицами и состоят в предоставлении своих ИР и услуг другим участникам ЕИП РФ, внешним пользователям и потребителям, а также в получении от них необходимых ИР. Внешние взаимодействия целесообразно рассматривать в виде иерархии четырех слоев: пользовательского (потребительского), функционального, информационного и коммуникационного. Внешние потребители, используя коммуникационный слой (рассматривается ниже), по соответствующим шинам получают необходимые им ИР через ЕИС (ЕИТКС) данной УЕИП и соответственно передают свои ИР данному участнику. Приведенные четыре слоя внешнего взаимодействия должны учитываться в архитектуре ЕИС (ЕИТКС) конкретной УЕИП.

Архитектура ЕИС (ЕИТКС) для УЕИП должна быть нацелена на обеспечение информационной, аналитической, документационной, инструментальной и технологической поддержки принятия решений и выполнения основных функций в сфере деятельности данного участника. При этом в обобщенной архитектуре ЕИС (ЕИТКС) любой УЕИП обязательно учитываются как внутренние, так и внешние взаимодействия.

2. Обобщенная архитектура ЕИС (ЕИТКС) конкретной УЕИП РФ, как показано в [2, 3], может включать в общем случае три сегмента: О-сегмент, К-сегмент и З-сегмент. Каждый из сегментов фактически является самостоятельной автоматизированной информационной системой (АИС). Любой из трех сегментов ЕИС (ЕИТКС) конкретной УЕИП осуществляет информационные

взаимодействия с пользователями различного ранга — с учетом их прав на доступ — по согласованным регламентам.

Каждый сегмент выполняет свойственные ему функции, в том числе:

- обеспечение надежного, достоверного, юридически значимого информационного взаимодействия в электронной и (или) бумажной форме, включая идентификацию участников взаимодействия в рамках ЕИП РФ;
- обнаружение и регистрацию участников взаимодействия, информационных объектов, данных (с помощью электронных каталогов, реестров, поисковых систем);
- интеграционные функции и функции преобразования (шлюзования) данных;
- вспомогательные функции, необходимые для обеспечения современного технологического уровня, удобства и комфорта использования сегментов ЕИС (ЕИТКС) конкретной УЕИП.

Выполнение любой функции должно быть единообразным при решении однотипных задач и определяться соответствующим нормативным документом.

3. Информация соответствующей сферы деятельности необходима множеству людей для различных целей. Следует выделить основные группы субъектов — физических лиц, получающих и использующих информацию конкретной сферы деятельности, ИР и сервисы ЕИС (ЕИТКС) конкретной УЕИП в качестве пользователей, а также персонала, обеспечивающего функционирование ЕИС (ЕИТКС) участника. К этим группам в общем случае относятся:

- (1) высшее руководство страны федерального и регионального уровня, которому необходимо получать соответствующую информацию для управления страной во все периоды времени (в мирное время, при чрезвычайных ситуациях, в особый период);
- (2) руководители, сотрудники и исполнители организации, выполняющие непосредственно свои внутренние работы и действия в своей сфере деятельности;
- (3) руководители, сотрудники и исполнители федерального и регионального уровня, заинтересованные в получении информации через ЕИС (ЕИТКС) конкретной УЕИП и ее использовании;
- (4) пользователи — государственные служащие федерального, регионального и муниципального уровня, использующие информацию сферы деятельности через ЕИС (ЕИТКС) конкретной УЕИП (в пределах установленных административных регламентов);
- (5) физические лица, так или иначе взаимодействующие с конкретной УЕИП и (или) получающие услуги в сфере деятельности данной организации. В их число входят граждане России и зарубежных стран, лица без гражданства и другие категории физических лиц;

- (6) персонал ЕИС (ЕИТКС) конкретной УЕИП — сегментов и компонентов данной системы. К этой группе относится инженерный, оперативно-диспетчерский и административный персонал, выполняющий все виды работ по эксплуатации ЕИС (ЕИТКС) и обеспечению заданных тактико-технических характеристик этой системы. В том числе эта группа совместно с разработчиками участвует в развитии, модернизации, сопровождении и эксплуатации ЕИС (ЕИТКС) своей УЕИП — ее сегментов и компонентов.

Для каждой из перечисленных групп существуют различные права доступа и ожидания от использования сервисов ЕИС (ЕИТКС) конкретной УЕИП, а также специфичные аспекты доступа в ЕИС (ЕИТКС), определяемые их ролевыми функциями.

4. К ЕИС (ЕИТКС) конкретной УЕИП целесообразно принять предметное понятие архитектуры, включающее следующие *три ключевых положения*: *фундаментальный подход к ЕИС (ЕИТКС), рациональное структурирование ЕИС (ЕИТКС), учет фактора внешней среды*.

*Фундаментальный подход* к ЕИС (ЕИТКС) конкретной УЕИП при определении архитектуры данной системы предусматривает ее проектирование и создание как интегрированной системы [1, 2]. Действительно, ЕИС (ЕИТКС) является интегратором. В соответствии с ее предназначением она может иметь в своем составе три взаимосвязанных сегмента: О-АИС, К-АИС и З-АИС. Также ЕИС (ЕИТКС) посредством своих сегментов консолидирует и обслуживает все функциональные системы конкретной УЕИП, в том числе для взаимодействия с различными информационными системами (АИС) органов государственной власти федерального, регионального и муниципального уровня.

Фундаментальный подход ориентирован на отказ от «лоскутной» информатизации организации [1] именно за счет создания единой системы — ЕИС или ЕИТКС. Такой подход учитывает следующий системный фактор: ЕИС (ЕИТКС) конкретной УЕИП постоянно, непрерывно обеспечивает поддержание важнейшего интегрального свойства — информационной устойчивости деятельности УЕИП, которая напрямую зависит от функционирования данной сложной системы. Прекращение функционирования по какой-либо причине ЕИС (ЕИТКС), или какого-либо сегмента, или отдельных функционально важных компонентов из состава сегментов может привести к следующему: во-первых, соответствующая информация и услуги станут недоступными множеству людей из различных групп (см. выше п. 3); во-вторых, реализация совокупности функций соответствующей сферы деятельности в электронной форме станет затрудненной или невозможной со всеми вытекающими отсюда комплексами негативных последствий.

*Рациональное структурирование* ЕИС (ЕИТКС) в конкретной УЕИП предусматривает и конкретное определение номенклатуры составных частей — сегментов и компонентов (систем, подсистем, сетей, комплексов и др.) такой системы.

Осуществляется также взаимодействие составных частей между собой на принципах консолидации и интеграции. Рациональное структурирование необходимо также ввиду сложности системы, в том числе возможности функционирования ЕИС (ЕИТКС) как объекта информатизации, территориально распределенного по России.

*Учет фактора внешней среды* предусматривает информационное взаимодействие ЕИС (ЕИТКС) конкретной УЕИП и ее определенных составных частей с информационными системами руководства страны федерального, регионального и муниципального уровня в целях эффективного государственного управления и исполнения своих функций в электронной форме.

5. Единая информационная система (ЕИТКС) конкретной УЕИП — это сложный объект информатизации, в котором должны быть постоянно и надежно защищены все ИР и информация в общем случае трех контуров (О-контур, К-контур и З-контур), а также обеспечен надлежащий допуск к ним. Соответственно, должно осуществляться комплексное обеспечение информационной безопасности в О-сегменте, К-сегменте и З-сегменте, каждый из которых фактически является автоматизированной системой. Это означает, что каждый сегмент должен соответствовать требованиям нормативных документов, определяющих порядок создания автоматизированных систем в защищенном исполнении и требования к ним [4, 5].

6. Построение ЕИС (ЕИТКС), как и любой другой аналогичной системы, предполагает формирование и функционирование двух элементов. Первый — информация как совокупность данных, необходимых для получения требуемого результата. Второй — процессы (веб-сервисы, технологические процессы), обеспечивающие обработку информации для получения требуемого результата в сфере своей деятельности. Эти два элемента тесно связаны, и изменения в первом, информационном, невозможны без второго — процессов работы с информацией, а работа второго невозможна без учета особенностей первого.

В соответствии с назначением ЕИС (ЕИТКС) конкретной УЕИП и в контексте работ по ее созданию требуется решение задачи формирования и функционирования второго, весьма важного элемента — организации процессов информационного взаимодействия в электронной форме, обеспечения эффективного информационного обмена. Действительно, информация сферы деятельности ЕИС (ЕИТКС) конкретной УЕИП необходима при реализации внутреннего взаимодействия в своей организации и внешнего взаимодействия в рамках ЕИП РФ, информационного обмена с множеством людей для различных целей (см. в п. 3 основные группы субъектов — физических лиц).

По сути, архитектура должна быть нацелена не только на внешнее функционирование в рамках ЕИП РФ, но и на создание и функционирование единого информационного пространства в составе своей УЕИП, которое в общем случае может включать три составляющие (подпространства контуров): О-ЕИП, К-ЕИП и З-ЕИП.

## **2 Общий подход к архитектурному построению единых информационных (информационно-телекоммуникационных) систем конкретных организаций — участников единого информационного пространства России**

1. В общем подходе к архитектурному построению весьма существенным является *рациональное структурирование ЕИС (ЕИТКС) конкретной УЕИП* с определением номенклатуры функциональных составных частей — сегментов и компонентов (систем, подсистем, сетей, комплексов и др.) ЕИС (ЕИТКС) и их взаимодействия между собой на принципах консолидации, интеграции и централизации.

Количество автономных информационных систем у каждого потенциального участника ЕИП РФ с переходом на основе единой технической политики к ЕИС или ЕИТКС должно сокращаться [1–3]. Прежде всего, это потребует кардинального пересмотра архитектуры ЕИС (ЕИТКС) в рамках УЕИП, разработки общего языка для ее определения и описания, а также определения возможности сопряжения и эффективного взаимодействия различных подсистем / компонентов в рамках единой системы. Обязательным условием является выработка модели процесса создания профиля среды открытой системы (СОС) — ЕИС (ЕИТКС) для конкретной УЕИП [2, 6].

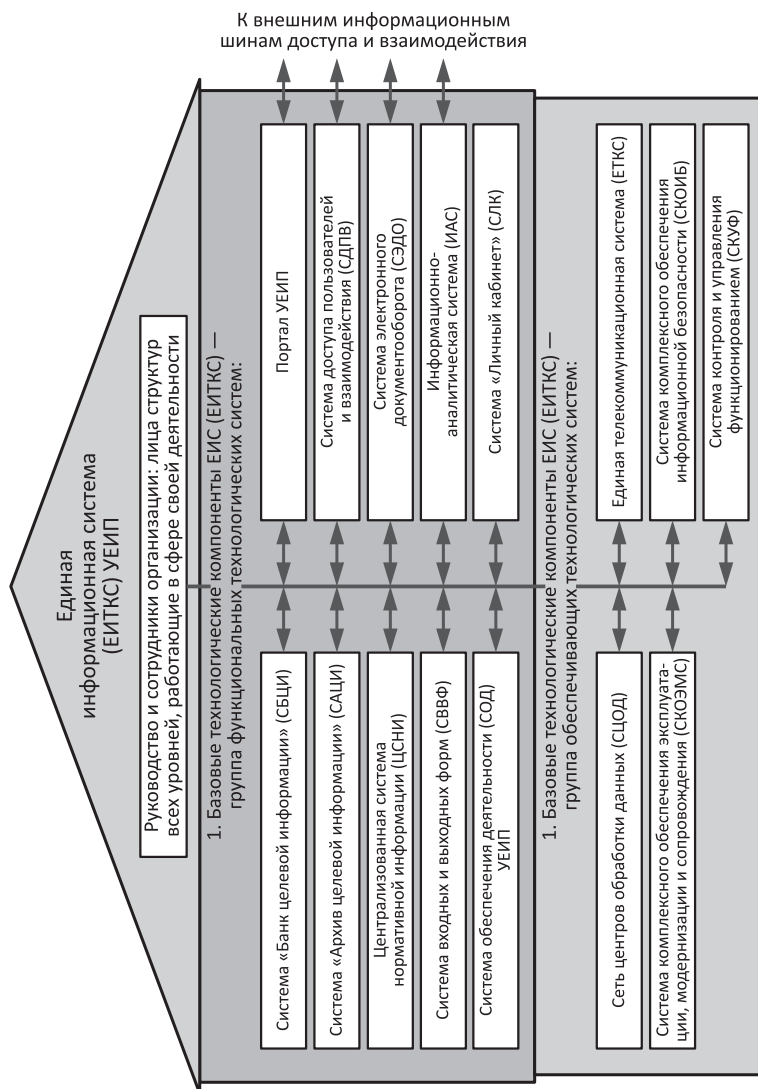
Рациональная компоновка реализуемых в ЕИС (ЕИТКС) функций и решаемых задач должна обеспечивать: минимизацию числа входящих в ЕИС (ЕИТКС) различных компонентов — систем, подсистем, сетей и комплексов; логическую законченность функций и задач каждого компонента; независимость компонентов по реализации целевых функций.

2. *Рациональное структурирование базируется на коллективном использовании ресурсов, услуг и возможностей ЕИС (ЕИТКС) конкретной УЕИП.* В основе такого структурирования лежит системный подход, согласно которому предлагается разделение сложной системы на две укрупненные составные части:

- (1) *группа функциональных технологических систем.* В эту группу входят базовые системы, которые непосредственно связаны с требованиями обеспечения необходимого функционала, возложенного на ЕИС (ЕИТКС) конкретной УЕИП, с выполнением соответствующих функций в электронной форме и оказанием требуемых услуг;
- (2) *группа обеспечивающих технологических систем.* Эта группа включает базовые системы, которые поддерживают и обеспечивают нормальную работу группы базовых функциональных систем.

На рисунке приведен предлагаемый обобщенный вариант архитектурного построения ЕИС (ЕИТКС) УЕИП на уровне базовых технологических систем. Ниже показан и кратко охарактеризован примерный типизированный состав двух указанных групп базовых систем.





1. Основная шина электронного доступа и взаимодействия — на базе общероссийской системы электронного взаимодействия (ОСЭВ)
2. Шина электронного документооборота
3. Шина общего пользования — на базе сети Интернет
4. Шина взаимодействия — на базе многофункциональных центров (МФЦ)
5. Шина бумажного документооборота — на базе Почты России, Спецсвязи и др.

Предлагаемый обобщенный вариант архитектурного построения ЕИС (ЕИТКС) УЕИП на уровне базовых технологических систем



3. *Группа функциональных технологических систем* определяется на основе требований по обеспечению эффективной информатизации конкретной ЕИС (ЕИТКС). Эта группа позволяет сформировать рациональное функциональное наполнение и фактически отражает одновременно сущность целевого функционирования информационной системы, специализацию функций и задач по определенным укрупненным направлениям функционирования каждого ее сегмента.

Анализ результатов исследований по аналогичным системам, следование принципу централизации, необходимость консолидации и коллективного использования ресурсов — все это в совокупности указывает на безусловную целесообразность минимизации числа первой группы систем. Такой подход ведет к сокращению объема аппаратно-программных средств, потребностей в информационных и телекоммуникационных ресурсах при их эффективном коллективном использовании, уменьшению числа обслуживающего персонала. Минимизация числа систем первой группы не накладывает каких-либо ограничений на развитие соответствующих функциональных возможностей ЕИС (ЕИТКС) конкретной УЕИП. Определение числа систем первой группы и их «функционального наполнения» должно осуществляться на основе индивидуального, гибкого подхода применительно к потребностям ЕИС (ЕИТКС) конкретной УЕИП, с учетом специфики ее деятельности и тенденций развития.

Предлагается *примерный типизированный состав первой группы*, включающий десять следующих функциональных технологических систем:

- (1) *система «Банк целевой информации» (БЦИ)*, в которой целевая информация непосредственно отражает сущность целевого функционирования ЕИС (ЕИТКС) своей УЕИП;
- (2) *система «Архив целевой информации» (АЦИ)*;
- (3) *информационно-аналитическая система (ИАС)*. В состав ИАС при необходимости может входить информационно-аналитический или ситуационный центр своей УЕИП;
- (4) *портал УЕИП*;
- (5) *система доступа пользователей и взаимодействия (СДПВ)*;
- (6) *система электронного документооборота (СЭДО)*;
- (7) *система входных и выходных форм (СВВФ)*;
- (8) *централизованная система нормативной информации (ЦСНИ)*. В состав ЦСНИ входят три составных части: нормативно-правовая, нормативно-техническая, нормативно-справочная;
- (9) *система «Личный кабинет» (СЛК)*. Потенциально объем хранимой (используемой) в СЛК информации может оказаться чрезмерно большим. В таком случае СЛК целесообразно использовать при необходимости выборочно — только назначенному контингенту из числа руководителей, сотрудников и исполнителей, выполняющих свои непосредственные работы в конкретной сфере деятельности;

- (10) *система обеспечения деятельности (СОД) УЕИП*. Сюда входит необходимая совокупность подсистем (кадровая, финансовая, материально-техническая и др.), которая определяется в процессе детального проектирования ЕИС (ЕИТКС) конкретной УЕИП.

Каждая из функциональных подсистем создается в составе трех сегментов: открытого (О), конфиденциального (К) и закрытого (З), а также общей подсистемы администрирования.

Выбор состава функциональных технологических систем для конкретной ЕИС (ЕИТКС) определяется потребностями своей УЕИП.

4. *Группа обеспечивающих технологических компонентов* в общей архитектуре О-сегмента, К-сегмента и З-сегмента ЕИС (ЕИТКС) конкретной УЕИП призвана решать три укрупненные задачи:

- (1) реализация эффективного, хорошо структурированного информационного обмена в пределах своих контуров;
- (2) реализация централизованной территориально распределенной обработки информации в пределах своих контуров;
- (3) обеспечение безопасного и надежного функционирования каждого из сегментов в течение жизненного цикла.

Основным компонентом для решения *первой задачи* является телекоммуникационная подсистема (ТКП) своего сегмента, которая обеспечивает два вида информационного обмена:

- (а) локальный информационный обмен в пределах объектового программно-технического комплекса (О-ПТК, К-ПТК и З-ПТК) внутри ПТК и данного ПТК с другими смежными ПТК в рамках своего объектового узла. Локальный обмен реализуется посредством локальной вычислительной сети (ЛВС);
- (б) удаленный обмен информацией (данными) между территориально разнесенными ПТК своего сегмента, который реализуется посредством ТКП с использованием каналов связи различных операторов.

Для решения *второй задачи* предлагается использовать сеть центров обработки данных (СЦОД) в качестве базисного компонента ЕИС (ЕИТКС) конкретной УЕИП. Такая СЦОД в общем случае создается как высоконадежная (катастрофоустойчивая) система, которая используется совместно всеми тремя контурами (О-сегмент, К-сегмент и З-сегмент), что позволит получать от этой сети для своего функционирования необходимые доли ресурсов. Сеть центров обработки данных берет на себя выполнение основной части задач, решаемых группой функциональных технологических компонентов. Центры обработки данных СЦОД также осуществляют резервное хранение и архивирование информации. Введение СЦОД в архитектуру соответствует требованиям современных технологий обработки данных (включая технологии облачных вычислений) и, безусловно, является перспективным для ЕИС (ЕИТКС) конкретной УЕИП.

Сеть центров обработки данных как базисный компонент ЕИС (ЕИТКС) конкретной УЕИП разделяется на три (по максимуму) подсегмента (подсистемы): О-ПЦОД, К-ПЦОД и З-ПЦОД. Кроме того, в СЦОД входит общая подсистема администрирования СЦОД-ОПА.

*Третья укрупненная задача* решается в общем случае посредством использования трех обеспечивающих технологических систем: системы комплексного обеспечения информационной безопасности (СКОИБ); системы контроля и управления функционированием (СКУФ); системы комплексного обеспечения эксплуатации, модернизации и сопровождения (СКОЭМС). В каждую из этих систем входят четыре составные части: три подсистемы сегментов О, К и З, а также общая подсистема администрирования.

Таким образом, решение трех укрупненных задач в общей архитектуре ЕИС (ЕИТКС) многочисленных УЕИП выполняется группой из пяти обеспечивающих компонентов: единой телекоммуникационной системой (ЕТКС); СЦОД; СКОИБ; СКУФ; СКОЭМС.

5. Единой информационной системе (ЕИТКС) конкретной УЕИП необходимо осуществлять внешние взаимодействия с другими УЕИП посредством указанного выше так называемого коммуникационного слоя, который включает набор необходимых внешних информационных шин доступа и взаимодействия.

С использованием таких шин соответствующие структуры РФ и физические лица получают через ЕИС (ЕИТКС) определенных УЕИП необходимую им информацию в соответствии со сферой их деятельности. По этим же шинам данная УЕИП получает необходимую ей информацию от других УЕИП.

Конфигурация коммуникационного слоя может быть различной для разных УЕИП. На рисунке внешние взаимодействия включают в свой состав несколько информационных шин доступа и взаимодействия. В группе функциональных систем ЕИС (ЕИТКС) конкретных УЕИП доступ к таким внешним шинам и осуществление информационных взаимодействий предлагается выполнять на основе следующих четырех систем: портала УЕИП; СДПВ; СЭДО; ИАС.

К числу внешних информационных шин доступа и взаимодействия, показанных на рисунке, относятся следующие:

- (1) основная шина электронного доступа и взаимодействия — на базе общероссийской системы электронного взаимодействия (ОСЭВ) [2]. Прототипом ОСЭВ должна стать система межведомственного электронного взаимодействия (СМЭВ), созданная в рамках ФЦП «Электронная Россия» [7]. Система межведомственного электронного взаимодействия в настоящее время активно развивается. Минкомсвязь России в 2014 г. провела успешное тестирование модернизированной СМЭВ версии 3.0 в продуктивной среде. Принято Постановление Правительства РФ от 19 ноября 2014 г. «О дальнейшем развитии Единой системы межведомственного электронного взаимодействия» (ЕСМЭВ), также издан приказ Минкомсвязи России № 510 от 30 декабря 2014 г. «Об утверждении Технических требований

к взаимодействию информационных систем в Единой системе межведомственного электронного взаимодействия» [8]. В ЕСМЭВ создана подсистема «Ситуационный центр», порядок работы этой подсистемы утвержден Постановлением Правительства РФ от 19 марта 2014 г. «О внесении изменений в Положение о Единой системе межведомственного электронного взаимодействия». К ситуационному центру электронного правительства в 2014 г. подключились 19 основных ведомств — участников межведомственного электронного взаимодействия. Ситуационный центр считается единой точкой контакта участников по всем возникающим проблемам. Отсюда следует, что ЕСМЭВ будет развиваться и далее от межведомственного уровня до уровня ОСЭВ;

- (2) шина электронного документооборота (желательно, чтобы она была общероссийского масштаба);
- (3) шина общего пользования — на базе сети Интернет;
- (4) шина взаимодействия — на базе многофункциональных центров (МФЦ), созданных в рамках ФЦП «Электронная Россия»;
- (5) шина бумажного документооборота (Почта России, Спецсвязь и др.).

Возможно, что некоторые из перечисленных внешних информационных шин доступа и взаимодействия частично могут использоваться для реализации внутренних взаимодействий в ЕИС (ЕИТКС) конкретных УЕИП на основе результатов технического проектирования и по согласованию с владельцами внешних информационных шин. Кроме того, определенные УЕИП могут взаимодействовать между собой с использованием средств из состава своих ЕИС (ЕИТКС) по соответствующим соглашениям.

### 3 Заключение

В данной статье на основе требований к архитектуре ЕИС (ЕИТКС), изложенных в статье [3], рассматриваются предложения по обобщенной архитектуре ЕИС УЕИП с ориентацией на ее возможную типизацию. Показано, что обобщенная архитектура ЕИС (ЕИТКС) может ориентироваться в общем случае на три взаимосвязанных сегмента — О-сегмент, К-сегмент и З-сегмент, каждый из которых может рассматриваться как самостоятельная автоматизированная система. Определены в общем виде группы пользователей ЕИС различных УЕИП, а также обслуживающего персонала, обеспечивающего функционирование таких систем. Предлагается предметное понятие архитектуры ЕИС различных УЕИП с базированием на трех ключевых положениях: фундаментальном подходе к системе, рациональному структурированию системы, учету фактора внешней среды.

В рамках рационального структурирования дается общий подход к определению номенклатуры составных частей единой системы (базовых систем, подсистем).

тем, сетей, комплексов и др.) и их взаимодействию между собой на принципах консолидации, интеграции и централизации.

Предлагается разделение ЕИС конкретной УЕИП на две укрупненные группы базовых технологических систем: функциональные и обеспечивающие системы. Показано, что примерный типизированный состав этих групп могут составить 10 функциональных и 5 обеспечивающих систем.

Материалы данной статьи могут быть полезны при формировании нормативно-технической базы построения единых информационных систем различных УЕИП.

## Литература

1. *Зацаринный А. А., Киселев Э. В.* Некоторые подходы к формированию нормативно-технической базы для создания единого информационного пространства России // Системы и средства информатики, 2014. Т. 24. № 4. С. 206–220.
2. *Зацаринный А. А., Киселев Э. В.* Некоторые подходы к формированию нормативно-технической базы единого информационного пространства России в части информационных ресурсов // Системы и средства информатики, 2015. Т. 25. № 1. С. 157–169.
3. *Зацаринный А. А., Киселев Э. В.* Некоторые подходы к формированию нормативно-технической базы в части требований к архитектурному построению информационных систем организаций — участников единого информационного пространства России // Системы и средства информатики, 2015. Т. 25. № 3. С. 179–194.
4. ГОСТ Р 51275-99. Защита информации. Объект информатизации. Факторы, воздействующие на информацию. Общие положения.
5. ГОСТ Р 51583-2000. Защита информации. Порядок создания автоматизированных систем в защищенном исполнении.
6. Информационные технологии: Руководство Р 50.1.041-2002 по проектированию профилей среды открытой системы (СОС) организации-пользователя. Утверждено Постановлением Госстандарта России от 14 ноября 2002 г. № 415-ст. Введено в действие с 1 января 2004 г.
7. Электронная Россия: Федеральная целевая программа. Последние изменения 2008 г.
8. Книга участника годовой расширенной коллегии Министерства связи и массовых коммуникаций РФ. — М.: Минкомсвязь России, 2015. С. 43–49.

*Поступила в редакцию 14.09.15*

## REGARDING SOME APPROACHES TO CREATION OF GENERALIZED ARCHITECTURE OF RUSSIAN INFORMATION SPACE MEMBERS OF INFORMATION SYSTEMS

*A. A. Zatsarinny<sup>1</sup> and E. V. Kiselev<sup>2</sup>*

<sup>1</sup>Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

<sup>2</sup>Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** The article continues a publications line about regulatory and technical base forming for creation of Russian Federation Unified Information Space (RF UIS). The requirements to unified information systems architecture are in the article foundation. There are shown general propositions of systemic approach about architecture of information systems of organizations which are the members of RF UIS. The propositions regarding generalized architecture of RF UIS member information system and its possible typing are also presented. It is proved that unified information system generalized architecture should be focused on three interrelated segments: open, confidential, and closed. The segments must be considered as independent automated systems. The RF UIS organizations user groups are generally determined, including maintenance staff providing unified system functioning. Objective determination of unified information system architecture is given. It is based on three key points: fundamental approach to system, system rational structuring, and environment factors consideration. In the second point, the authors propose a general approach to determination of unified system component nomenclature (base systems, subsystems, networks, complexes, etc.) and their interrelations on principles of consolidation, integration, and centralization. There is an approach which presents unified information system of RF UIS member as two groups of base technological systems: Functions and Maintenance. The thesis of the article can be used as a scientific and methodological ground during regulatory and technical base forming of RF UIS members information systems creation.

**Keywords:** Russian Federation Unified Information Space; Unified Information Space member; systemic approach; UIS member unified information system architecture; secured information resources; centralization; integration; information security; telecommunication provision; group of base functional technological systems; group of base maintenance technological systems

**DOI:** 10.14357/08696527150409

### References

1. Zatsarinny, A. A., and E. V. Kiselev. 2014. Nekotorye podkhody k formirovaniyu normativno-tekhnicheskoy bazy dlya sozdaniya edinogo informatsionnogo prostranstva

- Rossii [Regarding some approaches to creation of regulatory and technical base for Russian Unified Informational Space creation]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 24(4):206–220.
2. Zatsarinny, A. A., and E. V. Kiselev. 2015. Nekotorye podkhody k formirovaniyu normativno-tekhnicheskoy bazy edinogo informatsionnogo prostranstva Rossii v chasti informatsionnykh resursov [Regarding some approaches to creation of regulatory and technical base of Russian Unified Informational Space in the field of information resources]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 25(1):157–169.
  3. Zatsarinny, A. A., and E. V. Kiselev. 2015. Nekotorye podkhody k formirovaniyu normativno-tekhnicheskoy bazy v chasti trebovaniy k arkhitektturnomu postroeniyu informatsionnykh sistem organizatsiy — uchastnikov edinogo informatsionnogo prostranstva Rossii [Regarding some approaches to creation of regulatory and technical base concerning requirements to architectural development of Russian information space members information systems]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 25(3):179–194.
  4. GOST R 51275-99. Zashchita informatsii. Ob"ekt informatizatsii. Faktory, vozdeystvuyushchie na informatsiyu. Obshchie polozheniya [Information security. Informatization object. Affecting information factors. Generalized theses].
  5. GOST R 51583-2000. Zashchita informatsii. Poryadok sozdaniya avtomatizirovannykh sistem v zashchishchennom ispolnenii [Information security. An order of secured automated information systems creation].
  6. RF R 50.1.041-2002. Actual from 01.01.2004. Rekomendatsii po standartizatsii: Informatsionnye tekhnologii. Rukovodstvo po proektirovaniyu profiley sredey otkrytoy sistemy (SOS) organizatsii-pol'zovatelya [Standartization recommendation: Information technologies. Guidance on design of user-organization open system environment profiles].
  7. Elektronnaya Rossiya: Federal'naya tselevaya programma [Electronic Russia. Federal objective program]. Update of 2008.
  8. Kniga uchastnika godovoy rasshirennoy kollegii Ministerstva svyazi i massovykh kommunikatsiy RF [A book of participant of annual extended board of Ministry of Communications and Mass Media of Russian Federation]. 2015. Moscow: Minkomsviyz'. 43–49.

*Received September 14, 2015*

## Contributors

**Zatsarinny Alexander A.** (b. 1951) — Doctor of Science in technology, professor, Deputy Director, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; azatsarinny@ipiran.ru

**Kiselev Eduard V.** (b. 1938) — senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; ekiselev@ipiran.ru



## ОСНОВЫ ПРИМЕНЕНИЯ ОНТОЛОГИИ И КОМПЬЮТЕРНОЙ ЛИНГВИСТИКИ ПРИ ПРОЕКТИРОВАНИИ ПЕРСПЕКТИВНЫХ АВТОМАТИЗИРОВАННЫХ ИНФОРМАЦИОННЫХ СИСТЕМ

*И. И. Быстров<sup>1</sup>, Б. В. Тарасов<sup>2</sup>, А. А. Хорошилов<sup>3</sup>, С. И. Радманов<sup>4</sup>*

**Аннотация:** Рассматривается применение онтологии и компьютерной лингвистики (КЛ) в автоматизированных информационных системах (АИС). В части 1 статьи излагаются методологические основы комплексного применения онтологии и КЛ при проектировании перспективных АИС. Приводится описание онтологических процессов и КЛ.

**Ключевые слова:** семантика; онтология; компьютерная лингвистика; предметная область; декларативные и процедурные знания; автоматизированная информационная система; лингвистическое информационное обеспечение; логико-семантическое моделирование; концептуализация

**DOI:** 10.14357/08696527150410

### 1 Введение

В настоящее время в связи с возросшей ролью АИС в деятельности государственных и бизнес-организаций и превращения их в автоматизированные органы повысились требования к качеству обработки информации. До недавнего времени современная наука почти не касалась информационных процессов, в ходе которых возникают, а затем наращиваются по своему смысловому содержанию новые знания о реальном мире. Данные информационные процессы основаны на семантических свойствах информации и используют семантические принципы ее преобразования. В работе [1] эти процессы названы динамическими в отличие от

---

<sup>1</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, [ibystrov@ipiran.ru](mailto:ibystrov@ipiran.ru)

<sup>2</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, [tarasov@nrcmodul.ru](mailto:tarasov@nrcmodul.ru)

<sup>3</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, [khoroshilov@mail.ru](mailto:khoroshilov@mail.ru)

<sup>4</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, [radomanov@list.ru](mailto:radomanov@list.ru)



доминирующих ныне статических процессов, где смысл информации представляется в формализованном неизменном виде на уровне данных и метаданных. При таком статическом подходе нередко утрачиваются те качества информации, которые реализуют ее функцию познания окружающего мира и происходящих в нем событий, в том числе представленных в предметной области (ПО). В рамках АИС ПО реализуется в виде инфолингвистической и алгоритмической модели. Ориентация АИС на предоставление качественных семантически емких информационных и телекоммуникационных информационно-технологических (ИТ) услуг, на обработку динамической информации обуславливает необходимость поднять «планку» логико-семантической обработки формализованной и неформализованной информации для дальнейшего развития интеллектуальных способностей АИС, включая порождение новых данных и знаний [2, 3].

Концептуальной основой дальнейшего развития «интеллектуальных способностей» перспективных АИС является продуктивное осмысление декларативной и процедурной составляющей логико-семантической обработки информации, представленной на естественном языке (ЕЯ). Декларативная часть включает в себя информацию о сущностях ПО, ее свойствах и отношениях с другими сущностями, что является задачей онтологии. Процедурная часть содержит сведения о том, как можно использовать декларативную информацию, чтобы получить знания (данные), которые необходимы для дальнейшей обработки информации с целью формализации ее смыслового содержания. В результате такого продуктивного осмысления происходит моделирование процесса анализа и синтеза исходной декларативной информации в знания и обогащение процедурных знаний, что является главной задачей КЛ.

Важной методологической проблемой является создание таких АИС, в которых должно быть обеспечено необходимое соотношение между декларативными и процедурными компонентами системы автоматической обработки текстовой информации, т. е. мощными процедурами КЛ и более простыми процедурами онтологического моделирования. Решению этой проблемы посвящена теория и практика развития АИС, обладающих требуемым уровнем интеллекта для эффективной логико-семантической обработки информации и общения пользователя с системой на ЕЯ, а также создания систем, адекватно моделирующих человеческий интеллект. Проблемами использования ЕЯ в системах автоматической обработки занимается КЛ. Эта наука возникла в 1950-е гг. За прошедшее время были получены значительные теоретические и практические результаты в области машинного перевода, автоматизированного поиска информации и т. п. Однако остается еще много нерешенных проблем в области моделирования анализа и синтеза процесса понимания смысла текстов, в частности для семантического поиска и формирования новых знаний. В 2000-е гг. как зарубежными, так и российскими учеными было предложено использовать в моделях семантического поиска и обработки знаний онтологии ПО. Потребность в использовании онтологии обусловлена целым рядом причин, основными из которых являются следующие:

- возникла необходимость и возможность формировать новые знания на базе обработки статической и динамической информации больших объемов;
- возросла потребность как пользователей, так и обслуживающего персонала в общем понимании смысла ИТ-услуг, предоставляемых со стороны АИС;
- современные информационные технологии сделали актуальной задачу обработки распределенных знаний, а также отделения знаний ПО от оперативных знаний, получаемых в процессе обработки данных и логико-семантического анализа текстов, представленных на ЕЯ;
- появились и развиваются средства формального выражения семантики ПО для машинной обработки смыслового содержания текстов.

В настоящее время разработка онтологии переходит из мира лабораторий по искусственному интеллекту в область описания и моделирования смысла ПО и описывающего ее корпуса текстов как объекта реального мира. Возникает проблема обеспечения интероперабельности в АИС методов и средств онтологии с методами и средствами КЛ.

В литературе по АИС освещаются отдельные аспекты взаимосвязи КЛ и онтологии [4–7]. Однако в этих источниках, рассматривающих построение АИС, отсутствует комплексный подход к проблеме совместного использования инструментов онтологии и средств КЛ. К сожалению, мало практически значимых примеров построения АИС на базе совмещения этих двух инструментов. И хотя оба этих подхода в конечном счете решают (как правило, частично) общую проблему: извлечение смысла из полученных данных о той или иной ПО, — в большинстве случаев они рассматриваются как самостоятельные и независимые направления исследований. В данной статье предпринята попытка в какой-то мере восполнить этот пробел. В частности, рассматриваются вопросы взаимосвязи КЛ и онтологии при поддержке ИТ-услуг и при проектировании современных перспективных АИС класса информационно-телекоммуникационных систем (ИТС). По мнению авторов, автоматизированная обработка данных и знаний в перспективной АИС с целью получения новых знаний должна использовать такие подходы, инструменты и методы анализа и синтеза структурированной и неструктурированной динамической информации, которые предусматривают применение комплекса моделей как КЛ, так и онтологии.

## **2 Концептуальная модель предметной области**

Понятие ПО является базовым в информатике и процессах проектирования АИС. В современных АИС информация о внешнем мире и способах ее интерпретации представляется в виде данных и метаданных. Такой подход ограничивает возможность смысловой привязки информации к семантике той или иной ПО. Основной причиной этого является недостаточно глубокое представление и моделирование знаний о ПО с помощью существующих декларативных и процедурных

средств манипулирования, поэтому одной из важных задач компьютерных наук является разработка теории и практики концептуального моделирования ПО без потерь ее семантики.

