

# **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

**Учредитель:**  
**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 В. Ротарь (V. Rotar, San-Diego, USA)  
профессор М. Финкельштейн (M. Finkelstein, Bloemfontein, South Africa)  
профессор В. Хофкирхнер (W. Hofkirchner, Wien, Austria)

## **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. А. И. Зейфман  
проф., д.т.н. В. Д. Ильин  
проф., д.т.н. К. К. Колин  
проф., д.ф.-м.н. В. Ю. Королев  
д.ф.-м.н. Р. В. Разумчик  
д.ф.-м.н. В. И. Сеницын  
проф., д.т.н. И. Н. Сеницын  
проф., д.ф.-м.н. В. Г. Ушаков  
д.ф.-м.н. А. К. Горшенин — отв. секретарь  
к.ф.-м.н. С. А. Христочевский

## **Редакция**

к.ф.-м.н. Е. Н. Арутюнов  
С. Н. Стригина

© Федеральный исследовательский центр «Информатика  
и управление» Российской академии наук, 2024

Журнал включен в базу данных Russian Science Citation Index (RSCI),  
интегрированную с Web of Science

Журнал входит в систему Российского индекса научного цитирования (РИНЦ)  
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» ВИНИТИ  
и в системе Google Scholar

Журнал включен в сформированный 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 34 № 1 Год 2024

## СОДЕРЖАНИЕ

Аналитическое моделирование стохастических систем, не разрешенных относительно производных, со случайными параметрами <b>И. Н. Сеницын</b>	4
Задача разделения ресурса беспроводной сети по задержке передачи эластичного трафика <b>И. А. Кочеткова, А. С. Власкина, П. Л. Э. Эсенг Манге, В. С. Шоргин</b>	23
Методика десинхронизации при синтезе самосинхронных схем <b>Ю. А. Степченков, Д. В. Хилько, Ю. Г. Дьяченко, Н. В. Морозов, Д. Ю. Степченков, Г. А. Орлов</b>	33
Разработка информационной системы для хранения и обработки результатов мониторинга сельскохозяйственных угодий юга Дальнего Востока <b>Д. В. Федосеев, А. С. Степанов, Л. В. Илларионова, В. А. Елисеев, И. О. Прохорец, А. Н. Фролов, А. И. Шамурина</b>	44
Картографирование земель сельскохозяйственного назначения Хабаровского края методами машинного обучения с использованием изображений Sentinel-2 <b>И. О. Прохорец, А. С. Степанов</b>	57
Neural network architecture for artifacts detection in ZTF survey <b>Т. А. Semenikhin</b>	70
Жизненный цикл кибератак на уязвимости систем машинного обучения <b>А. П. Сучков</b>	80
Выработка искусственными интеллектуальными рефлексивными агентами тактики ведения переговоров <b>С. В. Листопад</b>	92

# СИСТЕМЫ И СРЕДСТВА ИНФОРМАТИКИ

Том 34 № 1 Год 2024

## СОДЕРЖАНИЕ

Некоторые сетевые аспекты программного определения «облачных» платформ <b><i>В. Б. Егоров</i></b>	<b>102</b>
Математическое моделирование оценки влияния глобальных изменений в мире на цифровой след российских университетов на примере аграрных <b><i>В. И. Будзко, В. И. Мединников</i></b>	<b>111</b>
Коллективное разрешение сущностей в технологии поддержки конкретно-исторических исследований <b><i>И. М. Адамович, О. И. Волков</i></b>	<b>128</b>
О параллелизме в развитии базисных информационных технологий, пространственных типов цивилизаций и типологии знаковых систем <b><i>С. Н. Гринченко</i></b>	<b>139</b>
Об авторах	<b>150</b>
Правила подготовки рукописей статей	<b>153</b>
Requirements for manuscripts	<b>157</b>

## АНАЛИТИЧЕСКОЕ МОДЕЛИРОВАНИЕ СТОХАСТИЧЕСКИХ СИСТЕМ, НЕ РАЗРЕШЕННЫХ ОТНОСИТЕЛЬНО ПРОИЗВОДНЫХ, СО СЛУЧАЙНЫМИ ПАРАМЕТРАМИ\*

*И. Н. Синицын*<sup>1</sup>

**Аннотация:** Статья посвящена нелинейным корреляционным методам аналитического моделирования (МAM) процессов в дифференциальных стохастических системах (СтС), не разрешенных относительно производных (НРОП). Приведен обзор работ в области аналитического моделирования СтС НРОП. Даны необходимые сведения из теории интегральных канонических представлений (ИКП) случайных процессов (СтП) и их линейных и нелинейных преобразований. Для скалярных, векторных, стационарных и нестационарных СтП приводятся необходимые и достаточные условия существования ИКП. Особое внимание уделяется многокомпонентным ИКП и их преобразованиям. Для существенно нелинейных преобразований на базе линейной регрессии для многокомпонентных ИКП разработаны два типа оптимальных по среднеквадратичному критерию моделей статистической линеаризации (МСЛ). Показано, как можно дифференциальные СтС НРОП привести к дифференциальным СтС. При фиксированном векторе случайных параметров рассматриваются гладкие относительно старших производных уравнения СтС НРОП, допускающие дифференциалы Ито определенного порядка, и разрывные, допускающие регрессионную линеаризацию. На основе метода нормальной аппроксимации (МНА) выводятся уравнения для условных векторов математического ожидания, ковариационной матрицы и матрицы ковариационных функций. Приведены алгоритмы МAM качества на основе ИКП для типовых СтС НРОП. Представлен пример, иллюстрирующий особенности СтС НРОП при нестационарных случайных параметрах, заданных ИКП. Даны направления дальнейших обобщений.

**Ключевые слова:** аналитическое моделирование; интегральное каноническое представление (ИКП); метод нормальной аппроксимации (МНА); стохастическая система, не разрешенная относительно производной (СтС НРОП); условные и безусловные корреляционные характеристики

**DOI:** 10.14357/08696527240101

**EDN:** ZPTXJI

### 1 Введение

В [1–4] рассмотрены вопросы аналитического моделирования процессов в СтС НРОП. Особое внимание в них уделено нормальным (гауссовским) СтП. В [5]

---

\* Работа выполнена при финансовой поддержке РАН (номер государственного учета НИОКТР АААА-А19-119091990037-5).

<sup>1</sup>Федеральный исследовательский центр «Информатика и управление» Российской академии наук; Московский авиационный институт, sinitsin@dol.ru

предложены методы нормализации систем, стохастически не разрешенных относительно производных. Теория распределений с инвариантной мерой в СтС НРОП развита в [6].

Рассмотрим обобщение [1–4] на случай СтС НРОП со случайными переменными параметрами, описываемых ИКП. В разд. 2 приводятся сведения из корреляционной теории ИКП и их линейных и нелинейных преобразований, допускающих линейную регрессионную линеаризацию. В разд. 3 рассматриваются скалярные дифференциальные СтС НРОП при фиксированном случайном векторе параметров, приводимые к дифференциальным СтС. Представлены уравнения МНА для 6 типов приведенных СтС. Основные результаты по аналитическому моделированию качества СтС НРОП со случайными параметрами, основанные на безусловных корреляционных характеристиках СтП, представлены в разд. 4. В разд. 5 приводится иллюстративный пример. Дан анализ новых эффектов в СтС НРОП, вызванных переменными случайными параметрами. Заключение содержит выводы и направления возможных обобщений.

## 2 Интегральные канонические представления и их преобразования

Как известно [7], для скалярного СтП  $X = X(t)$  ИКП называется его и его ковариационной функции интегральное представление через непрерывный скалярный белый шум  $V = V(\lambda)$  параметра  $\lambda \in \Lambda$ :

$$X(t) = m_x(t) + \int_{\Lambda} V(\lambda)x(t, \lambda) d\lambda; \quad (1)$$

$$K_x(t, t') = \mathbf{M} \left[ X(t)\overline{X(t')} \right] = \int_{\Lambda} G(\lambda)x(t, \lambda)\overline{x(t', \lambda)} d\lambda.$$

Здесь  $x(t, \lambda)$  — координатная функция;  $G(\lambda)$  — интенсивность  $V(\lambda)$ , причем  $K_v(\lambda, \lambda') = G(\lambda)\delta(\lambda - \lambda')$ . Формула для координатной функции ИКП имеет вид:

$$x(t, \lambda) = \frac{1}{G(\lambda)} K_{xv}(t, \lambda) = \frac{1}{G(\lambda)} \mathbf{M} \left[ X^0(t)\overline{V(\lambda)} \right].$$

Для построения ИКП используются следующие необходимые и достаточные условия:

$$V(\lambda) = \int_T \overline{a(t, \lambda)} X^0(t) dt; \quad (2)$$

$$x(t, \lambda) = \frac{1}{G(\lambda)} \int_T a(s, \lambda) K_x(t, s) ds; \quad (3)$$

$$\int_T \overline{a(t, \lambda)} x(t, \lambda') dt = \delta(\lambda - \lambda'); \quad (4)$$

$$\int_\Lambda \overline{a(t', \lambda)} x(t, \lambda) d\lambda = \delta(t - t'). \quad (5)$$

**Теорема 2.1.** Для того чтобы скалярный СтП  $X = X(t)$  в области  $T$  допускал ИКП (1), условия (3) и (4) необходимы, а условия (3)–(5) достаточны.