Понятие ПО было введено в начале 1980-х гг., когда была осознана необходимость использования семантических моделей для предоставления и обработки информации в АИС. К сожалению, оно не имеет точного определения, так как одни исследователи рассматривают ПО как часть действительности (онтологический аспект), другие — как знания об этой действительности (гносеологический аспект). В то же время извлечение знаний путем автоматической обработки информации, предоставляемой на формализованных или неформализованных ЕЯ, отгалкивается от лингвистических форм представления знаний о ПО и поэтому требует применения всего комплекса методов и средств КЛ, онтологии и гносеологии. Для дальнейшего рассмотрения введем рабочее определение ПО. Под ПО будем понимать целенаправленную трансформацию информационно-функциональной деятельности объекта информатизации в некоторую информационно-лингвистическую модель, которая фиксируется в АИС в качестве алгоритмической модели понимания смысла обрабатываемой информации. Из этого определения следует, что внешний мир отражается в деятельности объекта автоматизации через ПО и не воздействует непосредственно на АИС. Информация о внешнем мире в АИС поступает в виде знаний разработчиков системы через свойства, функции и структуру ПО. Эти знания поступают в виде текстовых описаний, документов и др. Существует целый ряд методик исследования ПО. При этом важно учитывать, что концептуальная модель ПО скорее описывает процессы, происходящие в ПО, и данные, используемые этими процессами. От качества моделирования ПО зависит уровень разработки программных приложений. Здесь важно отметить, что если онтология используется для спецификации знаний о внешнем мире без связи с реализацией в АИС, то это область философии (онтология верхнего уровня). Разработчики АИС сталкиваются с концептуализацией ПО в интересах построения модели решаемой практической задачи, поэтому они часто используют вместо термина «онтология» термин «концептуальная схема» (например, в теории построения баз данных (БД)). В общем случае методы построения АИС зависят от структуры, свойств ПО и качества отображения ее в информационно-лингвистическую модель обработки данных и знаний. В основе проектирования информационного и лингвистического обеспечения АИС лежит моделирование ПО. При этом под моделью ПО понимается система, имитирующая содержание и структуру ПО в интересах ее реализации в АИС. В рамках современных методологий разработки АИС описание ПО и, в частности, ее декларативной составляющей чаще всего имеет концептуальный характер, что предопределяет применение методов концептуального моделирования ПО. В результате концептуального моделирования создается модель ПО в виде множества понятий и связей между ними, определяющих смысловую структуру ПО. Концептуальное моделирование в ряде работ рассматривается как общая проблема для технологий БД, искусственного интел-

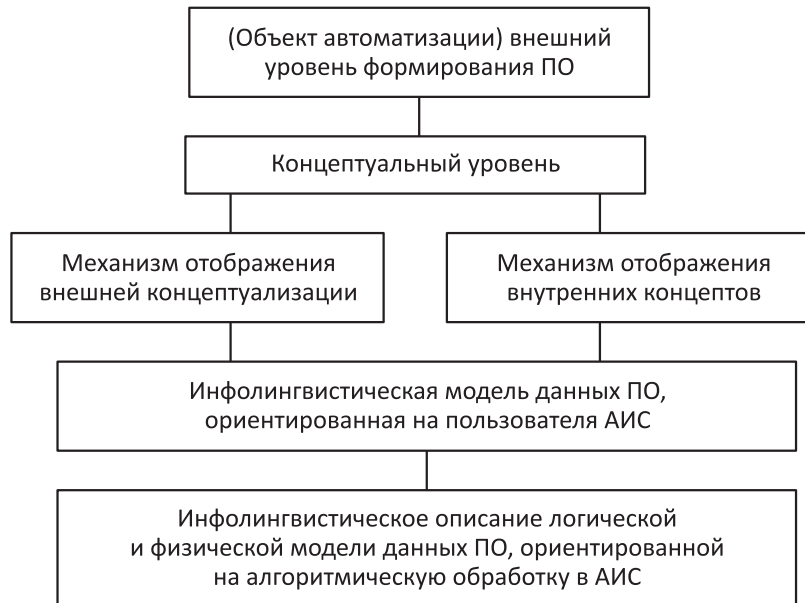


Рис. 1 Архитектура концептуального моделирования ПО

лекта и программного инжиниринга. В общем виде концептуальная модель — это наши знания о внешнем мире, которые необходимо преобразовать в модель ПО в виде системы понятий (концептов) и их взаимосвязей. Семантический уровень концептуальной модели зависит от выразительных средств, используемых для ее описания. Поскольку для описания внешнего мира, как правило, используется ЕЯ, то для концептуальной модели лучшим выразительным средством также является ЕЯ, но, к сожалению, он не структурирован и не обеспечивает однозначной интерпретации. Использование специального языка моделирования ПО является единственно возможным способом однозначного описания ПО на различных семантических уровнях. Концептуальное моделирование ПО широко используется в технологиях БД и в онтологическом моделировании. На рис. 1 показана архитектура концептуального моделирования ПО.

Для выполнения показанных на рис. 1 функций нужно иметь языки описания в форме, понятной человеку, и в форме, понятной компьютеру. Эту роль выполняют информационные и алгоритмические языки, которые предусматривают возможность однозначного описания ПО. Данные языки являются специально разработанными для удобства выражения человеком основного смыслового содержания информации и запросов, понимания смысла ИТ-услуг и обеспечения необходимой алгоритмизированной обработки данных в АИС. В ИПИ РАН в 1993 г. разработан язык высокого уровня «Синтез», который вопло-

щает семантическую модель данных для концептуального моделирования ПО, а также позволяет специфицировать концептуальную схему ПО [8]. Построение концептуальной модели ПО можно определить как процесс формализации зафиксированных в ПО концептов, логико-семантических отношений между ними и ограничений. При этом важно, чтобы эти элементы отражали существенные признаки информационно-функциональной деятельности организации и обеспечивали необходимую семантическую взаимосвязь с реальностями внешнего мира. Это не означает, что в модели ПО должны быть отражены все ее функциональные и информационные аспекты, а также их связи. Здесь важно понимать, что модель ПО необходима не сама по себе, а для однозначного семантического понимания решаемых задач, исходя из предназначения АИС. Это требует, чтобы разработчики лингвистического обеспечения руководствовались определенным критерием глубины детализации ПО исходя из семантической значимости задач, решаемых в АИС. В работе [9] предлагается при проведении моделирования ПО выделять следующие свойства модели ПО: зависимость от целей и задач, решаемых АИС, формальность, адекватность действительности, способность давать новую информацию, функциональная полнота и логическая целостность. Из этого перечня видно, что первичным для описания и построения модели ПО в АИС является класс решаемых задач, способность лингвистических и информационных компонент обеспечить семантическое понимание информации, предоставляемой пользователю в виде результатов обработки его запросов, т. е. в формате ИТ-услуг. Структура и свойства всех процессов обработки запросов пользователей могут быть эффективно описаны с помощью словаря терминов, адекватно опеределяющих атрибуты различных концептов: систем, объектов, свойств, отношений, процессов, ситуаций, сюжетов.

Концепт-системы — объекты, которые описываются списком собственных атрибутов, характеризующих структуру ПО, функциональность и критическую важность ресурсно-сервисных возможностей АИС.

Концепт-объекты описываются именами компонент АИС, каждое из которых имеет свои атрибуты, описывающие состав, структуру и функциональность. Описание аппаратной программной компоненты АИС соответствует концепт-свойству с отражением своей специфической ПО. Концепт-объектам в терминологии Resource Description Framework (RDF) соответствуют концепты предметов.

Концепт-свойство характеризует состояние устойчивости систем и объектов к негативным внешним и внутренним воздействиям (например, свойства отказо- и катастрофоустойчивости). Концепт-свойствами могут быть типы данных, описывающие значения и методы оценки устойчивости и рисков, которые позволяют судить о мере близости понятий при выработке решений на основе логического вывода на неточных и неполных данных. В RDF концепт-свойствам соответствуют предикаты.

Концепт-отношения используются для отображения парадигматических и синтагматических отношений, которые отражают связь между понятиями как внутри ПО, так и понятиями различных ПО.

Концепт-процессы характеризуют преобразование ресурсно-сервисных возможностей АИС в ИТ-услуги. С точки зрения обработки содержимого ПО в категорию концепт-процессов попадают понятия, атрибуты которых описывают требования бенефицианта/реципиента ИТ-услуги, технологические процессы ресурсно-сервисного формирования ИТ-услуги, состояние негативных факторов, сценарии их воздействия на компоненты АИС, каналы воздействия, результаты нарушения функциональности и доступности сервисов и ИТ-услуг, риски и др.

Ситуация — совокупность значений атрибутов, характеризующих состояние функционирования ИТС, доступность ИТ-услуг, непрерывность процессов их формирования, реакцию на внутренние и внешние воздействия и отношения между атрибутами в некоторый момент времени.

Сюжет — упорядоченная во времени последовательность ситуаций, имеющая системную и процессную пересекающуюся взаимосвязь. Сюжет может быть представлен в виде сценария, сетевого графика, диаграммы Ганта.

Все перечисленные концепты используются для описания ПО, а в процессе эксплуатации АИС — для отображения знаний при предоставлении ИТ-услуг. Все процессы предоставления ИТ-услуг, а также ресурсно-сервисного формирования последних в своей семантической интерпретации опираются на имеющиеся знания о ПО. В исходном виде ПО описывается множеством текстов (корпусом текстов), в результате анализа которых формируется модель понятий и связей между ними, определяющая смысловую структуру ПО. Анализируемые тексты содержат как наиболее общие понятия, которые не зависят от ПО, так и текстуральные понятия, определяющие специфическую терминологическую базу ПО. Таким образом, процесс концептуализации ПО представляет собой формирование множества понятий, а также информации о них (свойств, отношений, утверждений о них) путем анализа корпуса текстов, необходимых для описания и решения требуемых задач в рамках данной ПО.

Выделим два подхода к построению модели ПО на этапе концептуализации: признаковый (атрибутный) и структурный (когнитивный). Первый подход предполагает применение для описания информации метода элементарных триад (объект – атрибут – значение атрибута). Данный метод был предложен авторами в 1966 г. [10]. Второй подход основан на выделении сущностей ПО по семантике и характеру отношений между ними. Этот подход к построению модели ПО предполагает, в частности, выделение из корпуса текстов понятий, взаимосвязей понятий, метапонятий, логико-семантических отношений. Таким образом, концептуализация ПО выступает в качестве описания множества объектов и связей между ними, представленных в виде словарей, классификаторов и классификационных схем БД и знаний, извлекаемых из текстов ПО.

Процесс извлечения знаний из БД и их семантическая интерпретация основываются на упорядочении знаний путем их систематизации, создания единой иерархии понятий, унификации терминов и правил интерпретации в едином онтологическом словаре. В онтологическом словаре каждое слово или устойчивое словосочетание имеет несколько определений (толкований). При этом каждое

ключевое слово определения содержит гиперссылку на определение в этом же словаре. Данный словарь является средством устранения семантической неоднозначности как в ходе общения между специалистами, так и в человеко-машинном диалоге «запрос–ответ». Такой словарь является не просто тезаурусом ПО, но и средством построения лингвистического обеспечения АИС (машинных словарей и классификаторов). Таким образом, сбор, обработка и извлечение знаний из БД в АИС основывается на понимании семантики с учетом логико-понятийной системы представления ПО, что составляет основу общей онтологии. Эта онтология уточняется и расширяется с учетом результатов обработки текстов методами и средствами КЛ.

### 3 Онтология предметной области

В настоящее время существует несколько подходов к определению понятия «онтология»: философский, гуманитарный и компьютерный [5, 11–14]. В литературе часто при определении понятия «онтология предметной области» идет смешение этих подходов и стремление применить его в различных сферах деятельности. Термин «онтология» впервые появился в работе Г. Груберга [15], в которой рассматривались различные аспекты взаимодействия интеллектуальных систем между собой и с человеком. Онтология по Грубергу представляет собой описание декларативных знаний. Для обозначения работы по описанию декларативных знаний он ввел термин «концептуализация», а описание назвал «спецификацией», т. е. онтология стала определяться как «спецификация концептуализации». Это формальное представление, которое включает словарь понятий и соответствующих им терминов ПО, а также логические выражения (аксиомы), которые описывают множество отношений между понятиями. Для описания отношений в онтологиях используется весь арсенал формальных моделей и языков, разработанных в области искусственного интеллекта — исчисление предикатов, системы продукций, семантические сети, фреймы и т. п. Таким образом, модный сегодня термин «онтология» оказался близок по значению к термину «искусственный интеллект», а термин «онтологический инжиниринг» стал синонимом термина «инженерия знаний». Термин «онтология» в *слабом смысле* стал популярен в области систем машинного анализа текста, где в большинстве случаев используется в узком значении — в качестве синонима терминов «тезаурус» или «классификатор» — и представляет собой просто словарь понятий (концептов), каждому из которых соответствует синонимический ряд терминов, плюс иерархическую структуру взаимосвязей между ними типа «часть–целое» или «род–вид». Такие «онтологии в слабом смысле» используются для формулировки запросов к поисковой машине, для автоматической классификации (категоризации) текстов и, пожалуй, на этом все. Онтологии в *сильном смысле* предполагают, что работающие прикладные программы, относящиеся к классу систем извлечения знаний из текста и использующие «онтологии в сильном смысле», т. е. методы искусственного интеллекта, способны нетривиально перерабатывать извлеченные



из текста элементы знаний (интерпретировать, обобщать, выявлять зависимости, прогнозировать и т. п.) и порождать новые знания. Таких онтологий сегодня не существует, во всяком случае для русского языка.

Ограниченность в использовании онтологий обусловлена, на взгляд авторов, двумя факторами. Во-первых, слабым распространением методов компьютерного лингвистического анализа и синтеза текстов в АИС, способных интерпретировать синтаксические и логико-семантические отношения между словами и потому действительно извлекать знания как некие нетривиальные элементы, обладающие внутренней структурой, пригодные для нетривиальной смысловой обработки динамической информации. Такие системы на мировом и российском рынках только начали появляться в последние несколько лет (Net Owl, Attensity, RCO Fact Extractor) и еще не успели «обрасти» приложениями. Во-вторых, относительно низкой достоверностью автоматически извлекаемых из текста утверждений и фактов, что обусловлено как несовершенством алгоритмов анализа текста, так и качеством источников информации, поскольку для практики интересно извлечение знаний не из научной литературы, которая уже представляет конгломерат знаний, а из текстов любого уровня, к каковым относятся социальные сети Интернета, современные средства массовой информации и даже архивы научно-технических отчетов и т. п. Другая особенность применения онтологий в системах извлечения знаний из текста — необходимость иметь дополнительную лингвистическую составляющую как для распознавания различных способов обозначения понятий (омонимия и синонимия терминов), так и для семантической интерпретации разнообразных языковых конструкций в отношениях между этими понятиями (синонимичные лексико-грамматические конструкции). В итоге для систем извлечения знаний из текста наиболее типичной стала «онтология в слабом смысле» с относительно бедной концептуальной, но чрезвычайно богатой лингвистической составляющей.

Любая онтология для ее успешного использования должна обладать рядом свойств, выделенных Томом Грубером:

- прозрачностью — онтология должна эффективно передавать подразумеваемое значение определенного термина;
- связностью — онтология должна быть связанной, т. е. она должна позволять делать выводы, которые согласуются с исходными определениями понятий. По крайней мере, определяемые аксиомы должны быть логически согласованы между собой;
- расширяемостью — онтология должна быть разработана с возможностью использования разделяемого и пополняемого словаря;
- независимостью от синтаксиса — концептуализация должна быть специфицирована на уровне знания максимально независимо от представления понятий на уровне символов;
- минимальным базисом при высокой выразительности — онтология должна вводить минимальный базовый набор понятий, но их должно быть достаточно.



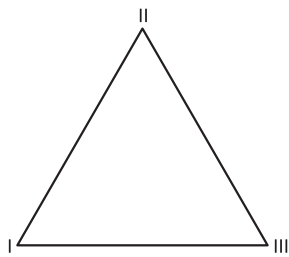
Компоненты, из которых состоят онтологии, зависят от парадигмы их представления. Но практически все модели онтологий в той или иной степени содержат концепты (понятия, классы, сущности, категории), свойства концептов (слоты, атрибуты, роли), отношения между концептами (связи, зависимости, функции) и дополнительные ограничения (определяются аксиомами, в некоторых парадигмах — фасетами). Термин «экземпляр» используется для представления элемента в ПО, т. е. элемента данного концепта. Онтология вместе с множеством отдельных экземпляров составляет базу знаний (БЗ).

В 1998 г. Н. Гуарино дал анализ возможных подходов к применению онтологии при создании информационных систем (ИС), в которых показал варианты использования онтологий в процессе проектирования ИС и в процессе их функционирования [16]. В рамках данной статьи содержание самого понятия «онтология» и состав функций, предписываемых онтологиям, следует трактовать несколько уже. Под онтологией АИС будем понимать формальное представление смыслового содержания информации в виде множества понятий ПО и связей между понятиями в виде формализованного описания на основе знаний, полученных в результате сбора, хранения и обработки статической и динамической информации, представленной на ЕЯ.

В информационных технологиях и компьютерных науках под онтологией подразумевается эксплицитная, т. е. явная, спецификация концептуализации, где в качестве концептуализации выступает описание множества объектов и связей между ними (*англ.* ontology is the theory of objects and their ties). Формально онтология состоит из понятий терминов, организованных в таксономию, их описаний и правил вывода. Онтология (в информатике) — это попытка всеобъемлющей и детальной формализации некоторой области знаний с помощью концептуальной схемы. Обычно такая схема состоит из структуры данных, содержащей все релевантные классы объектов, их связи и правила (теоремы, ограничения), принятые в этой области. Этот термин в информатике является производным от древнего философского понятия «онтология». Он включает в себя:

- метаонтологию, которая описывает наиболее общие понятия, не зависящие от ПО;
- онтологию ПО, которая формально описывает ПО, обычно применяется для того, чтобы уточнить понятия, определенные в метаонтологии (если используется) и/или определить общую терминологическую базу ПО;
- онтологию конкретной задачи, определяющую общую терминологическую базу задачи, проблемы;
- сетевую онтологию, которую часто используют для описания конечных результатов действий, выполняемых объектами ПО или задачи.

Основными понятиями онтологической модели являются: понятия, экземпляры, значения, атрибуты, отношения, парадигматические и синтагматические связи понятий, термины, синонимы и омонимы, глоссарии и тезаурусы.



**Рис. 2** Строение знака — треугольник Фреге

Для того чтобы ответить на сформулированный в предыдущем разделе вопрос о понимании семантики ПО, разберем сначала, что означают формулировки «понимать смысл» и «употреблять в смысле». Принято считать, что наиболее полное раскрытие понятия «смысл» было дано в известных работах Готлоба Фреге (1848–1925), который, следуя учению стоиков, разделял понятия «денотат» (экземпляр, предмет, значение), с одной стороны, «смысл» (понятие) — с другой и «знак» (слово, словосочетание) — с третьей. В более общем случае треугольник Фреге схематизируется

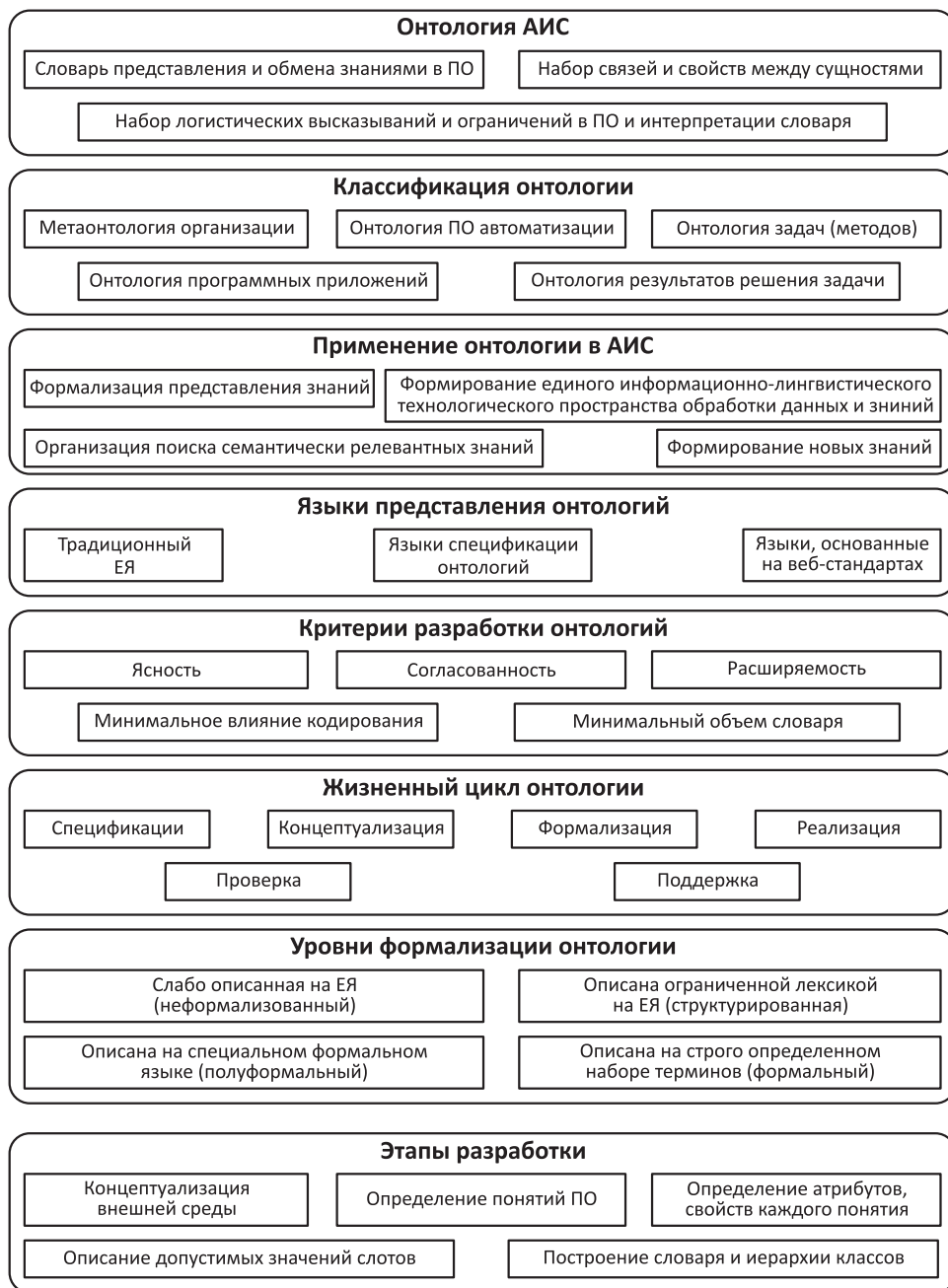
так, как это изображено на рис. 2. При этом Г. Фреге и его последователи фактически отождествляют «смысл» и «понятие», определяя «объем» последнего как некоторый класс предметов (сущностей), обладающих некоторым свойством (свойствами):

I — предмет, вещь, явление действительности, в математике — число и т. д. Иное название — денотат. Иногда этой вершиной треугольника обозначают не саму вещь, а ее восприятие или представление о ней, словом, ее отражение в сознании человека, называя это сигнификат. Сущность схемы-треугольника от этого не изменится;

II — знак: в лингвистике, например, фонетическое слово или написанное слово; в математике — математический символ; иное название, принятое особенно в философии и математической логике, — имя;

III — понятие о предмете, вещи. Иные названия: в лингвистике — десигнат, в математике — смысл имени, или концепт денотата.

Таким образом, онтология в современном понимании — это инструмент понимания смысла внешнего мира в интересах построения концептуальной модели ПО АИС, а также инструмент разработки и понимания смысла текстов на ЕЯ, фиксирование единого понимания смысла различных текстов ПО. Поскольку в каждой ПО могут существовать различные трактовки понимания одних и тех же терминов, онтология определяет соглашение о знании терминов и является посредником между человеком и машинно-ориентированным уровнем интеграции информации ПО. Иными словами, онтология является инструментом интеграции результатов КЛ для понимания смысла обрабатываемой в АИС информации. Здесь важно заметить, что кроме сходства между классификационными схемами и онтологией существуют и различия. При разработке онтологии ПО стремятся как можно к большей степени общности и независимости от информационных потребностей отдельных приложений и групп пользователей. Поэтому онтология включает относительно общие знания. В онтологиях, в отличие от концептуальных схем, используется более богатый набор логико-семантических отношений и используемых понятий. Если концептуальное моделирование в традиционных



**Рис. 3** Общая структура онтологической модели АИС

технологиях создания АИС используется только на стадии проектирования БД, то онтология используется на стадии обработки и восприятия ИТ-услуг. Иными словами, сфера применения онтологии шире концептуального моделирования БД, шире области проектирования алгоритмической инфраструктуры АИС. Они используются не только в области проектирования конкретных АИС и систем БД, но и в инженерии знаний, в программной инженерии и других областях информационных технологий.

Современные информационные технологии требуют решения задач распределенной семантики знаний. Онтология предоставляет средства для формального выражения этой семантики в АИС различного назначения. Область применения онтологии достаточно широка. Ее можно применять в качестве компоненты БЗ, схемы объектов в системах, схемы БД, структурированного глоссария взаимодействующих инженерных АИС, определять смысл терминов при ресурсно-сервисном подходе к формированию ИТ-услуг и т. д. На рис. 3 представлена общая структура онтологической модели АИС.

#### **4 Роль онтологии в проектировании автоматизированных информационных систем**

В рамках построения АИС возникает ряд проблем, связанных с обработкой знаний, содержащих информацию о процессах, угрозах, функционировании систем (сетей) и подсистем, методах и средствах организации превентивных мер обеспечения устойчивости функционирования системы.

Структура и состав знаний определяются назначением, особенностями функциональных процессов и методами проектирования АИС в каждом конкретном случае применения системы. Помимо технического аспекта решения этой проблемы центральное место занимают вопросы логико-семантического описания ПО, в рамках которого обеспечиваются сбор, хранение и обработка информации о текущей ситуации и поддержка принятия управленческих решений с учетом сложившейся ситуации. В основе проектирования информационной инфраструктуры АИС лежат процессы скоординированного взаимодействия онтологических и лингвистических подходов.

Однозначных механизмов решения проблемы онтолингвистического проектирования АИС, адекватной возникающим ситуациям, из-за семантического разнообразия ПО-процессов, функционально-конструкторских особенностей построения каждой системы (сети), подсистемы, их способности по обработке данных и реагированию на разнообразные угрозы на сегодня не существует. Это требует от разработчиков систем и эксплуатирующего персонала анализа содержания ПО каждой системы и условий их функционирования для обеспечения логико-семантического единства гетерогенных ПО. Классические методы анализа взаимодействия для решения задач технической интероперабельности не подходят. Системный подход предполагает, что входящие в систему методы ра-

ботаю по согласованному алгоритму (имеют синергетический смысловой эффект) и имеют одну точку управления. В то же время каждая система работает по независимым алгоритмам в рамках своей ПО (своих данных и метаданных). Поэтому при проектировании АИС используется семантически акцентированный подход к анализу спецификаций ПО, когда существует необходимость в обмене данными на смысловом уровне. Учитывая современные тенденции в семантической интероперабельности при решении задач проектирования АИС, целесообразно использовать онтологическое описание составных элементов системы. Тогда задача проектирования будет включать в себя формирование модели требований к ПО каждой системы исходя из вербальных спецификаций требований к процессам и решаемым задачам, разработку онтологий ПО и преобразование онтологии к концептуальным схемам.

Под онтологией проектирования конфигурационной схемы АИС будем понимать формальное описание содержания сервисов и ИТ-услуг в виде множества компонент, параметров, ограничений и понятий, описывающих ПО каждого типа сервисов и ИТ-услуг, а также связей между ними на основе формализованного описания процессов и компонент АИС, участвующих в их формировании. Из данного определения следует, что проектирование АИС строится на базе множества естественных и машинных языков. Онтология обеспечивает полное описание компонент системы вместе со значениями подходящих параметров и выбором необходимых понятий, т. е. с помощью которых описывается процесс функционирования каждой системы по предоставлению сервисов и их смысловому содержанию. Для каждой системы онтология может иметь различные формы, но она обязательно включает словарь терминов и некоторую спецификацию их смысла. Онтология проектирования АИС состоит из двух основных частей: общей онтологии, описывающей процессы семантической взаимосвязи пользователей с системой, и специализированной предметно-ориентированной онтологии, описывающей логико-семантическое содержание процессов сбора, хранения, обработки и передачи данных в формате ПО каждой системы (сети), подсистемы. Общая онтология содержит базовый набор терминов (гlossарий или тезаурус, классификаторы), необходимый для общения пользователя с ИТ (восприятия семантики ИТ-услуг, предоставляемых АИС). Общая онтология, как правило, представляет собой инструмент формализованного описания ЕЯ пользователя системы и является основным инструментом, определяющим семантическую доступность сервисов и ИТ-услуг пользователям АИС. Современная АИС представляет собой не только технически гетерогенную среду, но и среду, ориентированную на формализованное использование ЕЯ в системах автоматической обработки информации. В связи с этим она является инструментом преобразования данных, циркулирующих в процессах обработки различных систем в единую смысловую форму предоставления ИТ-услуг в интересах пользователя, средством преобразования машинных результатов в текстовые или речевые средства общения на ЕЯ. Применение общей онтологии позволяет изучать и сопоставлять между собой машинные знаковые системы с учетом трех аспектов: их структуры

(синтактики), их смысловой стороны (семантики) и их отношения к человеку (прагматики).

Спецификация концептуальной схемы помимо описания классов примитивов ПО и их связи с решаемой задачей содержит описание функций и процессов на формальном языке, т. е. является источником алгоритмизации, создания различного рода языков программирования и языков управления процессами преобразования ресурсно-сервисных возможностей АИС в ИТ-услуги, предоставляемые пользователю, с последующей их семантической интерпретацией. Способность отображать ПО, ее атрибуты и отношения между ними, описывать все эти знания на «языке смыслов» (ЕЯ), позволяющем понимать и анализировать информационные и телекоммуникационные ИТ-услуги в процессе функционирования АИС, сближает онтологию с КЛ, которая связана с решением задач автоматической обработки информации, представленной на ЕЯ [17]. Процесс проектирования лингвистического обеспечения АИС включает формирование модели требований к ПО исходя из вербальных спецификаций требований к задаче и преобразования онтологии к концептуальной схеме. Если в ПО определена совокупность задач, которые предстоит решать, используя информацию, содержащуюся в описаниях свойств ее сущностей, то такую совокупность называют пространством задач или проблемной областью [18]. Это связано с необходимостью проведения онтологического моделирования процессов проектирования больших систем, обеспечивающих процессно-ориентированный и структурно-семантический анализ ПО. В стандарте IDEF 5 отмечается, что основным средством спецификации систем и моделирования бизнес-процессов является онтологический инжиниринг. Онтологический инжиниринг — это основа концепции «управления знаниями» (knowledge management). При этом знания в процессе проектирования АИС трактуются следующим образом. Под знаниями понимается семантическая интерпретация данных о технологических и информационных процессах, а также о технологиях защиты этих процессов. Существуют различные подходы, модели и языки описания данных и знаний. Исследования показывают, что АИС состоит из множества разнородных компонентов, каждый из которых функционирует в своей онтологической среде. Следует отметить, что сама онтология часто носит описательный характер и никак не связана с реальными технологическими процессами обработки информации. Обращаясь к системе с запросом на ИТ-услуги извне, нельзя абсолютно точно предсказать, как именно она «сформулирует» свою реакцию на запрос. Организация интерфейса к онтологической системе АИС подразумевает абстрагирование от «внутренних знаний» (ресурсов систем). Семантическая интероперабельность подразумевает переход к обработке запросов пользователя на ИТ-услуги на базе метаданных, поддерживаемых в системе процессом формирования знаний о сервисах. Возникает проблема наличия в автоматизированной организации интеллектуального агента (службы типа *Service Desk*), который будет сопровождать семантическую обработку каждого запроса пользователя в случае возникновения непредвиденных обстоятельств, а также владеть всей информацией о деятельности организации (бизнес-процессах,

сервисной бизнес-логике, онтологической структуре метаданных, угрозах, функциональных процессах средств защиты). Все это превращает процесс обработки разнородных ПО в формализованное представление основных понятий и связей между ними. В результате деятельности организации ее информационная эффективность приобретает все большую зависимость от состояния АИС. Если оценивать критически концепцию управления знаниями (онтологию) в автоматизированной организации, то следует признать это в большей степени как лозунг менеджмента. Его связь с обработкой информации в АИС достаточно условна, так как онтология трактует знания свободно и широко, а в АИС знания трактуются на уровне ресурсов, БД, метаданных, различных процессов, сервисов и ИТ-услуг.

В инженерии знаний онтология позволяет детально описать ПО, которая используется для формального декларативного определения и концептуализации. Применение онтологического инжиниринга интерпретирует процессы управления различными категориями концептов, присущих каждой компоненте АИС. Структура онтологического инжиниринга каждой компоненты определена в процессе проектирования АИС своими концептами и представляет собой набор системных атрибутов. В формате перечисленных ПО онтологический инжиниринг обеспечивает управление разнообразными технологиями: процессами, БД, хранилищами метаданных, ресурсами и сервисами, системами информационной поддержки деятельности, локальными корпоративными системами автоматизации, системами документооборота, электронной почтой, системами управления АИС, системами индивидуального и группового доступа к сервисам и ИТ-услугам АИС, транспортными системами, системами обеспечения информационной безопасности, системами эксплуатации и поддержки ИТ-услуг. Каждая из этих технологий характеризуется своим набором концептов, которые описываются в рамках предметно-ориентированной онтологии, и своей проблемной областью. Для данной области характерно описание процессов концептуализации структур формализации. Это позволяет описать проблемную область каждой компоненты в виде онтологической модели ее функционирования посредством понятий, релевантных каждой системе и каждому процессу преобразования данных при формировании смысла сервисов и ИТ-услуг. Онтологические системы в АИС строятся на основе следующих принципов:

- формализации деятельности организации (внешнего мира) в единицах строго определенного формата (переменных, моделях);
- использования для построения ПО автоматизации строго определенного набора базовых терминов (сущностей), на основе которых конструируются все процессы автоматической обработки информации;
- внутренней полноты и логической непротиворечивости сущностей ПО;
- семантической интероперабельности предметно-ориентированных онтологий, специфичных для определенного типа технических компонент АИС и различных гетерогенных АИС.





**Рис. 4** Классификация онтологий АИС

Классифицировать онтологию в АИС можно по различным признакам (основаниям) в зависимости от стоящей цели. Например, по прикладным ПО: область описания содержания ИТ-услуг, область содержания процессов формирования информационных и телекоммуникационных ИТ-услуг, область управления, область оценки угроз и рисков и др. По уровню прикладных ПО, исходя из инфраструктуры АИС и процессов решения задач сервисного обслуживания конечных пользователей, целесообразно различать уровни онтологии, приведенные на рис. 4.

Верхний уровень — онтология предметной области организации — описывает наиболее общие концепты, которые ориентированы как на ПО организации, так и на концепты прикладных областей инфраструктуры АИС. Эта онтология предназначена для совместного использования концептов и отношений между ними как в организации, так и в АИС. Основой этой онтологии является БЗ, содержащая все понятия организации и АИС в виде единого отраслевого стандартного словаря. Например, системная онтология АИС Банка России строится на базе финансовых терминов и базовых терминов, описывающих область функционирования АИС.



Онтологии ИТ-услуг используются как при описании выходной функциональности ИТ-услуг, так и при описании требований к обеспечивающим их приложениям, предназначенным для формирования конкретных ИТ-услуг, поддержки пользователей, а также обеспечения информационной и операционной безопасности процессов деятельности организации. Данные онтологии отражают специфику ресурсно-сервисных процессов формирования ИТ-услуг, отношения между концептами внешней среды и их взаимосвязь с концептами уязвимостей АИС, безопасностью, порождаемыми АИС рисками и др.

Онтологии уровня приложений ориентированы на использование при описании конкретных задач как в составе процессов формирования ИТ-услуг, так и в процессах поддержки пользователей, а также в процессах обеспечения информационной и операционной безопасности. Одновременно онтологии данного уровня содержат специальные термины, присущие описанию процессов в аппаратных и программных средствах систем.

Уровень системных онтологий включает в себя онтологии функциональных, обеспечивающих систем, онтологию средств обеспечения устойчивости и безопасности АИС. Примером такой онтологии может быть БЗ, содержащая понятия функциональных информационных систем, телекоммуникационных систем, систем оперативно-технического управления АИС и т. п.

Одним из основных требований к онтологиям всех уровней является наличие терминологической и понятийной взаимосвязи между ними. Строение и свойства всех процессов непрерывного функционирования любой системы АИС, равно как и процесс решения той или иной прикладной задачи, могут быть эффективно исследованы при помощи словаря терминов, адекватно описывающего характеристики объектов, систем и процессов решения, при наличии точных и однозначных определений терминов и их классификации в виде логической иерархии. Применение онтологии при проектировании АИС позволит:

- создать унифицированное нормативное ядро понятий, которое обеспечит однозначное толкование смысла текстов ПО, ее объектов и процессов и станет семантической основой для порождения новых знаний;
- сформировать множество парадигматических и синтагматических отношений между понятиями ядра на всех уровнях применения онтологических атрибутов (концептов систем, объектов, свойств, ситуаций и сюжетов);
- преодолеть проблемы лингвистической несогласованности при взаимодействии гетерогенных АИС;
- создать благоприятные условия для коллективной работы в рамках единого электронного информационного пространства путем широкого применения онтологических словарей, обеспечивающих ретранслируемость понятий во всех случаях проявления полисинонимии.

Для реализации этих возможностей необходимо создать программные приложения, использующие нормативное ядро и множество отношений. Эти средства

должны обеспечить семантическую интероперабельность созданных или создаваемых на базе библиотеки онтологий систем моделирования свойств ПО и средства автоматической обработки информации, представленной на ЕЯ.

## **Литература**

1. *Гладких Н. Г.* Информационная динамика. — М.: АКСОН, 2004. 160 с.
2. *Смирнов А. В., Пашкин М. П., Шилов Н. Г., Левашова Т. В.* Онтологии в системах искусственного интеллекта: способы построения и организации (часть 1) // *Новости искусственного интеллекта*, 2002. Вып. 49. № 1. С. 3–13.
3. *Дубровин А. Д.* Интеллектуальные информационные системы. — М.: МГУКИ, 2008. 230 с.
4. *Бездудный А. Н., Гаврилова Э. А., Серебряков В. А., Шкатын А. В.* Место онтологий в единой интегрированной системе РАН. [http://www.benran.ru/Magazin/cgi-bin/Sb\\_03/pr03.exe?!15](http://www.benran.ru/Magazin/cgi-bin/Sb_03/pr03.exe?!15).
5. *Клещев А. С., Артемьева И. Л.* Математические модели онтологий предметных областей. Часть 1. Существующие подходы к определению понятия «онтология» // *Научно-техническая информация. Сер. 2: Информационные процессы и системы*. — М.: ВИНТИ, 2001. № 2. С. 20–27.
6. *Артемьева И. Л.* Онтологии предметных областей и их использование при создании программных систем // *Онтологическое моделирование: Тр. симпозиума / Под ред. Л. А. Калиниченко*. — М.: ИПИ РАН, 2008. С. 83–113.
7. *Лапшин В. А.* Онтологии в компьютерных системах. — М.: Научный мир, 2010. 224 с.
8. *Козаловский М. Р., Калиниченко Л. А.* Концептуальное моделирование в технологиях баз данных и онтологические модели // *Онтологическое моделирование: Тр. симпозиума / Под ред. Л. А. Калиниченко*. — М.: ИПИ РАН, 2008. С. 114–148.
9. *Быстров И. И., Веселов В. Н., Радоманов С. И.* Онтология поддержки непрерывности ИТ-услуг в информационно-телекоммуникационных системах // *Системы и средства информатики*, 2014. Т. 24. № 4. С. 221–241.
10. *Белоголов Г. Г., Быстров И. И.* О логической структуре информационных таблиц // *Информационные системы: Сб. научных тр.* — М.: МО СССР, 1966. С. 5–8.
11. *Смирнов С. В.* Онтологии в прикладных интеллектуальных системах: прагматический подход // *КИИ'2004: Тр. 9-й Национальной конф. по искусственному интеллекту с междунар. участием*. — М.: Физматлит, 2004. Т. 3. С. 1059–1067.
12. *Константинова Н. С., Митрофанова О. А.* Онтологии как системы хранения знаний // *Информационно-коммуникационные технологии в образовании*. — СПб: СПбГУ, 2008. 54 с.
13. *Онтологическое моделирование: Тр. симпозиума / Под ред. Л. А. Калиниченко*. — М.: ИПИ РАН, 2008. 303 с.
14. *Лапшин В. А.* Роль онтологии в современной компьютерной науке // *RSDN Magazin*, 2009. № 4.
15. *Gruber T.* Ontology // *The encyclopedia of database systems / Eds. L. Liu, M. Tamer Özsu*. — Berlin: Springer-Verlag, 2008. 4 p.

16. Guarino N. Formal ontology and information system // FOIS'98 Proceedings. — Amsterdam: IOS Press, 1998. P. 3–15.
17. Белоногов Г. Г., Быстров И. И., Козачук М. В., Новоселов А. П., Хорошилов А. А. Автоматический концептуальный анализ текстов // Научно-техническая информация. Сер. 2: Информационные процессы и системы. — М.: ВИНТИ, 2002. № 10. С. 26–32.
18. Востров Г. Н., Межуев В. И. Проблемы построения информационных систем над предметными областями // Искусственный интеллект, 2008. № 4. С. 736–746.

Поступила в редакцию 15.06.15

---

---

## THE APPLICATION BASIS OF ONTOLOGY AND COMPUTATIONAL LINGUISTICS IN THE DESIGN OF ADVANCED AUTOMATED INFORMATION SYSTEMS

*I. I. Bystrov, B. V. Tarasov, A. A. Khoroshilov, and S. I. Radomanov*

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** The application of ontologies and computational linguistics in automated information systems (AIS) is considered. Part 1 describes the methodological basis of complex application of ontologies and computational linguistics in the design of advanced AIS. The description of ontological processes and processes of computational linguistics is provided.

**Keywords:** semantics; ontology; computational linguistics; subject area; declarative and procedural knowledge; automated information system; linguistic information support; logical-semantic modeling; conceptualization

**DOI:** 10.14357/08696527150410

### References

1. Gladkikh, N. G. 2004. *Informatsionnaya dinamika*. [Information dynamics]. Moscow: AKSON. 160 p.
2. Smirnov, A. V., M. P. Pashkin, N. G. Shilov, and T. V. Levashova. 2002. Ontologii v sistemakh iskusstvennogo intellekta: Sposoby postroeniya i organizatsii (Chast' 1) [Ontology in artificial intelligence systems: Construction and organization methods (Pt. 1)]. *Novosti iskusstvennogo intellekta* [Artificial Intelligence News] 49(1):3–13.
3. Dubrovin, A. D. 2008. *Intellektual'nye informatsionnye sistemy* [Intellectual information systems]. Moscow: MGUKI. 230 p.
4. Bezduzhnyy, A. N., E. A. Gavrilova, V. A. Serebryakov, and A. V. Shkatin. Mesto ontologii v edinoj integrirovannoy sisteme RAN [The location of ontologies in a unified

- integrated system RAS]. Available at: <http://www.benran.ru/Magazin/cgi-bin/Sb-03/pr03/xe?!15> (accessed May 07, 2015).
5. Kleshchev, A. S., and I. L. Artem'eva. 2001. Matematicheskie modeli ontologii predmetnykh oblastey. Chast' 1. Sushchestvuyushchie podkhody k opredeleniyu ponyatiya "ontologiya" [Subject area ontology mathematical models. Part 1. Existing approaches to definition of concept "ontology"]. *Nauchno-tehnicheskaya informatsiya, ser. 2* "Informatsionnye protsessy i sistemy" [Scientific-technical information, ser. 2 "Information processes and systems"] 2:20–27.
  6. Artem'eva, I. L. 2008. Ontologii predmetnykh oblastey i ikh ispol'zovanie pri sozdanii programnykh sistem [Subject area ontology and their use when creating program systems]. *Tr. Simpoziuma "Ontologicheskoe Modelirovanie"* [Simposium "Ontological Modeling" Proceedings]. Ed. L. A. Kalinichenko. Moscow: IPI RAN. 83–113.
  7. Lapshin, V. A. 2010. *Ontologii v komp'yuternykh sistemakh* [Ontology in complex systems]. Moscow: Nauchnyy mir. 224 p.
  8. Kogalovskiy, M. R., and L. A. Kalinichenko. 2008. Kontseptual'noe modelirovanie v tekhnologiyakh baz dannykh i ontologicheskie modeli [Conceptual modeling in data base technologies and ontological models]. *Tr. Simpoziuma "Ontologicheskoe modelirovanie"* [Simposium "Ontological Modeling" Proceedings]. Ed. L. A. Kalinichenko. Moscow: IPI RAN. 114–148.
  9. Bystrov, I. I., V. N. Veselov, and S. I. Radomanov. 2014. Ontologiya podderzhki nepreryvnosti IT-uslug v informatsionno-telekommunikatsionnykh sistemakh [Ontology of IT-services continuity support in information-telecommunication systems]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 24(4):221–241.
  10. Belonogov, G. G., and I. I. Bystrov. 1966. O logicheskoy strukture informatsionnykh tablits [About information tables logical structures]. *Informatsionnye sistemy. Sb. nauchnykh tr.* [Information Systems Proceedings]. Moscow: MO SSSR. 5–8.
  11. Smirnov, S. V. 2004. Ontologii v prikladnykh intellektual'nykh sistemakh: Pragmaticheskiy podkhod [Ontology in applied intelligent systems: A pragmatic approach]. *Tr. 9-y Natsional'noy konf. po iskusstvennomu intellektu s mezhdunarodnym uchastiem KII'2004* [9th National Conference on Artificial Intelligence with International Participation Proceedings]. Moscow: Fizmatlit. 3:1059–1067.
  12. Konstantinova, N. S., and O. A. Mitrofanova. 2008. Ontologii kak sistemy khraneniya znaniy [Ontologies as knowledge storage systems] *Informatsionno-kommunikatsionnye tekhnologii v obrazovanii. Elektronnyy resurs* [Information-communication technologies in education. Electronic resource]. SPb: GU. 54 p.
  13. Kalinichenko, L. A., ed. 2008. *Tr. Simpoziuma "Ontologicheskoe Modelirovanie"* [Simposium "Ontological Modeling" Proceedings]. Moscow: IPI RAN. 303 p.
  14. Lapshin, V. A. 2009. Rol' ontologii v sovremennoy komp'yuternoy nauke [Ontology role in contemporary computer science]. *RSDN Magazin* 4.
  15. Gruber, T. 2008. *Ontology. The encyclopedia of database systems*. Eds. L. Liu and M. Tamer Özsu. Berlin: Springer-Verlag. 4 p.
  16. Guarino, N. 1998. Formal ontology and information system. *FOIS'98 Proceedings*. Amsterdam: IOS Press. 3–15.
  17. Belonogov, G. G., I. I. Bystrov, M. V. Kozachuk, A. P. Novoselov, and A. A. Khoroshilov. 2002. Avtomaticheskii kontseptual'nyy analiz tekstov [Texts automatic conceptual analysis]. *Nauchno-tehnicheskaya informatsiya. Ser. 2* [Scientific-Technical Information collection. Ser. 2]. Moscow: VINITI. 10:26–32.

18. Vostrov, G.N., and V. I. Mezhuiev. 2008. Problemy postroeniya informatsionnykh sistem nad predmetnymi oblastyami [Problems of an information systems construction on the subject area]. *Artificial Intelligence* 1(4):736-746. (In Russian.)

*Received June 15, 2015*

### **Contributors**

**Bystrov Igor I.** (b. 1931) — Doctor of Science in technology, professor, Head of Department, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; bystrov@ipiran.ru

**Tarasov Boris V.** (b. 1946) — Doctor of Science in technology, professor, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; tarasov@npcmodul.ru

**Khoroshilov Alexandr A.** (b. 1952) — Doctor of Science in technology, professor, leading scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; khoroshilov@mail.ru

**Radomanov Sergei I.** (b. 1953) — scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; radomanov@list.ru

## НЕКОТОРЫЕ ВОПРОСЫ ОЦЕНКИ НАДЕЖНОСТИ КОРПОРАТИВНОЙ ИНФОРМАЦИОННОЙ СЕТИ С УЧЕТОМ УПРАВЛЕНИЯ РЕЗЕРВИРОВАНИЕМ ЕЕ ЭЛЕМЕНТОВ

*В. В. Бородин<sup>1</sup>, Е. А. Растрелин<sup>2</sup>*

**Аннотация:** Рассмотрена модель надежности территориально распределенной корпоративной сети, в которой реализовано групповое резервирование однотипных элементов. Показана необходимость учета в модели надежности временного фактора, связанного с доставкой резервного оборудования к месту отказа, и создания системы управления резервным ресурсом, обеспечивающей сбор данных о надежности элементов сети, управление оперативным резервом запасных частей и формирование требований на поставку резервного оборудования с центрального склада.

**Ключевые слова:** надежность сети; информационная корпоративная сеть; территориально распределенная сеть; элементы сети; унификация; резервирование; модель надежности; коэффициент готовности; время восстановления; управление ресурсом

**DOI:** 10.14357/08696527150411

### 1 Основные понятия

Возросшая сложность вычислительных сетей, увеличение номенклатуры оборудования и вариантов его использования, а также способов организации эксплуатации автоматизированных систем обуславливают необходимость разработки новых методик оценки надежности распределенных компьютерных сетей. В то же время, если для сетей связи и передачи данных стандартом установлены некоторые общие правила, расчет надежности информационных сетей оставляет возможность поиска оптимального соотношения между бесперебойным функционированием и издержками на создание и эксплуатацию.

Построение систем анализа и оценки надежности распределенных компьютерных сетей на основе графов с последующей проверкой свойств этого графа на базе использования различных методов является преобладающим направлением в данной области. Большинство таких систем не применяются на практике, в отличие от основанных на сборе статистики методов прогнозирования, используемых пакетами программного обеспечения Hewlett-Packard или IBM, что делает разработку практически значимой методики весьма актуальной задачей.

---

<sup>1</sup>Московский авиационный институт (национальный исследовательский университет), doc\_bor1@mail.ru

<sup>2</sup>ЗАО «РАМЭК-ВС», erastrelin@mail.ru

Большинство корпоративных информационных сетей обладает общими ключевыми особенностями, важными для проектирования с точки зрения обеспечения надежности:

- сеть является территориально распределенной;
- каждый узел сети является многофункциональным, решает задачи приема, передачи, обработки и отображения информации;
- узлы содержат большое количество подсистем, номенклатура электронного оборудования содержит порядка сотни позиций;
- оборудование разных узлов отличается высоким уровнем совместимости (унификации);
- сеть должна функционировать в течение длительного времени;
- сеть должна обеспечивать высокий уровень готовности в течение всего времени эксплуатации.

Расчеты показывают [1], что механизмы обеспечения надежности, связанные с индивидуальным резервированием оборудования каждого узла, экономически нецелесообразны. В этой связи возникает задача выработки методов оптимального резервирования для обеспечения заданного уровня надежности при минимальных дополнительных затратах.

Унификация оборудования узлов позволяет обеспечивать так называемое групповое резервирование, при котором общий запас резервных элементов может использоваться для нескольких узлов [2, 3]. Однако их удаленность друг от друга в сочетании с большой номенклатурой оборудования приводит к необходимости организации централизованного хранения и оперативной доставки резервных элементов на узлы. При этом необходимо дополнительно учитывать временной фактор, связанный с доставкой резервного оборудования, так как несвоевременная поставка запасных частей может привести к отказу системы.

Для обеспечения надежности информационной сети необходимо наличие отдельной системы управления резервным ресурсом, которая должна выполнять следующие функции:

- хранение оперативного резерва запасных частей для каждого элемента каждого узла с обеспечением возможности замены отказавших элементов на резервные за минимальное время;
- формирование заявок на пополнение оперативного резерва запасных частей из централизованного склада;
- формирование запросов на поставку резервных элементов на централизованный склад по мере расходования запасов на нем;
- обеспечение доставки заказанных элементов с централизованного склада на соответствующие узлы.

Статья посвящена разработке модели надежности корпоративной информационной сети с управлением резервным ресурсом.



## 2 Модель надежности сети

Сеть содержит  $N$  узлов, на каждом  $n$ -м узле имеется  $M_n$  элементов, при этом каждый элемент имеет некоторый индивидуальный резерв.

При отказе основного элемента выполняется его замена на резервный (если имеется) и одновременно формируется запрос на доставку соответствующего резервного элемента на центральный склад. Через некоторое время, определяемое параметрами склада, формируется поставка требуемого оборудования на соответствующий узел. Центральный склад в целях своего пополнения регулярно формирует запросы на поставку нового оборудования и направляет эти запросы соответствующим поставщикам или изготовителям. Заказанное оборудование через определенное время поставляется на центральный склад.

Запросы на поставку нового оборудования возникают в случайные моменты времени, определяемые отказами элементов сети. Время удовлетворения запроса в общем случае является случайным, определяется наличием соответствующих элементов на складе и временем доставки элементов со склада на узел. Наличие случайных компонент может приводить к ситуациям несвоевременного выполнения запросов на поставку требуемого резервного элемента и в конечном итоге с определенной вероятностью к отказу сети в целом. Восстановление будет выполнено после того, как этот резервный элемент поступит на узел и будет обеспечена его установка вместо соответствующего отказавшего элемента.

Показателем надежности сети является коэффициент готовности  $K_c$ . Для расчета показателя надежности введем понятие резервной группы. Резервная группа — это сочетание основного и определенного числа резервных элементов в каждом узле. Резервные группы в рамках одного узла соединены последовательно, все узлы сети также соединены последовательно. В этих условиях коэффициент готовности сети равен произведению коэффициентов готовности  $K$  всех резервных групп [4]:

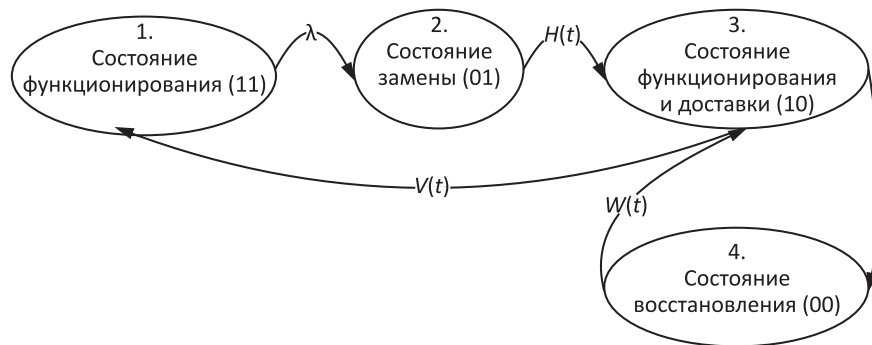
$$K_c = \prod_{n=1, i=1}^{n=N, i=M_n} K_{n,i}.$$

Таким образом, для оценки надежности сети необходимо определить коэффициент готовности произвольной резервной группы.

## 3 Модель надежности резервной группы

Рассмотрим дублированную систему, в которой основной и резервный элементы имеют экспоненциальную функцию надежности с параметром  $\Lambda$ . Замена отказавшего элемента на работоспособный резервный осуществляется через случайное время  $H(t)$  со средним значением  $\theta_3$ .





Граф переходов системы

После замены система продолжает функционировать и одновременно выполняется доставка нового элемента (или ремонт отказавшего элемента). Доставка продолжается в течение случайного времени, имеющего функцию распределения  $V(t)$  и среднее значение  $\theta_D$ . Отказ элемента переводит систему в состояние ожидания восстановления, в котором она находится в течение случайного времени  $W(t)$  со средним значением  $\theta_B$ .

Граф переходов системы приведен на рисунке. Каждое состояние дополнительно имеет код вида  $(ij)$ , где первое значение определяет состояние основного элемента (1 — исправное; 0 — отказ), а второе — состояние резервного элемента.

При составлении графа переходов учтено, что при отказе двух элементов сначала восстанавливается один (тот, который отказал первым), а затем другой. По этой причине длительность состояния  $C_4$  в среднем равна остаточному времени восстановления отказавшего элемента.

Поскольку переходы между состояниями в общем случае осуществляются по произвольным законам, моделью системы будет полумарковский процесс. Коэффициент готовности  $K_c$  будет равен сумме вероятностей нахождения в состояниях работоспособности, а именно:

$$K_c = P_1 + P_3,$$

где  $P_k$  — стационарная вероятность нахождения в состоянии  $C_k$ .

Обозначим через  $T_k$  среднее время нахождения в состоянии  $C_k$ . Тогда искомые вероятности определятся выражением [4, 5]:

$$P_k = \frac{T_k q_k}{\sum_{l=1}^4 T_l q_l},$$

где  $q_k$  — стационарные вероятности вложенной цепи Маркова. Эти вероятности являются решением следующей системы уравнений:

$$q_k = \sum_l q_l Q_{lk}; \quad \sum_l q_l = 1. \quad (1)$$

В свою очередь, переходные вероятности  $Q_{kl}$  согласно графу переходов равны:

$$\left. \begin{aligned} Q_{12} = Q_{23} = Q_{43} = 1; \\ Q_{31} = q = \int_0^{\infty} e^{-\Lambda t} dV(t) = V^*(\Lambda); \\ Q_{34} = 1 - V^*(\Lambda), \end{aligned} \right\} \quad (2)$$

где  $V^*(\Lambda)$  — преобразование Лапласа функции  $V(t)$ .

В свою очередь, средние времена нахождения в состояниях  $C_k$  равны:

$$\begin{aligned} T_1 &= \frac{1}{\Lambda} = T_0; \\ T_2 &= \theta_3; \\ T_4 &= \theta_B; \\ T_3 &= \int_0^{\infty} e^{-\Lambda t} (1 - V(t)) dt = \frac{1 - V^*(\Lambda)}{\Lambda}, \end{aligned}$$

где  $T_0$  — среднее время до отказа элемента.

Решением системы уравнений (1) с учетом (2) будет выражение:

$$K = \frac{1}{1 + \Lambda (Q_{31}\theta_3 + (1 - Q_{31})\theta_B)}. \quad (3)$$

Полученное выражение может быть упрощено при условии, что среднее время ремонта существенно меньше среднего времени до отказа. Это условие на практике обычно выполняется, и из (3) следует:

$$Q_{31} = \int_0^{\infty} e^{-\Lambda t} dV(t) = V^*(\Lambda) = 1 - \Lambda\theta_D.$$

Окончательно выражение для коэффициента готовности можно записать в виде:

$$K = \frac{1}{1 + \Lambda\theta_3 + \Lambda^2\theta_D\theta_B}.$$

Среднее время восстановления  $\theta_B$  есть остаточное время до окончания ремонта. Для получения нижней оценки коэффициента готовности в качестве времени

восстановления может быть использовано среднее время доставки  $\theta_D$ . С учетом этого выражение для коэффициента готовности резервной группы можно записать в виде:

$$K = \frac{1}{1 + \Lambda\theta_3 + \Lambda^2\theta_D^2}.$$

В предположении, что характеристики надежности всех элементов сети примерно одинаковы, из последнего выражения следует, что среднее время доставки должно удовлетворять следующему соотношению:

$$\theta_D \leq T \sqrt{\frac{1 - K_{\text{доп}} - K_{\text{доп}}\theta_3/T}{K_{\text{доп}}}},$$

где  $K_{\text{доп}}$  — допустимое значение коэффициента готовности сети;  $T$  — средняя наработка на отказ всех элементов сети.

Из полученных результатов следует, что для достаточно сложных систем со средним временем на отказ  $T = 2000\text{--}4000$  ч среднее время доставки не должно превышать 500 ч.

На основании представленной методики могут быть получены значения коэффициента готовности резервной группы, содержащей более одного резервного элемента. В частности, при двух резервных элементах значение коэффициента готовности равно:

$$K = \frac{1}{1 + \Lambda(\theta_3q + \theta_D(1 - q)^2)/(q^2 + 1 - q)}.$$

Как показал анализ результатов расчета, время доставки резервного оборудования в этом случае увеличивается до нескольких тысяч часов.

## 4 Выводы

Разработана аналитическая модель для определения коэффициента готовности корпоративной информационной сети с управлением резервным оборудованием, которая учитывает характеристики надежности элементов сети и параметры системы управления резервным оборудованием. Полученные выражения позволяют оценивать время доставки резервного оборудования с центрального склада для замены отказавших элементов узлов сети.

## Литература

1. *Анатьев А. Н.* Разработка и исследование методов расчета надежности корпоративных сетей региональных операторов связи // *Электросвязь*, 2002. № 10. С. 30–33.

2. Зацаринный А. А., Ггаранин А. И., Козлов С. В., Кондрашев В. А. Особенности расчета комплектов ЗИП в автоматизированных информационных системах в защищенном исполнении // Системы и средства информатики, 2013. Т. 23. № 1. С. 113–131.
3. Зацаринный А. А., Козлов С. В., Ггаранин А. И. Особенности формирования номенклатуры и количества компонентов ЗИП в автоматизированных информационных системах в защищенном исполнении // Системы и средства информатики, 2014. Т. 24. № 3. С. 144–154.
4. Барзилович Е. Ю., Беляев Ю. К., Каштанов В. А. и др. Вопросы математической теории надежности / Под ред. Б. В. Гнеденко. — М.: Радио и связь, 1983. 376 с.
5. Байхельт Ф., Франкен П. Надежность и техническое обслуживание. Математический подход / Пер с нем. — М.: Радио и связь, 1988. 434 с. (*Baichelt F., Franken P. Zuverlässigkeit und Instandhaltung. Mathematische Methoden.* — Berlin: VEB Verlag Technik, 1983. 434 p.)

*Поступила в редакцию 10.09.15*

---

## SOME ASPECTS OF RELIABILITY ASSESSMENT OF THE CORPORATIVE DATA NETWORK WITH CONSIDERATION OF REDUNDANCY MANAGEMENT

*V. V. Borodin<sup>1</sup> and E. A. Rastrelin<sup>2</sup>*

<sup>1</sup>Department of Infocommunications, Moscow Aviation Institute (National Research University), MAI, 4 Volokolamskoe Shosse, Moscow 125993, A-80, GSP-3, Russian Federation

<sup>2</sup>LLC “RAMEC-VS,” 2 Volgogradskiy Prospect, Moscow 109316, Russian Federation

**Abstract:** The paper considers the model of reliability of a spatially distributed corporative network with implementation of the batch redundancy of typical elements. The necessity of time factor, concerning the delivery of spare equipment, should be considered during the reliability modeling and, also, as a stockpiles management system, providing data collection about current condition and reliability of the network elements, operational stock management, and forming of the requests for spare parts from the central storage facility.

**Keywords:** network reliability; corporate computer network; wide area network; network elements; unification; reliability model; availability rate; restoration time; storage control

**DOI:** 10.14357/08696527150411

## References

1. Anan'ev, A. N. 2002. Razrabotka i issledovanie metodov rascheta nadezhnosti korporativnykh setey regionalnykh operatorov svyazi [Local service providers network reliability analysis research and development methods]. *Elekrosvyaz'* 10:30–33.
2. Zatsarinny, A. A., A. I. Garanin, S. V. Kozlov, and V. A. Kondrashev. 2013. Osobennosti rascheta komplektov ZIP v avtomatizirovannykh informatsionnykh sistemakh v zashchishchennom ispolnenii [About some particularities of spare parts sets calculation for secured information systems]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 23(1):113–131.
3. Zatsarinny, A. A., A. I. Garanin, and S. V. Kozlov. 2014. Osobennosti formirovaniya nomenklatury i kolichestva komponentov ZIP v avtomatizirovannykh informatsionnykh sistemakh v zashchishchennom ispolnenii [About some particularities of spare parts sets spectrum forming and quantity calculation for secured information systems]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 24(3):144–154.
4. Barzilovich, E. Yu., Yu. E. Belyaev, V. A. Kashtanov, I. N. Kovalenko, A. D. Solovyev, and I. A. Ushakov. 1983. *Voprosy matematicheskoy teorii nadezhnosti* [Issues of the mathematical theory of reliability]. Ed. B. V. Gnedenko. Moscow: Radio and Communication. 376 p. (In Russian.)
5. Baichelt, F., and P. Franken. 1983. *Zuverlassigkeit und Instandhaltung. Mathematische Methoden*. Berlin: VEB Verlag Technik. 434 p.

Received September 10, 2015

## Contributors

**Borodin Vyacheslav V.** (b. 1957) — Candidate of Science (PhD) in technology, Associate Professor, Department of Infocommunications, Moscow Aviation Institute (National Research University), MAI, 4 Volokolamskoe Shosse, Moscow 125993, A-80, GSP-3, Russian Federation; doc\_bor1@mail.ru

**Rastrelin Evgeniy A.** (b. 1972) — Deputy Director, LLC “RAMEC-VS,” 2 Volgogradskiy Prospect, Moscow 109316, Russian Federation; erastrelin@mail.ru

## ОЦЕНКА СТОЙКОСТИ КОДОВОГО ЗАШУМЛЕНИЯ В ЗАДАЧЕ РАСПРЕДЕЛЕННОГО ХРАНЕНИЯ ДАННЫХ

*Ю. В. Косолапов<sup>1</sup>, А. В. Поздняков<sup>2</sup>*

**Аннотация:** Рассматривается модель распределенного хранения данных в нескольких хранилищах в условиях, когда каждый из владельцев хранилищ может оказаться недобросовестным. Для защиты данных от несанкционированного ознакомления применяется метод кодового зашумления с последующим распределением частей кодовых слов по соответствующим хранилищам. Как правило, в работах исследуется теоретико-информационная стойкость этого метода в условиях защиты случайных и равновероятных данных. В прикладных задачах часто защищаемые данные не удовлетворяют этим условиям. В настоящей работе решается задача разработки методики для экспериментальной оценки стойкости кодового зашумления в случае, когда данные представляют собой фотографические изображения лиц и при этом наблюдатель обладает ограниченными возможностями анализа доступных ему данных. Приведены результаты экспериментов в случае использования бинарных кодов Рида–Маллера длины 64.

**Ключевые слова:** кодовое зашумление; стойкость; распределенное хранилище

**DOI:** 10.14357/08696527150412

### 1 Введение

Обеспечение конфиденциальности является одной из основных задач в области защиты информации. Особенно остро эта проблема возникает в случае, когда владелец информации не является владельцем хранилищ, в которых она записана. В этом случае владелец хранилища рассматривается как потенциальный наблюдатель, который может нарушить конфиденциальность сохраненной информации. Одним из некриптографических способов защищенной от наблюдателя хранения является способ распределенного хранения данных с предварительным преобразованием их по методу кодового зашумления [1]. Однако такая схема хранения исследована, как правило, в случае, когда, с одной стороны, защищаемые данные являются случайными и равновероятными, а с другой стороны, наблюдатель обладает неограниченными вычислительными способностями [1–3]. Заметим, что в прикладных задачах условие случайности и равновероятности данных не всегда выполняется. Например, биометрические шаблоны не обладают этими свойствами [4, 5]; при этом с точки зрения защищен-

<sup>1</sup>Южный федеральный университет, itaim@mail.ru

<sup>2</sup>Южный федеральный университет, al.vldr@mail.ru

ности требуется, чтобы доступных наблюдателю данных было недостаточно для восстановления биометрического шаблона с целью прохождения аутентификации [6], хотя при этом в теоретико-информационном смысле наблюдатель может обладать ненулевой информацией о шаблоне. В таких случаях представляется, что критерий защищенности конфиденциальности данных зависит, во-первых, от самих данных, а во-вторых, от возможностей наблюдателя.

В настоящей работе в качестве защищаемых данных рассматриваются цифровые фотографические изображения лиц, которые могут применяться, например, в системах биометрической аутентификации [7]. В качестве критерия защищенности конфиденциальности рассматривается возможность определить наличие лица на изображении с помощью исключительно автоматизированной обработки. В рамках этих условий ставится задача подбора подходящей меры для оценки стойкости кодового зашумления и разработки методики экспериментальной оценки.

Кроме введения работа включает три раздела. Во втором разделе в удобном виде приводятся необходимые сведения о методе кодового зашумления и его свойствах; строятся модель случайного декодирования частично известных кодовых слов и модель распределенного хранения данных, в которой для защиты конфиденциальности применяется кодовое зашумление. Третий раздел содержит описание методики экспериментальной оценки стойкости кодового зашумления; вводится и исследуется мера стойкости, зависящая от применяемой метрики на множестве информационных сообщений. Там же доказан ряд утверждений, упрощающих вычисление введенной меры стойкости при использовании метрики Хэмминга и исследованной в работе новой байтовой метрики. Результаты экспериментов изложены в четвертом разделе, где для реализации метода кодового зашумления выбраны двоичные коды Рида–Маллера длины 64.

## 2 Модель распределенного хранения данных

### 2.1 Схема частичного наблюдения и факторные коды

Введем необходимые для дальнейшего обозначения. Пусть  $\mathbb{F} = \mathbb{F}_2$  — поле Галуа мощности 2;  $\underline{n} = \{1; \dots; n\}$ ;  $\chi(D) = \{i : \exists(x_1, \dots, x_n) \in D, x_i \neq 0\}$  — носитель множества  $D(\subseteq \mathbb{F}^n)$ . Для  $S \subseteq \underline{n}$  будем использовать обозначения:  $\mathbb{F}^S = \{\mathbf{y} \in \mathbb{F}^n : \chi(\mathbf{y}) \subseteq S\}$  — координатное подпространство;  $\pi_S : \mathbb{F}^n \rightarrow \mathbb{F}^S$  — оператор линейной проекции пространства  $\mathbb{F}^n$  на координатное подпространство  $\mathbb{F}^S$  параллельно координатному подпространству  $\mathbb{F}^{\overline{S}}$ ,  $\overline{S} = \underline{n} \setminus S$ . Если  $A$  — некоторая  $(k \times n)$ -матрица, то символом  $A_S$  обозначим  $(k \times n)$ -матрицу, полученную из матрицы  $A$  действием оператора  $\pi_S$  на каждую строку. Далее при использовании оператора  $\pi_S$  размерность проецируемого пространства  $\mathbb{F}^n$  будет ясна из контекста. Транспонированную матрицу  $A$  будем обозначать, как обычно,  $A^\top$ . Для случайного и равновероятного выбора элемента  $\mathbf{r}$  из множества  $\mathcal{R}$  будет использоваться обозначение  $\mathbf{r} \stackrel{\$}{\leftarrow} \mathcal{R}$ .

Опишем рассмотренную в [8] схему частичного наблюдения в канале. Пусть по *главному (легальному) каналу* от отправителя получателю передаются векторы длины  $n$  над полем  $\mathbb{F}$ , а наблюдателю для подслушивания доступны координаты этих векторов с номерами из множества наблюдаемых координат  $\tau$ ,  $|\tau| = \mu \leq n$ . Отметим, что частичное наблюдение данных в этом случае можно рассматривать как передачу данных от отправителя наблюдателю по *каналу наблюдения*, в котором стираются  $n - \mu$  координат с номерами из множества  $\bar{\tau}$ . Другими словами, если по главному каналу передается вектор  $\mathbf{z}$ , то наблюдатель получает вектор  $\mathbf{z}_\tau$ , при этом множество  $\tau$  наблюдателю известно. Для защиты от частичного наблюдения в [8] предложен метод кодового зашумления (метод факторного кодирования), состоящий в следующем. Пусть  $C$  — линейный  $[n, n - k]$ -код с порождающей  $(n - k \times n)$ -матрицей  $G$  и проверочной  $(k \times n)$ -матрицей  $H$ ,  $G^*$  — матрица линейного оператора  $\mathcal{G}^* : \mathbb{F}^k \rightarrow \mathbb{F}^n$ , причем  $C \cap \text{Im}(\mathcal{G}^*) = \mathbf{0}$ . Кодовое зашумление информационного блока  $\mathbf{x} (\in \mathbb{F}^k)$  выполняется по правилу

$$\mathbf{x}G^* + \mathbf{r}G = \mathbf{z} \in \mathbb{F}^n, \quad (1)$$

где  $\mathbf{r} \stackrel{\S}{\leftarrow} \mathbb{F}^{n-k}$ . Отметим, что правило (1) задает отображение

$$FC : \mathbb{F}^k \times \mathbb{F}^{n-k} \rightarrow \mathbb{F}^n,$$

которое каждой паре  $(\mathbf{x}, \mathbf{r}) \in \mathbb{F}^k \times \mathbb{F}^{n-k}$  однозначно ставит в соответствие вектор  $\mathbf{z} \in \mathbb{F}^n$ . Информационный вектор  $\mathbf{x}$  может быть закодирован любым случайно и равновероятно выбранным вектором  $\mathbf{z}$  из фактор-класса  $C_{\mathbf{x}} \in \mathbb{F}^n / C$ . Кодирование вектора  $\mathbf{x}$  по правилу (1) будем обозначать  $FC(\mathbf{x})$ . Извлечение (декодирование) информационного вектора  $\mathbf{x}$  из  $\mathbf{z}$  осуществляется умножением  $\mathbf{z}$  на специально подобранную проверочную матрицу  $H$  кода  $C$  [3, 8]:

$$\mathbf{x} = \mathbf{z}H^\top, \quad (2)$$

что определяет отображение  $FC^{-1} : \mathbb{F}^n \rightarrow \mathbb{F}^k$ ,  $\mathbf{x} = FC^{-1}(\mathbf{z})$ . Под неопределенностью  $\Delta_\tau$  будем понимать, в соответствии с [8], энтропию наблюдателя при подслушивании координат из множества  $\tau$ . Как показано в [8], в случае, когда кодируемые данные случайные и равновероятные, выполняется равенство

$$\Delta_\tau = \text{rank}(H_{\bar{\tau}}). \quad (3)$$

При  $\Delta_\tau = k$  говорят, что обеспечивается *совершенная защита*. В частности, в [8] показано, что  $\text{rank}(H_{\bar{\tau}}) = k$  при

$$|\tau| < d(C^\perp), \quad (4)$$

где  $C^\perp$  — код, ортогональный к коду  $C$ . Для факторного кода с правилами кодирования и декодирования (1) и (2) соответственно будем далее использовать обозначение  $FC(C, k)$ .



## 2.2 Случайное декодирование факторных кодов

Пусть  $FC(C, k)$  — факторный код,  $\tau$  — множество наблюдаемых координат в кодовом векторе  $\mathbf{z} = FC(\mathbf{x})$ , при этом информационный вектор  $\mathbf{x}$  наблюдателю не известен. Наблюдатель не может применить к вектору  $\mathbf{z}_\tau$  правило декодирования (2), так как координаты из множества  $\bar{\tau}$  ему не известны. Однако информационное сообщение  $\mathbf{x}$  принадлежит множеству претендентов  $\mathcal{X}_{FC}(\mathbf{z}_\tau)$ , которое может построить наблюдатель по вектору  $\mathbf{z}_\tau$ , множеству  $\tau$  и факторному коду  $FC(C, k)$ :

$$\mathcal{X}_{FC}(\mathbf{z}_\tau) := \{FC^{-1}(\mathbf{y}) : \mathbf{y} \in \pi_\tau^{-1}(\mathbf{z}_\tau)\} = \left\{ \mathbf{z}_\tau H_\tau^\top + \mathcal{L}\left(H_{\bar{\tau}}^\top\right) \right\},$$

где  $\mathcal{L}(H_{\bar{\tau}}^\top)$  — линейная оболочка, натянутая на строки матрицы  $H_{\bar{\tau}}^\top$ . Очевидно, что

$$\mathcal{X}_{FC}(\mathbf{z}_\tau) = \mathbf{x}' + \mathcal{L}\left(H_{\bar{\tau}}^\top\right) \quad \forall \mathbf{x}' \in \mathcal{X}_{FC}(\mathbf{z}_\tau). \quad (5)$$

Из (5) следует, что если  $\mathbf{z} = FC(\mathbf{x})$ ,  $\mathbf{y} = FC(\mathbf{x})$  и  $\mathbf{z} \neq \mathbf{y}$ , то  $\mathcal{X}_{FC}(\mathbf{z}_\tau) = \mathcal{X}_{FC}(\mathbf{y}_\tau)$ . Таким образом, для фиксированного  $\mathbf{x}$  множество претендентов не зависит от значений наблюдаемых координат, а зависит от множества  $\tau$ . Отметим, что этот факт также был доказан в [2] в контексте многократного частичного наблюдения. В дальнейшем множество претендентов при заданном  $\tau$  и заданном информационном векторе  $\mathbf{x}$  будем обозначать  $\mathcal{X}_{FC,\tau}(\mathbf{x})$ . Далее будем полагать, что к вектору  $\mathbf{z}_\tau$  наблюдатель применяет *случайное декодирование*, заключающееся в случайном и равновероятном выборе вектора  $\mathbf{u}$  из множества претендентов  $\mathcal{X}_{FC}(\mathbf{z}_\tau) = \mathcal{X}_{FC,\tau}(\mathbf{x})$ :

$$\mathbf{u} \stackrel{\S}{\leftarrow} \mathcal{X}_{FC,\tau}(\mathbf{x}). \quad (6)$$

Так как в множестве претендентов только один вектор верный, а остальные ошибочные, то подпространство  $\mathcal{L}(H_{\bar{\tau}}^\top) = \mathcal{X}_{FC}(\mathbf{0}_\tau)$  назовем *кодом ошибок* и обозначим  $V_{FC,\tau}$ . Заметим, что  $|V_{FC,\tau}| = 2^{\Delta_\tau}$ , как следует из (3).