**Теорема 2.2.** Если СтП стационарен и обладает спектральной плотностью  $s^x = s^x(\omega)$ , то его ИКП имеет показательные координатные функции  $x(t, \omega) = e^{i\omega t}$ , причем

$$k_x(\tau) = \int_{-\infty}^{\infty} s_x(\omega) e^{-i\omega\tau} d\omega;$$

$$s_x(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} k_x(\tau) e^{i\omega\tau} d\tau;$$

$$k_x(t - t') = \int_{-\infty}^{\infty} s_x(\omega) e^{i\omega t} \overline{e^{i\omega t'}} d\omega;$$

$$X(t) = m_x + \int_{-\infty}^{\infty} V(\omega) e^{i\omega t} d\omega.$$

**Замечание 2.1.** Интегральное каноническое представление произвольных СтП впервые рассматривались Каруненом, который доказал общую теорему, заключающуюся в том, что необходимое и достаточное условие возможности представления СтП  $X(t)$  формулой (2) заключается в представлении его ковариационной функции формулой (3). Однако Карунен не доказал существования ИКП для произвольного СтП и не дал способов определения функций  $x(t, \lambda)$  и  $G(\lambda)$ . До Карунена частный вид ИКП стационарного СтП был изучен А. Н. Колмогоровым.

Аналогично для векторного СтП  $X(t) = [X_1(t) \cdots X_n(t)]^T$  имеем:

$$X(t) = m^x(t) + \int_\Lambda V(\lambda) x(t, \lambda) d\lambda,$$

$$X_h(t) = m_h^x(t) + \int_\Lambda V(\lambda) x_h(t, \lambda) d\lambda \quad (h = \overline{1, n}). \quad (6)$$

Здесь  $m^x(t) = [m_1^x(t) \cdots m_n^x(t)]^T$ . При этом ИКП матрицы ковариационных функций  $K^x(t, t') = [K_{hl}^x(t, t')]$  имеют вид:

$$K_{hl}^x(t, t') = \int_{\Lambda} G(\lambda) x_h(t, \lambda) \overline{x_l(t', \lambda)} d\lambda \quad (h, l = \overline{1, n}), \quad (7)$$

где белый шум  $V(\lambda)$  определяется формулой

$$V(\lambda) = \sum_{h=1}^n \int_T \overline{a_h(t, \lambda)} X_h^0(t) dt,$$

а его интенсивность  $G(\lambda)$  равна

$$G(\lambda) = \sum_{h,l=1}^n \int_T \int_T \int_T \overline{a_h(t, \lambda)} a_l(t', \lambda') K_{hl}^x(t, t') dt dt' d\lambda'.$$

Для вычисления координатных функций  $x_h(t, \lambda)$  и функций  $a_h(t, \lambda)$  получим вместо (3), (4) и (5) уравнения

$$x_h(t, \lambda) = \frac{1}{G(\lambda)} \sum_{l=1}^n \int_T a_l(t', \lambda) K_{hl}^x(t, t') dt' \quad (h = \overline{1, n}); \quad (8)$$

$$\begin{aligned} \sum_{h=1}^n \int_T \overline{a_h(t, \lambda)} x_h(t, \lambda') dt &= \delta(\lambda - \lambda'); \\ \int_{\Lambda} \overline{a_l(t', \lambda)} x_h(t, \lambda) d\lambda &= \delta_{hl}(t - t'). \end{aligned} \quad (9)$$

Если параметр  $\lambda$  принимает все возможные значения, принадлежащие нескольким областям  $\Lambda_1, \dots, \Lambda_r$ , то ИКП векторного СтП  $X(t) = [X_1(t) \cdots X_n(t)]^T$  и матрицы ковариационных функций  $K^x(t, t') = [K_{hl}^x(t, t')]$  будут иметь вид:

$$\begin{aligned} X(t) &= m^x(t) + \sum_{\rho=1}^r \int_{\Lambda_{\rho}} V_{\rho}(\lambda) x_{\rho}(t, \lambda) d\lambda, \\ X_h(t) &= m_h^x(t) + \sum_{\rho=1}^r \int_{\Lambda_{\rho}} V_{\rho}(\lambda) x_{\rho h}(t, \lambda) d\lambda \quad (h = \overline{1, n}); \end{aligned} \quad (10)$$

$$K_{hl}(t, t') = \sum_{\rho=1}^r \int_{\Lambda_\rho} G_\rho(\lambda) x_{\rho h}(t, \lambda) \overline{x_{\rho l}(t', \lambda)} d\lambda \quad (h, l = \overline{1, n}), \quad (11)$$

где  $V_1(\lambda), \dots, V_r(\lambda)$  — некоррелированные белые шумы, определяемые формулой

$$V_\rho(\lambda) = \sum_{h=1}^n \int_T \overline{a_{\rho h}(t, \lambda)} X_h^0(t) dt \quad (\lambda \in \Lambda_\rho; \rho = \overline{1, r}).$$

Интенсивности белых шумов  $V_\rho(\lambda)$  определяются формулой:

$$G_\rho(\lambda) = \sum_{h, l=1}^n \int_T \int_T \overline{a_{\rho h}(t, \lambda)} a_{\rho l}(t', \lambda') K_{hl}^x(t, t') dt dt' d\lambda' \quad (\lambda \in \Lambda_\rho; \rho = \overline{1, r}).$$

Для вычисления координатных функций  $x_{\rho h}(t', \lambda')$  и функций  $a_{\rho h}(t, \lambda)$  используются формулы:

$$x_{\rho h}(t, \lambda) = \frac{1}{G_\rho(\lambda)} \sum_{l=1}^n \int_T a_{\rho l}(t', \lambda) K_{hl}^x(t, t') dt' \quad (\lambda \in \Lambda_\rho; \rho = \overline{1, r}; h = \overline{1, n}); \quad (12)$$

$$\sum_{h=1}^n \int_T \overline{a_{\mu h}(t, \lambda)} a_{\rho h}(t, \lambda') dt = \delta_{\rho\mu} \delta(\lambda - \lambda') \quad (\lambda \in \Lambda_\mu; \lambda' \in \Lambda_\rho; \rho, \mu = \overline{1, r});$$

$$\sum_{\rho=1}^r \int_{\Lambda_\rho} \overline{a_{\rho l}(t', \lambda)} x_{\rho h}(t, \lambda) d\lambda = \delta_{hl} \delta(t - t') \quad (h, l = \overline{1, n}), \quad (13)$$

выражающие необходимые и достаточные условия представления векторного СтП посредством ИКП.

Имеют место следующие утверждения [7].

**Теорема 2.3.** В условиях (8), (9) или (12), (13) из ИКП (6) или (10) вытекает ИКП матрицы его ковариационных функций (7) или (11).

**Теорема 2.4.** Если векторный стационарный СтП обладает матрицей спектральных плотностей  $s_x(\omega) = [s_{hl}(\omega)]$ , то в основе ИКП такого СтП лежат формулы:

$$k^x(\tau) = [k_{hl}^x(\tau)] = \int_{-\infty}^{\infty} s^x(\omega) e^{i\omega\tau} d\omega, \quad k_{hl}^x(\tau) = \int_{-\infty}^{\infty} s_{hl}^x(\omega) e^{i\omega\tau} d\omega;$$



$$s^x(\omega) = [s_{hl}^x(\omega)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} k^x(\tau) e^{-i\omega\tau} d\omega, \quad s_{hl}^x(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} k_{hl}^x(\tau) e^{-i\omega\tau} d\omega;$$

$$k_{hl}^x(t-t') = \sum_{r=1}^n \int_{-\infty}^{\infty} G_r(\omega) a_{rh}(\omega) e^{i\omega t} \overline{a_{rl}(\omega) e^{i\omega t'}} d\omega \quad (h, l = \overline{1, n})$$

ИЛИ

$$X_h(t) = m_h^x + \int_{-\infty}^{\infty} U_h(\omega) e^{i\omega t} d\omega \quad (h = \overline{1, n});$$

$$X(t) = m_x(t) + \sum_{\rho=1}^r \int_{\Lambda_\rho} V_\rho(\lambda) x_\rho(t, \lambda) d\lambda.$$

**Теорема 2.5.** Если известно ИКП векторного СтП (6), то вектор математического ожидания и матрица ковариационных функций линейного преобразования  $Y_t(t) = A_t X_t$  допускают ИКП, определяемые формулами:

$$X_p(t) = m_p^x(t) + \sum_{\rho=1}^r \int_{\Lambda_\rho} V_\rho(\lambda) x_{p\rho}(t, \lambda) d\lambda \quad (p = \overline{1, n});$$

$$Y_p(s) = m_p^y(s) + \sum_{\rho=1}^r \int_{\Lambda_\rho} V_\rho(\lambda) y_{p\rho}(s, \lambda) d\lambda,$$

$$y_{p\rho}(s, \lambda) = \sum_{h=1}^n A_{ph} x_{\rho h}(t, \lambda) \quad (p = \overline{1, m});$$

$$K^y(s, s') = [K_{pq}^y(s, s')],$$

$$K_{pq}^y(s, s') = \sum_{\rho=1}^r \int_{\Lambda_\rho} G_\rho(\lambda) y_{p\rho}(s, \lambda) \overline{y_{q\rho}(s', \lambda)} d\lambda \quad (p, q = \overline{1, m}).$$

В задачах нелинейной корреляционной теории [7–9] невырожденные безынерционные скалярные и векторные существенно нелинейные преобразования  $Y_t = \varphi_t(X_t)$  заменяют оптимальными (в среднеквадратичном смысле) линейными регрессионными преобразованиями. Задача эквивалентной регрессионной линеаризации детерминированной векторной нелинейной функции  $Y = \varphi(X)$  при использовании критерия минимума среднеквадратичной ошибки совпадает с классической задачей линейного регрессионного анализа. В этом случае

оптимальная линейная среднеквадратичная регрессия вектора  $Y$  на вектор  $X$  определяется формулой:

$$m^y(X) = gX, \quad g = \Gamma^{yx}(\Gamma^x)^{-1} \quad (14)$$

или, с учетом смещения  $a$ ,

$$m^y(X) = gX + a, \quad g = K^{yx}(K^x)^{-1}, \quad a = m^y - gm^x. \quad (15)$$

Если векторы  $Y$  и  $X$  действительны и их совместное распределение нормально, регрессии всегда линейны. Поэтому формулы (14) и (15) определяют регрессию проекции  $Y$  нормально распределенного вектора  $[Y^T X^T]^T$  на его проекцию  $X$ .

Пусть  $f(y, x)$  — совместная плотность случайных векторов  $Y$  и  $X$ ;  $m^x$  и  $K^x$  — математическое ожидание и ковариационная матрица вектора  $x$ ,  $\det |K^x| \neq 0$ . Формула (14) для  $g$  при этом принимает вид:

$$g = K^{yx}(K^x)^{-1} = \int_{-\infty}^{\infty} [m^y(x) - m^y] (x - m^x)^T (K^x)^{-1} f_1(x) dx, \quad (16)$$

где  $f_1(x)$  — плотность случайного вектора  $X$ . Эта формула вместе с приближенной формулой

$$m^y(X) \approx m^y + g(X - m^x) \quad (17)$$

дает статистическую линейнизацию регрессии  $m^y(X)$  по Казакову.

Аналогично в случае (14) имеем

$$g = \Gamma^{yx}(\Gamma^x)^{-1} = \int_{-\infty}^{\infty} m^y(x) X^T (\Gamma^x)^{-1} f_1(x) dx. \quad (18)$$

Формула (17) вместе с приближенной зависимостью

$$m^y(X) \approx gX \quad (19)$$

дает статистическую линейнизацию регрессии  $m^y(X)$  по Бутону.

**Теорема 2.6.** *Если существуют конечные моменты первого и второго порядка векторного СтП  $X = X(t)$ , то векторное нелинейное преобразование  $Y = \varphi(X)$  допускает линейную среднеквадратичную регрессию  $Y = Y(t)$  на  $X = X(t)$ , определяемую по Казакову формулами (16) и (17) и формулами (18) и (19) по Бутону.*

Первый подход к линеаризации основан на использовании формул для ИКП  $X_t$  в теореме 2.3 в случае одной области  $\Lambda$ , а второй подход — для нескольких областей  $\Lambda_r$ . При этом имеют место следующие утверждения.

**Теорема 2.7.** *Оптимальная среднеквадратичная линеаризация посредством ИКП (теорема 2.3) определяется следующими формулами:*

$$\varphi_t(X_t) \approx m_t^{(1)y}(X_t) = \varphi_{0t}^{(1)}(m_t^x, K_t^x) + g_t^{(1)}(m_t^x, K_t^x) \left[ \sum_{\rho=1}^r \int_{\Lambda_\rho} V_\rho(\lambda) x_t(\lambda) d\lambda - m_t^x \right],$$

где

$$\varphi_{0t}^{(1)}(m_t^x, K_t^x) = M_N [\varphi(X_t)]; \quad g_t^{(1)}(m_t^x, K_t^x) = K_t^{yx} (K_t^x)^{-1}.$$

Здесь  $M_N$  – символ математического ожидания для гауссовского (нормального) распределения.

**Теорема 2.8.** *Оптимальная среднеквадратичная линеаризация посредством ИКП (теорема 2.3) второго рода определяется следующими формулами:*

$$\varphi_t(X_t) \approx m_t^{(2)y}(X_t) = \varphi_{0t}^{(2)}(m_t^x, K_{\rho t}^x) + \sum_{\rho=1}^r \int_{\Lambda_\rho} g_{\rho t}^{(2)}(m_t^x, K_{\rho t}^x) V_\rho(\lambda) x_{\rho t} d\lambda = g_t^{(2)}(m_t^x, K_t^x) m_t^x,$$

где

$$\varphi_{0t}^{(2)}(m_t^x, K_{\rho t}^x) = M_N [\varphi(X_t)]; \quad g_t^{(2)}(m_t^x, K_t^x) = K_t^{yx} (K_t^x)^{-1}.$$

**Замечание 2.2.** Аналогично выписываются формулы теорем 2.7 и 2.8 в терминах вероятностных начальных моментов  $\Gamma_t^x$ ,  $\Gamma_t^{yx}$ ,  $\Gamma_{\rho t}^x$  и  $\Gamma_{\rho t}^{yx}$

Для типовых нелинейных преобразований  $\varphi$  соответствующие таблицы можно найти в [7–9].

### 3 Приводимые дифференциальные стохастические системы, не разрешенные относительно производных

Следуя [1, 2], при фиксированном векторе случайных параметров  $\Theta = \Theta(t)$  рассмотрим скалярную дифференциальную СтС НРОП с нелинейностями, описываемыми гладкими функциями

$$\Phi = \Phi(t, \Theta, Y_t, \dot{Y}_t, \dots, Y_t^{(k)}, U_t) = 0, \\ Y(t_0) = Y_0, \quad \dot{Y}(t_0) = \dot{Y}_0, \dots, Y^{(n)}(t_0) = Y_0^{(n)}, \quad (20)$$

при этом векторное уравнение нелинейного формирующего фильтра ( $\Phi\Phi$ ) возьмем в следующем виде, разрешенном относительно возмущений:

$$\dot{U}_t = a^U(t, \Theta, U_t) + b^U(t, \Theta, U_t) V_t^U, \quad U(t_0) = U_0. \quad (21)$$

Здесь  $a^U = a^U(t, \Theta, U_t)$  и  $b^U(t, \Theta, U_t)$  —  $(n^Y \times 1)$ - и  $(n^Y \times n^V)$ -мерные функции;  $V_t^U$  — белый шум в строгом смысле [8, 9], допускающий представление

$$V_t^U = \dot{W}_t^U, \quad W_t^U = W_0^U(t, \Theta) + \int_{R_0^q} c^U(\Theta, \rho) P^0(t, \Theta, d\rho), \quad (22)$$

где  $\nu_t$  — его интенсивность:

$$\nu_t = \nu_t^W = \nu_t^{W_0} + \int_{R_0^q} c^U(\Theta, \rho) [c^U(\Theta, \rho)]^T \nu_P(t, \Theta, \rho) d\rho; \quad (23)$$

$c^U = c^U(\Theta, \rho)$  — известная векторная функция той же размерности, что и  $W_t^0$ , а интеграл при любом  $t \geq t_0$  представляет собой стохастический интеграл по центрированной пуассоновской мере  $P^0(t, \Theta, \mathcal{A})$ , независимой от  $W_0^U$  и имеющей независимые значения на попарно непересекающихся множествах;  $\mathcal{A}$  — борелевское множество пространства  $R_0^q$  с выколотым началом;  $\nu_t^W$ ,  $\nu_t^{W_0}$  и  $\nu_P$  — интенсивности СтП  $W_t^U$ ,  $W_0^U$  и  $P^0$ . Уравнение (21) понимается в смысле Ито и имеет единственное решение в среднем квадратическом [8, 9].