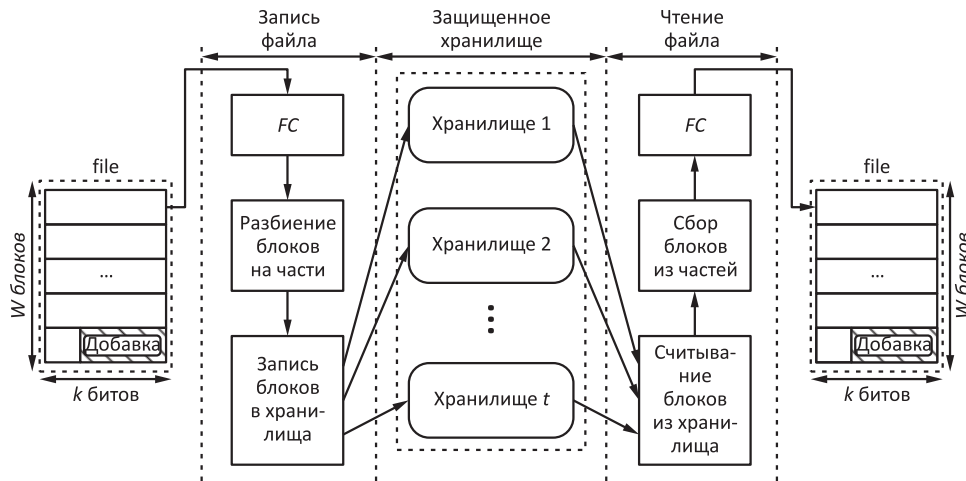
## 2.3 Использование факторных кодов при распределенном хранении файлов

Рассмотрим *защищенное хранилище*, построенное на основе  $t (\in \mathbb{N})$  хранилищ (серверов), владельцы которых не могут образовывать коалиции. Опишем процесс записи одного файла  $\text{file}$  в защищенное хранилище. Защищаемый файл  $\text{file}$  представляется в виде последовательности  $n$ -битных блоков, при этом для дополнения файла до кратного  $n$  числа битов используется, например, способ формирования добавки, описанный в [9]. Каждый блок разбивается на  $t$  частей, которые затем сохраняются в соответствующих хранилищах. Разбиение множества номеров координат  $\underline{n}$  осуществляется так, что в  $i$ -е хранилище записываются

координаты с номерами из множества  $\tau_i \subseteq \underline{n}$ ,  $i = 1, \dots, t$ , причем  $\bigcup_{i=1}^t \tau_i = \underline{n}$  и  $\tau_i \cap \tau_j = \emptyset$  для  $i \neq j$ . Для чтения файла file из защищенного хранилища пользователь получает из  $t$  хранилищ все части каждого блока, собирает сформированные блоки и отбрасывает добавленные биты.

Пусть владельцы хранилищ выступают в роли пассивных наблюдателей, т. е. по доступным в соответствующих хранилищах данным пытаются получить информацию о содержимом файла file. В этом случае, как предложено в [1], данные могут быть защищены с использованием подходящего факторного кода  $FC(C, k)$ . Записываемый в защищенное хранилище файл file представляется в виде последовательности  $k$ -битных блоков, каждый из которых затем преобразуется в  $n$ -битные кодовые слова ( $n > k$ ) с помощью кодера  $FC$  факторного кода; кодовые слова далее распределяются по  $t$  хранилищам в соответствии с описанной выше схемой. Считывание файла из защищенного хранилища также выполняется по описанной ранее схеме, после чего каждый  $n$ -битный блок преобразуется в соответствующий информационный  $k$ -битный блок с помощью отображения  $FC^{-1}$ . Операции чтения и записи в защищенном факторным кодом хранилище схематично изображены на рис. 1.

Если у наблюдателей (владельцев хранилищ) нет априорной информации о записываемом файле file и они обладают неограниченными вычислительными возможностями, то, как следует из (4), совершенная защита достигается при условии  $|\tau_i| < d(C^\perp)$  для всех  $i \in \underline{t}$ . В этом случае минимальное число хранилищ определяется из неравенства  $t \geq \lceil n / (d(C^\perp) - 1) \rceil$ . В случае, когда у наблюдателей имеется априорная информация о файле file и их вычислитель-



**Рис. 1** Чтение/запись файла в защищенном хранилище с использованием факторного кодирования

ные способности ограничены и заранее известны, применение условия (4) для выбора множеств  $\tau_i, i = 1, \dots, t$ , представляется необоснованным. В следующем разделе решается задача экспериментальной оценки стойкости факторного кода  $FC(C, k)$  на множестве  $\tau$  в рамках отмеченных ограничений на случайность данных и вычислительные способности наблюдателей.

### 3 Оценка практической стойкости

#### 3.1 Методика оценки практической стойкости

Рассмотрим множество  $\mathcal{D}$  — некоторое абстрактное множество файлов, содержащих однотипные данные, например цифровые фотографические изображения лиц. Предполагается, что файлы из  $\mathcal{D}$  записываются в защищенное факторным кодом  $FC(C, k)$  хранилище, построенное на основе  $t$  хранилищ (см. рис. 1), при этом в  $i$ -е хранилище записываются значения координат с номерами из множества  $\tau_i, i = 1, \dots, t$ . Каждый из наблюдателей (владельцев хранилищ) применяет случайное декодирование по правилу (6) к частично доступным данным. Будем полагать, что пригодность декодированных данных определяется наблюдателями с помощью некоторого известного алгоритма  $\mathcal{A}$ , результатом работы которого могут быть два значения: 0, если данные не пригодны для использования, и 1, если пригодны. Пусть  $D = \{\text{file}_1; \dots; \text{file}_L\}$  — набор файлов из  $\mathcal{D}$ , записанных в защищенное хранилище,  $D'_{\tau_i}$  — набор из  $L$  соответствующих файлов, полученных  $i$ -м наблюдателем после случайного декодирования частично доступных ему данных. Будем говорить, что пара  $(FC(C, k), \tau_i)$  обладает  $\mathcal{A}$ -стойкостью, если  $\mathcal{A}(D'_{\tau_i}) = 0, i = 1, \dots, t$ . Однако использование алгоритма  $\mathcal{A}$  и выборки  $D$  для оценки применимости произвольного множества  $\tau$  может быть неприемлемым ввиду, например, чрезмерной длительности такого эксперимента. Предварительное вычисление по всем возможным подмножествам  $\tau$  с использованием алгоритма  $\mathcal{A}$  и выборки  $D$  также представляется неприемлемым при больших  $n$ , так как потребуются  $2^n$  запусков алгоритма  $\mathcal{A}$  для выборки  $D$ . Поэтому представляет интерес поиск альтернативного способа оценки стойкости факторного кода  $FC(C, k)$  на множестве  $\tau$ , опирающегося только на параметры кода  $FC(C, k)$  и на множество  $\tau$ .

Пусть  $2^{\underline{n}}$  — множество всех подмножеств множества  $\underline{n}$ , из которого случайным и равновероятным образом выбрано подмножество  $T \subset 2^{\underline{n}}, |T| \ll 2^n$ . Построим алгоритм  $\mathcal{B}$ , на вход которого подается пара  $(FC(C, k), \tau), \tau \in \underline{n}$ , а на выходе получается одно из двух значений: 1 или 0. При этом алгоритм  $\mathcal{B}$  будем строить так, чтобы он никогда не возвращал значение 1 для множества  $\tau \in T$  и набора  $D$ , если  $\mathcal{A}(D'_{\tau}) = 1$ . Другими словами, если  $\mathcal{B}(FC(C, k), \tau) = 1$  на наборе  $D$ , то  $\mathcal{A}(D'_{\tau}) = 0$ . В случае  $\mathcal{B}(FC(C, k), \tau) = 1$  для  $\tau \in \underline{n}$  будем говорить, что пара  $(FC(C, k), \tau)$  является  $\mathcal{B}$ -стойкой. Таким образом, при построении алгоритма  $\mathcal{B}$  будем требовать, чтобы для пары  $(FC(C, k), \tau), \tau \in T$ , на множестве файлов  $D$  выполнялась импликация:

$$\text{пара } (FC(C, k), \tau) \text{ — } \mathcal{B}\text{-стойкая} \implies \text{пара } (FC(C, k), \tau) \text{ — } \mathcal{A}\text{-стойкая.} \quad (7)$$

Для построения алгоритма  $\mathcal{B}$  необходимо ввести меру. Примером такой меры может быть величина  $\Delta_\tau$ , вычисленная для  $\tau$  по формуле (3). Тогда соответствующий алгоритм  $\mathcal{B} = \mathcal{B}_\Delta(FC, \tau)$  определения стойкости пары  $(FC(C, k), \tau)$  примет вид:

$$\mathcal{B}_\Delta(FC, \tau) = \begin{cases} 1, & \Delta_\tau \geq \alpha_\Delta; \\ 0, & \Delta_\tau < \alpha_\Delta, \end{cases}$$

где параметр  $\alpha_\Delta$  определяется в соответствии с (7) на множествах  $D$  и  $T$ :  $\alpha_\Delta$  выбирается так, чтобы при  $\mathcal{A}(D'_\omega) = 1$  выполнялось неравенство  $\Delta_\omega < \alpha_\Delta$ ,  $\omega \in T$ .

Кроме  $\Delta_\tau$  введем еще одну меру стойкости пары  $(FC(C, k), \tau)$ . Определим эту меру как среднее расстояние в метрике  $d : \mathbb{F}^k \times \mathbb{F}^k \rightarrow \mathbb{Z}_+$  между информационным сообщением  $\mathbf{x} \in \mathbb{F}^k$  и сообщениями из соответствующего множества претендентов наблюдателя  $\mathcal{X}_{FC, \tau}(\mathbf{x})$ :

$$m_d(FC, \tau) = \frac{1}{2^k} \sum_{\mathbf{x} \in \mathbb{F}^k} \frac{1}{|\mathcal{X}_{FC, \tau}(\mathbf{x})|} \sum_{\mathbf{x}' \in \mathcal{X}_{FC, \tau}(\mathbf{x})} d(\mathbf{x}, \mathbf{x}'). \quad (8)$$

Алгоритм  $\mathcal{B} = \mathcal{B}_d(FC, \tau)$ , использующий меру  $m_d(FC, \tau)$ , определим следующим образом:

$$\mathcal{B}_d(FC, \tau) = \begin{cases} 0, & |m_d(FC, \tau) - \alpha_d| \geq \varepsilon_d; \\ 1, & |m_d(FC, \tau) - \alpha_d| < \varepsilon_d, \end{cases}$$

где  $\varepsilon_d$  — величина, подбираемая так, чтобы на множествах  $D$  и  $T$  выполнялось требование (7), а  $\alpha_d$  — среднее расстояние между двумя векторами из  $\mathbb{F}^k$  в метрике  $d$ :

$$\alpha_d = \frac{1}{2^{2k}} \sum_{\mathbf{a} \in \mathbb{F}^k} \sum_{\mathbf{b} \in \mathbb{F}^k} d(\mathbf{a}, \mathbf{b}). \quad (9)$$

Отметим, что для метрики Хэмминга  $d_H$  выполняется равенство  $\alpha_{d_H} = k/2$ .

Детально способ определения с помощью алгоритма  $\mathcal{A}$  параметров  $\alpha_\Delta$

и  $\varepsilon_d$  для соответствующих алгоритмов  $\mathcal{B}_\Delta(FC, \tau)$  и  $\mathcal{B}_d(FC, \tau)$  описан в разд. 4. В общем виде взаимосвязь алгоритмов  $\mathcal{A}$  и  $\mathcal{B}$  представлена на рис. 2. Построенные алгоритмы  $\mathcal{B}_\Delta(FC, \tau)$  и  $\mathcal{B}_d(FC, \tau)$ , как будет показано далее, могут быть

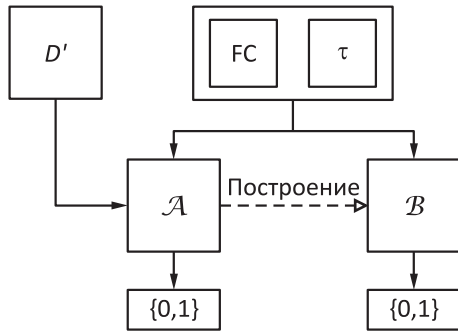


Рис. 2 Взаимосвязь алгоритмов  $\mathcal{A}$  и  $\mathcal{B}$

использованы для оценки стойкости факторного кода  $FC(C, k)$  на произвольном подмножестве  $\tau \in \underline{n}$  вместо применения алгоритма  $\mathcal{A}$  и выборки  $D$ .

Далее приводятся некоторые свойства меры  $m_d(FC, \tau)$ .

**Лемма 1.** Пусть  $FC(C, k)$  — факторный код,  $\tau$  — множество наблюдаемых координат. Тогда

$$m_d(FC, \tau) = \frac{1}{2^{k+\Delta_\tau}} \sum_{\mathbf{x} \in \mathbb{F}^k} \sum_{\mathbf{t} \in V_{FC, \tau}} d(\mathbf{x}, \mathbf{x} + \mathbf{t}). \quad (10)$$

Доказательство. Верность леммы следует из (5) и равенства  $|V_{FC, \tau}| = 2^{\Delta_\tau}$ .

Распределением весов множества  $K \subset \mathbb{F}^k$  назовем набор  $W(K) = (w_0, w_1, w_2, \dots, w_n)$ , где  $w_i = |\{\mathbf{c} \in K \mid w(\mathbf{c}) = i\}|$ ,  $w(\mathbf{c}) := d(\mathbf{0}, \mathbf{c})$ . Аналогично распределением расстояний множества  $K \subset \mathbb{F}^k$  до вектора  $\mathbf{b} \in \mathbb{F}^k$  будем называть набор  $D(C, \mathbf{b}) = (d_0, d_1, \dots, d_n)$ , где  $d_i = |\{\mathbf{c} \in K \mid d(\mathbf{c}, \mathbf{b}) = i\}|$ . Из леммы 1 следует, что

$$m_d(FC, \tau) = \frac{1}{2^{k+\Delta_\tau}} \sum_{\mathbf{x} \in \mathbb{F}^k} \sum_{d_i \in D(\mathcal{X}_{FC, \tau}(\mathbf{x}), \mathbf{x})} id_i. \quad (11)$$

Отсюда вытекает

**Лемма 2.** Если для метрики  $d : \mathbb{F}^k \times \mathbb{F}^k \rightarrow \mathbb{Z}_+$  для любых  $\mathbf{a}(\in \mathbb{F}^k)$  и  $\mathbf{b}(\in \mathbb{F}^k)$  выполняется равенство  $d(\mathbf{a}, \mathbf{a} + \mathbf{b}) = w(\mathbf{b})$ , то

$$m_d(FC, \tau) = \frac{1}{2^{\Delta_\tau}} \sum_{w_i \in W(V_{FC, \tau})} iw_i. \quad (12)$$

Заметим, что для метрики Хэмминга  $d_H$  равенство (12) выполняется.

**Теорема 1.** Пусть  $FC(C, k)$  — факторный код,  $V_{FC, \tau} \subset \mathbb{F}^k$  — код ошибок. Тогда

$$m_{d_H}(FC, \tau) = \frac{|\chi(V_{FC, \tau})|}{2}. \quad (13)$$

Доказательство. Согласно [10] (см. доказательство леммы 1, с. 120), для любого линейного  $[n, k]$ -кода  $C$  ( $|\chi(C)| = n$ ) выполняется равенство  $|C|^{-1} \sum_{\mathbf{x} \in C} w(\mathbf{x}) = n/2$ . Отсюда с учетом (12) вытекает равенство (13).

В следующем подразделе вводится и исследуется байтовая метрика  $d_B$ , которая, в отличие от метрики Хэмминга  $d_H$ , позволяет учитывать старшинство битов в байтах.

### 3.2 Байтовая метрика

Метрика Хэмминга не учитывает позиции битов, в которых различаются векторы  $\mathbf{a}$  и  $\mathbf{b}$ . Так, в пространстве  $\mathbb{F}^8$  расстояние Хэмминга между векторами  $(1, 1, 1, 1, 1, 1, 1, 1)$  и  $(0, 1, 1, 1, 1, 1, 1, 1)$  и расстояние Хэмминга между векторами  $(1, 1, 1, 1, 1, 1, 1, 1)$  и  $(1, 1, 1, 1, 1, 1, 1, 0)$  будет одинаковым. В то же время, если эти векторы представляют собой двоичные записи чисел из  $\mathbb{Z}_{256}$ , характеризующих интенсивность цвета пиксела, то разница в цвете в соответствующих парах будет значительной. Поэтому для оценки пригодности случайно декодированных данных необходима метрика, учитывающая позиции различающихся битов.

Пусть  $\mathbf{x} = (x_0, x_1, \dots, x_7, x_8, \dots, x_{8k'-1}) \in \mathbb{F}^{8k'}$ ,  $k' \in \mathbb{N}$ . Сопоставим вектору  $\mathbf{x}$  вектор  $\tilde{\mathbf{x}} = (\tilde{x}_1, \dots, \tilde{x}_{k'}) \in \mathbb{Z}_{256}^{k'}$ , где числу  $\tilde{x}_j \in \mathbb{Z}_{256}$  соответствует вектор  $(x_{8j+0}, x_{8j+1}, \dots, x_{8j+7}) \in \mathbb{F}^8$ . Для  $\mathbf{a} \in \mathbb{F}^{8k'}$  и  $\mathbf{b} \in \mathbb{F}^{8k'}$  определим байтовую метрику по формуле:

$$d_B(\mathbf{a}, \mathbf{b}) = \sum_{j=1}^{k'} \tilde{d}_B(\tilde{a}_j, \tilde{b}_j),$$

где  $\tilde{d}_B(\tilde{a}, \tilde{b}) := |\tilde{a} - \tilde{b}|$ ,  $\tilde{a}, \tilde{b} \in \mathbb{Z}$ . Отметим, что  $d_B(\mathbf{a}, \mathbf{b})/k'$  характеризует среднее расстояние между байтами соответствующих векторов  $\tilde{\mathbf{a}}$  и  $\tilde{\mathbf{b}}$ ,  $0 \leq d_B(\mathbf{a}, \mathbf{b})/k' < 256$ .

**Лемма 3.** Пусть  $n \in \mathbb{N}$ ,  $a \in \mathbb{Z}_+$ ,  $0 < a < 2^n$ . Тогда

$$\frac{1}{2^n} \sum_{i=0}^{2^n-1} |i - (i \oplus a)| = 2^{\lfloor \log_2 a \rfloor}. \quad (14)$$

*Доказательство.* Под длиной числа в двоичном представлении будем понимать количество битовых разрядов от младшего до старшего ненулевого разряда включительно. Докажем утверждение (14) методом математической индукции по длине  $n$  числа  $a$  в двоичном представлении. Для  $n = 1$  имеется только одно ненулевое число длины один в двоичном представлении:  $a = 1$ . В этом случае равенство (14) выполняется:

$$\frac{1}{2} (|0 - (0 \oplus 1)| + |1 - (1 \oplus 1)|) = \frac{1}{2} \cdot 2 = 1 = 2^0 = 2^{\lfloor \log_2 1 \rfloor}.$$

Пусть для  $n$  утверждение (14) выполняется. Проверим для длины  $n + 1$ . Представим  $i$  ( $a$ ) как сумму старшей степени  $\bar{i} = 2^n i_n$  ( $\bar{a} = 2^n a_n$ ) и числа  $\underline{i} = i - \bar{i}$  ( $\underline{a} = a - \bar{a}$ ), где  $i_n$  ( $a_n$ ) — бит числа  $i$  ( $a$ ), стоящий на  $n$ -й позиции (начиная с нуля) в двоичном представлении числа. Получим

$$\frac{1}{2^{n+1}} \sum_{i=0}^{2^{n+1}-1} |i - (i \oplus a)| = \frac{1}{2^{n+1}} \sum_{i=0}^{2^{n+1}-1} |(\bar{i} - (\bar{i} \oplus \bar{a})) + (\underline{i} - (\underline{i} \oplus \underline{a}))|. \quad (15)$$

Определим знак каждого слагаемого под модулем. Отметим, что  $a_n = 1$ , так как если длина числа  $a$  меньше  $n + 1$ , то применяется предположение индукции для длины  $n$ . Для  $b \in \{0; 1\}$  определим множество  $V_b = \{i \in \mathbb{Z}_{2^{n+1}} : i_n = b\}$ ,  $|V_b| = 2^n$ . Тогда

$$\begin{aligned} & \frac{1}{2^{n+1}} \sum_{i=0}^{2^{n+1}-1} |(\bar{i} - (\bar{i} \oplus \bar{a})) + (\underline{i} - (\underline{i} \oplus \underline{a}))| = \\ & = \frac{1}{2^{n+1}} \left( \sum_{i \in V_1} (2^n + (\underline{i} - (\underline{i} \oplus \underline{a}))) - \sum_{i \in V_0} (-2^n + (\underline{i} - (\underline{i} \oplus \underline{a}))) \right) = 2^n = 2^{\log_2 a}. \end{aligned}$$

Из лемм 1 и 3 следует

**Лемма 4.** Пусть  $d_B$  — байтовая метрика,  $k = 8k'$ ,  $k, k' \in \mathbb{N}$ . Тогда

$$m_{d_B}(FC, \tau) = \frac{1}{2^{\Delta_\tau}} \sum_{\mathbf{t} \in V_{FC, \tau}} \sum_{j=1}^{k'} 2^{\lfloor \log_2 \tilde{t}_j \rfloor}. \quad (16)$$

Пусть RREF (row reduced echelon form) — алгоритм, который  $(v \times m)$ -матрицу  $A$  ранга  $v' (\leq v)$  приводит к  $(v' \times m)$ -матрице  $A'$  ступенчатого вида по строчкам:

$$A' = \begin{pmatrix} 1 & * & * & * & * & \cdots & * & * & * & * \\ 0 & 1 & * & * & * & \cdots & * & * & * & * \\ 0 & 0 & 0 & 1 & * & \cdots & * & * & * & * \\ \vdots & \dots & \dots & \dots & \dots & \dots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 \end{pmatrix} \quad (17)$$

и возвращает подмножество номеров столбцов  $\nu (\subseteq \underline{m}, |\nu| = v')$ , образующих верхнетреугольную  $(v' \times v')$ -матрицу с единицами на главной диагонали,  $|\omega| = v'$ . Рассмотрим код ошибок  $V_{FC, \tau}$  с порождающей  $(\Delta_\tau \times n)$ -матрицей  $U$  вида

$$U = (U_1 | \cdots | U_{k'}), \quad (18)$$

где  $U_i$  —  $(\Delta_\tau \times 8)$ -матрица,  $i = 1, \dots, k'$ .

**Теорема 2.** Пусть  $U$  — порождающая матрица вида (18) для кода ошибок  $V_{FC, \tau}$ ,  $w_i = \text{rank}(U_i)$ ,  $\nu_i = \{\nu_{i,1}; \dots; \nu_{i,w_i}\} = \text{RREF}(U_i)$ ,  $i = 1, \dots, k'$ . Тогда

$$m_{d_B}(FC, \tau) = \sum_{j=1}^{k'} \sum_{l=1}^{w_j} 2^{-l} 2^{8-\nu_{j,l}}. \quad (19)$$

Доказательство. Поменяем порядок суммирования в (16):

$$m_{d_B}(FC, \tau) = \frac{1}{|V_{FC, \tau}|} \sum_{j=1}^{k'} \sum_{\mathbf{t} \in V_{FC, \tau}} 2^{\lfloor \log_2 \tilde{t}_j \rfloor}. \quad (20)$$

Для некоторого  $i \in \underline{k'}$  рассмотрим множество номеров координат  $\omega_i = \{8(i - 1) + 1; \dots; 8i\}$ . Очевидно, что матрица  $U_i$  является порождающей матрицей кода  $\mathcal{U}_i := \pi_{\omega_i}(V_{FC, \tau})$ . Так как  $\Delta_\tau$  — размерность кода  $V_{FC, \tau}$ , то векторы кода  $\mathcal{U}_i$  встретятся в коде  $V_{FC, \tau}$  на координатах  $\omega_i$  ровно  $2^{\Delta_\tau - w_i}$  раз. С другой стороны, из представления произвольной матрицы в виде (17) следует, что в  $\mathcal{U}_i$  для любого  $s \in \underline{w_i}$  ровно  $2^{w_i - s}$  векторов, у которых крайние слева  $\nu_{i, s} - 1$  координат нулевые, а координата с номером  $\nu_{i, s}$  ненулевая, а также один нулевой вектор. Тогда (20) примет вид:

$$m_{d_B}(FC, \tau) = \frac{1}{|V_{FC, \tau}|} \sum_{j=1}^{k'} 2^{\Delta_\tau - w_j} \left( 0 \cdot 2^{w_j - w_j} + \sum_{l=1}^{w_j} 2^{8 - \nu_{j, l}} \cdot 2^{w_j - l} \right). \quad (21)$$

Из равенства  $|V_{FC, \tau}| = 2^{\Delta_\tau}$  и (21) следует (19).

Теорема 2 позволяет быстро вычислять значение меры  $m_{d_B}$  по порождающей матрице кода ошибок, сложность которого линейна по  $\dim(V_{FC, \tau})$ . Отметим, что для метрики  $d_B$  величина  $\alpha_{d_B}$  равна  $(k/8) \cdot 85,33203125 \approx (k/8) \cdot (256/3)$  (см. (9)).

#### 4 Результаты эксперимента

Целью эксперимента является, во-первых, построение алгоритмов  $\mathcal{B}_d$  и  $\mathcal{B}_\Delta$  для конкретных факторных кодов в соответствии с описанной выше методикой, а во-вторых, последующая проверка этих алгоритмов для произвольных  $\tau \subseteq \underline{n}$ . В качестве набора файлов  $D = \{\text{file}_1; \dots; \text{file}_L\}$  взяты  $L = 200$  фотографических изображений лиц из базы «FEI face database» [11]. В рамках данной работы для представления изображений выбран формат bmp, содержащий изображения глубины 24 бита, где каждые 3 Бита представляют 3 оттенка одного пиксела: красный, зеленый и синий. В эксперименте используется несколько факторных кодов таких, что матрица  $\begin{pmatrix} G^* \\ G \end{pmatrix}$  (см. (1)) образует порождающую матрицу [64, 64]-кода Рида–Маллера  $\mathcal{RM}(6, 6)$ ; отличаются факторные коды только размерностью  $k$  матрицы  $G^*$ :  $k \in \{24; 32; 40; 56\}$ . Обозначим соответствующие факторные коды символами  $FC_{24}$ ,  $FC_{32}$ ,  $FC_{40}$ ,  $FC_{56}$ .

Для случайного выбора в процессе факторного кодирования по правилу (1) и в процессе случайного декодирования по правилу (6) используется энтропийный



источник случайных чисел из стандартной библиотеки языка Java: `java.security.SecureRandom`.

Набор  $T$  множеств наблюдаемых координат формируется следующим образом. Для каждого  $\mu \in \underline{n}$  случайно и равновероятно выбираются такие множества, чтобы при использовании их в модели защищенного хранилища можно было корректно восстановить закодированную информацию. Иначе говоря, для каждого  $\mu$  найдутся множества размера  $\mu$ , объединение которых дает множество  $\underline{n}$ ; всего при построении алгоритмов  $\mathcal{B}$  рассмотрено 336 ( $|T| = 336$ ) шаблонов стираний из  $2^{64}$  всего возможных. Опишем далее эксперимент для фиксированного факторного кода  $FC \in \{FC_{24}; FC_{32}; FC_{40}; FC_{56}\}$ . Для каждого  $\tau \in T$  набор файлов  $D$ , закодированный по правилу (1), декодируется в соответствии с (6) в набор файлов  $D'_\tau = \{\text{file}_{i,FC,\tau}\}_{i=1}^{200}$ . Для каждого случайно декодированного файла  $\text{file}_{i,FC,\tau} \in D'_\tau$  запускается метод определения лиц `detectMultiScale` из библиотеки OpenCV [12]. Этот метод возвращает массив прямоугольников `objects`, каждый из которых содержит определяемый объект (лицо).

Алгоритм  $\mathcal{A}$ , введенный в подразд. 3.1, выполняется следующим образом: если массив `objects` не пуст и лицо определилось, то  $\mathcal{A}(\text{file}_{i,FC,\tau}) = 1$ , иначе  $\mathcal{A}(\text{file}_{i,FC,\tau}) = 0$ . Для каждого  $\tau$  вычисляется доля файлов, в которых лицо на декодированном изображении распознается:  $F(FC, \tau) = L^{-1} \sum_{i=1}^L \mathcal{A}(\text{file}_{i,FC,\tau})$ . Помимо этого для  $\tau$  вычисляются меры  $m_{d_H}(FC, \tau)$  и  $m_{d_B}(FC, \tau)$ , используя теоремы 1 и 2 соответственно, а также вычисляется значение  $\Delta_\tau$  в соответствии с (3). По вычисленным значениям составляется таблица, где строка содержит информацию о  $\tau$  и значения  $F(FC, \tau)$ ,  $m_{d_H}(FC, \tau)$ ,  $m_{d_B}(FC, \tau)$ ,  $\Delta_\tau$ . Опираясь на данные из таблицы определяются параметры  $\varepsilon_{d_H}$ ,  $\varepsilon_{d_B}$  и  $\alpha_\Delta$  для соответствующих алгоритмов оценки практической стойкости  $\mathcal{B}_{d_H}$ ,  $\mathcal{B}_{d_B}$  и  $\mathcal{B}_\Delta$  так, чтобы выполнялось требование (7), а именно: отдельно для каждой метрики  $d$  строки таблицы упорядочиваются по возрастанию значения  $|m_d(FC, \tau) - \alpha_d|$  и находится первая такая строка, что  $F(FC, \tau) \neq 0$ , т. е. такая строка, пара  $(FC, \tau)$  в которой не  $\mathcal{A}$ -стойка. По значению  $m_d(FC, \tau)$  в этой строке определяется  $\varepsilon_d$ :  $\varepsilon_d = |m_d(FC, \tau) - \alpha_d|$ . Аналогично определяется и  $\alpha_\Delta$ . По построенным параметрам для каждого алгоритма  $\mathcal{B}$  находятся два соответствующих непересекающихся множества  $T_0$  и  $T_1$  ( $T_0 \cup T_1 = T$ ):  $T_0$  — те наборы  $\tau$ , которые являются  $\mathcal{B}$ -стойкими,  $T_1 = T \setminus T_0$ . Далее каждое множество  $T_i$ ,  $i \in \{0; 1\}$ , также разбивается на два подмножества  $T_{i0}$  и  $T_{i1}$ :  $T_{i0}$  — те наборы  $\tau$ , которые являются  $\mathcal{A}$ -стойкими,  $T_{i1} = T_i \setminus T_{i0}$ . По полученным результатам строятся

**Таблица 1** Таблица вероятностей стойкости набора  $(FC, \tau)$  по  $\mathcal{A}$  и по  $\mathcal{B}$

Стойкость	$\mathcal{B}$ -стойко	Не $\mathcal{B}$ -стойко
$\mathcal{A}$ -стойко	$ T_{00} / T_0  = 1$	$ T_{10} / T_1  = \delta$
Не $\mathcal{A}$ -стойко	$ T_{01} / T_0  = 0$	$ T_{11} / T_1  = 1 - \delta$

**Таблица 2** Таблицы вероятностей для кодов  $FC_{24}$ ,  $FC_{32}$ ,  $FC_{40}$  и  $FC_{56}$ 

Код	Стойкость	Не		Не		Не	
		$\mathcal{B}_{d_H}$ -стойко	$\mathcal{B}_{d_H}$ -стойко	$\mathcal{B}_{d_B}$ -стойко	$\mathcal{B}_{d_B}$ -стойко	$\mathcal{B}_{\Delta}$ -стойко	$\mathcal{B}_{\Delta}$ -стойко
$FC_{24}$	$\mathcal{A}$ -стойко	1	0,120	1	0,185	1	0,577
	Не $\mathcal{A}$ -стойко	0	0,880	0	0,815	0	0,423
$FC_{32}$	$\mathcal{A}$ -стойко	1	0,226	1	0,077	1	0,4
	Не $\mathcal{A}$ -стойко	0	0,774	0	0,923	0	0,6
$FC_{40}$	$\mathcal{A}$ -стойко	1	0,286	1	0,038	1	0,519
	Не $\mathcal{A}$ -стойко	0	0,714	0	0,962	0	0,481
$FC_{56}$	$\mathcal{A}$ -стойко	1	0,300	1	0,325	1	0,692
	Не $\mathcal{A}$ -стойко	0	0,700	0	0,675	0	0,308

**Таблица 3** Результаты проверки алгоритмов  $\mathcal{B}$  для кодов  $FC_{24}$ ,  $FC_{32}$ ,  $FC_{40}$  и  $FC_{56}$ 

Код	Стойкость	Не		Не		Не	
		$\mathcal{B}_{d_H}$ -стойко	$\mathcal{B}_{d_H}$ -стойко	$\mathcal{B}_{d_B}$ -стойко	$\mathcal{B}_{d_B}$ -стойко	$\mathcal{B}_{\Delta}$ -стойко	$\mathcal{B}_{\Delta}$ -стойко
$FC_{24}$	$\mathcal{A}$ -стойко	0,990	0,048	0,990	0,2	0,986	0,630
	Не $\mathcal{A}$ -стойко	0,010	0,952	0,010	0,8	0,014	0,370
$FC_{32}$	$\mathcal{A}$ -стойко	0,978	0,162	0,947	0,097	0,942	0,300
	Не $\mathcal{A}$ -стойко	0,022	0,838	0,053	0,903	0,058	0,700
$FC_{40}$	$\mathcal{A}$ -стойко	1	0,314	0,952	0,136	1	0,538
	Не $\mathcal{A}$ -стойко	0	0,686	0,048	0,864	0	0,462
$FC_{56}$	$\mathcal{A}$ -стойко	0,965	0,268	0,978	0,162	0,955	0,622
	Не $\mathcal{A}$ -стойко	0,035	0,732	0,022	0,838	0,045	0,378

таблицы вероятностей вида табл. 1. Значение «0» в левой нижней ячейке отражает выполнение требования (7) к алгоритму. Величина  $\delta$  показывает долю неверно отвергнутых  $\tau$ , т. е. таких наборов, которые стойки по алгоритму  $\mathcal{A}$ , но не стойки по  $\mathcal{B}$ . Чем меньше  $\delta$ , тем меньше неверно отвергнутых  $\tau$ , т. е. тем более «близок» алгоритм  $\mathcal{B}$  к алгоритму  $\mathcal{A}$ . В табл. 2 обобщены таблицы вероятностей для рассмотренных факторных кодов; вычисления для четырех рассмотренных кодов производились в течение 96 ч на компьютере с восьмиядерным процессором AMD FX-8150 Zambezi и 12 ГБайт оперативной памяти.

После построения алгоритмов  $\mathcal{B}$  был проведен проверочный эксперимент для тестирования найденных параметров  $\varepsilon_{d_H}$ ,  $\varepsilon_{d_B}$  и  $\alpha_{\Delta}$ . Для проверочного эксперимента выбрана база лиц «BAO face database» [13], отличная от базы [11], использованной при построении алгоритмов. Тест проводился на первых 40 изображениях лиц из базы [13], для каждого  $\mu \in \underline{64}$  дважды случайно и равновероятно выбирался набор  $\tau$  и определялась стойкость как с помощью алгоритма  $\mathcal{A}$ , так и с помощью построенных алгоритмов  $\mathcal{B}$ . Результаты проверочного эксперимента для рассмотренных факторных кодов приведены в табл. 3. В силу более высокого

разрешения фотографий из базы [13] по сравнению с фотографиями из базы [11], результаты проверочного эксперимента вычислялись также порядка 96 ч. Анализ табл. 3 позволяет сделать ряд выводов. Во-первых, построенные алгоритмы  $\mathcal{B}$  редко неверно определяют множество  $\tau$  стойким в то время, когда оно не является  $\mathcal{A}$ -стойким: от 1% до 6% всех рассмотренных случаев, когда наборы  $\tau$  определяются как  $\mathcal{B}$ -стойкие. Это дает возможность оценивать стойкость факторных кодов на произвольном множестве наблюдаемых координат  $\tau$  с помощью построенных алгоритмов оценки стойкости  $\mathcal{B}$ . В модели распределенного хранения данных это позволит выбирать разбиения множества всех координат кодового вектора на такие подмножества, что факторный код будет стойким на каждом из них: для владельцев хранилищ наблюдаемые данные не будут содержать полезной, в смысле алгоритма  $\mathcal{A}$ , информации. Во-вторых, алгоритмы  $\mathcal{B}_{d_H}$  и  $\mathcal{B}_{d_B}$  реже определяют набор  $\tau$  как нестойкий в то время, когда он является  $\mathcal{A}$ -стойким, по сравнению с алгоритмом  $\mathcal{B}_\Delta$ : алгоритм  $\mathcal{B}_{d_H}$  ошибается в 5%–30%, алгоритм  $\mathcal{B}_{d_B}$  ошибается в 10%–20%, а алгоритм  $\mathcal{B}_\Delta$  ошибается в 30%–70% случаев, когда набор  $\tau$  определяется как нестойкий. В-третьих, нельзя однозначно отдать предпочтение какой-либо из метрик  $d_H$  и  $d_B$ , так как для одних кодов реже не  $\mathcal{A}$ -стойкий набор принимается за стойкий с помощью алгоритма  $\mathcal{B}_{d_H}$ , а для других кодов — реже с помощью алгоритма  $\mathcal{B}_{d_B}$ . Поэтому имеет смысл при исследовании защитных свойств конкретного факторного кода  $FC$  строить оба алгоритма:  $\mathcal{B}_{d_H}$  и  $\mathcal{B}_{d_B}$  — и считать пару  $(FC, \tau)$  стойкой в случае, когда она  $\mathcal{B}_{d_H}$ - и  $\mathcal{B}_{d_B}$ -стойкая. Более того, для усиления требований к стойкости можно использовать только те множества  $\tau$ , для которых пара  $(FC, \tau)$  стойка по всем трем алгоритмам:  $\mathcal{B}_{d_B}$ ,  $\mathcal{B}_{d_H}$  и  $\mathcal{B}_\Delta$ . Отметим, что одновременное построение трех алгоритмов практически не отличается по длительности от построения какого-либо одного алгоритма в силу простоты соответствующих вычислений (см. теоремы 1, 2 и формулу (3)).