Для гладких функций в (20), допускающих стохастические производные Ито до  $h$ -го порядка, выполним следующие преобразования. Будем дифференцировать сполна по  $t$  левые части уравнений (20) по обобщенной формуле Ито [8, 9] до тех пор, пока не появятся производные белого шума. В результате получим следующую систему нелинейных дифференциальных уравнений:

$$\Phi = 0, \quad \dot{\Phi} = 0, \quad \dots, \quad \Phi^{(h)} = 0. \quad (24)$$

Далее введем вектор

$$Z_t = \left[ Z_t'^T Z_t''^T \right]^T,$$

составленный из

$$Z_t' = \left[ Y_t^T \dot{Y}_t^T \dots Y_t^{(k-1)T} \right]^T$$

и вспомогательного вектора  $Z_t''$ , определяемого уравнениями (24). В результате придем к уравнениям, разрешенным относительно дифференциалов, следующего вида:

$$dZ_t = a^Z dt + b^Z dW_0 + \int_{R_0^q} c^Z P^0(t, \Theta, du), \quad (25)$$

где  $a^Z = a^Z(t, \Theta, Z_t)$ ;  $b^Z = b^Z(t, \Theta, Z_t)$ ;  $c^Z = c^Z(t, \Theta, Z_t, u)$ .

Таким образом, имеем следующие утверждения.

**Теорема 3.1.** Пусть нелинейная негауссовская СтС НРОП (20), (21), не разрешенная относительно производных  $k$ -го порядка, удовлетворяет условиям:

- 1<sup>0</sup> функции (20) допускают обобщенные стохастические дифференциалы Ито вплоть до  $h$ -го порядка включительно;
- 2<sup>0</sup> уравнение ФФ (21), разрешено относительно возмущений  $U_t$  и имеет единственное среднеквадратичное решение. Тогда система (20), (21) приводима к системе, разрешенной относительно производных (25).

**Теорема 3.2.** Пусть для нелинейной негауссовской СтС НРОП (20), (21) выполнены условия:

- 1<sup>0</sup> правые части уравнений (20) имеют стохастические производные Ито до порядка  $h$ ;
- 2<sup>0</sup> уравнения (20) совместно с (24) имеют единственное решение, понимаемое в среднем квадратическом.

Тогда в основе МАМ на базе МНА для (20)–(25) лежат следующие обыкновенные дифференциальные уравнения [8, 9]:

$$\left. \begin{aligned} \dot{m}_t^Z &= A_t^m(t, \Theta, m_t^Z, K_t^z), & m_0^Z &= m_{t_0}^Z; \\ \dot{K}_t^Z &= A_t^K(t, \Theta, m_t^Z, K_t^z), & K_0^Z &= K_{t_0}^Z; \\ \frac{\partial K^Z(t_1, t_2)}{\partial t_2} &= A_{t_1, t_2}^K(t_1, t_2, \Theta, m_{t_2}^Z, K_{t_2}^z), & K^Z(t_1, t_2) &= K_{t_1}^Z. \end{aligned} \right\} \quad (26)$$

Здесь

$$A_t^m(t, \Theta, m_t^Z, K_t^Z) = M_{\mathcal{N}}[a^Z]; \quad (27)$$

$$\begin{aligned} A_t^K(t, \Theta, m_t^Z, K_t^Z) &= \\ &= M_{\mathcal{N}} A_{1t}(t, \Theta, m_t^Z, K_t^Z) + A_{1t}^T(t, \Theta, m_t^Z, K_t^Z) + A_{2t}(t, \Theta, m_t^Z, K_t^Z), \end{aligned} \quad (28)$$

где

$$A_{1t}(t, \Theta, m_t^Z, K_t^Z) = M_{\mathcal{N}}[a^Z(Z_t - m_t^Z)^T],$$

$$A_{2t}(t, \Theta, m_t^Z, K_t^Z) = M_{\mathcal{N}}[\sigma(t, \Theta, Z_t)],$$

$$\sigma(t, \Theta, Z_t) = \sigma_0(t, \Theta, Z_t) + \int_{R_0^q} c^Z(t, \Theta, Z_t, u) c^Z(t, \Theta, Z_t, u)^T \nu_P(t, \Theta, du),$$

$$\sigma_0(t, \Theta, Z_t) = b^Z(t, \Theta, Z_t) \nu_0(t, \Theta) b^Z(t, \Theta, Z_t)^T;$$

$$A_{t_1, t_2}^K(t_1, t_2, \Theta, m_{t_2}^Z, K_{t_2}^Z) = K^Z(t_1, t_2) (K_{t_2}^Z)^{-1} A_{1t}^T. \quad (29)$$

В стационарном случае уравнения МНА имеют следующий вид:

$$\left. \begin{aligned} A_*^m(\Theta, m_*^Z, K_*^Z) = 0; \quad A_*^K(\Theta, m_*^Z, K_*^Z) = 0; \\ \frac{dk^Z(\tau)}{dt} = A_*^k(k^Z(\tau), \Theta, m_*^Z, K_*^Z), \quad k^Z(0) = K_*^Z. \end{aligned} \right\} \quad (30)$$

Здесь

$$A_*^m(\Theta, m_*^Z, K_*^Z) = M_{\mathcal{N}}[a^Z]; \quad (31)$$

$$\begin{aligned} A_*^K(\Theta, m_*^Z, K_*^Z) = \\ = A_{1*}(\Theta, m_*^Z, K_*^Z) + A_{1*}(\Theta, m_*^Z, K_*^Z)^T + A_{2*}(\Theta, m_*^Z, K_*^Z), \end{aligned} \quad (32)$$

где

$$\begin{aligned} A_{1*}(\Theta, m_*^Z, K_*^Z) &= M_{\mathcal{N}}[a^Z(Z_t - m_*^Z)], \\ A_{2*}(\Theta, m_*^Z, K_*^Z) &= M_{\mathcal{N}}[\sigma_*(\Theta, Z_t)], \\ \sigma_*(\Theta, Z_t) &= \sigma_{0*}(\Theta, Z_t) + \int_{R_0^q} c^Z(\Theta, Z_t, u) c^Z(\Theta, Z_t, u)^T \nu_P(t, \Theta, du), \\ \sigma_{0*}(\Theta, Z_t) &= b^Z(\Theta, Z_t) \nu_{0*}(\Theta) b^Z(\Theta, Z_t)^T; \\ A_{\tau}^k(k^Z(\tau), \Theta, m_*^Z, K_*^Z) &= k^Z(\tau) (K_*^Z)^{-1} A_{1*}^T, \quad \tau = t_1 - t_2. \end{aligned} \quad (33)$$

**Теорема 3.3.** Пусть в условиях теоремы 3.2 параметры СтС НРОП не зависят от времени и имеет место ковариационно-стационарный СтП. Тогда его математические ожидания, ковариационная матрица и матрица ковариационных функций определяются уравнениями (30) при условиях конечности интегралов (31), (32) и (33) и асимптотической устойчивости матрицы  $A_{1*}$ .

**Замечание 3.1.** Для СтС с аддитивными шумами (30)–(33) уравнения МНА переходят в уравнения МСЛ [8, 9]. В этом случае требование к гладкости функций (20) можно снять. Приходим к уравнениями МСЛ для СтС (30):

$$\begin{aligned} \varphi_0 = 0; \quad \dot{K}_t^Z &= a^Z K_t^Z + K_t^Z (a^Z)^T + b_0^Z \nu (b_0^Z)^T, \quad K_{t_0}^Z = K_0^Z; \\ \frac{\partial K^Z(t_1, t_2)}{\partial t_2} &= K^Z(t_1, t_2) (a_{t_2}^Z)^T, \quad K^Z(t_1, t_1) = K_{t_1}^Z \quad (\text{теорема 3.4}); \end{aligned}$$

$$\begin{aligned} \varphi_{0*} = 0; \quad a_*^Z K_*^Z + K_*^Z (a_*^Z)^T + b_{0*}^Z \nu_* (b_{0*}^Z)^T &= 0; \\ \frac{dk^Z(\tau)}{d\tau} &= a_* k^Z(\tau), \quad k^Z(0) = K_*^Z \quad (\text{теорема 3.5}). \end{aligned}$$

Для существенно нелинейных СтС НРОП в основе приведения лежат следующие утверждения.

**Теорема 3.6.** Пусть нелинейная гауссовская СтС НРОП (20), (21) удовлетворяет условиям:

- 1<sup>0</sup> уравнения (20) допускают регрессионную линеаризацию относительно производной  $l$ -го порядка;
- 2<sup>0</sup> уравнение ФФ (21) разрешено относительно возмущений  $U_t$  и имеет единственное среднеквадратичное решение. Тогда система (20), (21) приводима к системе (25) при  $c^Z \equiv 0$ .

**Теорема 3.7.** В условиях теоремы 3.4 уравнения МАМ имеют вид (26)–(33) при  $c^Z \equiv 0$ .

## 4 Основные результаты

Для оценки качества СтС НРОП в [2] в случае гауссовских случайных параметров  $\Theta$ , представляющих собой нормальные (гауссовские) случайные величины, использовалась функция потерь, допускающая квадратичную аппроксимацию [10]

$$\rho = \rho(\Theta) = \rho(m^\Theta) + \sum_{i=1}^{p^\Theta} \rho'_i(m^\Theta) \Theta_i^0 + \sum_{i,j=1}^{p^\Theta} \rho''_{ij}(m^\Theta) \Theta_i^0 \Theta_j^0, \quad (34)$$

и показатель  $\varepsilon$ , равный

$$\varepsilon = \varepsilon_2^{1/4}, \quad \varepsilon_2 = M_N[\rho(\Theta)]^2 - \rho(m^\Theta)^2, \quad (35)$$

где

$$M_N[\rho(\Theta)^2] = \rho(m^\Theta)^2 + \rho'(m^\Theta)^T K^\Theta \rho'(m^\Theta) + 2\rho(m^\Theta) \text{tr}[\rho''(m^\Theta) K^\Theta] + \left\{ \text{tr}[\rho''(m^\Theta) K^\Theta] \right\}^2 + 2 \text{tr}[\rho''(m^\Theta) K^\Theta]^2. \quad (36)$$

Для переменных  $\Theta = \Theta(t)$ , определяемых ИКП (теоремы 2.1–2.5), естественно использовать формулы (34)–(36) для безусловных корреляционных характеристик. Для их вычисления используются статистически линеаризованные посредством ИКП (теоремы 2.6 и 2.7) уравнения МНА (теоремы 3.3–3.5) для условных характеристик.

Если качество СтС НРОП характеризовать безусловным математическим ожиданием  $\tilde{m}_t^z$  и использовать первое уравнение системы (26) для условного математического ожидания  $\tilde{\dot{m}}_t^z$ , то получим следующие соотношения:

$$\begin{aligned} \dot{m}_t^z = A_t^m \left( t, \tilde{m}_t^z, \tilde{K}_t^z, \Theta \right) &\approx A_{0t}^m \left( \tilde{m}_t^z, \tilde{K}_t^z, m_t^\Theta, K_t^\Theta \right) + \\ &+ \sum_{\rho=1}^r \int_{\Lambda_\rho} V_\rho(\lambda) x_{\rho t}^{A^m} \left( \lambda, \tilde{m}_t^z, \tilde{K}_t^z, m_t^\Theta, K_t^\Theta \right) d\lambda; \end{aligned} \quad (37)$$

$$\tilde{\dot{m}}_t^z = \tilde{m}_t^z + (\delta M_V \tilde{m}_t^z)^0; \quad (38)$$

$$\tilde{K}_t^z = \tilde{K}_t^z + M_V (\delta M_V \tilde{m}_t^z)^0 (\delta M_V \tilde{m}_t^z)^{0T}; \quad (39)$$

$$(M_V m_t^z)^\bullet = A_{0t}^m \left( \tilde{m}_t^z, \tilde{K}_t^z, m_t^\Theta, K_t^\Theta \right); \quad (40)$$

$$(\delta M_V \tilde{m}_t^z)^\bullet = \sum_{\rho=1}^r \int_{\Lambda_\rho} V_\rho(\lambda) x_t^{A^m} \left( \lambda, \tilde{m}_t^z, \tilde{K}_t^z, m_t^\Theta, K_t^\Theta \right) d\lambda. \quad (41)$$

Здесь  $M_V$  — символ математического ожидания для гауссовых белых шумов  $V_\rho = V_\rho(\lambda)$ ;  $A_{0t}^m(\tilde{m}_t^z, \tilde{K}_t^z, m_t^\Theta, K_t^\Theta)$  и  $x^{A^m}(\lambda, \tilde{m}_t^z, \tilde{K}_t^z, m_t^\Theta, K_t^\Theta)$  — коэффициенты регрессионной линеаризации векторной функции в (37). Решая линейные уравнения (41) методом ИКП, найдем ковариационную матрицу  $\tilde{K}_t^z$ :

$$\begin{aligned} \tilde{K}_t^z = \tilde{K}_t^z + \sum_{\rho=1}^r \int_{\Lambda_\rho} G_\rho(\lambda) x_{\rho t}^{A^m} \left( \lambda, \tilde{m}_t^z, \tilde{K}_t^z, m_t^\Theta, K_t^\Theta \right) \times \\ \times \left[ x_{\rho t}^{A^m} \left( \lambda, \tilde{m}_t^z, \tilde{K}_t^z, m_t^\Theta, K_t^\Theta \right) \right]^T d\lambda, \end{aligned} \quad (42)$$

где  $\tilde{K}_t^z$  определяется вторым уравнением (26) при условии  $K_t^\Theta = 0$ , когда

$$\dot{K}_t^z = A_t^k \left( t, \tilde{m}_t^z, \tilde{K}_t^z, m_t^\Theta, K_t^\Theta \right) \approx A_{0t}^k \left( t, \tilde{m}_t^z, \tilde{K}_t^z, m_t^\Theta \right). \quad (43)$$

Таким образом, имеем следующие утверждения.

**Теорема 4.1.** *Если СтС НРОП приводима к дифференциальной согласно теоремам 3.1 и 3.2 и уравнению (22), то аналитическое моделирование ее качества проводится на основе алгоритма (38)–(40), (42), (43).*

**Теорема 4.2.** *Если СтС НРОП приводима к дифференциальной согласно теоремам 3.1, 3.2, 3.3, то аналитическое моделирование ее качества проводится на основе конечных стационарных уравнений (30).*



Аналогично формулируются результаты в условиях теорем 3.4–3.5 для типовых СтС НРОП.

**Теорема 4.3.** Если СтС НРОП приводима к дифференциальной согласно теоремам 3.6, 3.7 и уравнению (22) при  $c^U \equiv 0$ , то аналитическое моделирование ее качества проводится на основе алгоритма (38)–(40), (42), (43).

**Теорема 4.4.** Если СтС НРОП приводима к дифференциальной согласно теоремам 3.6 и 3.7 и уравнению (22) при  $c^U \equiv 0$ , то аналитическое моделирование ее качества проводится на основе конечных стационарных уравнений (30).

**Замечание 4.1.** Вторые и третьи уравнения системы (26) могут использоваться при анализе качества СтС НРОП на основе следа  $R = \text{tr } \tilde{K}_t^z$ . В этом случае проводится линейная регрессионная линеаризация первого и второго уравнения (26), затем находятся диагональные компоненты  $\tilde{K}_t^z$  и, наконец, вычисляется  $\text{tr } \tilde{K}_t^z$ .