Представляется, что разработанная методика также может быть полезна при защите звуковых файлов и видеофайлов, для которых критериями защищенности (фактически алгоритмом  $\mathcal{A}$ ) являются, например, соответственно возможность выделения речи [14] и возможность выделения объектов [15].

## Литература

1. *Subramanian A., McLaughlin S. W.* MDS codes on the erasure-erasure wiretap channel. 2009. <http://arxiv.org/abs/0902.3286>.
2. *Деундяк В. М., Косолапов Ю. В.* Об одном методе снятия неопределенности в канале с помехами в случае применения кодового зашумления // Известия ЮФУ. Технические науки, 2014. № 2. С. 197–208.
3. *Косолапов Ю. В.* Коды для обобщенной модели канала с подслушиванием // Проблемы передачи информации, 2015. Т. 51. № 1. С. 23–28.
4. *Cohen G., Zemor G.* The wiretap channel applied to biometrics // Symposium (International) on Information Theory and Applications Proceedings, 2004. <https://hal.archives-ouvertes.fr/hal-00359822/document>.

5. *Balakirsky V. B., Ghazaryan A. R., Han Vinck A. J.* Secrecy of permutation block coding schemes designed for biometric authentication // 30th Symposium on Information Theory in the Benelux Proceedings. — Delft, the Netherlands: Werkgemeenschap Informatie- en Communicatietheorie, 2009. P. 11–18.
6. *Bromba M.* On the reconstruction of biometric raw data from template data. <http://www.bromba.com/knowhow/temppriv.htm>.
7. International Biometric Group. Comparative Biometric Testing Round 7 Public Report, 2009. <http://l.b5z.net/i/u/6084428/i/CBT7JBGReport.pdf>.
8. *Ozarov H., Wyner A. D.* Wire-tap channel II // Bell Labs Tech. J., 1984. Vol. 63. No. 10. P. 2135–2157.
9. ISO/IEC 9797-1:1999. Information technology — Security techniques — Message Authentication Codes (MACs). Part 1: Mechanisms using a block cipher.
10. *Деундяк В. М., Косолапов Ю. В.* О стойкости кодового зашумления к статистическому анализу наблюдаемых данных многократного повторения // Моделирование и анализ информационных систем, 2012. Т. 19. № 2. С. 110–127.
11. FEI face database. <http://fei.edu.br/cet/facedatabase.html>.
12. *Bradski G.* The opencv library // Dr. Dobbs J., 2000. Vol. 25. No. 11. P. 120–126.
13. BAO face database. <https://facedetection.com/wp-content/uploads/BaoDataBase.zip>.
14. *Sohn J., Kim N. S., Sung W.* A statistical model-based voice activity detection // IEEE Signal Proc. Lett., 1999. Vol. 6. No. 1. P. 1–3.
15. *Nadimi S., Bhanu B.* Physical models for moving shadow and object detection in video // IEEE Trans. Pattern Anal. Mach. Intelligence, 2004. Vol. 26. No. 8. P. 1079–1087.

*Поступила в редакцию 13.07.15*

---

---

## **EVALUATION OF RESISTANCE OF CODE NOISING IN THE DISTRIBUTED DATA STORAGE**

*Yu. V. Kosolapov and A. V. Pozdnyakov*

Institute for Mathematics, Mechanics, and Computer Science named after I. I. Vorovich, Southern Federal University, 105/42 Bol'shaya Sadovaya Str., Rostov-on-Don 344006, Russian Federation

**Abstract:** This paper considers the model of distributed data storage across multiple storages in an environment where each of the owners of storage facilities can be unfair. To protect data from unauthorized access, the method of code noising is used with a subsequent distribution of code words parts to the corresponding storages. Generally, this method is examined, provided that the protected data are random and uniform. Often, in applied problems, protected data do not satisfy this condition. This paper solves the problem of development of experimental techniques to estimate the resistance of code noising in a case

where data are the photographic images of faces and the observer has limited abilities for analysis of data available to him. Experimental results are presented for the factor codes that are based on binary Reed–Muller codes of length 64.

**Keywords:** code noising; resistance; distributed storage

**DOI:** 10.14357/08696527150412

## References

1. Subramanian, A., and S. W. McLaughlin. 2009. MDS codes on the erasure-erasure wiretap channel. Available at: <http://arxiv.org/abs/0902.3286> (accessed July 7, 2015).
2. Deundyak, V. M., and Yu. V. Kosolapov. 2014. Ob odnom metode snyatiya neopredelennosti v kanale s pomekhami v sluchae primeneniya kodovogo zashumleniya [One method of removing the uncertainty in the channel with errors in the case of code noising]. *Izvestiya YuFU. Tekhnicheskie nauki* [SFU News. Technical sciences] 2:197–208.
3. Kosolapov, Yu. V. 2015. Codes for a generalized wire-tap channel model. *Prob. Inf. Transm.* 51(1):20–24.
4. Cohen, G., and G. Zemor. 2004. The wiretap channel applied to biometrics. *Symposium (International) on Information Theory and Applications Proceedings*. Available at: <https://hal.archives-ouvertes.fr/hal-00359822/document> (accessed July 7, 2015).
5. Balakirsky, V. B., A. R. Ghazaryan, and A. J. Han Vinck. 2009. Secrecy of permutation block coding schemes designed for biometric authentication. *30th Symposium on Information Theory in the Benelux*. Delft, the Netherlands: Werkgemeenschap Informatie- en Communicatietheorie. 11–18.
6. Bromba, M. On the reconstruction of biometric raw data from template data. Available at: <http://www.bromba.com/knowhow/temppriv.htm> (accessed July 7, 2015).
7. International Biometric Group. 2009. Comparative Biometric Testing Round 7 Public Report. Available at: <http://l.b5z.net/i/u/6084428/i/CBT7JBGReport.pdf> (accessed July 7, 2015).
8. Ozarov, H., and A. D. Wyner. 1984. Wire-tap channel II. *Bell Labs Tech. J.* 63(10):2135–2157.
9. ISO/IEC 9797-1:1999. Information technology — Security techniques — Message Authentication Codes (MACs). Part 1: Mechanisms using a block cipher.
10. Deundyak, V. M., and Yu. V. Kosolapov. 2012. O stoykosti kodovogo zashumleniya k statisticheskomu analizu nablyudaemykh dannykh mnogokratnogo povtoreniya [On the firmness code noising to the statistical analysis of the observable data of repeated repetition]. *Modelirovanie i analiz informatsionnykh sistem* [Information Systems Modeling and Analysis] 19(4):110–127.
11. FEI face database. Available at: <http://fei.edu.br/cet/facedatabase.html> (accessed July 7, 2015).
12. Bradski, G. 2000. The opencv library. *Dr. Dobbs J.* 25(11):120–126.
13. BAO face database. Available at: <https://facedetection.com/wp-content/uploads/BaoDataBase.zip> (accessed July 7, 2015).

14. Sohn, J., N. S. Kim, and W. Sung. 1999. A statistical model-based voice activity detection. *IEEE Signal Proc. Lett.* 6(1):1–3.
15. Nadimi, S., and B. Bhanu. 2004. Physical models for moving shadow and object detection in video. *IEEE Trans. Pattern Anal. Mach. Intelligence* 26(8):1079–1087.

*Received July 13, 2015*

## **Contributors**

**Kosolapov Jury V.** (b. 1982)—Candidate of Science (PhD) in technology, associate professor, Institute for Mathematics, Mechanics, and Computer Science named after I. I. Vorovich, Southern Federal University, 105/42 Bol'shaya Sadovaya Str., Rostov-on-Don 344006, Russian Federation; itaim@mail.ru

**Pozdnyakov Alexey V.** (b. 1992)—student, Institute for Mathematics, Mechanics, and Computer Science named after I. I. Vorovich, Southern Federal University, 105/42 Bol'shaya Sadovaya Str., Rostov-on-Don 344006, Russian Federation; adelier@mail.ru.

## СОЦИАЛЬНАЯ ИНФОРМАТИКА: РОССИЙСКАЯ НАУЧНАЯ ШКОЛА И ПЕРСПЕКТИВНЫЕ НАПРАВЛЕНИЯ ИССЛЕДОВАНИЙ

*К. К. Колин*<sup>1</sup>

**Аннотация:** Рассматривается история становления, современное состояние и перспективы развития социальной информатики как актуального направления в науке и образовании. Показаны отличительные особенности российской научной школы социальной информатики и ее приоритет в формировании этого направления. Определены перспективные направления исследований в области социальной информатики в условиях становления глобального информационного общества. Аргументирована необходимость более глубокого изучения проблем социальной информатики в системе образования и подготовки научных кадров.

**Ключевые слова:** информатизация общества; информационные ресурсы; информационное общество; информационная безопасность; социальная информатика

**DOI:** 10.14357/08696527150413

### Введение

Социальная информатика является одним из важных и перспективных направлений в науке и образовании, в котором изучаются процессы информационного взаимодействия в обществе в условиях его глобализации и комплексной информатизации. Как научное направление социальная информатика развивается в России уже более 30 лет. В настоящее время — это один из важных разделов информатики, который можно рассматривать как научную базу формирования информационного общества, изучения новых возможностей и проблем человека в условиях становления принципиально новой информационной среды обитания и профессиональной деятельности.

Исследования показывают, что одной из наиболее важных тенденций развития цивилизации в XXI в. является переход от индустриального к постиндустриальному и далее — к *информационному обществу*, в котором основными объектами и результатами труда для большинства занятого населения будут информационные продукты и услуги. Поэтому существует объективная потребность в систематическом изучении проблем становления информационного общества, а также формирования в этой области научных знаний и распространения их

---

<sup>1</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, kolinkk@mail.ru

через систему образования. Закономерной реакцией науки на эту потребность стало формирование отдельного направления исследований в области информатики, которое появилось в России в середине 1980-х гг. и получило название *социальной информатики*.

## **1 Социальная информатика как наука: предмет изучения, задачи и методы**

Первой работой, в которой социальная информатика была представлена как самостоятельное научное направление, исследующее закономерности информатизации общества, и предложен концептуальный подход к определению основных целей и задач этого направления, стала монография А. Д. Урсула «Информатизация общества. Введение в социальную информатику» [1]. До этого термин «социальная информатика» в России также использовался, но только для целей исследования научной информации [2]. Поэтому основоположником российской научной школы социальной информатики, в ее современном понимании, является именно А. Д. Урсул. Его философский подход связал этот раздел информатики с глобальной проблемой информатизации общества, в процессе которой и осуществляется «социализация» теоретических достижений и практических приложений информатики. Им было показано, что процесс информатизации общества должен быть гуманистически ориентированным по самой своей сути. При этом средства информатики и новые информационные технологии должны выступать лишь как инструментальные средства, которые служат человеку для того, чтобы он мог более эффективно *овладевать информацией* и использовать ее в целях социального прогресса.

В указанной монографии особо отмечена принципиальная важность гуманистической ориентации процесса информатизации общества, который не должен развиваться стихийно, так как он создает для человека и общества не только новые возможности, но и новые угрозы. Именно это и определяет актуальность проведения комплекса научных исследований, направленных на изучение закономерностей процесса информатизации общества и разработки на этой основе научно обоснованных концепций, прогнозов и рекомендаций в области проведения международной, государственной и региональной научно-технической, экономической, социальной, образовательной и культурной политики.

Решению этих задач и должна содействовать социальная информатика как наука. Она должна стать научной базой формирующегося глобального информационного общества, которое многие ученые рассматривают как новую, более высокую стадию развития цивилизации [3–5].

### **Объект и предмет исследований социальной информатики**

Социальная информатика формируется на стыке ряда естественных и гуманитарных наук под воздействием интегративных факторов, обусловленных не-



обходимостью формирования научной базы постиндустриальной цивилизации — *информационного общества*. В работах специалистов Института проблем информатики РАН, которые опубликованы еще в начале 1990-х гг. [6–11], обоснована правомерность выделения социальной информатики в качестве самостоятельной научной дисциплины. Критериями здесь являются наличие собственного объекта и предмета исследований, а также характерной для данной дисциплины методологии их проведения.

**Объектом исследования** социальной информатики являются *проявления информационной реальности в социальной сфере общества*. Иными словами, эта дисциплина изучает совокупность всех видов информационных ресурсов, процессов, технологий, систем и коммуникаций, которые имеют социальную значимость для жизнеобеспечения и развития общества.

**Предметом исследований** являются *свойства и закономерности* информационных ресурсов, процессов, технологий, систем и коммуникаций в социальной сфере. При этом особо важным является комплексное изучение процесса *информатизации общества*, его воздействия на социальные структуры, а также на роль и положение в обществе самого человека. Анализ показал, что это воздействие является двухсторонним, т. е. социальные структуры общества и отношение людей к процессу информатизации, в свою очередь, оказывают существенное воздействие на характер и темпы развития этого процесса [5].

По мнению А. Д. Урсула, одной из фундаментальных проблем социальной информатики является проблема согласованного развития общества и процесса его глобальной информатизации. Эта проблема является в настоящее время особенно важной, так как она самым тесным образом связана с проблемой преодоления кризиса цивилизации и ее перехода на путь устойчивого и безопасного развития на основе «овладения информацией» (термин академика Н. Н. Моисеева), использования новых знаний и информационных технологий. Поэтому приоритетными задачами исследований в этой области следует считать выявление и анализ возможностей человека и общества эффективно использовать информационные ресурсы и технологии с учетом новых угроз, которые уже существуют сегодня в связи с глобальной информатизацией общества и могут появиться в будущем.

### **Концепции и методы социальной информатики**

Проблемы формирования новой среды обитания человека в условиях становления информационного общества настолько значимы, что должны систематически изучаться различными научными дисциплинами. В их числе необходимо в первую очередь выделить культурологию, социологию, экономику и политологию, а также психологию и педагогику.

Принципиальное отличие в подходе к изучению этих проблем в рамках социальной информатики состоит в том, что используется основная научная гипотеза информатики. Ее суть в том, что окружающий нас мир обладает свойством

*информационного единства*, и поэтому закономерности проявления информации как всеобщего свойства материи в природе и обществе должны иметь общую первооснову [12]. Научной базой для развития социальной информатики являются достижения в области философии информации и теоретической информатики, которая сегодня рассматривается в России как фундаментальная наука, приобретающая все большее междисциплинарное и мировоззренческое значение [13, 14].

Что же касается собственной научной методологии социальной информатики, то она находится в стадии формирования. Важным отличительным свойством этой методологии является использование совокупности различных методов научного подхода к проблемам анализа и синтеза изучаемых проблем и процессов. Основными из них являются *информационный, системный и коэволюционный подходы*, а также подходы социальных отраслей знания. Какого-то особого, отличного от других, метода в социальной информатике пока нет, но определенной спецификой обладает комплекс используемых в этой науке методов, который в дальнейшем может стать ее методологической системой.

**Информационный подход.** Сущность информационного подхода заключается в том, что при изучении некоторого процесса или явления в первую очередь выделяются его информационные аспекты [13]. Опыт показал, что это позволяет увидеть многие, казалось бы, привычные и хорошо знакомые процессы и явления в совершенно новом свете, раскрыть их информационную сущность, которая во многих случаях и является главной причиной того или иного направления развития этих процессов.

**Коэволюционный подход** к проблемам социальной информатики заключается в необходимости учета неразрывного единства и взаимовлияния процессов развития общества и его информатизации, их *коэволюции*. Этот подход нашел свое отражение в Окинавской Хартии глобального информационного общества, которую некоторые специалисты называют «Международной конституцией» информационного общества.

**Социальный подход** состоит в том, что исследования ориентируются, главным образом, на социальные аспекты процесса информатизации, его гуманистическую направленность. Сюда, в первую очередь, относятся проблемы становления и развития информационного общества, новые возможности и проблемы его демократизации, обеспечения информационной свободы и информационной безопасности человека и общества в целом.

Методологический аппарат социальной информатики, помимо указанных выше подходов, включает в себя методы *информационного моделирования* изучаемых процессов, а также традиционные методы мониторинга и прогнозирования. Следует ожидать, что состав и мощность методологического аппарата этой науки будут активно развиваться по мере расширения и углубления ее предметной области и появления необходимости в решении новых задач, выдвигаемых практикой развития информационного общества.

## 2 Структура предметной области социальной информатики

Структура предметной области социальной информатики как научной дисциплины была определена в России еще в начале 1990-х гг. [6–11]. Ее обсуждение проводилось на ряде научных форумов. Наиболее значимым стал II Международный конгресс ЮНЕСКО «Образование и информатика» (Москва, 1996). В Национальном докладе России на этом Конгрессе [15] раздел «*Социальная информатика*» был выделен в качестве одного из четырех основных разделов нового курса «*Фундаментальные основы информатики*», который был рекомендован в итоговых документах Конгресса для перспективной системы образования.

В современной структуре предметной области социальной информатики выделены следующие основные направления исследований:

- изучение общих закономерностей процесса *информатизации общества*;
- исследования *информационных ресурсов* общества (их видов, свойств, структуры и топологии). Изучение потребностей общества в информационных ресурсах для достижения целей общественного развития;
- исследования *информационного потенциала* общества, определяющего его возможности по формированию и эффективному использованию информации как стратегического ресурса развития. Сюда относятся проблемы, связанные с формированием *информационной инфраструктуры* и *информационной среды* общества, создания и функционирования центров генерации, хранения и распространения информационных ресурсов, а также развития инструментальных средств и технологий, обеспечивающих активизацию и эффективное использование информационных ресурсов;
- исследования проблем формирования *глобального информационного общества*, закономерностей и особенностей его становления и перспектив развития. Сюда входят проблемы *информационной экономики*, изменения структуры занятости населения, а также проблема *информационной демократии*, в которой особо важны доступность информации и ее достоверность как важнейшие условия практической реализации людьми своих гражданских прав и свобод. Существенное значение для решения этих проблем имеет развитие *интеллектуального потенциала общества*, т. е. его способности продуцировать и усваивать новые знания, а также развитие информационной культуры, характеризующей возможности адаптации общества к новой информационной среде обитания человека;
- комплексное изучение *проблем человека в информационном обществе*. Здесь должны изучаться новые возможности и проблемы развития личности в информационном обществе, образования и воспитания человека, развития его творческих способностей на основе достижений информатики и креативных информационных технологий. Актуальными являются также проблемы *информационной свободы* и *информационной безопасности* человека, пре-

одоления им технологического и лингвистического барьеров, возникающих в новом информационном пространстве.

Практика показала, что приведенная выше структура проблемной области оказалась достаточно удачной, а перечисленные проблемы, составляющие эту область, до сих пор продолжают оставаться важными и актуальными. Об этом, в частности, свидетельствует большое количество ссылок на публикации российских ученых. Так, количество ссылок на монографии [16, 17] в базе данных Google Scholar превысило 240 и продолжает расти.

### **3 Особенности изучения проблем социальной информатики в России и других странах**

#### **Украинская научная школа информатики и ее социальные аспекты**

Украинские специалисты в области информатики всегда уделяли повышенное внимание социально значимым аспектам ее развития. Достаточно указать на концепцию «безбумажной информатики» академика В. М. Глушкова [18], которая была ориентирована на информатизацию документооборота в системах организационного управления.

В своих работах по теоретической информатике украинские специалисты подчеркивают принципиальную важность социальных аспектов использования достижений этой науки. Так, например, в работе [19] указано: «...мы связываем информатику с закономерностями движения и переработки информации в общественной среде, с общими принципами построения и функционирования информационных систем социальной природы, использующих информацию в виде человеческих знаний. Иными словами, речь идет в данной работе о социальной информатике». И далее: «Таким образом, информатика на сегодня предстает перед нами в качестве самостоятельной науки, изучающей информационные процессы и системы, относящиеся к области социальной практики. Один из главных моментов в понимании теоретической информатики, смысла ее основных категорий и практической важности состоит в следующем: эта наука открывает ресурсы перспективного социально-экономического развития, сложной, качественно более совершенной организации жизни и деятельности общества, перехода на новый виток эволюции современной цивилизации — к информационному обществу».

#### **Социальная информатика в странах Запада**

Изучение проблем социальной информатики в странах Запада началось несколько позднее, чем в России. Так, например, с 1996 г. термин «социальная информатика» стал использоваться в ходе общественных дискуссий по проблемам социальных последствий компьютеризации и использования информационных систем в социуме. В создании этого направления в США приняли участие такие американские ученые, как П. Агр, А. Бишоп, Р. Клинг, С. Соьер и др.

Однако даже в таких крупных университетах США, как Стэнфордский университет, социальная информатика в те годы необходимого развития не получила [20], что представляется довольно странным, учитывая большую практическую значимость этого направления информатики. В других же университетах такие исследования были развернуты. Так, в Университете штата Индиана был создан Центр социальной информатики, который провел комплекс исследований по программе «Информационные технологии и социальные изменения». Аналогичные исследования проводятся и в Великобритании (Университет Уэлса), а в Университете г. Квебек (Канада) даже создан факультет социальной информатики. Заметим, что первый такой факультет был ранее создан в России в составе Московского социального университета [21].

Социальные проблемы становления глобального информационного общества в настоящее время привлекают к себе внимание и общественных международных научных организаций. Одной из них является Международное общество по изучению информации (ISIS — International Society for Information Studies), имеющее свою штаб-квартиру в Австрии и объединяющее ученых из 20 стран мира. В июле 2015 г. это общество провело в Вене крупную международную конференцию, посвященную выработке концепции дальнейшего развития глобального информационного общества. Материалы и рекомендации этой конференции публикуются в международном журнале «Information» и активно обсуждаются в сети Интернет.

### **Изучение проблем социальной информатики в Китае**

Лидером в области изучения проблем социальной информатики в Китае является Хуаджонгский государственный университет науки и технологий (Провинция Ухань), где в 2006 г. был создан Институт социальной информационной науки, который возглавил вице-президент этого университета профессор Оуян Кун. Уже в 2007 г. этот Институт провел в Китае первую национальную конференцию по проблемам социальной информатики, а в настоящее время продолжает исследования в этом направлении, используя при этом и результаты специалистов ИПИ РАН, с которыми он поддерживает научное сотрудничество. Результатами этого сотрудничества стали:

- участие специалистов ИПИ РАН в работе IV Международной конференции по фундаментальным основам информационной науки (FIS-2010) в Пекине, на которой был сделан пленарный доклад по проблемам социальной информатики [22];
- участие китайских ученых в аналогичной конференции (FIS-2013) в Москве;
- публикация статей и монографий ученых ИПИ РАН в Китае.

Следует отметить, что китайские специалисты признают лидерство ученых ИПИ РАН как в области изучения социальных аспектов информатики, так и ее философских и научно-методологических оснований. Они охотно публикуют

наши работы на китайском языке [23] и предлагают проведение совместных исследований.

## 4 Социальная информатика как образовательная дисциплина

### Социальная информатика и научное мировоззрение

Комплексное изучение процессов информатизации общества и возникающей при этом новой информационной реальности методами социальной информатики является исключительно важным не только с прагматической точки зрения, но также и в гуманитарном аспекте, так как содействует формированию современного научного мировоззрения. В условиях становления информационного общества существенным образом изменяются привычные стереотипы поведения людей, а также и их традиционные представления о пространстве и времени, богатстве и бедности, равенстве и неравенстве, свободе и ответственности, качестве жизни [24]. Исследования показывают, что качество жизни в значительной степени определяется доступностью необходимой человеку информации и информационных коммуникаций, а также уровнем потребления информационных продуктов и услуг [25]. Поэтому появились и все более широко используются такие понятия, как *информационное неравенство*, *информационная бедность*, *информационная культура личности*, *электронное правительство* и многие другие.

Быстро развивается *информационный сектор экономики*. В ряде стран, например в Индии, он становится одним из важнейших источников экспорта продукции и создания новых рабочих мест. Во многих странах, в том числе и в России, получают развитие системы открытого образования, основанные на использовании цифровых образовательных ресурсов и технологий дистанционного обучения. Становление «умного образования» (Smart education), интенсивно развиваемого в странах Запада, предполагает перенос образовательного процесса преимущественно в электронную среду, обеспечивающую эффективное обучение с использованием *цифрового контента*, находящегося в свободном доступе. Такой подход делает обучение доступным везде и всегда, объединяя учебные заведения и профессорско-преподавательский состав для реализации совместной образовательной деятельности в сети Интернет (являющейся одним из основных источников знаний для студентов). Это так называемое «гибкое обучение» на базе общих технологий, стандартов и соглашений.

*Информационная безопасность* также становится одним из важнейших условий обеспечения национальной и международной безопасности, личной безопасности каждого человека [26]. Охвативший многие страны глобальный финансово-экономический кризис является здесь очень наглядным примером. На современном этапе формирования общества, основанного на знаниях, начинает осознаваться стратегическая важность развития лингвистической и филологической культуры личности и общества [27]. Ведь язык — это не только средство общения, но также и средство познания, это «пространство мысли», необ-

ходимое для информационного моделирования окружающего мира в сознании человека.

### **Актуальность информационной ориентации системы образования**

Социальная информатика активно развивается в России уже более 25 лет. В настоящее время это один из важных разделов информатики, являющийся научной базой для формирования информационного общества, а также для изучения новых возможностей и новых проблем человека в условиях становления принципиально новой информационной среды его жизни и профессиональной деятельности. Многие проблемы социальной информатики уже изучаются в средней и высшей школе России, а также в системе подготовки научных кадров и повышения квалификации дипломированных специалистов. Однако перед системой образования сегодня встает новая актуальная проблема — своевременно подготовить миллионы граждан и, в особенности, новое поколение, к жизни и профессиональной деятельности в условиях глобального информационного общества [27].

Анализ показывает, что для решения этой проблемы необходима существенно бóльшая, чем это имеет место сегодня, *информационная ориентация содержания системы образования* [28]. Конструктивным решением здесь может явиться широкое внедрение в систему образования общеобразовательной дисциплины «Социальная информатика», имеющей своей целью обеспечить формирование у людей новых знаний и умений, которые потребуются им в высокоавтоматизированной информационной среде обитания [29].

### **Опыт изучения курса социальной информатики в России**

Россия является мировым лидером в области изучения проблем социальной информатики в системе образования. Достаточно указать, что этот процесс начался в России еще в конце 1980-х гг., когда начала осознаваться социальная значимость процесса информатизации общества. В этот период в ряде российских вузов стали создаваться специализированные *кафедры социальной информатики*, а в 1990 г. был создан *Институт социальной информатики* в Академии общественных наук при ЦК КПСС. Его директором стал доктор философских наук, профессор А. Д. Урсул. К сожалению, этот Институт просуществовал недолго из-за последующих социальных трансформаций в нашей стране после распада СССР. Была также создана и российская Ассоциация кафедр социальной информатики, которая продолжает свою работу до настоящего времени. Ее руководителем был избран доктор социологических наук, профессор Б. А. Сулаков.

В середине 1990-х гг. в Московском государственном социальном университете (ныне Российский государственный социальный университет) по инициативе Института проблем информатики РАН был создан первый в мире *факультет со-*



циальной информатики. Он достаточно успешно функционировал в течение пяти лет и подготовил значительное количество специалистов с квалификацией *социолог-информатик*. Автор настоящей работы поставил и прочел на этом факультете инновационный курс «Философия и история образования» [28], в котором была показана важность изучения информационных аспектов развития цивилизации.

Опыт образовательной деятельности данного факультета докладывался российскими учеными на II Международном конгрессе ЮНЕСКО «Образование и информатика» [21], получил одобрение участников этого Конгресса и до сих пор успешно используется в ряде образовательных учреждений высшей школы России.

### **Методическое обеспечение учебного курса**

Первая экспериментальная программа учебного курса «Социальная информатика» для системы высшего образования, в которой были учтены рекомендации Конгресса ЮНЕСКО, была разработана в 1997 г. в Институте проблем информатики РАН при активном содействии Института ЮНЕСКО по информационным технологиям в образовании (директор Института — академик РАО В. Г. Кинелев). За последующие годы эта программа была развита до уровня полноценной Базовой модульной программы, которая представлена в сети Интернет и хорошо известна специалистам, работающим в сфере образования [30].

В соответствии с содержанием этой программы было разработано и учебное пособие по курсу «Социальная информатика» для системы высшего образования [16, 17]. В сети Интернет имеются многочисленные ссылки на эти издания, что свидетельствует о высокой востребованности данного курса в российской системе образования.

## **5 Перспективы развития социальной информатики**

Проблематика социальной информатики не ограничивается только проблемами информатизации и становления информационного общества. Ведь многие социально значимые проявления феномена информации необязательно связаны с процессами информатизации общества, а социальная информатика должна изучать также и эти проявления информационной реальности, поэтому к числу наиболее перспективных направлений дальнейшего развития социальной информатики можно отнести следующие:

- системные исследования проблем становления глобального информационного общества, основанного на знаниях, и связанных с этим задач науки, экономики, образования и культуры;
- комплексное изучение проблем формирования электронных информационных ресурсов, методов их сохранения и эффективного использования в различных сферах жизнедеятельности общества;



- изучение новых возможностей для развития человека в информационном обществе, включая его интеллект, творческие способности и моральные качества;
- исследование гуманитарных проблем информационной безопасности, включая такие проблемы, как *информационное неравенство, информационная преступность, киберболезни, манипуляция сознанием, виртуализация общества*, а также противоборство в информационной сфере и *информационные войны*.

Некоторые из этих проблем уже изучаются современной наукой [31–37]. Эти исследования, безусловно, получают свое дальнейшее развитие и в будущем.

Информационная среда обитания человека сегодня развивается так стремительно, что можно уже в ближайшие годы ожидать радикальных изменений во всех сферах жизнедеятельности общества. По имеющимся прогнозам, условия жизни и деятельности людей в развитых странах через два десятилетия будут так же отличаться от современных условий, как наши времена отличаются от эпохи времен Ивана Грозного, причем эти изменения произойдут не только в научно-технической, экономической и информационной областях, но также и в сферах культуры, искусства, в условиях быта и отдыха людей. Получит развитие новая, уже активно формирующаяся *информационная культура* общества. Возникнут новые виды искусств, новая медицина и новые виды коммуникаций между людьми, основой которых станут информационно-телекоммуникационные системы, информационные сервисы и дружественные интерфейсы.

Другими словами, на нашей планете возникает и станет все более широко распространяться новая среда обитания, которую все чаще называют *информационным киберпространством*. Воздействие киберпространства на человека и способы адаптации людей в этом пространстве порождают новые проблемы для фундаментальной науки и системы образования.

Что же касается изменения социальных аспектов развития общества, то здесь следует ожидать не только позитивных результатов, но и возникновения новых глобальных проблем. Одной из них является проблема усиливающегося *информационного неравенства* людей, стран и регионов в новой высокоавтоматизированной информационной среде. Уменьшить социальную остроту этой проблемы должна система образования, информационная ориентация которой должна быть существенным образом усилена. При этом информатизация образования должна быть переведена из инструментально-технологической плоскости в содержательную. Ее главной целью должно стать формирование у людей информационного мировоззрения и новой *информационной культуры* [31, 36]. Поэтому комплексное изучение проблем социальной информатики не только в академической науке, но и в системе образования, является сегодня исключительно актуальной и социально значимой проблемой. Ее решение требует совместных усилий мирового научно-образовательного сообщества и, конечно же, соответствующей модернизации системы образования и подготовки научных и педагогических кадров.

## **Социальная информатика как область науки с интенсивным использованием данных**

Социальная информатика изучает информационные процессы в обществе, которые под воздействием информатизации становятся все более интенсивными. Достаточно указать, что количество пользователей Интернет уже превышает 3 млрд чел., а по прогнозам ООН мировой уровень проникновения интернет-технологий к 2020 г. достигнет 62%. Динамика этого процесса такова, что в 2017 г. этот уровень превысит 50%, т. е. можно будет говорить о наступлении эпохи глобального информационного общества, что ранее прогнозировалось в середине XXI в.

К сожалению, сегодня этот глобальный социально-технологический феномен еще недостаточно осознан и не находит своего отражения ни в программах научных исследований, ни в системе образования. В то же время анализ показывает, что в глобальном информационном обществе открываются принципиально новые возможности для научного анализа и прогноза многих социальных процессов, т. е. для изучения динамики самого общества с использованием средств и методов социальной информатики. Достаточно указать на возможность использования для этих целей данных об информационных коммуникациях в социальных сетях, которые представляют собой новую область жизнедеятельности миллионов людей.

Анализ потоков информации в сети Интернет уже сегодня позволяет производить крупномасштабные социологические исследования, не прибегая к традиционным методам социологических опросов, причем по мере дальнейшего увеличения количества пользователей этой сети результаты исследований будут становиться все более достоверными.

Кроме того, за последние десятилетия сформированы массивные коллекции данных о различных направлениях жизнедеятельности общества. Характерным примером здесь могут служить коллекции данных о различных аспектах деятельности в сфере культуры, полученные в результате выполнения европейского проекта «Обсерватория культуры». Они содержат сведения о тематике различных мероприятий в области культуры: кинофестивалей, театральных постановок, эстрадных концертов и т. п. Анализ этой информации необходим для того, чтобы получить адекватные представления об основных тенденциях развития современной культуры, которая, по оценкам многих специалистов, находится в состоянии системного кризиса, а его дальнейшее развитие представляет собой угрозу для глобальной безопасности [32–34]. Однако для этого необходимо обеспечить соответствующую ориентацию средств и методов социальной информатики, а также подготовить необходимые кадры для работы с интенсивными потоками информации и массивными коллекциями данных. Это именно тот случай, когда количество переходит в качество, т. е. мы имеем дело с принципиально новой проблемой интенсивного использования данных большого объема, которая уже получила название Big Data.

## Социальная информатика и ее междисциплинарные приложения

Для решения перечисленных выше задач анализа данных большого объема необходимо проведение междисциплинарных исследований и разработок, при этом необходимо обеспечить опережающее развитие средств и методов информатики для семантического поиска и анализа текстов в многоязычных коллекциях данных социальной и культурологической информации. Эта задача становится приоритетным направлением развития социальной информатики на ближайшие десятилетия [38, 39].

В то же время необходимо указать, что сами исследования общества и его культуры с использованием средств и методов информатики — это задачи других научных дисциплин, которые уже формируются в настоящее время и которые можно назвать, соответственно, *информационной социологией* и *информационной культурологией* [31, 36].

## 6 Заключение

Настоящая работа является одним из результатов научно-исследовательской работы (НИР) «*Разработка и исследование научно-методологических основ информатизации общества с учетом реализации стратегических национальных приоритетов социально-экономического развития Российской Федерации*». Эта НИР выполнена в ИПИ РАН в рамках Программы фундаментальных научных исследований государственных академий наук на 2013–2020 гг. Проведенные исследования показали, что глобальная информатизация является одной из доминирующих тенденций социально-экономического, научно-технического и культурного развития общества в XXI в., а уровень информационного развития страны в значительной степени определяет качество жизни населения, развитие науки, образования и культуры, международный рейтинг страны в мировом сообществе, а также *национальную безопасность*.

Анализ современной геополитической ситуации и тенденций ее дальнейшего развития [40] показывает, что в ближайшие годы приоритетными задачами обеспечения национальной и глобальной безопасности станут *кибербезопасность* [35], а также гуманитарные аспекты информационной безопасности, обусловленные целенаправленной деформацией общественного сознания с использованием средств массовой информации и социальных сетей в Интернет [37]. Эти угрозы нашли свое отражение в новой Военной доктрине России [41] и, безусловно, будут включены в текст новой Доктрины информационной безопасности РФ, которая уже разработана и должна быть принята в 2016 г.

Для эффективного противодействия этим угрозам необходимо активно развивать новое направление культуры общества — *культуру информационной безопасности*. Такую задачу Совет Безопасности РФ поставил еще в конце 2013 г., однако действенных мер по ее решению пока не принято. Поэтому важно, чтобы данная стратегически важная задача нашла свое отражение в Стра-

тегии государственной культурной политики России, которая в настоящее время разрабатывается в соответствии с поручением Президента РФ.

## Литература

1. *Урсул А. Д.* Информатизация общества. Введение в социальную информатику. — М.: АОН при ЦК КПСС, 1990. 189 с.
2. *Соколов А. В.* Информационный подход к документальной информации. — Л.: ЛГИК, 1988. 85 с.
3. *Кастельс М.* Информационная эпоха. — М.: Изд-во ГУ ВШЭ, 2001. 606 с.
4. *Колин К. К.* Инновационное развитие в информационном обществе и качество образования // Открытое образование, 2005. № 3. С. 63–72.
5. *Колин К. К.* Информатизация общества и глобализация. — Красноярск: Изд-во СФУ, 2011. 52 с.
6. *Урсул А. Д.* Информатизация общества и социальная информатика // Сб. научн. тр. «Социальная информатика». — М.: Высшая комсомольская школа при ЦК КПСС, 1990. С. 3–12.
7. *Колин К. К.* О структуре научных исследований по комплексной проблеме «Информатика» // Сб. научн. тр. «Социальная информатика». — М.: Высшая комсомольская школа при ЦК КПСС, 1990. С. 19–33.
8. *Колин К. К.* Актуальные проблемы социальной информатики // Сб. научн. тр. «Социальная информатика-93». — М.: Институт молодежи, 1993. С. 5–18.
9. *Колин К. К.* Социальная информатика — научная база постиндустриального общества // Сб. научн. тр. «Социальная информатика-94». — М.: Институт молодежи, 1994. С. 4–23.
10. *Колин К. К.* Социальная информатика — новое направление научных исследований по комплексной проблеме «Информатика» // Системы и средства информатики, 1995. Вып. 7. С. 20–37.
11. *Колин К. К.* Наука для будущего: Социальная информатика // Информационные ресурсы России, 1995. № 3. С. 8–15.
12. *Колин К. К.* Становление информатики как фундаментальной науки и комплексной научной проблемы // Системы и средства информатики, 2006. Спец. вып. Научно-методологические проблемы информатики. С. 7–57.
13. *Колин К. К.* Информационный подход как фундаментальный метод познания // Межотраслевая информационная служба, 1998. № 1. С. 3–17.
14. *Колин К. К.* Философские проблемы информатики. — М.: БИНОМ, Лаборатория знаний, 2010. 264 с.
15. Политика в сфере образования и новые информационные технологии. Национальный доклад России // 2-й Междунар. конгресс ЮНЕСКО «Образование и информатика». Москва, 1996.
16. *Колин К. К.* Фундаментальные основы информатики: социальная информатика. — М.: Академический Проект, 2000. 350 с.
17. *Колин К. К.* Социальная информатика. — М.: Академический Проект, 2003. 432 с.
18. *Глушков В. М.* Основы безбумажной информатики. — М.: Наука, 1987. 552 с.

19. *Каныгин Ю. М., Калитич Г. И.* Основы теоретической информатики. — Киев: Наукова думка, 1990. 232 с.
20. *Коллин К. К.* Опыт изучения проблем информатики в Стэнфордском университете // Открытое образование, 2007. № 2. С. 52–63.
21. *Коллин К. К., Соколова И. В., Суслаков Б. А.* Опыт изучения проблем социальной информатики в системе образования России. Доклад на II Междунар. конгрессе ЮНЕСКО «Образование и информатика». Москва, 1996.
22. *Kolin K.* Social Informatics today and tomorrow: Status, problems and prospects of development of complex lines in the field of Science and Education // TripleC: Cognition, Communication, Co-operation. Int. Peer Reviewed Open Access J. Global Sustainable Inform. Soc., 2011. Vol. 9. No. 2. P. 460–465.
23. *Kolin K.* Philosophical problems in Informatics Science. — Beijing: China Social Sciences Press, 2012. 288 p. (На китайском яз.)
24. *Коллин К. К.* Информационная глобализация общества и гуманитарная революция // Сб. научн. тр. «Глобализация: синергетический подход». — М.: Изд-во РАГС, 2002. С. 323–334.
25. *Коллин К. К.* Качество жизни в информационном обществе // Человек и труд, 2010. № 1. С. 39–43.
26. *Соколов И. А., Коллин К. К.* Развитие информационного общества в России и актуальные проблемы информационной безопасности // Информационное общество, 2009. № 4–5. С. 98–106.
27. *Коллин К. К.* Модернизация России и стратегические приоритеты образования // Модернизация России: информационный, экономический, политический и социокультурный аспекты. — М.: Изд-во МосГУ, 2012. С. 3–16.
28. *Коллин К. К.* Философия и история образования. — М.: Изд-во МГСУ, 2003. 48 с.
29. *Коллин К. К.* Человек в информационном обществе: новые задачи для образования, науки и культуры // Открытое образование, 2007. № 5. С. 40–46.
30. *Коллин К. К.* Социальная информатика. Базовая модульная программа учебного курса для системы высшего образования. — М.: ИПИ РАН, 2001. 79 с.
31. *Коллин К. К., Урсул А. Д.* Информационная культурология: предмет и задачи нового научного направления. — Saarbrucken, Germany: LAP LAMBERT Academic Publishing, 2011. 249 с.
32. *Дашкевич В. С.* Великое культурное одичание: Арт-анализ. — М.: Russian Chess House, 2013. 720 с.
33. *Кошкин Р. П.* Культура в системе национальной и глобальной безопасности // Стратегические приоритеты, 2014. № 3. С. 3–5.
34. *Коллин К. К.* Системный кризис культуры: структура и содержание проблемы // Стратегические приоритеты, 2014. № 3. С. 6–27.
35. *Роговский Е. А.* Кибер-Вашингтон: глобальные амбиции. — М.: Межд. отношения, 2014. 848 с.
36. *Коллин К. К., Урсул А. Д.* Информация и культура. Введение в информационную культурологию. — М.: Стратегические приоритеты, 2015. 300 с.
37. *Коллин К. К.* Гуманитарные аспекты проблем национальной и международной безопасности // Стратегические приоритеты, 2015. № 1(5). С. 84–92.
38. *Коллин К. К.* Информационное пространство культуры: проблема многоязычия в информационном обществе // Вестник КемГУКИ, 2011. № 15. С. 8–17.

39. *Коллин К. К., Хорошилов А. А.* Проблема многоязычия в информационном обществе и новые интеллектуальные переводческие технологии // Информационное общество, 2012. № 2. С. 62–67.
40. *Кошкин Р. П.* Россия и мир: новые приоритеты геополитики. — М.: Стратегические приоритеты, 2015. 300 с.
41. *Коллин К. К.* Новая военная доктрина и гуманитарные приоритеты национальной безопасности России // Стратегические приоритеты, 2015. № 1. С. 30–47.

*Поступила в редакцию 14.08.15*

---

---

## SOCIAL INFORMATICS: RUSSIAN SCIENTIFIC SCHOOL AND FUTURE RESEARCH DIRECTIONS

*K. K. Kolin*

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** The paper describes the history of formation, present state, and prospects of development of Social Informatics as a current trend in science and education. The paper shows distinctive features of the Russian scientific school of Social Informatics and its priority in formation and development of this direction. The paper defines perspective directions of research in the conditions of formation of the Global Information Society. The paper explains the necessity of a broader consideration of Social Informatics in education and research training.

**Keywords:** informatics; information; information resources; information society; information security; social informatics

**DOI:** 10.14357/08696527150413

### References

1. Ursul, A. D. 1990. *Informatizatsiya obshchestva. Vvedenie v sotsial'nyuyu informatiku* [Informatization of society. Introduction to social informatics]. Moscow: AON pri TsK KPSS [AON at the Central Committee of CPSU]. 189 p.
2. Sokolov, A. V. 1988. *Informatsionnyy podkhod k dokumental'noy informatsii* [Information approach to documentary information]. Leningrad: LGIK. 85 p.
3. Kastel's, M. 2001. *Informatsionnaya epokha* [Information era]. Moscow: Izd-vo GU VShE [GU HSE Publishing House]. 606 p.
4. Kolin, K. K. 2005. *Innovatsionnoe razvitie v informatsionnom obshchestve i kachestvo obrazovaniya* [Innovative development in information society and quality of education]. *Otkrytoe Obrazovanie* [Open Education] 3:63–72.



5. Kolin, K. K. 2011. *Informatizatsiya obshchestva i globalizatsiya* [Informatization of society and globalization]. Krasnoyarsk: Izdatel'stvo SFU [SFU Publishing House]. 52 p.
6. Ursul, A. D. 1990. Informatizatsiya obshchestva i sotsial'naya informatika [Informatization of society and social informatics]. *Sb. nauchn. tr. "Sotsial'naya informatika"* [Collection of scientific works "Social informatics"]. Moscow: Vysshaya komsomol'skaya shkola pri TsK KPSS [The Higher Komsomol School at the Central Committee of CPSU]. 3–12.
7. Kolin, K. K. 1990. O strukture nauchnykh issledovaniy po kompleksnoy probleme "Informatika" [About structure of scientific researches on a complex problem of "Information scientist"]. *Sb. nauchn. tr. "Sotsial'naya informatika"* [Collection of scientific works "Social informatics"]. Moscow: Vysshaya komsomol'skaya shkola pri TsK KPSS [The higher Komsomol School at the Central Committee of CPSU]. 19–33.
8. Kolin, K. K. 1993. Aktual'nye problemy sotsial'noy informatiki [Actual problems of social informatics]. *Sb. nauchn. tr. "Sotsial'naya informatika-93"* [Collection of scientific works "Social informatics-93"]. Moscow: Institut molodezhi [Institute of Youth]. 5–18.
9. Kolin, K. K. 1994. Sotsial'naya informatika — nauchnaya baza postindustrial'nogo obshchestva [Social informatics — scientific base of postindustrial society]. *Sb. nauchn. tr. "Sotsial'naya informatika-94"* [Collection of scientific works "Social informatics-94"]. Moscow: Institut molodezhi [Institute of Youth]. 4–23.
10. Kolin, K. K. 1995. Sotsial'naya informatika — novoe napravlenie nauchnykh issledovaniy po kompleksnoy probleme "Informatika" [Social informatics — a new direction of scientific researches on a complex problem of "Information scientist"]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 7:20–37.
11. Kolin, K. K. 1995. Nauka dlya budushchego: Sotsial'naya informatika [Science for the future: Social informatics]. *Informatsionnye Resursy Rossii* [Information Resources of Russia] 3:8–15.
12. Kolin, K. K. 2006. Stanovlenie informatiki kak fundamental'noy nauki i kompleksnoy nauchnoy problemy [Formation of informatics as fundamental science and complex scientific problem]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics*. Spets. vyp. Nauchno-Metodologicheskie Problemy Informatiki [Special Issue. Scientific and Methodological Problems of Informatics] 7–57.
13. Kolin, K. K. 1998. Informatsionnyy podkhod kak fundamental'nyy metod poznaniya [Information approach as fundamental method of knowledge]. *Mezhotraslevaya Informatsionnaya Sluzhba* [Interindustry Information Service] 1:3–17.
14. Kolin, K. K. 2010. *Filosofskie problemy informatiki* [Philosophical problems of informatics]. Moscow: BINOM. 264 p.
15. Politika v sfere obrazovaniya i novye informatsionnye tekhnologii. Natsional'nyy doklad Rossii [Policy in education and new information technologies. National report of Russia]. 1996. *2-y Mezhdunar. kongress YuNESKO "Obrazovanie i informatika"* [2nd Congress (International) of UNESCO "Education and Informatics"]. Moscow.
16. Kolin, K. K. 2000. *Fundamental'nye osnovy informatiki: Sotsial'naya informatika* [Fundamental bases of informatics: Social informatics]. Moscow: Akademicheskii Proekt [Academic Project]. 350 p.
17. Kolin, K. K. 2003. *Sotsial'naya informatika* [Social informatics]. Moscow: Akademicheskii Proekt [Academic Project]. 432 p.

18. Glushkov, V. M. 1987. *Osnovy bezbumazhnoy informatiki* [Fundamentals of paperless informatics]. Moscow: Nauka. 552 p.
19. Kanygin, Yu. M., and G. I. Kalitich. 1990. *Osnovy teoreticheskoy informatiki* [Fundamentals of theoretical informatics]. Kiev: Naukova dumka [Scientific thought]. 232 p.
20. Kolin, K. K. 2007. Opyt izucheniya problem informatiki v Stendforskom universitete [Experience of research of problems of informatics at Stanford University]. *Otkrytoe Obrazovanie* [Open Education] 2:52–63.
21. Kolin, K. K., I. V. Sokolova, and B. A. Suslakov. 1996. Opyt izucheniya problem sotsial'noy informatiki v sisteme obrazovaniya Rossii [Experience of research of problems of social informatics in an education system of Russia]. *2-y Mezhdunar. kongress YuNESKO "Obrazovanie i informatika"* [2nd Congress (International) of UNESCO "Education and informatics"]. Moscow.
22. Kolin, K. 2011. Social Informatics today and tomorrow: Status, problems and prospects of development of complex lines in the field of Science and Education. *Triple-C: Cognition, Communication, Co-operation. Int. Peer Reviewed Open Access J. Global Sustainable Inform. Soc.* 9(2):460–465.
23. Kolin, K. 2012. *Philosophical problems in Informatics Science*. Beijing, China: Social Sciences Press. 288 p.
24. Kolin, K. K. 2002. Informatsionnaya globalizatsiya obshchestva i gumanitarnaya revolyutsiya [Information globalization of society and humanitarian revolution]. *Sb. nauchn. tr. "Globalizatsiya: Sinergeticheskiy podkhod"* [Collection of scientific works "Globalization: Synergetic approach"]. Moscow: Izd-vo RAGS [Publishing House of RAGS]. 323–334.
25. Kolin, K. K. 2010. Kachestvo zhizni v informatsionnom obshchestve [Quality of life in information society]. *Chelovek i Trud* [Person and Labour] 1:39–43.
26. Sokolov, I. A., and K. K. Kolin. 2009. Razvitie informatsionnogo obshchestva v Rossii i aktual'nye problemy informatsionnoy bezopasnosti [Development of information society in Russia and actual problems of information security]. *Informatsionnoe Obshchestvo* [Information Society] 4-5:98–106.
27. Kolin, K. K. 2012. Modernizatsiya Rossii i strategicheskie priority obrazovaniya [Modernization of Russia and strategic priorities of education]. *Sb. nauchn. tr. "Modernizatsiya Rossii: Informatsionnyy, ekonomicheskiy, politicheskiy i sotsiokul'turnyy aspekty"* [Collection of scientific works "Modernization of Russia: Information, economic, political, and sociocultural aspects"]. Moscow: Izd-vo MosGU [Publishing House of MOSGU]. 3–16.
28. Kolin, K. K. 2003. *Filosofiya i istoriya obrazovaniya* [Philosophy and history of education]. Moscow: Izd-vo MGSU [Publishing House of MGSU]. 48 p.
29. Kolin, K. K. 2007. Chelovek v informatsionnom obshchestve: Novye zadachi dlya obrazovaniya, nauki i kul'tury [The person in information society: New tasks for education, sciences, and culture]. *Otkrytoe Obrazovanie* [Open Education] 5:40–46.
30. Kolin, K. K. 2001. *Sotsial'naya informatika. Bazovaya modul'naya programma uchebnogo kursa dlya sistemy vysshego obrazovaniya* [Social informatics. Basic modular program of a training course for system of the higher education]. Moscow: Institute of Informatics Problems RAS. 79 p.
31. Kolin, K. K., and A. D. Ursul. 2011. *Informatsionnaya kul'turologiya: Predmet i zadachi novogo nauchnogo napravleniya* [Information cultural science: Subject and



- problems of the new scientific direction]. Saarbrucken, Germany: LAP LAMBERT Academic Publishing. 249 p.
32. Dashkevich, V. S. 2013. *Velikoe kul'turnoe odichanie: Art-analiz* [Great cultural running wild: Art-analysis]. Moscow: Russian Chess House. 720 p.
  33. Koshkin, R. P. 2014. Kul'tura v sisteme natsional'noy i global'noy bezopasnosti [Culture in system of national and global security]. *Strategicheskie Prioritety* [Strategic Priorities] 3:3–5.
  34. Kolin, K. K. 2014. Sistemnyy krizis kul'tury: Struktura i sodержanie problemy [System crisis of culture: Structure and maintenance of a problem]. *Strategicheskie Prioritety* [Strategic Priorities] 3:6–27.
  35. Rogovskiy, E. A. 2014. *Kiber-Vashington: Global'nye ambitsii* [Kiber-Washington: Global ambitions]. Moscow: Mezhdunarodnye otnosheniya [International Relations]. 848 p.
  36. Kolin, K. K., and A. D. Ursul. 2015. *Informatsiya i kul'tura. Vvedenie v informatsionnyu kul'turologiyu* [Information and culture. Introduction to information cultural science]. Moscow: Izd-vo Strategicheskie Prioritety [Publishing House Strategic Priorities]. 300 p.
  37. Kolin, K. K. 2015. Gumanitarnye aspekty problem natsional'noy i mezhdunarodnoy bezopasnosti [Humanitarian aspects of problems of the national and international security]. *Strategicheskie Prioritety* [Strategic Priorities] 1(5):84–92.
  38. Kolin, K. K. 2011. Informatsionnoe prostranstvo kul'tury: Problema mnogoyazychiya v informatsionnom obshchestve [Information space of culture: A multilingualism problem in information society]. *Vestnik KemGUKI* [Messenger of KEMGUKI] 15:8–17.
  39. Kolin, K. K., and A. A. Khoroshilov. 2012. Problema mnogoyazychiya v informatsionnom obshchestve i novye intellektual'nye perevodcheskie tekhnologii [Problem of multilingualism in information society and new intellectual translation technologies]. *Informatsionnoe Obshchestvo* [Information Society] 2:62–67.
  40. Koshkin, R. P. 2015. *Rossiya i mir: Novye prioritety geopolitiki* [Russia and world: New priorities of geopolitics]. Moscow: Izd-vo Strategicheskie Prioritety [Publishing House Strategic Priorities]. 300 p.
  41. Kolin, K. K. 2015. Novaya voennaya doktrina i humanitarnye prioritety natsional'noy bezopasnosti Rossii [New military doctrine and humanitarian priorities of national security of Russia]. *Strategicheskie Prioritety* [Strategic Priorities] 1:30–47.

Received August 14, 2015

## Contributor

**Kolin Konstantin K.** (b. 1935) — Doctor of Science in technology, professor, Honored scientist of RF, principal scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; kolinkk@mail.ru

## СИМВОЛЬНОЕ МОДЕЛИРОВАНИЕ И СРЕДСТВА ПОДДЕРЖКИ ИНТЕЛЛЕКТУАЛЬНОЙ ДЕЯТЕЛЬНОСТИ

*В. Д. Ильин*<sup>1</sup>

**Аннотация:** Рассмотрена роль символического моделирования произвольных объектов в становлении и развитии интеллектуальных способностей. Представлена точка зрения на проблему построения интеллектуальных роботов. Предложен подход к построению человеко-машинной среды поддержки интеллектуальной деятельности, методологическим основанием которого служит теория s-моделирования.

**Ключевые слова:** символическое моделирование; s-моделирование; интеллект; искусственный интеллект; интеллектуальные роботы

**DOI:** 10.14357/08696527150414

### 1 Введение

Изобретение символов (*жестовых, графических* и др.) и построенных из них символических моделей сообщений, представление и накопление таких моделей во внешней среде стали ключевыми средствами формирования и развития разумного человека [1]. Доминирующая роль символических моделей в интеллектуальной деятельности определяется не только их компактностью и выразительностью, но и тем, что не существует ограничений на типы носителей, применяемых для их хранения. Носителями могут быть память человека, бумажный лист, матрица цифровой фотокамеры, память цифрового диктофона или еще что-то. Затраты на построение, копирование, передачу, сохранение и накопление символических моделей несопоставимо меньше, чем аналогичные затраты, связанные с несимволическими моделями (например, макетами судов, зданий и др.). Примерами символических моделей могут служить видеоклипы, компьютерные программы, чертежи машин, записи музыкальных композиций, шахматных партий и т. д.

#### 1.1 Об эволюции символического моделирования

Видимо, до какого-то времени разнообразие моделируемых объектов ограничивалось тем, что принято называть объектами окружающей среды, и модели

---

<sup>1</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, vdilyin@ipiran.ru

этих объектов были физическими. Развитие *звуковых, жестовых и других средств символьного моделирования смыслов*, вызванное потребностями сообщать об опасности, о размещении объектов охоты и других объектах наблюдения, способствовало совершенствованию механизмов познания, взаимопонимания и обучения. Стали формироваться *языки сообщений*, включающие звуковые и жестовые символы. Стремление *моделировать поведение* (включая собственное) поставило новые задачи. Можно предположить, что изначально это стремление было связано с обучением рациональному поведению на охоте, в быту, при стихийных бедствиях. На определенном этапе задумались о создании таких средств моделирования, которые позволяли бы создавать *модели, допускающие их хранение, копирование и передачу*.

Этапным событием в развитии символьного моделирования стали *двумерные графические модели* (в виде рисунков) при уже освоенном изготовлении трехмерных (в виде лепных и резных фигурок). Особая роль принадлежит графическим моделям, обозначающим некоторые ситуации, свойства предметов и другие вещи, не имеющие видимых прообразов в окружающей среде. Переход от примитивных рисунков с натуры к изображениям того, что выдает сознание, приблизил изобретение *графических схем*. Это повлияло на развитие *жестозвуковых средств построения сообщений* и способствовало возникновению *речи*, ставшей важным средством создания и передачи сообщений.

Стремление повысить эффективность пояснений, сопровождающих показ, приводило к совершенствованию понятийного аппарата и средств его речевого воплощения. Развитие символьных моделей в виде графических схем и одновременное совершенствование речи привели к *графической модели речи*. Появилась *письменность*. Она стала не только важным этапом в становлении символьного моделирования, но и мощным инструментом развития интеллектуальной деятельности. Теперь описание объектов моделирования и связей между ними могли быть представлены композициями рисунков, схем и текстов. Была создана возможность фиксировать наблюдения, рассуждения и планы в виде *символьных моделей, которые можно было хранить и передавать*. С этого времени актуальными стали задачи изобретения *носителей*, инструментов для рисования и письма, красящих средств и др. Это были первые задачи на пути построения *среды символьного моделирования*.

Потребность в количественных оценках при обмене (охотничьей добычей, плодами земледелия, орудиями охоты и труда, изделиями ремесленников и т. д.) привела к *изобретению счета* и соответствующих систем жестовых, а затем и графических символов. Сначала количественные оценки, видимо, выражались с помощью жестовых символов (показом пальцев рук и др.). Когда жестовых символов стало не хватать, начали изобретать графические. Формирование понятия *числа* и идея экономии символов привела к изобретению *систем счисления*. Одной из них (двоичной) суждено было сыграть ключевую роль в изобретении *цифровой программируемой машины* и *числовом кодировании символических моделей*, реализуемых с помощью программируемых машин.

В [1] такие машины (суперкомпьютеры, планшеты, смартфоны и др.) были названы *s-машинами*, а символьные модели, реализуемые с помощью *s-машин*, — *s-моделями*.

Важный этап в графическом моделировании связан с *моделями схематических изображений* (прародителей чертежей) — основы проектирования. Представление проектируемого трехмерного объекта в трех двумерных проекциях, на которых показаны размеры и наименования деталей, сыграло решающую роль в развитии инженерного дела.

## 1.2 Выделение фрагментов текста статьи

Для выделения определений, замечаний, примеров, имен понятий и отдельных частей описания используются следующие средства:

□ (фрагмент описания) □ ≈ утверждение (определение) (здесь и далее символ ≈ заменяет слово «означает»);

◇ (фрагмент описания) ◇ ≈ замечание;

○ (фрагмент описания) ○ ≈ пример.

Курсивом выделены первые вхождения названий понятий (определяемых или определенных, последние могут быть гиперссылками) и фрагменты описания, к которым автор хочет привлечь внимание.

## 2 Интеллект и интуиция

□ *Интеллект* — комплекс способностей, обладатели которого умеют:

- выбирать цели;
- познавать себя и окружение;
- обучаться;
- формировать системы правил;
- решать задачи (распознавания образов, изобретения символьных моделей систем понятий и др.);
- изобретать искусственные усилители природных способностей человека (энергетические и другие машины, *s-среду* [1] и др.);
- действовать интуитивно (по обстановке, в условиях неполной информированности).

Этот список способностей не является исчерпывающим. □

Человек входит в систему *человечество*, связан с нею многими естественными и искусственными средствами взаимодействия. Его интеллектуальные способности развиваются и реализуются на основе знаний и умений, накопленных человечеством. Природный сенсорный комплекс (зрение, слух, осязание, обоняние, вкус) человека дополнен изобретенными сенсорами, рассчитанными на восприятие *визуальных, аудио-, тактильных и запаховых символов* [1]. Чело-

век с интеллектуальными способностями не только пользуется известными методами символьного моделирования сущностей, но и участвует в изобретении новых и совершенствовании существующих методов. В этой деятельности сформировалось и продолжает развиваться *абстрактное мышление*.

□ *Интуиция* — основание интеллекта, механизм изобретения моделей познаваемых сущностей, построения и сортировки гипотетических методов решения задач. □

Конечным продуктом интуиции является идея построения модели, постановки задачи или метода решения. Интуитивно решаемая задача может принадлежать произвольной предметной области и иметь любую степень сложности и определенности.

*Подсознание и сознание* — части комплекса, от функциональных характеристик которого зависят интеллектуальные способности. Подсознание, которому принадлежит основная часть работы при получении и сортировке интуитивных решений, играет ведущую роль и в поиске решений под контролем сознания. Продуктивность функционирования этого комплекса существенно зависит от символьного представления задач. Подсознание распоряжается хранилищами систем правил, определяющих поведение человека (в том числе и при решении задач), моделей систем понятий, механизмов интерпретации сообщений на моделях и других стратегических ресурсов. Подсознание и сознание можно представить в виде иерархии клиент-серверных архитектур. При этом масштабы и производительность всего, что относится к подсознанию, гораздо значительнее того, что относится к сознанию.

### 3 Об искусственном интеллекте

○ Рентгеновские аппараты, электронные микроскопы и радиотелескопы, авиалайнеры и космические корабли, радио- и телевизионные системы — примеры вещей, не имеющих прямых аналогов в среде обитания человека и в нем самом. Известно, что наиболее эффективные энергетические, коммуникационные и другие искусственные средства, удачно дополнившие естественные, изобретены не путем копирования естественных средств. Не являются подобной имитацией и *человеко-машинная среда символьного моделирования (s-среда [1])*, и *электронные сервисы*, реализованные на ее основе. ○

◇ Рассуждать об изобретении *искусственного интеллекта [2]*, не учитывая вышесказанное, — опрочечивое занятие. Продолжающееся связывание систем искусственного интеллекта с интерфейсом на естественном языке (имеется в виду неформализованный язык взаимодействия человека и s-машины) основано на ошибке, так как неформализованные языки не могут быть реализованы в s-среде [1]. ◇

Вслед за А. Тьюрингом многие продолжают задаваться вопросом: можно ли создать «думающую машину», поведение которой невозможно отличить от поведения разумного человека [3]?

Зададим встречные вопросы:

- какого человека (по интеллектуальному потенциалу, зависящему от способностей познавать, изобретать, обучаться, использовать информационные ресурсы и т. д.)?
- к каким предметным областям будут относиться тестовые вопросы?

У разных людей весьма различны способности:

- выбирать цели;
- разворачивать их в комплексы задач, решение которых позволяет достичь поставленной цели;
- формировать системы правил и следовать им;
- маскировать истинные намерения и распознавать подобный маскарад, применяемый другими, и т. д.

Полагаем, что для создания *s-машинного комплекса, имитирующего поведение человека с определенными интеллектуальными способностями*, необходимо изобрести, реализовать и объединить следующие *s-машинные системы*:

- восприятия внешних сообщений (по каналам, аналогичным человеческим: зрение, слух, осязание, обоняние, вкус);
- порождения и передачи внутренних сообщений, связанных с инстинктами самосохранения и др.;
- интерпретации внутренних и внешних сообщений;
- реагирования на результаты интерпретации сообщений;
- порождения целей и развертки их в комплексы задач.

И это далеко не исчерпывающий список того, что требуется для построения функциональной аналогии человеческого интеллекта.

### 3.1 О тесте А. Тьюринга

Полагаем, что в спецификациях подобных тестов необходимо указывать классы тестовых задач, которые предлагаются людям и «думающим машинам». Естественно, что люди, участвующие в тесте, должны быть подготовлены к решению задач этих классов.

Приведем пару поясняющих примеров.

○ Шахматные программы все чаще выигрывают даже у чемпионов мира. Любая из таких программ с успехом пройдет тест А. Тьюринга, если речь идет об обладателях способностей, необходимых для игры в шахматы. ○

○ Компания IBM создала суперкомпьютер Watson, который умеет отвечать на вопросы типа тех, что задают в известной ТВ-игре «Своя игра». Недавно Watson победил лучших «знатоков» США (в США эта игра называется Jeopardy). Успешно пройдет тест А. Тьюринга и суперкомпьютер Watson, если тестировать обладателей способностей отвечать на вопросы типа задаваемых в упомянутой игре. ○

### 3.2 О человеко-машинной среде интеллектуальной деятельности

Развитие s-среды [1], ядром которой в наши дни является Интернет, и реализуемых в ней электронных сервисов,  $\bigcirc$  таких как *интернет-сервисы планирования ресурсов* [4],  $\bigcirc$  — наиболее целесообразное направление совершенствования комплекса искусственных средств поддержки интеллектуальной деятельности. Результаты исследований механизма естественного интеллекта необходимы, прежде всего, как часть методологического обеспечения построения этого комплекса. Нарращивание функционала, удобства и надежности сервисов s-среды — постоянно актуальная задача. Ее решение во многом зависит от результативности исследований и разработок в области автоматизации программирования (создания языков, инструментальных систем программирования, связанных с системами знаний о программируемых задачах, и др.).

## 4 Об интеллектуальных роботах-партнерах

Естественный комплекс, обеспечивающий способности, отнесенные к интеллектуальным, включает не только мозг человека. Нет оснований полагать, что, имитируя этот комплекс, можно получить нечто эквивалентное ему по функциональности.

Полагаем, что роботы-партнеры — наиболее целесообразное семейство специализированных s-машинных помощников человека. Примерами роботов, предназначенных для решения хорошо определенных задач, могут служить шахматные машины. Гораздо сложнее изобрести и реализовать роботы, умеющие решать не только хорошо определенные задачи.

## 5 Заключение

Символьное моделирование, являясь средством описания смыслов, представленных системами понятий и знаний, не только сопровождает абстрактное мышление, но и служит инструментом его совершенствования. Компактность и выразительность символьных моделей позволяют эффективно сочетать детализацию и обобщение в процессе рассуждений. Символьные модели — испытанный инструментальный механизм ассоциаций, от продуктивности которого зависят судьбы изобретений и научных открытий.

Методология *символьного моделирования в s-среде (s-моделирования)* [1] является теоретическим основанием информатизации различных видов деятельности (научной, инженерной, экономической и др.).

## Литература

1. *Ильин А. В., Ильин В. Д.* Символьное моделирование в информатике. — М.: ИПИ РАН, 2011. 204 с. <https://smodeling.files.wordpress.com/2014/12/sm-v-informatike-2011.pdf>.
2. *McCarthy J.* Artificial intelligence, logic and formalizing common sense. — Stanford: Computer Science Department, Stanford University, 1990. 28 p.



3. Turing A. Computing machinery and intelligence // Mind, 1950. Vol. LIX. No. 236. P. 433–460.
4. Ильин А. В. Интернет-сервис планирования расходов // Системы и средства информатики, 2015. Т. 25. № 2. С. 111–122.

Поступила в редакцию 05.10.15

---

---

## SYMBOL MODELING AND MEANS OF SUPPORTING INTELLECTUAL ACTIVITY

V. D. Ilyin

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** The article describes the role of symbol modeling of arbitrary objects in forming and development of intellectual faculties. The point of view on the problem of building intelligent robots is presented. The approach to building the human-machine environment of supporting intellectual activity is suggested.

**Keywords:** symbol modeling; intelligence; artificial intelligence; intelligent robots

**DOI:** 10.14357/08696527150414

### References

1. Ilyin, A. V., and V. D. Ilyin. 2011. *Simvol'noe modelirovanie v informatike* [The symbol modeling in informatics]. Moscow: Institute of Informatics Problems of RAS. 204 p. Available at: <https://smodeling.files.wordpress.com/2014/12/sm-v-informatike-2011.pdf> (accessed September 15, 2015).
2. McCarthy, J. 1990. *Artificial intelligence, logic and formalizing common sense*. Stanford: Computer Science Department, Stanford University. 28 p. Available at: <http://www-formal.stanford.edu/jmc/ailogic.pdf> (accessed September 15, 2015).
3. Turing A. 1950. Computing machinery and intelligence. *Mind* LIX(236):433–460. Available at: <http://www.loebner.net/Prizef/TuringArticle.html> (accessed September 15, 2015).
4. Ilyin, A. V. 2015. Internet-servis planirovaniya raskhodov [The online service for cost planning]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 25(2):111–122. Available at: <http://www.mathnet.ru/links/bab8ebe7e21744d1809379c2e7b64bab/ssi410.pdf> (accessed September 15, 2015).

Received October 5, 2015

### Contributor

**Ilyin Vladimir D.** (b. 1937)— Doctor of Science in technology, Head of Laboratory, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilova Str., Moscow 119333, Russia; [vdilyin@ipiran.ru](mailto:vdilyin@ipiran.ru)



## НОРМАЛИЗАЦИЯ ПОКАЗАТЕЛЕЙ В СИСТЕМЕ ИНФОРМАЦИОННОГО МОНИТОРИНГА НАЦИОНАЛЬНОЙ БЕЗОПАСНОСТИ

*Г. В. Лукьянов<sup>1</sup>, Д. А. Никишин<sup>2</sup>, Г. Ф. Веревкин<sup>3</sup>, В. В. Косарик<sup>4</sup>*

**Аннотация:** Статья затрагивает проблематику получения обобщенных (интегральных) оценок на основе свертывания экспертами разнородных количественных и качественных показателей и разработки информационных технологий, поддерживающих процесс формирования обобщенной оценки в различных сферах деятельности. Предметом исследования является классификация и нормализация показателей национальной безопасности в интересах создания автоматизированной системы мониторинга, одной из задач которой является компьютерная поддержка процесса формирования экспертами обобщенной оценки на основе свертывания показателей. Показано, что нормализация показателей, включающая согласование единиц и масштабов измерения, а также использование единой системы форматов данных, во многом обеспечивает когнитивную интероперабельность совместной деятельности экспертов в процессе формирования обобщенной оценки и может служить основой для разработки информационных технологий, поддерживающих процесс формирования обобщенной оценки национальной безопасности.

**Ключевые слова:** информационные технологии; формирование обобщенной оценки; национальная безопасность; нормализация; информационный мониторинг; индикаторное оценивание; информационная система

**DOI:** 10.14357/08696527150415

### 1 Введение

Одной из основных задач системы информационного мониторинга национальной безопасности Российской Федерации (далее — НБРФ) является получение обобщенной оценки на основе пошагового свертывания экспертами разнородных количественных и качественных показателей, характеризующих состояние защищенности личности, общества и государства от внутренних и внешних угроз. Решение этой задачи необходимо для поддержки принятия решений по ключевым

<sup>1</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, gena-mslu@mail.ru

<sup>2</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, dmnik@a170.ipi.ac.ru

<sup>3</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, gennadij.verevckin2012@yandex.ru

<sup>4</sup>Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, valery@a\_170.ipi.ac.ru

вопросам национальной безопасности в рамках стратегического планирования на федеральном уровне [1, 2].

Конечный результат процесса формирования экспертами обобщенной оценки на основе свертывания разнородных количественных и качественных показателей определяется двумя основными факторами: значениями свертываемых показателей и применяемой методологией свертывания. Следовательно, обобщенная оценка должна сопровождаться достаточно подробным описанием методологии, раскрывающим ее влияние на полученный конечный результат. Например, результат свертывания одних только количественных показателей, относящихся к разным сферам оценивания национальной безопасности, будет зависеть от выбора весовых коэффициентов. В этом случае ухудшение показателей национальной безопасности в какой-либо сфере будет в разной степени влиять на формирование обобщенной оценки, так как будет зависеть от выбора весовых коэффициентов.

В процессе создания методологии свертывания необходимо учитывать, что исходные данные для формирования обобщенной оценки, поступающие из органов государственной статистики и иных ведомств, как правило, относятся к разным классификаторам и вычисляются в широком спектре абсолютных и относительных единиц измерения.

С учетом принадлежности к разным классификаторам, различий в применяемых единицах и масштабах измерения, методология свертывания должна предусматривать приведение показателей к единой системе согласованных единиц измерения.

Еще одна особенность методологии свертывания вытекает из применения качественных показателей, значения которых могут относиться к разным номинативным шкалам, и необходимости представления в численной форме, если обобщенная оценка является количественной, или приведения их к единой номинативной шкале, если обобщенная оценка является качественной. К таким показателям, например, относятся индикаторы оценивания дипломатической деятельности, положительные, либо негативные последствия которой для национальной безопасности трудно оценить численно.

Данная статья посвящена лишь отдельным сторонам методологии свертывания, касающимся вопросов приведения показателей к единой системе согласованных единиц измерения и применения качественных показателей.

## **2 Сопоставление единиц измерения в системах информационного мониторинга**

Обычно для решения задач автоматизированного сбора, обработки и представления информации применяется единый классификатор объектов, т. е. систематизированный перечень подлежащих учету объектов реального или абстрактного мира с указанием единиц измерения, масштабов, форматов, правил

формирования производных единиц измерения и порядка вывода относительных (безразмерных) единиц измерения. Такой классификатор в общем случае должен удовлетворять следующим требованиям:

- полноте — все требуемые для заполнения поля содержат значения в установленных единицах измерения, вычисленные по унифицированным методикам;
- непротиворечивости — данные в разных полях и записях не противоречат друг другу;
- унификации — формат представления един для всех однотипных данных;
- идентифицируемости — каждая позиция в классификаторе уникальна и может быть определена однозначно, что к тому же предполагает отсутствие дублей;
- актуальности — в любой момент времени такой классификатор должен содержать данные, отвечающие целям и задачам обработки информации в информационной системе.

В Стратегии НБРФ до 2020 г. определены следующие сферы, которые подлежат мониторингу и оцениванию в аспекте национальной безопасности [1]:

- экономическая;
- социальная;
- науки и образования;
- международная;
- духовная;
- информационная;
- военная;
- оборонно-промышленная;
- экологическая;
- общественная.

В этом документе для оценивания состояния дел в каждой сфере и в целом на национальном уровне определены следующие обобщенные показатели состояния национальной безопасности:

- уровень безработицы (доля от экономически активного населения);
- децильный коэффициент (ДК) (соотношение доходов 10% наиболее и 10% наименее обеспеченного населения);
- уровень роста потребительских цен;
- уровень государственного внешнего и внутреннего долга в процентном отношении к валовому внутреннему продукту (ВВП);
- уровень обеспеченности ресурсами здравоохранения, культуры, образования и науки в процентном отношении к ВВП;

- уровень ежегодного обновления вооружения, военной и специальной техники;
- уровень обеспеченности военными и инженерно-техническими кадрами.

Для сбора данных и их формализованной обработки с целью получения значений таких разнородных показателей в процессе информационного мониторинга и оценивания национальной безопасности необходима разработка формата представления обрабатываемых данных. В основе этого формата должны лежать унифицированные и сопоставимые единицы измерения исходных, промежуточных и итоговых данных. Однако в настоящее время разработка такого формата остается сложной задачей. Ее сложность демонстрирует, например, сопоставление следующих показателей экономической сферы [3, 4]:

- уровень безработицы (доля от экономически активного населения, %);
- ВВП на душу населения (руб. или долл. США);
- производительность труда (объем ВВП по отношению к численности населения, занятого в экономике, тыс. руб./чел.);
- уровень инфляции (%).

Даже этот простой пример демонстрирует, насколько трудно «совместить» разные показатели в процессе вычисления обобщенной оценки только в пределах одной экономической сферы, так как приходится сопоставлять проценты и денежные единицы.

Отдельного внимания заслуживает в процессе сопоставления выбор денежной единицы и ее номинала, так как рубль в этом качестве вряд ли может служить мерой оценки хотя бы при определении уровня доходов населения Российской Федерации и иных показателей, связанных с ВВП<sup>1</sup>.

В ряде случаев для оценки некоторых аспектов национальной безопасности применяются такие единицы измерения, которые являются относительными (безразмерными), но трудно сопоставимыми. Пример подобных данных, отражающих заболеваемость по разным классам болезней и входящих в социальную сферу, представлен в табл. 1.