## 5 Пример

Рассмотрим скалярную СтС НРОП вида

$$\Phi(\dot{Y}_t) + \alpha Y_t + \beta U_t = 0, \quad Y_{t_0} = Y_0; \quad (44)$$

$$\dot{U}_t = \alpha_1 U_t + \beta_1 V, \quad U_{t_0} = 0, \quad (45)$$

где  $V = V(\Theta)$  — гауссовский белый шум интенсивности  $\nu = \nu(\Theta)$ . При фиксированном  $\Theta$  выполним с учетом (44) следующие преобразования. Проведем статистическую линеаризацию кубической нелинейности:

$$\begin{aligned} \Phi(\dot{Y}_t) &\approx \Phi_0(m_t^{\dot{Y}}, D_t^{\dot{Y}}) + k_1(m_t^{\dot{Y}}, D_t^{\dot{Y}}) \dot{Y}_t^0, \quad \dot{Y}_t^0 = \dot{Y}_t - m_t^{\dot{Y}}, \\ \Phi_0(m_t^{\dot{Y}}, D_t^{\dot{Y}}) &= m_t^{\dot{Y}} D_t^{\dot{Y}} \left[ 3 + \frac{(m_t^{\dot{Y}})^2}{D_t^{\dot{Y}}} \right], \quad k_1(m_t^{\dot{Y}}, D_t^{\dot{Y}}) = 3D_t^{\dot{Y}} \left[ 1 + \frac{(m_t^{\dot{Y}})^2}{D_t^{\dot{Y}}} \right]. \end{aligned}$$

В результате при  $k_1 = k_1(m_t^{\dot{Y}}, D_t^{\dot{Y}}) \neq 0$  из уравнения (44) для  $m_t^{\dot{Y}}$  и  $\dot{Y}_t^0$  получим соотношения

$$\Phi_0(m_t^{\dot{Y}}, D_t^{\dot{Y}}) + \alpha m_t^{\dot{Y}} = 0; \quad (46)$$

$$\dot{Y}_t^0 + \alpha' Y_t^0 + \beta' U_t^0 = 0; \quad \dot{U}_t^0 = \alpha_1 U_t^0 + \beta_1 V,$$

где  $\alpha' = \alpha k_1^{-1}$ ;  $\beta' = \beta k_1^{-1}$ . В силу (44) и (45) имеем следующую систему взаимосвязанных уравнений для условных дисперсий и ковариаций  $D_t^U$ ,  $D_t^{\dot{Y}}$ ,  $d_t^{\dot{Y}}$  и  $K_t^{YU}$ :

$$\dot{D}_t^U = 2\alpha_1 D_t^U + \nu(\Theta); \quad (47)$$

$$\dot{D}_t^Y = 2(\alpha' D_t^Y + \beta' K_t^{YU}); \quad (48)$$

$$\dot{K}_t^{YU} = (\alpha_1 + \alpha') K_t^{YU} + \beta' D_t^U; \quad (49)$$

$$D_t^{\dot{Y}} = \alpha'^2 D_t^Y + \beta'^2 D_t^U + 2\alpha'\beta' K_t^{YU} \quad (50)$$

при соответствующих начальных условиях. Далее зададим интенсивность гауссовского белого шума в виде ИКП:

$$\nu(\Theta) = \nu_0 + \nu_1 \Theta, \quad \Theta = \int_{t_0}^t V(\lambda) x(t, \lambda) d\lambda. \quad (51)$$

Здесь координатные функции, математическое ожидание и ковариационная функция соответственно равны:

$$x(t, \lambda) = q_1(t)q_1(\lambda)^{-1} \mathbf{1}(t - \lambda), \quad q_1(t) = e^{-\eta t}, \quad q_2(t) = q_1(t) \int_{t_0}^t q_1^{-2}(\lambda) d\lambda;$$

$$m^\Theta(t) = \nu_0,$$

$$K^\Theta(t_1, t_2) = \begin{cases} q_1(t_2)q_2(t_1) & \text{при } t_1 < t_2; \\ q_1(t_1)q_2(t_2) & \text{при } t_1 > t_2. \end{cases} \quad (52)$$

Безусловные вероятностные характеристики находятся путем численного решения системы взаимосвязанных нелинейных дифференциальных уравнений (47)–(49) и конечного уравнения (50).

В частности, при  $\nu_0 \gg \nu_1$  после линеаризации уравнений (46)–(50) относительно стационарных значений, определяемых соотношениями

$$2\alpha_1 D_*^U + \nu_0 = 0;$$

$$\Phi_0(m_*^{\dot{Y}}, D_*^{\dot{Y}}) + \alpha m_*^Y = 0, \quad D_*^{\dot{Y}} = \alpha'^2 D_*^Y + \beta'^2 D_*^U + 2\alpha'_* \beta'_* K_*^{YU},$$

$$\alpha'_* D_*^Y + \beta'_* K_*^{YU} = 0;$$

$$(\alpha + \alpha'_*) K_*^{YU} + \beta'_* D_*^U = 0,$$

для переменных

$$\xi = [\xi_1 \ \xi_2 \ \xi_3 \ \xi_4 \ \xi_5 \ \xi_6]^T,$$

где

$$\begin{aligned}\xi_1 &= \delta m_t^{\dot{Y}} = \tilde{m}_t^{\dot{Y}} - m_*^{\dot{Y}}, & \xi_2 &= \delta m_t^Y = \tilde{m}_t^Y - m_*^Y, \\ \xi_3 &= \delta D_t^{\dot{Y}} = \tilde{D}_t^{\dot{Y}} - D_*^{\dot{Y}}, & \xi_4 &= \delta D_t^Y = \tilde{D}_t^Y - D_*^Y, \\ \xi_5 &= \delta D_t^U = \tilde{D}_t^U - D_*^U, & \xi_6 &= \delta K_t^{YU} = \tilde{K}_t^{YU} - K_*^{YU},\end{aligned}$$

в силу (47)–(50) получим следующие линейные уравнения:

$$\dot{\xi} = A \xi + B \nu_1 \Theta.$$

Обозначая через  $w^A(t - \tau)$  весовую функцию, отвечающую постоянной матрице  $A$ , получим

$$\left(m_t^\xi\right)_1 = 0; \quad \left(K_t^\xi\right)_1 = \int_{t_0}^{\infty} \int_{t_0}^{\infty} w^A(t - \tau_1) w^A(t - \tau_2)^T K^\Theta(\tau_1, \tau_2) d\tau_1 d\tau_2,$$

где  $K^\Theta(\tau_1, \tau_2)$  определено (52).

Таким образом, в линейном приближении при малых интенсивностях  $\nu_1 \ll \nu_0$  стохастичность белого шума в (44), (45), (51) приводит только к флуктуациям условных вероятностных характеристик. В квадратичном приближении уже  $(m_t^\xi)_2 \neq 0$ .

## 6 Заключение

Для аналитического моделирования качества типовых СтС НРОП, содержащих стационарные и нестационарные гауссовские (нормальные) случайные параметры, описываемые ИКП, разработаны нелинейные корреляционные методы, основанные на методах нормальной аппроксимации и статистической линеаризации для условных вероятностных характеристик качества и среднеквадратичной регрессионной линеаризации.

Результаты могут быть использованы и для негауссовских СтС НРОП, приводимых к негауссовским дифференциальным СтС.

Представляет практический интерес, во-первых, развитие методов для других типов [3–5] регрессионной линеаризации уравнений для условных вероятностных характеристик, во-вторых, когда в основу уравнений для условных вероятностных характеристик принимаются уравнения методов моментов, квазимоментов и ортогональных разложений плотностей одно- и многомерных распределений и, в третьих, применение комбинированных методов аналитического (вероятностного) статистического моделирования.

## Литература

1. *Синицын И. Н.* Аналитическое моделирование широкополосных процессов в стохастических системах, не разрешенных относительно производных // Информатика

- и её применения, 2017. Т. 11. Вып. 1. С. 3–10. doi: 10.14357/19922264170101. EDN: YOCMVL.
2. *Синицын И. Н.* Параметрическое аналитическое моделирование процессов в стохастических системах, не разрешенных относительно производных // Системы и средства информатики, 2017. Т. 27. № 1. С. 21–45. doi: 10.14357/08696527170102. EDN: YODCZL.
  3. *Sinitsyn I. N.* Analytical modeling and estimation of normal processes defined by stochastic differential equations with unsolved derivatives // J. Mathematics Statistics Research, 2021. Vol. 3. Iss. 1. Art. 139. P. 1–7. doi: 10.36266/JMSR/139.
  4. *Синицын И. Н.* Аналитическое моделирование и оценивание нестационарных нормальных процессов в стохастических системах, не разрешенных относительно производных // Системы и средства информатики, 2022. Т. 32. № 2. С. 58–71. doi: 10.14357/08696527220206. EDN: YMGERJ.
  5. *Синицын И. Н.* Нормализация систем, стохастически не разрешенных относительно производных // Информатика и её применения, 2022. Т. 16. Вып. 1. С. 32–38. doi: 10.14357/19922264220105. EDN: LDFJJB.
  6. *Синицын И. Н.* Аналитическое моделирование распределений с инвариантной мерой в стохастических системах, не разрешенных относительно производных // Информатика и её применения, 2023. Т. 17. Вып. 1. С. 2–10. doi: 10.14357/19922264230101. EDN: QWXVXC.
  7. *Синицын И. Н.* Канонические представления случайных функций. Теория и применения. — М.: ТОРУС ПРЕСС, 2023. 816 с.
  8. *Пугачёв В. С., Синицын И. Н.* Стохастические дифференциальные системы. Анализ и фильтрация. — М.: Наука, 1990. 632 с.
  9. *Пугачёв В. С., Синицын И. Н.* Теория стохастических систем. — М.: Логос, 2000; 2004. 1000 с.
  10. *Евланов Л. Г., Константинов В. М.* Системы со случайными параметрами. — М.: Наука, 1976. 568 с.