Во-первых, из приведенных в табл. 1 данных следует, что заболеваемость по разным классам болезней отличается друг от друга как минимум на целый порядок. Во-вторых, эти относительные единицы все же являются семантически разнородными, а именно: относительное количество вполне определенных классов заболеваний (у больных с диагнозом, установленным впервые в жизни).

---

<sup>1</sup>Покупательная способность рубля несколько раз изменялась скачкообразно в десятки раз. Например, в 1961 г. она была искусственно увеличена в десять раз. При этом если стоимость доллара до «деноминации» составляла 4 рубля, то после реформы американская валюта стала стоить не 40 копеек, а 90, т. е. реальная покупательная способность рубля упала более чем в 2 раза. Сильный инфляционный скачок в нашей стране произошел в 1990-х гг. Кроме того, если инфляция в США не превышает 0,4%, то в России этот показатель устойчиво «переваливает» за 10%, т. е. в 25 раз выше. Такое положение дел с рублем, в отличие от доллара, не позволяет взять в качестве исходной точки некоторый год и к нему приравнивать цены в рублях с учетом инфляции.

**Таблица 1** Сведения о заболеваемости

Классы заболеваний	Заболеваемость на 1000 чел. населения				
	1995 г.	1996 г.	1997 г.	1998 г.	1999 г.
Инфекционные и паразитные	121,5	108,1	92,2	110,5	127,5
Нервной системы	251,0	258,9	266,1	285,8	277,4
Крови и кроветворных тканей	8,8	11,1	11,6	13,2	16,0

Таким образом, в интересах разработки информационных технологий, поддерживающих процесс формирования обобщенной оценки в различных сферах деятельности, необходимо предварительно создать методологию свертывания семантически разнородных количественных и качественных показателей. При этом должны использоваться следующие классификаторы [5]:

- классификатор административно-территориальных единиц Российской Федерации<sup>1</sup>;
- классификатор временных периодов<sup>2</sup>;
- классификаторы семантических значений, каждый из которых представляет собой таблицу сопоставления численных значений позициям номинативных шкал<sup>3</sup>.

### 3 Методология нормализации данных

При определении обобщенной оценки НБРФ происходит свертывание сотен разнородных показателей, поступающих из десятков различных ведомств и организаций. Принимая во внимание различие способов представления данных и применяемых единиц измерения, перед свертыванием необходимо привести показатели к единой системе единиц измерения, что позволит осуществлять их сопоставление.

Из проведенного анализа статистических данных, предоставляемых органами государственной статистики, следует, что значительная часть показателей вычисляется в относительных единицах [6]. Чаще всего речь идет об отношении некоторой абсолютной величины (в числителе) к общему числу оцениваемых этой

<sup>1</sup>Например, в соответствии с Общероссийским классификатором объектов административно-территориального деления ОК 019-95 (ОКАТО) и Общероссийским классификатором стран мира (ОКСМ) ОК (МК (ИСО 3166) 004-97) 025-2001, а также ГОСТом 7.67-2003 «Коды для представления названий стран» и аутентичным ему ISO 3166-88, устанавливающими буквенные и цифровые обозначения названий стран в кодированной форме, единые для различных систем обработки информации, ее хранения и обмена.

<sup>2</sup>Перечисляющий все задействованные в системе временные периоды таких стандартных периодов, как месяц, квартал, год и т. п.

<sup>3</sup>Например, 0%–35% — «низко», 35%–65% — «средне», 65%–100% — «высоко».

**Таблица 2** Примеры показателей и их нормализации

Показатель	Формула вычисления (единицы измерения)	Коэффициент нормализации
1. ВВП на душу населения	Количество произведенного ВВП (выраженное, например, в рублях), отнесенное к численности населения	Численность населения в момент вычисления показателя
2. Число персональных компьютеров (отнесенное к числу сотрудников)	Число единиц компьютеров, приходящихся на 100 сотрудников	100
3. Уровень аварийности на автотранспорте (на дорогах различной ведомственной принадлежности)	Число дорожно-транспортных происшествий различной категории на 100 км уличной сети	100

величиной объектов, например численности населения, транспортных средств, километражу дорожной сети и т. п. (в знаменателе). Часто эту величину делят на соответствующий коэффициент (на 100 сотрудников, на 100 км уличной сети, 1000 единиц автотранспорта и т. д.), т. е. осуществляют ее нормализацию. Примеры таких показателей представлены в табл. 2.

Из приведенных в табл. 2 примеров видно, что значения различных показателей в исходной форме могут быть не сопоставимы между собой в процессе свертывания с целью получения обобщенной оценки. Так, объем оказанной населению стационарной медицинской помощи представлен из расчета на 1000 чел., а относительный уровень заболеваемости — на 100 тыс. чел., т. е. коэффициенты нормализации этих показателей отличаются друг от друга в сто раз. Еще более сложная ситуация возникает в процессе свертывания при попытке сопоставить уровень аварийности на автотранспорте на дорогах и в населенных пунктах (так как первая величина отнесена к километражу, а вторая — к числу транспортных средств).

Из приведенных примеров следует, что в системе информационного мониторинга НБРФ наряду с единой формой представления должны быть предусмотрены функции нормализации свертываемых величин, т. е. приведения показателей к некоторой единой шкале измерения.

В первом примере преобразование можно выполнить путем приведения величин оценок показателей к одному «знаменателю» (например, из расчета на 1 чел.). Если первая задача решается достаточно просто, то в случае с уровнями аварийности на дорогах и в населенных пунктах необходимо обращение к абсолютным величинам этих показателей и последующее их отнесение к некоторому единому знаменателю (к общему километражу дорог/улиц, к количеству транспортных средств или иному единому выбранному знаменателю).

#### 4 Унифицированное представление показателей

Принимая во внимание изложенные выше обстоятельства, для автоматизированной обработки данных в системах информационного мониторинга целесообразно найти способ преобразования исходных значений показателей в единообразно нормированные единицы (например, проценты). Применение таких единиц обеспечит универсальность и вычислений, и оценки.

Такую возможность можно продемонстрировать на примере показателей, приведенных в табл. 3, первые два из которых представляют собой относительные показатели, выраженные в размерных единицах, а ДК представляет собой безразмерную величину, характеризующую совокупность упорядоченных по величине значений и представляющую собой отношение 10%-ных частей этой совокупности, соответствующих областям максимальных и минимальных значений.

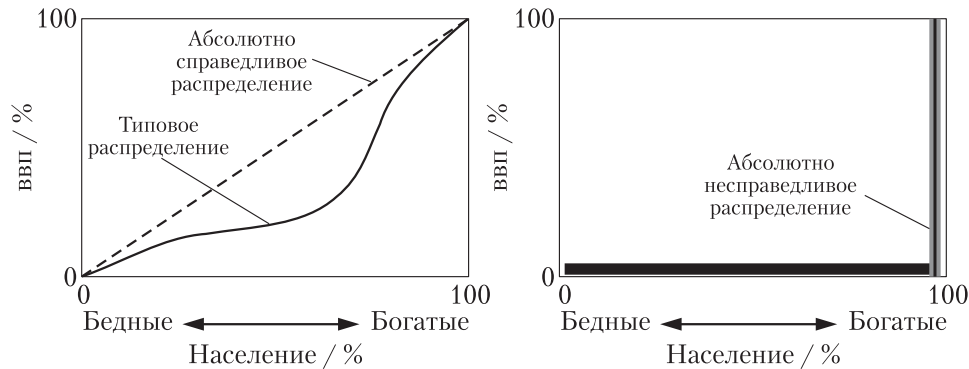
Первые два показателя в табл. 3 можно нормализовать в проценты, рассчитывая их по отношению к некоторому стандартному (целевому, желательному) значению. Например, в соответствии с «Концепцией долгосрочного социально-экономического развития Российской Федерации на период до 2020 года» в качестве целевого значения ВВП на душу населения по паритету покупательной способности рассматривается уровень, достигнутый в государствах — членах Организации экономического сотрудничества и развития [7]. Такое представление ВВП на душу населения на заданный временной период будет с максимальной объективностью отражать реальное состояние дел в социально-экономической сфере по этому важнейшему показателю и демонстрировать степень достижения целевых значений.

Для нормализации ДК целесообразно учесть его сущность и предназначение, а также варианты распределения национального богатства (ВВП) между гражданами страны, которые графически демонстрирует рисунок.

Если допустить, что идеалом в борьбе с социальным неравенством является абсолютно справедливое распределение национального богатства, т. е. его разделение поровну между всеми гражданами страны, как это представлено прямой

**Таблица 3** Примеры показателей

Показатель	Единица измерения
1. Относительная производительность труда в экономике	Объем ВВП по отношению к численности населения, занятого в экономике (рублей или долларов на человека)
2. Доля ВВП на душу населения	Объем ВВП по отношению к численности всего населения (рублей или долларов на человека)
3. Децильный коэффициент	Соотношение доходов 10% наиболее и 10% наименее обеспеченного населения



Распределение национального богатства

линией на левом графике рисунка, то такое значение ДК, равное «единице», при нормализации следует принять за 100%. Другая крайняя форма распределения национального богатства (ВВП), т. е. абсолютно несправедливое, состоит в том, что все богатство страны принадлежит одному человеку, а всем остальным не достается ничего, как это представлено на правом графике рисунка<sup>1</sup>. Такое значение ДК при нормализации следует принять за 0%.

Таким образом, нормализованное значение ДК будет лежать в пределах от 0% до 100%.

## 5 Формализация представления качественных показателей

Самую большую трудность с точки зрения аналитической обработки представляют все качественные показатели, значения которых принадлежат той или иной номинативной шкале, например относящиеся к международной сфере. Любая номинативная оценка представляет значительную трудность в процессе свертывания с целью получения обобщенной оценки, так как базируется на вербализованном субъективном мнении экспертов [8, 9].

Вместе с тем, чтобы использовать эти номинативные оценки, можно использовать, например, десятибалльную шкалу<sup>1</sup> для их представления, которая позволит вербальным характеристикам показателя назначить количественные значения.

Пример такой номинативной шкалы для показателя «Результативность мероприятий по формированию позитивного образа Российской Федерации, созда-

<sup>1</sup>Такое гипотетическое распределение национального богатства никак не противоречит смыслу ДК. Так, если к одному самому богатому человеку добавить необходимое число бедняков, то по формальным признакам получится ДК.

<sup>1</sup>В федеральных органах исполнительной власти США предпочтение отдается пятибалльной шкалам оценки подобных показателей (т. е. с 25%-ным шагом). Представленная же здесь шкала (с 10%-ным шагом) фактически содержит 11 позиций, включая нулевое значение, что позволяет более точно отслеживать динамику показателя.



**Таблица 4** Пример шкалы оценки показателя

Значение	Описание содержания показателя
10 (100%)	На международном культурно-информационном пространстве сформирован исключительно положительный образ Российской Федерации, который не вызывает сомнения
9 (90%)	На международном культурно-информационном пространстве сформирован исключительно положительный образ Российской Федерации, который в 10% случаев подвергается критике (по данным независимого мониторинга отечественных и зарубежных средств массовой информации и только по бытовым вопросам
8 (80%)	На международном культурно-информационном пространстве сформирован исключительно положительный образ Российской Федерации, который в 20% случаев подвергается критике и только по профессиональным и/или бытовым вопросам
7 (70%)	На международном культурно-информационном пространстве сформирован положительный образ Российской Федерации, который в 30% случаев подвергается критике (возможны любые аспекты)
6 (60%)	На международном культурно-информационном пространстве сформирован в целом положительный образ Российской Федерации, который подвергается критике в 40% случаев, т. е. положительное отношение превалирует
5 (50%)	На международном культурно-информационном пространстве сформирован баланс отношений к образу Российской Федерации, характеризующийся равенством уровня положительных и негативных оценок
4 (40%)	На международном культурно-информационном пространстве баланс отношений к образу Российской Федерации нарушен в негативную сторону, т. е. негативное отношение превалирует
3 (30%)	На международном культурно-информационном пространстве на всех уровнях сложился отрицательный образ Российской Федерации. Однако имеется 30% положительного отношения, возможности диалога и выбора области взаимного интереса
2 (20%)	На международном культурно-информационном пространстве на всех уровнях сложился преимущественно отрицательный образ Российской Федерации, имеется 20% положительного отношения с ограниченными возможностями для диалога с ограниченной областью взаимного интереса
1 (10%)	На международном культурно-информационном пространстве на всех уровнях сложился преимущественно отрицательный образ Российской Федерации, создающий серьезные препятствия для диалога и поиска точек взаимного интереса
0 (0%)	На международном культурно-информационном пространстве на всех уровнях сложилось исключительно враждебное отношение к Российской Федерации, которое практически исключает любой диалог и взаимодействие

ваемого на международном культурно-информационном пространстве» приведен в табл. 4.

Разработка подобных таблиц по каждому показателю международной сферы, естественно, может вестись только совместно со специалистами по вопросам международной деятельности и обеспечения национальной безопасности [10].

## 6 Заключение

Рассмотренные в статье вопросы нормализации исходных данных и значений показателей представляют собой только начальный этап разработки методологии свертывания. В результате проведенных исследований были разработаны отдельные модели и методы организации и проведения информационного мониторинга. Полученные результаты позволяют уточнить постановку проблемы создания методологии свертывания.

Основной итог проведенных исследований заключается в следующем: нормализация исходных данных и значений показателей является краеугольным камнем методологии свертывания, так как обеспечивает формирование единого информационного пространства в интересах информационного мониторинга НБРФ за счет применения одних и тех же классификаторов, единиц измерения и типов данных.

## Литература

1. Стратегия национальной безопасности Российской Федерации до 2020 года. Утверждена Указом Президента Российской Федерации от 12.05.2009 № 537.
2. О стратегическом планировании в Российской Федерации: Федеральный закон Российской Федерации от 28 июня 2014 г. № 172-ФЗ.
3. Лукьянов Г. В., Никишин Д. А., Веревкин Г. Ф. Информационное обеспечение мониторинга национальной безопасности в региональном разрезе // Системы и средства информатики, 2014. Т. 24. № 2. С. 193–204.
4. Лукьянов Г. В., Никишин Д. А., Веревкин Г. Ф., Косарик В. В. Специфика показателей национальной безопасности в контексте ее информационного мониторинга // Системы и средства информатики, 2014. Т. 24. № 4. С. 186–205.
5. Зеленков М. Ю. Теоретико-методологические проблемы теории национальной безопасности Российской Федерации. — М.: Юридический ин-т МИИТа, 2013. 196 с.
6. Национальная безопасность России в оценках экспертов: Аналитический отчет по результатам экспертного опроса (версия 2.4 от 28.0.2011). — М.: Институт социологии РАН, иссл. группа ЦИРКОН, 2010. 56 с.
7. Концепция долгосрочного социально-экономического развития Российской Федерации на период до 2020 года. Утверждена Распоряжением Правительства Российской Федерации от 17.11.2008 № 1662-р.
8. Зацман И. М., Дурново А. А. Моделирование процессов формирования экспертных знаний для мониторинга программно-целевой деятельности // Информатика и ее применения, 2011. Т. 5. Вып. 4. С. 59–73.

9. *Зацман И. М., Дурново А. А.* Разработка и применение программно-ориентированных индикаторов в сфере науки // Системы и средства информатики, 2012. Т. 22. Вып. 1. С. 110–120.
10. *Зацаринный А. А., Киселев Э. В.* Некоторые подходы к формированию нормативно-технической базы для создания единого информационного пространства России // Системы и средства информатики, 2014. Т. 24. № 4. С. 206–220.

*Поступила в редакцию 29.09.15*

---

---

## INDICES NORMALIZATION PROBLEMS IN THE NATIONAL SECURITY MONITORING SYSTEM

*G. V. Lukyanov, D. A. Nikishin, G. F. Verevkin, and V. V. Kosarik*

Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation

**Abstract:** The article deals with the normalization of indices of the national security of the Russian Federation in the interest of creating an automated system of monitoring and evaluation of this complex activity. It is shown that the normalization of indicators, including their classification, harmonization of units and scales of measurement, and conversion to unified formats of data provides cognitive interoperability and serves as a basis for adequate, objective, and reliable generalized assessment of the national security of the Russian Federation. The article also describes the features of relative indices and options of their creation. It is demonstrated that the use of relative measures avoids distortion in the development of a common statistical pattern and calculating a generalized evaluation. The article describes a new approach to the formalization of qualitative indicators that allows objective analysis and applications in the automated information system.

**Keywords:** national security; normalization; information monitoring; indicator evaluation; information systems; units and scales of measurement; relative indices

**DOI:** 10.14357/08696527150415

### References

1. Strategiya natsional'noy bezopasnosti Rossiyskoy Federatsii do 2020 goda. Utverzhdena Ukazom Prezidenta Rossiyskoy Federatsii ot 12.05.2009 № 537 [Strategy of national security of the Russian Federation till 2020, No. 537. Approved by the Decree of the President of the Russian Federation of 12.05.2009].
2. Federal'nyy zakon Rossiyskoy Federatsii ot 28 iyunya 2014 g. No. 172-FZ “O strategicheskom planirovanii v Rossiyskoy Federatsii” [The Federal Law of the Russian

- Federation of June 28, 2014, No. 172-FZ “About strategic planning in the Russian Federation”].
3. Luk'yanov, G. V., D. A. Nikishin, and G. F. Verevkin. 2014. Informatsionnoe obespechenie monitoringa natsional'noy bezopasnosti v regional'nom razreze [Information support of monitoring of national security in a regional aspect]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 24(2):193–203.
  4. Luk'yanov, G. V., D. A. Nikishin, G. F. Verevkin, and V. V. Kosarik. 2014. Spetsifika pokazateley natsional'noy bezopasnosti v kontekste ee informatsionnogo monitoringa [Specificity of indicators of national security in the context of its information monitoring]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 24(4):186–205.
  5. Zelenkov, M. Yu. 2013. *Teoretiko-metodologicheskie problemy teorii natsional'noy bezopasnosti Rossiyskoy Federatsii* [Theoretical and methodological problems of the theory of national security of the Russian Federation]. — Moscow: MIIT's Legal institute. 196 p.
  6. *Natsional'naya bezopasnost' Rossii v otsenkakh ekspertov* [National security of Russia in estimates of experts]. 2010. Analiticheskiy otchet po rezul'tatam ekspertnogo oprosa (versiya 2.4 ot 28.0.2011). [The analytical report on results of expert poll (version 2.4 from 28.0.2011)]. — Moscow: Institute of sociology of the Russian Academy of Sciences, CIRCON research group. 56 p.
  7. Kontseptsiya dolgosrochnogo sotsial'no-ekonomicheskogo razvitiya Rossiyskoy Federatsii na period do 2020 goda. Utverzhdena Rasporyazheniem Pravitel'stva Rossiyskoy Federatsii ot 17.11.2008 № 1662-r [The concept of long-term social and economic development of the Russian Federation for the period till 2020, No. 1662-r. Approved by the Order of the Government of the Russian Federation of 17.11.2008].
  8. Zatsman, I. M., and A. A. Durnovo. 2011. Modelirovanie protsessov formirovaniya ekspertnykh znaniy dlya monitoringa programmno-tselevoy deyatel'nosti [Modeling of processes of formation of expert knowledge for monitoring of program and target management]. *Informatika i ee Primeneniya — Inform. Appl.* 5(4):59–73.
  9. Zatsman, I. M., and A. A. Durnovo. 2012. Razrabotka i primeneniye programmno-orientirovannykh indikatorov v sfere nauki [Development and use of the program focused indicators in the sphere of science]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 22(1):110–120.
  10. Zatsarinny, A. A., and E. V. Kisilev. 2014. Nekotorye podkhody k formirovaniyu normativno-tekhnicheskoy bazy dlya sozdaniya edinogo informatsionnogo prostranstva Rossii [Some approaches to formation of normative and technical base for creation of a uniform information space of Russia]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 24(4):206–220.

Received September 29, 2015

## Contributors

**Lukyanov Gennady V.** (b. 1952) — PhD, associate professor, Head of Laboratory, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; gena-mslu@mail.ru

**Nikishin Dmitry A.** (b. 1976) — Candidate of Science (PhD) in technology, Head of Laboratory, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; dmnikishin@mail.ru

**Verevkin Gennady F.** (b. 1934) — Candidate of Science (PhD) in technology, senior scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; gennadij.verevckin2012@yandex.ru

**Kosarik Valerii V.** (b. 1970) — scientist, Institute of Informatics Problems, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; valery@a\_170.ipi.ac.ru

**Бородин Вячеслав Васильевич** (р. 1957) — кандидат технических наук, доцент кафедры инфокоммуникации Московского авиационного института (национальный исследовательский университет)

**Быстров Игорь Иванович** (р. 1931) — доктор технических наук, профессор, заведующий отделом Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Веревкин Геннадий Федорович** (р. 1934) — кандидат технических наук, старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Гончаров Алексей Владимирович** (р. 1995) — студент Московского физико-технического института

**Грушо Александр Александрович** (р. 1946) — доктор физико-математических наук, профессор, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Дьяченко Юрий Георгиевич** (р. 1958) — кандидат технических наук, старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Забезайло Михаил Иванович** (р. 1956) — кандидат физико-математических наук, доцент, старший научный сотрудник Всероссийского института научной и технической информации Российской академии наук (ВИНИТИ РАН)

**Захарова Татьяна Валерьевна** (р. 1962) — кандидат физико-математических наук, старший преподаватель кафедры математической статистики факультета вычислительной математики и кибернетики МГУ имени М. В. Ломоносова; старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Зацаринный Александр Алексеевич** (р. 1951) — доктор технических наук, профессор, заместитель директора Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Зейфман Александр Израилевич** (р. 1954) — доктор физико-математических наук, профессор, заведующий кафедрой прикладной математики Вологодского

государственного университета; старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук, главный научный сотрудник ИСЭРТ РАН

**Ильин Владимир Дмитриевич** (р. 1937) — доктор технических наук, заведующий лабораторией Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Киселев Эдуард Васильевич** (р. 1938) — старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Колин Константин Константинович** (р. 1935) — доктор технических наук, профессор, заслуженный деятель науки РФ, главный научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Коновалов Михаил Григорьевич** (р. 1950) — доктор технических наук, заведующий сектором Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Корепанов Эдуард Рудольфович** (р. 1966) — кандидат технических наук, заведующий сектором Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Королев Виктор Юрьевич** (р. 1954) — доктор физико-математических наук, профессор, заведующий кафедрой математической статистики факультета вычислительной математики и кибернетики МГУ имени М. В. Ломоносова; ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Коротышева Анна Владимировна** (р. 1988) — кандидат физико-математических наук, доцент Вологодского государственного университета; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Косарик Валерий Валентинович** (р. 1970) — научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Косолапов Юрий Владимирович** (р. 1982) — кандидат технических наук, доцент Института математики, механики и компьютерных наук им. И. И. Воровича Южного федерального университета

**Лукьянов Геннадий Викторович** (р. 1952) — кандидат военных наук, доцент, заведующий сектором Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Морозов Николай Викторович** (р. 1956) — старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Никишин Дмитрий Александрович** (р. 1976) — кандидат технических наук, заведующий сектором Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Николаев Андрей Владимирович** (р. 1973) — кандидат физико-математических наук, научный сотрудник Института химической физики им. Н. Н. Семенова Российской академии наук

**Писковский Виктор Олегович** (р. 1963) — кандидат физико-математических наук, ведущий программист Некоммерческого партнерства «Центр прикладных исследований компьютерных сетей» (НП «ЦПИ КС»)

**Поздняков Алексей Владимирович** (р. 1992) — студент Института математики, механики и компьютерных наук им. И. И. Воровича Южного федерального университета

**Попова Мария Сергеевна** (р. 1994) — студентка Московского физико-технического института

**Радоманов Сергей Иванович** (р. 1946) — научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Разумчик Ростислав Валерьевич** (р. 1984) — кандидат физико-математических наук, старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук; доцент Российского университета дружбы народов

**Растрелин Евгений Анатольевич** (р. 1972) — заместитель генерального директора ЗАО «РАМЭК-ВС»

**Сатин Яков Александрович** (р. 1978) — доктор физико-математических наук, доцент Вологодского государственного университета; Институт проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Синицын Владимир Игоревич** (р. 1968) — доктор физико-математических наук, доцент, заведующий отделом Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Синицын Игорь Николаевич** (р. 1940) — доктор технических наук, профессор, заслуженный деятель науки РФ, заведующий отделом Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Степченков Юрий Афанасьевич** (р. 1951) — кандидат технических наук, заведующий отделом Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук



**Стрижов Вадим Викторович** (р. 1967) — доктор физико-математических наук, ведущий научный сотрудник Вычислительного центра им. А. А. Дородницына Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Тарасов Борис Васильевич** (р. 1946) — доктор технических наук, профессор, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Хилько Дмитрий Владимирович** (р. 1987) — научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Хорошилов Александр Алексеевич** (р. 1952) — доктор технических наук, профессор, ведущий научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Шикунов Юрий Игорьевич** (р. 1995) — стажер-исследователь Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Щемирова Анна Алексеевна** (р. 1995) — студентка кафедры математической статистики факультета вычислительной математики и кибернетики МГУ имени М. В. Ломоносова

**Щенявская Елена Викторовна** (р. 1993) — студентка магистратуры кафедры математической статистики факультета вычислительной математики и кибернетики МГУ имени М. В. Ломоносова

---

## АВТОРСКИЙ УКАЗАТЕЛЬ ЗА 2015 г.

---

	№	Стр.
<b>Адамович И. М., Волков О. И.</b> Система извлечения биографических фактов из текстов исторической направленности	3	235
<b>Андреев А. М., Березкин Д. В., Козлов И. А., Симаков К. В.</b> Многокритериальный метод выявления нечетких дубликатов в потоке текстовых сообщений	1	34
<b>Белоусов В. В.</b> см. Сеницын И. Н.		
<b>Белоусов В. В.</b> см. Сеницын И. Н.		
<b>Березкин Д. В.</b> см. Андреев А. М.		
<b>Бородин В. В., Растрелин Е. А.</b> Некоторые вопросы оценки надежности корпоративной информационной сети с учетом управления резервированием ее элементов	4	150
<b>Босов А. В., Гресс Е. С., Наумов А. В.</b> Об одном подходе к автоматизации оценивания успеваемости студентов	2	123
<b>Быстров И. И., Тарасов Б. В., Хорошилов А. А., Радоманов С. И.</b> Основы применения онтологии и компьютерной лингвистики при проектировании перспективных автоматизированных информационных систем	4	128
<b>Веревкин Г. Ф.</b> см. Лукьянов Г. В.		
<b>Веревкин Г. Ф.</b> см. Лукьянов Г. В.		
<b>Воинов Н. В., Дробинцев П. Д., Котляров В. П., Никифоров И. В.</b> Анализ покрытия UCM-модели тестовыми сценариями	1	74
<b>Воинов Н. В., Дробинцев П. Д., Котляров В. П., Никифоров И. В., Селин И. А.</b> Генерация тестового набора на основе потока управления	1	54
<b>Волков О. И.</b> см. Адамович И. М.		
<b>Газизуллина Р. К., Стенина М. М., Стрижов В. В.</b> Прогнозирование объемов железнодорожных грузоперевозок по парам веток	1	142
<b>Гилязов Р. Р., Грушо А. А.</b> Проблемы взаимодействия вредоносного кода и программ защиты в архитектуре современных операционных систем	3	94
<b>Гончаров А. В., Попова М. С., Стрижов В. В.</b> Метрическая классификация временных рядов с выравниванием относительно центроидов классов	4	52
<b>Гресс Е. С.</b> см. Босов А. В.		
<b>Григорьев С. В., Рагозина А. К.</b> Обобщенный табличный LL-анализ	1	89

	№	Стр.
<b>Гридин В. Н., Карпухин Е. О., Евдокимов И. А.</b> Оценка эффективности способа управления доставкой сетевых пакетов на стороне получателя	3	150
<b>Гринченко С. Н.</b> см. Соколов И. А.		
<b>Грушо А. А., Грушо Н. А., Тимонина Е. Е., Шоргин С. Я.</b> Возможности построения безопасной архитектуры для динамически изменяющейся информационной системы	3	78
<b>Грушо А. А., Забежайло М. И., Зацаринный А. А., Николаев А. В., Писковский В. О.</b> Архитектура стенда для экспериментального исследования моделей, алгоритмов и решений по обеспечению информационной безопасности в облачных вычислительных средах	4	65
<b>Грушо А. А.</b> см. Гилязов Р. Р.		
<b>Грушо Н. А.</b> см. Грушо А. А.		
<b>Долгополов В. С.</b> см. Заикин М. Ю.		
<b>Дробинцев П. Д.</b> см. Воинов Н. В.		
<b>Дробинцев П. Д.</b> см. Воинов Н. В.		
<b>Дулин С. К., Розенберг И. Н., Уманский В. И.</b> Географические стандарты для обеспечения геоинтероперабельности	3	109
<b>Дьяченко Ю. Г.</b> см. Хилько Д. В.		
<b>Евдокимов И. А.</b> см. Гридин В. Н.		
<b>Забежайло М. И.</b> см. Грушо А. А.		
<b>Заикин М. Ю., Долгополов В. С., Обухова О. Л., Соловьев И. В.</b> Технология предотвращения дублирования библиографических описаний в базе данных научных публикаций БИАС ИПИ РАН	1	168
<b>Зарядов И. С., Мейханаджян Л. А., Милованова Т. А., Разумчик Р. В.</b> Метод нахождения стационарного распределения очереди в двухканальной системе с упорядоченным входом конечной емкости	3	44
<b>Захаров В. Н.</b> см. Соколов И. А.		
<b>Захарова Т. В., Королев В. Ю., Щемирова А. А.</b> Исследование особенностей сигналов миограммы	4	91
<b>Захарова Т. В.</b> см. Щенявская Е. В.		
<b>Зацаринный А. А., Ионенков Ю. С.</b> К вопросу оценки эффективности автоматизированных систем с использованием метода анализа иерархий	3	161
<b>Зацаринный А. А., Киселев Э. В.</b> Некоторые подходы к формированию нормативно-технической базы в части требований к архитектурному построению информационных систем организаций — участников единого информационного пространства России	3	179

	№	Стр.
<b>Зацаринный А. А., Киселев Э. В.</b> Некоторые подходы к формированию нормативно-технической базы единого информационного пространства России в части информационных ресурсов	1	155
<b>Зацаринный А. А., Киселев Э. В.</b> Некоторые подходы к формированию обобщенной архитектуры информационных систем организаций — участников единого информационного пространства России	4	114
<b>Зацаринный А. А.</b> см. Грушо А. А.		
<b>Зацаринный А. А.</b> см. Соколов И. А.		
<b>Зацаринный А. А.</b> см. Соколов И. А.		
<b>Зацман И. М.</b> см. Соколов И. А.		
<b>Зейфман А. И.</b> см. Разумчик Р. В.		
<b>Ильин А. В.</b> Интернет-сервис планирования расходов	2	111
<b>Ильин В. Д.</b> Нормализованные деньги в системе имущественных статусов	3	206
<b>Ильин В. Д.</b> Символьное моделирование и средства поддержки интеллектуальной деятельности	4	194
<b>Илюшин Г. Я., Лиманский В. И.</b> Построение системы управления потоками пациентов	1	186
<b>Ионенков Ю. С.</b> см. Зацаринный А. А.		
<b>Карнаков В. В.</b> см. Карпухин Е. О.		
<b>Карпухин Е. О., Карнаков В. В., Филипков С. В.</b> Специализированная программная модель для имитации процесса передачи данных на транспортном уровне с использованием алгоритмов управления перегрузкой	2	71
<b>Карпухин Е. О.</b> см. Гридин В. Н.		
<b>Кириков И. А., Колесников А. В., Листопад С. В., Румовская С. Б.</b> Виртуальные гетерогенные коллективы, поддерживающие принятие решений	3	126
<b>Кириков И. А., Колесников А. В., Листопад С. В., Солдатов С. А.</b> Гетерогенные интеллектуальные компьютерные системы поддержки принятия решений: модели координации и согласованности	2	96
<b>Киселев Э. В.</b> см. Зацаринный А. А.		
<b>Киселев Э. В.</b> см. Зацаринный А. А.		
<b>Киселев Э. В.</b> см. Зацаринный А. А.		
<b>Козлов И. А.</b> см. Андреев А. М.		
<b>Колесников А. В.</b> см. Кириков И. А.		
<b>Колесников А. В.</b> см. Кириков И. А.		
<b>Колин К. К.</b> Социальная информатика: российская научная школа и перспективные направления исследований	4	175

	№	Стр.
<b>Коновалов М. Г., Разумчик Р. В.</b> Приближенная оптимизация стратегии распределения вычислительных ресурсов на примере системы андеррайтинга	4	31
<b>Кореньков В. В., Нечаевский А. В., Ососков Г. А., Пряхина Д. И., Трофимов В. В., Ужинский А. В.</b> Моделирование грид и облачных сервисов как важный этап их разработки	1	4
<b>Корепанов Э. Р.</b> см. Сеницын И. Н.		
<b>Корепанов Э. Р.</b> см. Сеницын И. Н.		
<b>Корепанов Э. Р.</b> см. Сеницын И. Н.		
<b>Корепанов Э. Р.</b> см. Сеницын И. Н.		
<b>Королев В. Ю.</b> см. Захарова Т. В.		
<b>Коротышева А. В.</b> см. Разумчик Р. В.		
<b>Корчагин А. Ю.</b> О сходимости случайных сумм независимых случайных векторов к многомерным обобщенным дисперсионным гамма-распределениям	1	127
<b>Косарик В. В.</b> см. Лукьянов Г. В.		
<b>Косарик В. В.</b> см. Лукьянов Г. В.		
<b>Косолапов Ю. В., Поздняков А. В.</b> Оценка стойкости кодового зашумления в задаче распределенного хранения данных	4	158
<b>Котляров В. П.</b> см. Воинов Н. В.		
<b>Котляров В. П.</b> см. Воинов Н. В.		
<b>Кружков М. Г.</b> Информационные ресурсы контрастивных лингвистических исследований: типологические базы данных	1	198
<b>Кружков М. Г.</b> Информационные ресурсы контрастивных лингвистических исследований: электронные корпуса текстов	2	140
<b>Кудрявцев А. А.</b> Программный комплекс анализа байесовских моделей в теории массового обслуживания и надежности	3	251
<b>Кузнецов В. В.</b> Новый метод получения устойчивого ключа из динамической биометрической подписи	2	85
<b>Леонтьев Н. Д., Ушаков В. Г.</b> Исследование систем обслуживания с дискретным временем, входящим потоком авторегрессионного типа и обратной связью	2	60
<b>Лиманский В. И.</b> см. Илюшин Г. Я.		
<b>Листопад С. В.</b> см. Кириков И. А.		
<b>Листопад С. В.</b> см. Кириков И. А.		
<b>Лукашевич Н. В., Четверкин И. И.</b> Комбинирование тезаурусных и корпусных знаний для извлечения оценочных слов	1	20
<b>Лукьянов Г. В., Никишин Д. А., Веревкин Г. Ф., Косарик В. В.</b> Нормализация показателей в системе информационного мониторинга национальной безопасности	4	201

- Лукьянов Г. В., Никишин Д. А., Веревкин Г. Ф., Косарик В. В.**  
Нормативные и методологические аспекты организации информационного мониторинга национальной безопасности 3 218
- Мейханаджян Л. А.** см. Зарядов И. С.
- Милованова Т. А.** см. Зарядов И. С.
- Морозов Н. В.** см. Хилько Д. В.
- Наумов А. В.** см. Босов А. В.
- Нечаевский А. В.** см. Кореньков В. В.
- Никифоров И. В.** см. Воинов Н. В.
- Никифоров И. В.** см. Воинов Н. В.
- Никишин Д. А.** см. Лукьянов Г. В.
- Никишин Д. А.** см. Лукьянов Г. В.
- Николаев А. В.** см. Грушо А. А.
- Обухова О. Л.** см. Заикин М. Ю.
- Ососков Г. А.** см. Кореньков В. В.
- Писковский В. О.** см. Грушо А. А.
- Поздняков А. В.** см. Косолапов Ю. В.
- Попова М. С., Стрижов В. В.** Построение нейронных сетей глубокого обучения для классификации временных рядов 3 60
- Попова М. С.** см. Гончаров А. В.
- Пряхина Д. И.** см. Кореньков В. В.
- Рагозина А. К.** см. Григорьев С. В.
- Радоманов С. И.** см. Быстров И. И.
- Разумчик Р. В., Зейфман А. И., Коротышева А. В., Сатин Я. А.**  
Анализ энергоэффективности вычислительного комплекса, моделируемого с помощью системы обслуживания с пороговым управлением и интенсивностями, зависящими от времени 4 19
- Разумчик Р. В.** см. Зарядов И. С.
- Разумчик Р. В.** см. Коновалов М. Г.
- Растрелин Е. А.** см. Бородин В. В.
- Розенберг И. Н.** см. Дулин С. К.
- Румовская С. Б.** см. Кириков И. А.
- Сатин Я. А.** см. Разумчик Р. В.
- Селин И. А.** см. Воинов Н. В.
- Сергеев И. В.** см. Сеницын И. Н.
- Сергеев И. В.** см. Сеницын И. Н.
- Сибиряков П. Г.** см. Соколов И. А.
- Симаков К. В.** см. Андреев А. М.
- Сеницын В. И.** см. Сеницын И. Н.
- Сеницын В. И.** см. Сеницын И. Н.
- Сеницын В. И.** см. Сеницын И. Н.

	№	Стр.
<b>Синицын И. Н.</b> Методы моментов в задачах аналитического моделирования распределений в нелинейных стохастических системах на многообразиях	3	24
<b>Синицын И. Н.</b> Применение ортогональных разложений для аналитического моделирования многомерных распределений в нелинейных стохастических системах на многообразиях	3	4
<b>Синицын И. Н., Корепанов Э. Р.</b> Синтез устойчивых линейных фильтров и экстраполяторов Пугачёва для стохастических систем с мультипликативными широкополосными шумами	1	108
<b>Синицын И. Н., Сергеев И. В., Синицын В. И., Корепанов Э. Р., Белоусов В. В., Шоргин В. С.</b> Математическое обеспечение синтеза дискретных фильтров Пугачёва для обработки нормальных процессов в эрдитарных стохастических системах	2	20
<b>Синицын И. Н., Синицын В. И., Корепанов Э. Р.</b> Аналитическое моделирование нормальных процессов в стохастических системах со сложными дробно-рациональными нелинейностями	4	4
<b>Синицын И. Н., Синицын В. И., Сергеев И. В., Корепанов Э. Р., Белоусов В. В., Шоргин В. С.</b> Математическое обеспечение моделирования нормальных процессов в стохастических системах со сложными иррациональными нелинейностями	2	3
<b>Соколов И. А., Гринченко С. Н., Зацаринный А. А., Зацман И. М., Сибиряков П. Г.</b> К 80-летию заслуженного деятеля науки Российской Федерации К. К. Колина	2	160
<b>Соколов И. А., Зацаринный А. А., Захаров В. Н.</b> О вкладе академика И. А. Мизина в теорию и практику создания отечественных информационно-телекоммуникационных систем: к 80-летию со дня рождения	1	213
<b>Солдатов С. А.</b> см. Кириков И. А.		
<b>Соловьев И. В.</b> см. Заикин М. Ю.		
<b>Стенина М. М.</b> см. Газизуллина Р. К.		
<b>Степченков Ю. А.</b> см. Хилько Д. В.		
<b>Стрижов В. В.</b> см. Газизуллина Р. К.		
<b>Стрижов В. В.</b> см. Гончаров А. В.		
<b>Стрижов В. В.</b> см. Попова М. С.		
<b>Сучков А. П.</b> Некоторые подходы к интеграции аналитических данных существующих и перспективных систем поддержки принятия решений	3	195
<b>Тарасов Б. В.</b> см. Быстров И. И.		
<b>Тимонина Е. Е.</b> см. Грушо А. А.		
<b>Трофимов В. В.</b> см. Кореньков В. В.		

<b>Ужинский А. В.</b> см. Кореньков В. В.		
<b>Уманский В. И.</b> см. Дулин С. К.		
<b>Ушаков В. Г.</b> см. Леонтьев Н. Д.		
<b>Филипков С. В.</b> см. Карпухин Е. О.		
<b>Хилько Д. В., Степченко Ю. А., Дьяченко Ю. Г., Шикунов Ю. И., Морозов Н. В.</b> Аппаратно-программное моделирование и тестирование рекуррентного операционного устройства	4	78
<b>Хорошилов А. А.</b> см. Быстров И. И.		
<b>Четверкин И. И.</b> см. Лукашевич Н. В.		
<b>Шикунов Ю. И.</b> см. Хилько Д. В.		
<b>Шоргин В. С.</b> см. Сеницын И. Н.		
<b>Шоргин В. С.</b> см. Сеницын И. Н.		
<b>Шоргин С. Я.</b> см. Грушо А. А.		
<b>Щемирова А. А.</b> см. Захарова Т. В.		
<b>Щенявская Е. В., Захарова Т. В.</b> Метод синхронизации сигналов магнитоэнцефалограмм и миограмм	4	101



---

## 2015 AUTHOR INDEX

---

	<b>No.</b>	<b>Page</b>
<b>Adamovich I. M. and Volkov O. I.</b> The system of facts extraction from historical texts	3	235
<b>Andreev A. M., Berezkin D. V., Kozlov I. A., and Simakov K. V.</b> Multicriteria method for detecting near-duplicates in a stream of text messages	1	34
<b>Belousov V. V.</b> see Sinitsyn I. N.		
<b>Belousov V. V.</b> see Sinitsyn I. N.		
<b>Berezkin D. V.</b> see Andreev A. M.		
<b>Borodin V. V. and Rastrelin E. A.</b> Some aspects of reliability assessment of the corporative data network with consideration of redundancy management	4	150
<b>Bosov A. V., Gress E. S., and Naumov A. V.</b> An approach for assessment of automatization of students academic performance	2	124
<b>Bystrov I. I., Tarasov B. V., Khoroshilov A. A., and Radomanov S. I.</b> The application basis of ontology and computational linguistics in the design of advanced automated information systems	4	128
<b>Chetviorkin I. I.</b> see Loukachevitch N. V.		
<b>Chshenyavskaya E. V. and Zakharova T. V.</b> Method of synchronization of magnetoencephalograms and myograms signals	4	101
<b>Diachenko Yu. G.</b> see Khilko D. V.		
<b>Dolgopolov V. S.</b> see Zaikin M. Yu.		
<b>Drobintsev P. D.</b> see Voinov N. V.		
<b>Drobintsev P. D.</b> see Voinov N. V.		
<b>Dulin S. K., Rozenberg I. N., and Umanskiy V. I.</b> Geographical standards to ensure geointeroperability	3	109
<b>Evdokimov I. A.</b> see Gridin V. N.		
<b>Filipkov S. V.</b> see Karpukhin E. O.		
<b>Gazizullina R. K., Medvednikova M. M., and Strijov V. V.</b> Large capacity of railway cargo transportation forecasting	1	142
<b>Giliazov R. R. and Grusho A. A.</b> Problems of interaction of the malicious code and protection programs in architecture of modern operating systems	3	94
<b>Goncharov A. V., Popova M. S., and Strijov V. V.</b> Metric time series classification using dynamic warping relative to centroids of classes	4	52

	<b>No.</b>	<b>Page</b>
<b>Gress E. S.</b> see Bosov A. V.		
<b>Gridin V. N., Karpukhin E. O., and Evdokimov I. A.</b> Evaluating the effectiveness of method for controlling the delivery of network packets on receiver	3	150
<b>Grigorev S. V. and Ragozina A. K.</b> Generalized table-based LL-parsing	1	89
<b>Grusho A. A., Grusho N. A., Timonina E. E., and Shorgin S. Ya.</b> Possibilities of secure architecture creation for dynamically changing information systems	3	78
<b>Grusho A. A., Zabezhailo M. I., Zatsarinny A. A., Nikolaev A. V., and Piskovski V. O.</b> Architecture of the stand for the pilot study of models, algorithms, and solutions of information security in cloud computing environments	4	65
<b>Grusho A. A.</b> see Giliazov R. R.		
<b>Grusho N. A.</b> see Grusho A. A.		
<b>Ilushin G. Y. and Limansky V. I.</b> Creation of patient flows control system	1	186
<b>Ilyin A. V.</b> The online service for cost planning	2	112
<b>Ilyin V. D.</b> Normalized money in the system of property statuses	3	206
<b>Ilyin V. D.</b> Symbol modeling and means of supporting intellectual activity	4	194
<b>Ionenkov Y. S.</b> see Zatsarinny A. A.		
<b>Karnacov V. V.</b> see Karpukhin E. O.		
<b>Karpukhin E. O., Karnacov V. V., and Filipkov S. V.</b> Specialized software model to simulate data transfer in transport level with the congestion control algorithm	2	72
<b>Karpukhin E. O.</b> see Gridin V. N.		
<b>Khilko D. V., Stepchenkov Yu. A., Diachenko Yu. G., Shikunov Yu. I., and Morozov N. V.</b> Hardware and software modeling and testing of the recurrent operational device	4	78
<b>Khoroshilov A. A.</b> see Bystrov I. I.		
<b>Kirikov I. A., Kolesnikov A. V., Listopad S. V., and Rumovskaya S. B.</b> Virtual heterogeneous collectives for supporting decision-making	3	126
<b>Kirikov I. A., Kolesnikov A. V., Listopad S. V., and Soldatov S. A.</b> Heterogeneous intelligent computer decision support systems: Models of coordination and consistency	2	97
<b>Kiselev E. V.</b> see Zatsarinny A. A.		
<b>Kiselev E. V.</b> see Zatsarinny A. A.		
<b>Kiselev E. V.</b> see Zatsarinny A. A.		
<b>Kolesnikov A. V.</b> see Kirikov I. A.		
<b>Kolesnikov A. V.</b> see Kirikov I. A.		

	<b>No.</b>	<b>Page</b>
<b>Kolin K. K.</b> Social Informatics: Russian scientific school and future research directions	4	175
<b>Konovalov M. G. and Razumchik R. V.</b> Approximate optimization of resource allocation strategy: The case of bank underwriting system	4	31
<b>Korchagin A. Yu.</b> On convergence of random sums of independent random vectors to multivariate generalized variance-gamma distributions	1	127
<b>Korenkov V. V., Nechaevskiy A. V., Ososkov G. A., Pryakhina D. I., Trofimov V. V., and Uzhinskiy A. V.</b> Grid and cloud services simulation as an important step of their development	1	4
<b>Korepanov E. R.</b> see Sinitsyn I. N.		
<b>Korepanov E. R.</b> see Sinitsyn I. N.		
<b>Korepanov E. R.</b> see Sinitsyn I. N.		
<b>Korepanov E. R.</b> see Sinitsyn I. N.		
<b>Korolev V. Yu.</b> see Zakharova T. V.		
<b>Korotysheva A. V.</b> see Razumchik R. V.		
<b>Kosarik V. V.</b> see Lukyanov G. V.		
<b>Kosarik V. V.</b> see Lukyanov G. V.		
<b>Kosolapov Yu. V. and Pozdnyakov A. V.</b> Evaluation of resistance of code noising in the distributed data storage	4	158
<b>Kotlyarov V. P.</b> see Voinov N. V.		
<b>Kotlyarov V. P.</b> see Voinov N. V.		
<b>Kozlov I. A.</b> see Andreev A. M.		
<b>Kruzhkov M. G.</b> Information resources for contrastive studies: Electronic text corpora	2	141
<b>Kruzhkov M. G.</b> Information resources for contrastive studies: Typological databases	1	198
<b>Kudryavtsev A. A.</b> Software complex for analysis of Bayesian models in queueing and reliability theories	3	251
<b>Kuznetsov V. V.</b> Novel method for robust key generation from online biometric signature	2	86
<b>Leontyev N. D. and Ushakov V. G.</b> A study of queueing systems with discrete time, autoregressive arrivals, and feedback	2	61
<b>Limansky V. I.</b> see Ilushin G. Y.		
<b>Listopad S. V.</b> see Kirikov I. A.		
<b>Listopad S. V.</b> see Kirikov I. A.		
<b>Loukachevitch N. V. and Chetviorkin I. I.</b> Combining corpus and thesaurus information for extracting sentiment words	1	20
<b>Lukyanov G. V., Nikishin D. A., Verevkin G. F., and Kosarik V. V.</b> Indices normalization problems in the national security monitoring system	4	201

	No.	Page
<b>Lukyanov G. V., Nikishin D. A., Verevkin G. F., and Kosarik V. V.</b> Normative and methodological aspects of information monitoring organization of national security	3	218
<b>Medvednikova M. M.</b> see Gazizullina R. K.		
<b>Meykhanadzhyan L. A.</b> see Zaryadov I. S.		
<b>Milovanova T. A.</b> see Zaryadov I. S.		
<b>Morozov N. V.</b> see Khilko D. V.		
<b>Naumov A. V.</b> see Bosov A. V.		
<b>Nechaevskiy A. V.</b> see Korenkov V. V.		
<b>Nikiforov I. V.</b> see Voinov N. V.		
<b>Nikiforov I. V.</b> see Voinov N. V.		
<b>Nikishin D. A.</b> see Lukyanov G. V.		
<b>Nikishin D. A.</b> see Lukyanov G. V.		
<b>Nikolaev A. V.</b> see Grusho A. A.		
<b>Obuhova O. L.</b> see Zaikin M. Yu.		
<b>Ososkov G. A.</b> see Korenkov V. V.		
<b>Piskovski V. O.</b> see Grusho A. A.		
<b>Popova M. S. and Strijov V. V.</b> Building superposition of deep learning neural networks for solving the problem of time series classification	3	60
<b>Popova M. S.</b> see Goncharov A. V.		
<b>Pozdnyakov A. V.</b> see Kosolapov Yu. V.		
<b>Pryakhina D. I.</b> see Korenkov V. V.		
<b>Radomanov S. I.</b> see Bystrov I. I.		
<b>Ragozina A. K.</b> see Grigorev S. V.		
<b>Rastrelin E. A.</b> see Borodin V. V.		
<b>Razumchik R. V., Zeifman A. I., Korotysheva A. V., and Satin Ya. A.</b> Energy efficiency estimation of the computational cluster using a queueing system with threshold control and time-dependent service and arrival rates	4	19
<b>Razumchik R. V.</b> see Konovalov M. G.		
<b>Razumchik R. V.</b> see Zaryadov I. S.		
<b>Rozenberg I. N.</b> see Dulin S. K.		
<b>Rumovskaya S. B.</b> see Kirikov I. A.		
<b>Satin Ya. A.</b> see Razumchik R. V.		
<b>Schemirova A. A.</b> see Zakharova T. V.		
<b>Selin I. A.</b> see Voinov N. V.		
<b>Sergeev I. V.</b> see Sinitsyn I. N.		
<b>Sergeev I. V.</b> see Sinitsyn I. N.		
<b>Shikunov Yu. I.</b> see Khilko D. V.		
<b>Shorgin S. Ya.</b> see Grusho A. A.		
<b>Shorgin V. S.</b> see Sinitsyn I. N.		
<b>Shorgin V. S.</b> see Sinitsyn I. N.		

	<b>No.</b>	<b>Page</b>
<b>Simakov K. V.</b> see Andreev A. M.		
<b>Sinitsyn I. N.</b> Applications of orthogonal expansions for analytical modeling of multidimensional distributions in stochastic systems on manifold	3	4
<b>Sinitsyn I. N.</b> Moments methods for analytical modeling of stochastic systems on manifolds	3	24
<b>Sinitsyn I. N. and Korepanov E. R.</b> Synthesis of stable linear Pugachev filters and extrapolators for stochastic systems with wide band multiplicative noises	1	108
<b>Sinitsyn I. N., Sergeev I. V., Sinitsyn V. I., Korepanov E. R., Belousov V. V., and Shorgin V. S.</b> Software for synthesis of discrete Pugachev filters for normal processes in hereditary stochastic systems	2	21
<b>Sinitsyn I. N., Sinitsyn V. I., and Korepanov E. R.</b> Analytical modeling of normal processes in stochastic systems with complex fraction-rational nonlinearities	4	4
<b>Sinitsyn I. N., Sinitsyn V. I., Sergeev I. V., Korepanov E. R., Belousov V. V., and Shorgin V. S.</b> Mathematical software for normal processes modeling in stochastic systems with complex irrational nonlinearities	2	4
<b>Sinitsyn V. I.</b> see Sinitsyn I. N.		
<b>Sinitsyn V. I.</b> see Sinitsyn I. N.		
<b>Sinitsyn V. I.</b> see Sinitsyn I. N.		
<b>Sokolov I. A., Zatsarinny A. A., and Zakharov V. N.</b> About academician I. A. Mizin's contribution to theory and practice of domestic information-telecommunication systems creation: To the 80th anniversary	1	213
<b>Soldatov S. A.</b> see Kirikov I. A.		
<b>Soloviev I. V.</b> see Zaikin M. Yu.		
<b>Stepchenkov Yu. A.</b> see Khilko D. V.		
<b>Strijov V. V.</b> see Gazizullina R. K.		
<b>Strijov V. V.</b> see Goncharov A. V.		
<b>Strijov V. V.</b> see Popova M. S.		
<b>Suchkov A. P.</b> Some approaches to analytical data integration of existing and prospective decision support systems	3	195
<b>Tarasov B. V.</b> see Bystrov I. I.		
<b>Timonina E. E.</b> see Grusho A. A.		
<b>Trofimov V. V.</b> see Korenkov V. V.		
<b>Umanskiy V. I.</b> see Dulin S. K.		
<b>Ushakov V. G.</b> see Leontyev N. D.		
<b>Uzhinskiy A. V.</b> see Korenkov V. V.		
<b>Verevkin G. F.</b> see Lukyanov G. V.		
<b>Verevkin G. F.</b> see Lukyanov G. V.		

	<b>No.</b>	<b>Page</b>
<b>Voinov N. V., Drobintsev P. D., Nikiforov I. V., and Kotlyarov V. P.</b> Analysis of UCM-model coverage by test scenarios	1	74
<b>Voinov N. V., Drobintsev P. D., Nikiforov I. V., Kotlyarov V. P., and Selin I. A.</b> Control flow based test suite generation	1	54
<b>Volkov O. I.</b> see Adamovich I. M.		
<b>Zabezhaiko M. I.</b> see Grusho A. A.		
<b>Zaikin M. Yu., Dolgoplov V. S., Obuhova O. L., and Soloviev I. V.</b> Technology for prevention of duplication of bibliographic descriptions in the scientific database BIAS IPI RAS	1	168
<b>Zakharov V. N.</b> see Sokolov I. A.		
<b>Zakharova T. V., Korolev V. Yu., and Schemirova A. A.</b> Signal myogram features research	4	91
<b>Zakharova T. V.</b> see Chshenyavskaya E. V.		
<b>Zaryadov I. S., Meykhanadzhyan L. A., Milovanova T. A., and Razumchik R. V.</b> On the method of calculating the stationary distribution in the finite two-channel system with ordered input	3	44
<b>Zatsarinny A. A. and Ionenkov Y. S.</b> Regarding automated systems efficiency evaluation using analytic hierarchy process	3	161
<b>Zatsarinny A. A. and Kiselev E. V.</b> Regarding some approaches to creation of generalized architecture of Russian information space members of information systems	4	114
<b>Zatsarinny A. A. and Kiselev E. V.</b> Regarding some approaches to creation of the regulatory and technical base concerning requirements to architectural development of information systems of Russian information space members	3	179
<b>Zatsarinny A. A. and Kiselev E. V.</b> Some approaches to forming the regulatory and technical base for the unified information space of Russia in the field of information resources	1	155
<b>Zatsarinny A. A.</b> see Grusho A. A.		
<b>Zatsarinny A. A.</b> see Sokolov I. A.		
<b>Zeifman A. I.</b> see Razumchik R. V.		

---

## Правила подготовки рукописей статей для публикации в журнале «Системы и средства информатики»

---

Журнал «Системы и средства информатики» публикует теоретические, обзорные и дискуссионные статьи, посвященные научным исследованиям и разработкам в области информационных технологий.

Журнал издается на русском языке. По специальному решению редколлегии отдельные статьи могут печататься на английском языке.

Тематика журнала охватывает следующие направления:

- информационно-телекоммуникационные системы и средства их построения;
- архитектура и программное обеспечение вычислительных машин, комплексов и сетей;
- методы и средства защиты информации.

1. В журнале печатаются статьи, содержащие результаты, ранее не опубликованные и не предназначенные к одновременной публикации в других изданиях.

Публикация не должна нарушать закон об авторских правах.

Направляя рукопись в редакцию, авторы сохраняют все права собственников данной рукописи и при этом передают учредителям и редколлегии неисключительные права на издание статьи на русском языке (или на языке статьи, если он отличен от русского) и на ее распространение в России и за рубежом. Авторы должны предоставить в редакцию письмо в следующей форме:

**Соглашение о передаче права на публикацию:**

*«Мы, нижеподписавшиеся, авторы рукописи «. . .», передаем учредителям и редколлегии журнала «Системы и средства информатики» неисключительное право опубликовать данную рукопись статьи на русском языке как в печатной, так и в электронной версиях журнала. Мы подтверждаем, что данная публикация не нарушает авторских прав других лиц или организаций, а также не содержит сведений, запрещенных к опубликованию в открытой печати.*

*Подписи авторов: (ф. и. о., дата, адрес)».*

Это соглашение может быть предоставлено в бумажном виде или в виде отсканированной копии (с подписями авторов).

Редколлегия вправе запросить у авторов экспертное заключение о возможности публикации представленной статьи в открытой печати.

2. К статье прилагаются данные автора (авторов) (см. п. 8). При наличии нескольких авторов указывается фамилия автора, ответственного за переписку с редакцией.
3. Редакция журнала осуществляет экспертизу присланных статей в соответствии с принятой в журнале процедурой рецензирования.

Возвращение рукописи на доработку не означает ее принятия к печати.

Доработанный вариант с ответом на замечания рецензента необходимо прислать в редакцию.

4. Решение редколлегии о публикации статьи или ее отклонении сообщается авторам. Редколлегия может также направить авторам текст рецензии на их статью. Дискуссия по поводу отклоненных статей не ведется.

5. Редактура статей высылается авторам для просмотра. Замечания к редакции должны быть присланы авторами в кратчайшие сроки.
6. Рукопись предоставляется в электронном виде в форматах MS WORD (.doc или .docx) или L<sup>A</sup>T<sub>E</sub>X (.tex), дополнительно — в формате .pdf, на дискете, лазерном диске или электронной почтой. Предоставление бумажной рукописи необязательно.
7. При подготовке рукописи в MS Word рекомендуется использовать следующие настройки.

Параметры страницы: формат — А4; ориентация — книжная; поля (см): внутри — 2,5, снаружи — 1,5, сверху и снизу — 2, от края до нижнего колонтитула — 1,3.

Основной текст: стиль — «Обычный», шрифт — Times New Roman, размер — 14 пунктов, абзацный отступ — 0,5 см, 1,5 интервала, выравнивание — по ширине.

Рекомендуемый объем рукописи — не свыше 20 страниц указанного формата.

Сокращения слов, помимо стандартных, не допускаются. Допускается минимальное количество аббревиатур.

Все страницы рукописи нумеруются.

Шаблоны примеров оформления представлены в Интернете:

<http://www.ipiran.ru/publications/collected/template.doc>

8. Статья должна содержать следующую информацию на **русском и английском языках**:

- название статьи;
- Ф.И.О. авторов, на английском можно только имя и фамилию;
- место работы, с указанием города и страны и электронного адреса каждого автора;
- сведения об авторах, в соответствии с форматом, образцы которого представлены на страницах:  
[http://www.ipiran.ru/journal/collected/2012\\_22\\_02\\_rus/authors.asp](http://www.ipiran.ru/journal/collected/2012_22_02_rus/authors.asp) и  
[http://www.ipiran.ru/journal/collected/2012\\_22\\_02\\_eng/authors.asp](http://www.ipiran.ru/journal/collected/2012_22_02_eng/authors.asp);
- аннотация (не менее 100 слов на каждом из языков). Аннотация — это краткое резюме работы, которое может публиковаться отдельно. Она является основным источником информации в информационных системах и базах данных. Английская аннотация должна быть оригинальной, может не быть дословным переводом русского текста и должна быть написана хорошим английским языком. В аннотации не должно быть ссылок на литературу и, по возможности, формул;
- ключевые слова — желательно из принятых в мировой научно-технической литературе тематических тезаурусов. Предложения не могут быть ключевыми словами.
- источники финансирования работы (ссылки на гранты, проекты, поддерживающие организации и т. п.

9. Требования к спискам литературы.

Ссылки на литературу в тексте статьи нумеруются (в квадратных скобках) и располагаются в каждом из списков литературы в порядке первых упоминаний.



Списки литературы представляются в двух вариантах:

- (1) **Список литературы к русскоязычной части.** Русские и английские работы — на языке и в алфавите оригинала.
- (2) **References.** Русские работы и работы на других языках — в латинской транслитерации с переводом на английский язык; английские работы и работы на других языках — на языке оригинала.

Необходимо для составления списка “References” пользоваться размещенной на сайте <http://www.translit.ru/> бесплатной программой транслитерации русского текста в латиницу, при этом в закладке «варианты. . . » следует выбрать опцию BNG.

Список литературы “References” приводится полностью отдельным блоком, повторяя все позиции из списка литературы к русскоязычной части, независимо от того, имеются или нет в нем иностранные источники. Если в списке литературы к русскоязычной части есть ссылки на иностранные публикации, набранные латиницей, они полностью повторяются в списке “References”.

Примеры ссылок на различные виды публикаций в списке “References”:

**Описание статьи из журнала:**

Zhang, Z., and D. Zhu. 2008. Experimental research on the localized electrochemical micromachining. *Rus. J. Electrochem.* 44(8):926–930. doi:10.1134/S1023193508080077.

**Описание статьи из электронного журнала:**

Swaminathan, V., E. Lepkoswka-White, and B. P. Rao. 1999. Browsers or buyers in cyberspace? An investigation of electronic factors influencing electronic exchange. *JCMC* 5(2). Available at: <http://www.ascusc.org/jcmc/vol5/issue2/> (accessed April 28, 2011).

**Описание материалов конференций:**

Usmanov, T. S., A. A. Gusmanov, I. Z. Mullagalin, R. Ju. Muhametshina, A. N. Chervyakova, and A. V. Sveshnikov. 2007. Osobennosti proektirovaniya razrobotki mestorozhdeniy s primeneniem gidrorazryva plasta [Features of the design of field development with the use of hydraulic fracturing]. *Trudy 6-go Mezhdunarodnogo Simpoziuma “Novye resursosberegayushchie tekhnologii nedropol’zovaniya i povyshe-niya neftegazootdachi”* [6th Symposium (International) “New Energy Saving Subsoil Technologies and the Increasing of the Oil and Gas Impact” Proceedings]. Moscow. 267–272.

**Описание книги (монографии, сборники):**

Lindorf, L. S., and L. G. Mamikonians, eds. 1972. *Ekspluatatsiya turbogeneratorov s neposredstvennym okhlazhdeniem* [Operation of turbine generators with direct cooling]. Moscow: Energy Publs. 352 p.

**Описание переводной книги** (в списке литературы к русскоязычной части необходимо указать: / Пер. с англ. — после названия книги, а в конце ссылки указать оригинал книги в круглых скобках):

1. В русскоязычной части:

*Тимошенко С. П., Янг Д. Х., Уивер У.* Колебания в инженерном деле / Пер. с англ. — М.: Машиностроение, 1985. 472 с. (*Timoshenko S. P., Young D. H., Weaver W.* Vibration problems in engineering. — 4th ed. — N.Y.: Wiley, 1974. 521 p.)

2. В англоязычной части:

Timoshenko, S. P., D. H. Young, and W. Weaver. 1974. *Vibration problems in engineering*. 4th ed. N.Y.: Wiley. 521 p.

**Описание неопубликованного документа:**

Latypov, A. R., M. M. Khasanov, and V. A. Baikov. 2004. Geology and production (NGT GiD). Certificate on official registration of the computer program No. 2004611198. (In Russian, unpubl.)

**Описание интернет-ресурса:**

Pravila tsitirovaniya istochnikov [Rules for the citing of sources]. Available at: <http://www.scribd.com/doc/1034528/> (accessed February 7, 2011).

**Описание диссертации или автореферата диссертации:**

Semenov, V. I. 2003. Matematicheskoe modelirovanie plazmy v sisteme kompaktnyy tor [Mathematical modeling of the plasma in the compact torus]. D.Sc. Diss. Moscow. 272 p.

Kozhunova, O. S. 2009. Tekhnologiya razrabotki semanticheskogo slovary informatzionnogo monitoringa [Technology of development of semantic dictionary of information monitoring system]. PhD Thesis. Moscow: IPI RAN. 23 p.

**Описание ГОСТа:**

GOST 8.586.5-2005. 2007. Metodika vypolneniya izmereniy. Izmerenie raskhoda i kolichestva zhidkostey i gazov s pomoshch'yu standartnykh suzhayushchikh ustroystv [Method of measurement. Measurement of flow rate and volume of liquids and gases by means of orifice devices]. Moscow: Standardinform Publ. 10 p.

**Описание патента:**

Bolshakov, M. V., A. V. Kulakov, A. N. Lavrenov, and M. V. Palkin. 2006. Sposob orientirovaniya po krenu letatel'nogo apparata s opticheskoy golovkoy samonavedeniya [The way to orient on the roll of aircraft with optical homing head]. Patent RF No. 2280590.

10. Присланные в редакцию материалы авторам не возвращаются.

11. При отправке файлов по электронной почте просим придерживаться следующих правил:

- указывать в поле subject (тема) название журнала и фамилию автора;
- использовать attach (присоединение);
- в состав электронной версии статьи должны входить: файл, содержащий текст статьи, и файл(ы), содержащий(е) иллюстрации.

12. Журнал «Системы и средства информатики» является некоммерческим изданием. Плата за публикацию не взимается, гонорар авторам не выплачивается.

**Адрес редакции журнала «Системы и средства информатики»:**

Москва 119333, ул. Вавилова, д. 44, корп. 2, ФИЦ ИУ РАН

Тел.: +7 (499) 135-86-92 Факс: +7 (495) 930-45-05

e-mail: [rust@ipiran.ru](mailto:rust@ipiran.ru) (Сейфуль-Мулюков Рустем Бадриевич)

<http://www.ipiran.ru/journal/collected>

---

## Requirements for manuscripts submitted to Journal “Systems and Means of Informatics”

---

Journal “Systems and Means of Informatics” publishes theoretical, review, and discussion articles on the research and development in the field of information technology.

The journal is published in Russian. By a special decision of the editorial board, some articles can be published in English.

Topics covered include the following areas:

- information and communication systems and tools of their design;
- architecture and software of computational complexes and networks; and
- methods and tools of information protection.

1. The Journal publishes original articles which have not been published before and are not intended for simultaneous publication in other editions. An article submitted to the Journal must not violate the Copyright law. Sending the manuscript to the Editorial Board, the authors retain all rights of the owners of the manuscript and transfer the nonexclusive rights to publish the article in Russian (or the language of the article, if not Russian) and its distribution in Russia and abroad to the Founders and the Editorial Board. Authors should submit a letter to the Editorial Board in the following form:

***Agreement on the transfer of rights to publish:***

*“We, the undersigned authors of the manuscript “. . . ,” pass to the Founder and the Editorial Board of the Journal “Systems and Means of Informatics” the nonexclusive right to publish the manuscript of the article in Russian (or in English) in both print and electronic versions of the Journal. We affirm that this publication does not violate the Copyright of other persons or organizations.*

*Author(s) signature(s): (name(s), address(es), date).”*

This agreement should be submitted in paper form or in the form of a scanned copy (signed by the authors).

The Editorial Board has the right to request from the authors an official expert conclusion that the submitted article has no classified data prohibited for publication.

2. A submitted article should be attached with **the data on the author(s)** (see item 8). If there are several authors, the contact person should be indicated who is responsible for correspondence with the Editorial Board and other authors about revisions and final approval of the proofs.
3. The Editorial Board of the Journal examines the article according to the established reviewing procedure. If authors receive their article for correction after reviewing, it does not mean that the article is approved to be published. The corrected article should be sent to the Editorial Board for the subsequent review and approval.
4. The decision on the article publication or its rejection is communicated to the authors. The Editorial Board may also send the reviews on the submitted articles to the authors. Any discussion upon the rejected articles is not possible.
5. The edited articles will be sent to the authors for proofread. The comments of the authors to the edited text of the article should be sent to the Editorial Board as soon as possible.

6. The manuscript of the article should be presented electronically in the MS WORD (.doc or .docx) or L<sup>A</sup>T<sub>E</sub>X (.tex) formats, and additionally in the .pdf format. All documents may be sent by e-mail or provided on a CD or diskette. A hard copy submission is not necessary.
7. The recommended typesetting instructions for manuscript.

Pages parameters: format A4, portrait orientation, document margins (cm): left — 2.5, right — 1.5, above — 2.0, below — 2.0, footer 1.3.

Text: font — Times New Roman, font size — 14, paragraph indent — 0.5, line spacing — 1.5, justified alignment.

The recommended manuscript size: not more than 20 pages of the specified format.

Use only standard abbreviations. Avoid abbreviations in the title and abstract. The full term for which an abbreviation stands should precede its first use in the text unless it is a standard unit of measurement.

All pages of the manuscript should be numbered.

The templates for the manuscript typesetting are presented on site:  
<http://www.ipiran.ru/publication/collected/template.doc>
8. Articles should enclose data both in **Russian and English**:
  - title;
  - author's name and surname;
  - affiliation — organization, its address with ZIP code, city, country, and official e-mail address;
  - data on authors according to the format (see site):  
[http://www.ipiran.ru/journal/collected/2012\\_22\\_02\\_rus/authors.asp](http://www.ipiran.ru/journal/collected/2012_22_02_rus/authors.asp) and  
[http://www.ipiran.ru/journal/collected/2012\\_22\\_02\\_eng/authors.asp](http://www.ipiran.ru/journal/collected/2012_22_02_eng/authors.asp);
  - abstract (not less than 100 words) both in Russian and in English. Abstract is a short summary of the article that can be published separately. The abstract is the main source of information on the article and it could be included in leading information systems and data bases. The abstract in English has to be an original text and should not be an exact translation of the Russian one. Good English is required. In abstracts, avoid references and formulae.
  - Indexing is performed on the basis of keywords. The use of keywords from the internationally accepted thematic Thesauri is recommended.  
Important! Keywords must not be sentences.
  - Acknowledgments.
9. References. Russian references have to be presented both in English translation and in Latin transliteration (refer <http://www.translit.ru>, option BGN).

Please take into account the following examples of Russian references appearance:

**Article in journal:**  
Zhang, Z., and D. Zhu. 2008. Experimental research on the localized electrochemical micromachining. *Rus. J. Electrochem.* 44(8):926–930. doi:10.1134/S1023193508080077.

**Journal article in electronic format:**  
Swaminathan, V., E. Lepkoswka-White, and B. P. Rao. 1999. Browsers or buyers in cyberspace? An investigation of electronic factors influencing electronic

exchange. *JCMC* 5(2). Available at: <http://www.ascusc.org/jcmc/vol5/issue2/> (accessed April 28, 2011).

**Conference proceedings:**

Usmanov, T. S., A. A. Gusmanov, I. Z. Mullagalin, R. Ju. Muhametshina, A. N. Chervyakova, and A. V. Sveshnikov. 2007. Osobennosti proektirovaniya razrabotki mestorozhdeniy s primeneniem gidrorazryva plasta [Features of the design of field development with the use of hydraulic fracturing]. *Trudy 6-go Mezhdunarodnogo Simpoziuma "Novye resursosberegayushchie tekhnologii nedropol'zovaniya i povyshe-niya neftegazootdachi"* [6th Symposium (International) "New Energy Saving Subsoil Technologies and the Increasing of the Oil and Gas Impact" Proceedings]. Moscow. 267–272.

**Books and other monographs:**

Lindorf, L. S., and L. G. Mamikonians, eds. 1972. *Ekspluatatsiya turbogenera-torov s neposredstvennym okhlazhdeniem* [Operation of turbine generators with direct cooling]. Moscow: Energy Publs. 352 p.

**Dissertation and Thesis:**

Kozhunova, O. S. 2009. Tekhnologiya razrabotki semanticheskogo slovarya informat-sionnogo monitoringa [Technology of development of semantic dictionary of information monitoring system]. PhD Thesis. Moscow: IPI RAN. 23 p.

**State standards and patents:**

GOST 8.586.5-2005. 2007. Metodika vypolneniya izmereniy. Izmerenie raskhoda i kolichestva zhidkostey i gazov s pomoshch'yu standartnykh suzhayushchikh ustroystv [Method of measurement. Measurement of flow rate and volume of liquids and gases by means of orifice devices]. M.: Standardinform Publs. 10 p.

Bolshakov, M. V., A. V. Kulakov, A. N. Lavrenov, and M. V. Palkin. 2006. Sposob orientirovaniya po krenu letatel'nogo apparata s opticheskoy golovkoy samonavedeniya [The way to orient on the roll of aircraft with optical homing head]. Patent RF No. 2280590.

References in Latin transcription are presented in the original language.

References in the text are numbered according to the order of their first appearance; the number is placed in square brackets. All items from the reference list should be cited.

10. Manuscripts and additional materials are not returned to Authors by the Editorial Board.
11. Submissions of files by e-mail must include:
  - the journal title and author's name in the "Subject" field;
  - an article and additional materials have to be attached using the "attach" function;
  - an electronic version of the article should contain the file with the text and a separate file with figures.
12. "System and Means of Informatics" journal is not a profit publication. There are no charges for the authors as well as there are no royalties.

**Editorial Board address:**

FRC CSC RAS, 44, block 2, Vavilov Str., Moscow 119333, Russia

Ph.: +7 (499) 135 86 92, Fax: +7 (495) 930 45 05

e-mail: [rust@ipiran.ru](mailto:rust@ipiran.ru) (to Prof. Rustem Seyful-Mulyukov)

[http://www.ipiran.ru/english/journal\\_systems.asp](http://www.ipiran.ru/english/journal_systems.asp)

# SYSTEMS AND MEANS OF INFORMATICS (SISTEMY I SREDSTVA INFORMATIKI)

SCIENTIFIC JOURNAL

**Volume 25 No.4 Year 2015**

**Editor-in-Chief and Chair of Editorial Council**  
Academician I. A. Sokolov

## IN THIS ISSUE:

ANALYTICAL MODELING OF NORMAL PROCESSES IN STOCHASTIC SYSTEMS WITH COMPLEX FRACTION-RATIONAL NONLINEARITIES <i>I. N. Sinitsyn, V. I. Sinitsyn, and E. R. Korepanov</i>	4
ENERGY EFFICIENCY ESTIMATION OF THE COMPUTATIONAL CLUSTER USING A QUEUEING SYSTEM WITH THRESHOLD CONTROL AND TIME-DEPENDENT SERVICE AND ARRIVAL RATES <i>R. V. Razumchik, A. I. Zeifman, A. V. Korotysheva, and Ya. A. Satin</i>	19
APPROXIMATE OPTIMIZATION OF RESOURCE ALLOCATION STRATEGY: THE CASE OF BANK UNDERWRITING SYSTEM <i>M. G. Konovalov and R. V. Razumchik</i>	31
METRIC TIME SERIES CLASSIFICATION USING DYNAMIC WARPING RELATIVE TO CENTROIDS OF CLASSES <i>A. V. Goncharov, M. S. Popova, and V. V. Strijov</i>	52
ARCHITECTURE OF THE STAND FOR THE PILOT STUDY OF MODELS, ALGORITHMS, AND SOLUTIONS OF INFORMATION SECURITY IN CLOUD COMPUTING ENVIRONMENTS <i>A. A. Grusho, M. I. Zabezhailo, A. A. Zatsarinny, A. V. Nikolaev, and V. O. Piskovski</i>	65
HARDWARE AND SOFTWARE MODELING AND TESTING OF THE RECURRENT OPERATIONAL DEVICE <i>D. V. Khilko, Yu. A. Stepchenkov, Yu. G. Diachenko, Yu. I. Shikunov, and N. V. Morozov</i>	78
SIGNAL MYOGRAM FEATURES RESEARCH <i>T. V. Zakharova, V. Yu. Korolev, and A. A. Schemirova</i>	91
METHOD OF SYNCHRONIZATION OF MAGNETOENCEPHALOGRAMS AND MYOGRAMS SIGNALS <i>E. V. Chshenyavskaya and T. V. Zakharova</i>	101