*Поступила в редакцию 04.09.23*

---

---

## **ANALYTICAL MODELING OF STOCHASTIC SYSTEMS WITH RANDOM PARAMETERS AND UNSOLVED DERIVATIVES**

*I. N. Sinitsyn*<sup>1,2</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>Moscow State Aviation Institute (National Research University), 4 Volokolamskoe Shosse, Moscow 125933, Russian Federation

**Abstract:** The paper is devoted to nonlinear correlation methods for analytical modeling in differential stochastic systems with unsolved derivatives (StS USD) and random parameters. Survey is given. Necessary notations concerning integral canonical expansions (ICE) and its linear and nonlinear transforms

are presented. It is shown how differential StS USD can be reduced to differential StS. Basic quality analysis algorithms for reducible StS USD are described. Special attention is paid to multicomponent ICE theory of stochastic processes and StS USD reducible to the differential ones. Two types of nonlinear transforms based on linear ICE regression are developed. Normal approximation method is used for ordinary differential equations for conditional probabilistic characteristics: mathematical expectations, covariance matrix, and matrix of covariance functions. For unconditional characteristics, ICE method is implemented. Analytical modeling methods are presented both for stationary and nonstationary regimes. An illustrative example is given. Directions of future generalizations are given.

**Keywords:** analytical modeling; integral canonical expansion (ICE); normal approximation method (NAM); stochastic systems with unsolved derivatives (StS USD); conditional and unconditional characteristics

**DOI:** 10.14357/08696527240101

**EDN:** ZPTXJI

## Acknowledgments

The research was financially supported by the Russian Academy of Sciences (state registration number of R&D AAAA-A19-119091990037-5).

## References

1. Sinitsyn, I. N. 2017. Analiticheskoe modelirovanie shirokopolosnykh protsessov v stokhasticheskikh sistemakh, ne razreshennykh otnositel'no proizvodnykh [Analytical modeling of wide band processes in stochastic systems with unsolved derivatives]. *Informatika i ee Primeneniya — Inform. Appl.* 11(1):3–10. doi: 10.14357/19922264170101. EDN: YOCMVL.
2. Sinitsyn, I. N. 2017. Parametricheskoe analiticheskoe modelirovanie protsessov v stokhasticheskikh sistemakh, ne razreshennykh otnositel'no proizvodnykh [Parametric analytical modeling of wide band processes in stochastic systems with unsolved derivatives]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 27(1):21–45. doi: 10.14357/08696527170102. EDN: YODCZL.
3. Sinitsyn, I. N. 2021. Analytical modeling and estimation of normal processes defined by stochastic differential equations with unsolved derivatives. *J. Mathematics Statistics Research* 3(1):139. 7 p. doi: 10.36266/JMSR/139.
4. Sinitsyn, I. N. 2022. Analiticheskoe modelirovanie i otsenivanie nestatsionarnykh normal'nykh protsessov v stokhasticheskikh sistemakh, ne razreshennykh otnositel'no proizvodnykh [Analytical modeling and estimation of nonstationary normal processors with unsolved derivatives]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 32(2):58–71. doi: 10.14357/08696527220206. EDN: YMGERJ.
5. Sinitsyn, I. N. 2022. Normalizatsiya sistem, stokhasticheski ne razreshennykh otnositel'no proizvodnykh [Normalization of systems with stochastically unsolved derivatives]. *Informatika i ee Primeneniya — Inform. Appl.* 16(1):32–38. doi: 10.14357/19922264220105. EDN: LDFJJB.

6. Sinitsyn, I. N. 2023. Analiticheskoe modelirovanie raspredeleniy s invariantnoy meroy v stokhasticheskikh sistemakh, ne razreshennykh odnositel'no proizvodnykh [Analytical modeling of distributions with invariant measure in stochastic systems with unsolved derivatives]. *Informatika i ee Primeneniya — Inform. Appl.* 17(1):2–10. doi: 10.14357/19922264230101. EDN: QWXVXC.
7. Sinitsyn, I. N. 2023. *Kanonicheskie predstavleniya sluchaynykh funktsiy. Teoriya i primeneniya* [Canonical expansion of random functions. Theory and application]. Moscow: TORUS PRESS. 816 p.
8. Pugachev, V. S., and I. N. Sinitsyn. 1987. *Stochastic differential systems. Analysis and filtering*. Chichester–New York: John Wiley & Sons. 549 p.
9. Pugachev, V. S., and I. N. Sinitsyn. 2001. *Stochastic systems. Theory and applications*. Singapore: World Scientific. 908 p.
10. Evlanov, L. G., and V. M. Konstantinov. 1976. *Sistemy so sluchaynymi parametrami* [Systems with random parameters]. Moscow: Nauka. 568 p.

*Received September 4, 2023*

## **Contributor**

**Sinitsyn Igor N.** (b. 1940) — Doctor of Science in technology, professor, Honored scientist of RF, principal scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; professor, Moscow State Aviation Institute (National Research University), 4 Volokolamskoe Shosse, Moscow 125933, Russian Federation; sinitsin@dol.ru

## ЗАДАЧА РАЗДЕЛЕНИЯ РЕСУРСА БЕСПРОВОДНОЙ СЕТИ ПО ЗАДЕРЖКЕ ПЕРЕДАЧИ ЭЛАСТИЧНОГО ТРАФИКА\*

*И. А. Кочеткова<sup>1</sup>, А. С. Власкина<sup>2</sup>, П. Л. Э. Эсенг Манге<sup>3</sup>, В. С. Шоргин<sup>4</sup>*

**Аннотация:** Ожидается, что в будущих сетях шестого поколения (6G), или сетях-2030, получит распространение эластичное управление ресурсами. Оно подразумевает не только адаптивное занятие ресурса, но и учет в настройке меняющихся потребностей в услугах и поведения пользователей и непрерывно изменяемого состояния канала связи. Один из вариантов разделения ресурса между сегментами сети — справедливое не относительно требуемой скорости передачи данных, а относительно задержки передачи данных. В статье формализована задача разделения ресурса по задержке передачи данных в виде системы массового обслуживания с дисциплиной дискриминаторного разделения процессора между классами эластичного трафика. Параметры дисциплины — веса каждого класса — предлагается выбирать, ориентируясь на остаточное время передачи данных, а справедливость разделения ресурса оценивать по индексу Джейна по среднему времени передачи.

**Ключевые слова:** беспроводная сеть; сеть-2030; 6G; разделение ресурсов; справедливость по задержке передачи; эластичный трафик; система массового обслуживания; разделение процессора

**DOI:** 10.14357/08696527240102

**EDN:** WXYRQV

### 1 Введение

В настоящее время разрабатывается новое, шестое, поколение сетей, или сетей-2030 [1, 2]. Ключевым аспектом в построении сети 6G становится обеспечение требуемого качества обслуживания пользователей для новых услуг, например голографической коммуникации, цифровых двойников, и при нарезке сети [3, 4]. Требования к качеству такие, что существующие технологические и системные решения сетей предыдущих поколений не могут его обеспечить, поэтому прорабатывается архитектура будущих сетей, решения, характеристики [5]. Среди подходов выделяется так называемое эластичное управление ресурсами — не

---

\* Публикация выполнена в рамках проекта № 025319-2-000 Системы грантовой поддержки научных проектов РУДН.

<sup>1</sup> Российский университет дружбы народов имени Патриса Лумумбы; Федеральный исследовательский центр «Информатика и управление» Российской академии наук, kochetkova-ia@rudn.ru

<sup>2</sup> Российский университет дружбы народов имени Патриса Лумумбы, vlaskina-as@rudn.ru

<sup>3</sup> Российский университет дружбы народов имени Патриса Лумумбы, pedroluis\_eyamaeseng@yahoo.es

<sup>4</sup> Федеральный исследовательский центр «Информатика и управление» Российской академии наук, vshorgin@ipiran.ru

только адаптивное занятие ресурса, но и учет в настройке гибко меняющихся потребностей в услугах и поведения пользователей и непрерывно изменяемого состояния канала связи миллиметрового диапазона длин волн и терагерцевого диапазона радиочастот. Один из вариантов разделения ресурса — обеспечивать справедливость относительно задержки передачи данных [6, 7].

Структура работы следующая. В разд. 2 представлены основные особенности и задачи будущих сетей-2030, включая эластичное управление радиоресурсами. В разд. 3 описан ресурс сети, а в разд. 4 формализована задача разделения ресурса по задержке передачи эластичного трафика в виде системы массового обслуживания.

## **2 Сети-2030 и эластичное управление ресурсами**

Сети связи стремительно и динамично развиваются, появляются новые проводные и беспроводные технологии, благодаря которым скорости передачи данных измеряются уже не в мегабитах и гигабитах в секунду, а стремятся к значениям в терабитах в секунду, приложения требуют не только все более высокую пропускную способность, но и низкие задержки передачи данных и высокий уровень надежности. Ожидается, что число подключенных устройств вырастет с 28,5 млрд (2020 г.) до 100 млрд в 2025 г., а трафик увеличится с 33 ЗБайт (2018 г.) до 175 ЗБайт к 2025 г. Спектр предоставляемых пользователям услуг расширяется, растет уровень использования методов искусственного интеллекта и машинного обучения. Все перечисленные тенденции приводят к необходимости разработки нового поколения сетей, которые бы соответствовали будущим услугам и приложениям — сетей шестого поколения (6G) [1], или сетей-2030 (Network 2030) [2].

Сеть 6G представляет собой систему, компоненты системы и связанные с ними аспекты, которые позволяют организовать интегрированное, с высокой степенью автоматизации, с использованием интеллектуальных методов разделение инфраструктуры — разнородные коммуникации, вычислительные ресурсы, ресурсы для хранения данных, ресурсы для развертывания сетевых услуг и приложений — по всем сегментам сети: проводная и беспроводная связь, ядро сети, сеть граничных вычислений, спутниковая связь, комбинация перечисленных сегментов, — к которой пользователь может получить доступ из одного или нескольких местоположений [5].

Ключевой аспект в построении сети 6G — обеспечение требуемого качества обслуживания (Quality of Service, QoS) пользователей, включая сверхнизкую задержку передачи данных, сверхбольшой объем данных и т. д., причем по большей части данные требования должны быть выполнены для каждого «деления» «нарезанной» инфраструктуры. Концепция такого деления — нарезка сети (Network Slicing) — охватывает все сетевые и облачные сегменты: сеть радиодоступа, ядро сети, транспортную сеть, сеть граничных вычислений — и позволяет одновременно развертывать несколько логических, автономных и независимых



разделенных сетевых ресурсов и группы сетевых и сервисных функций на единой инфраструктурной платформе [3, 4].

Новые сценарии использования сетей 6G — голографическая коммуникация, тактильный интернет для удаленных операций, цифровые двойники и др. — выдвигают такие требования к скорости передачи данных, задержке, безопасности, поддержке алгоритмов на базе искусственного интеллекта, гетерогенности сети, что существующие технологические и системные решения сетей предыдущих поколений не могут обеспечить необходимое качество обслуживания. В настоящее время международные организации по стандартизации МСЭ (Международный союз электросвязи), 3GPP (3rd Generation Partnership Project), мировое научно-исследовательское сообщество, включая исследовательские центры компаний, прорабатывают детально архитектуру будущих сетей, их сценарии использования, требования, характеристики, технологические решения и пр.

Одной из важных характеристик сетей 6G выступает так называемая эластичность ресурса и эластичное управление ресурсами (Elastic Resource Management) — новое понятие, вобравшее в себя принципы динамического управления радиоресурсами (Dynamic Radio Resource Management) в беспроводных сетях и эластичных облачных вычислениях (Elastic Cloud Computing) и масштабировавшее их на всю инфраструктуру — сетевые ресурсы сети, включая принцип нарезки ресурсов [1]. Так, динамическое управление радиоресурсами направлено на гибкую настройку параметров сети радиодоступа под текущие нагрузки трафика, местоположение и мобильность пользователей, требования к качеству обслуживания, плотность базовых станций и т. п. для увеличения спектральной эффективности. Эластичность облачных вычислений выражается в способности адаптироваться к изменениям нагрузки за счет подключения и отключения вычислительных ресурсов для максимального соответствия запрашиваемых и выделяемых ресурсов в каждый момент времени.

Ресурс в эластичном управлении ресурсами — это не только ресурс сети радиодоступа и вычислительный ресурс, а все его перечисленные выше виды. Эластичность в управлении ресурсами — это не только адаптивное занятие ресурса, но и учет при этом в настройке гибко меняющихся потребностей в услугах и поведения пользователей и непрерывно изменяемого состояния канала связи, например миллиметрового диапазона длин волн и терагерцевого диапазона радиочастот.

### **3 Описание ресурса в беспроводной сети**

Рассмотрим соту беспроводной сети радиуса  $R$  [м] с базовой станцией с мощностью сигнала  $P_t$  [Вт] и шириной полосы пропускания  $B$  [Гц]. Для вычисления максимальной пропускной способности канала связи с учетом шума используется формула Шеннона–Хартли [5]. Пусть мощность шума составляет  $N_0$  [Вт], тогда пиковая пропускная способность соты

$$C = B \log_2 \left( 1 + \frac{P_t}{N_0} \right),$$

что фактически соответствует максимальной скорости передачи данных для единственного в соте пользователя, находящегося максимально близко к базовой станции, в отсутствие затухания мощности сигнала.

Учитывая большее число пользователей в соте и их разноудаленность от базовой станции, из формулы Шеннона–Хартли запишем максимальную скорость передачи данных для одного пользователя в виде

$$c = y_t y_f \log_2 \left( 1 + \frac{y'_p}{N_0} \right),$$

где  $y_f$  соответствует выделенной для пользователя ширине полосы пропускания,  $y'_p$  — мощности полученного сигнала, а  $y_t$  отражает долю слота времени, который также выделен для данного пользователя.

Мощность  $y'_p$  полученного сигнала отличается от мощности  $y_p$  передаваемого сигнала, а тройка  $(y_f, y_t, y_p)$  фактически представляет собой ресурс базовой станции, который выделяется для обеспечения необходимой скорости передачи данных. Связь между  $y'_p$  и  $y_p$  задается по формуле передачи сигнала Фрииса [5] и модели затухания сигнала:

$$y'_p = y_p \frac{G_t G_r}{y_L},$$

где  $G_t$  и  $G_r$  — коэффициенты усиления передающей и приемной антенн соответственно;  $y_L$  — коэффициент затухания сигнала.

В качестве моделей затухания сигнала можно использовать, например, те, которые представлены в спецификации [8]. Предположим случайное положение пользователей в соте, тогда случайная величина  $\xi_L$  затухания сигнала будет функцией от случайной величины  $\xi_d$  расстояния от пользователя до базовой станции. В данном случае искомой характеристикой выступает распределение  $p(n)$  вероятностей того, что в соте достаточно ресурсов для обслуживания  $n$  пользователей [5], как это сделано в статье [9].

#### 4 Формализация задачи разделения ресурса по задержке передачи эластичного трафика

Перейдем к постановке задачи разделения ресурса по задержке передачи эластичного трафика. Рассмотрим систему массового обслуживания с двумя классами эластичного трафика. Пусть потоки запросов на передачу блоков данных — пуассоновские с параметрами  $\lambda_1$  и  $\lambda_2$  [с<sup>-1</sup>], а объемы блоков данных

распределены по экспоненциальному закону со средними  $\mu_1^{-1}$  и  $\mu_1^{-2}$  [бит] соответственно. При передаче блока данных должна обеспечиваться минимальная скорость  $b$  [бит/с] [4].

Для того чтобы ориентироваться на справедливое (равное) с точки зрения задержки передачи блоков данных (Lantency Fairness) разделение ресурса [6, 7], воспользуемся дисциплиной дискриминаторного разделения процессора (Discriminatory Processor Sharing, DPS) или дисциплиной взвешенного циклического обслуживания (Weighted Round Robin, WRR) [10]. В отличие от дисциплины справедливого разделения процессора (Egalitarian Processor Sharing, EPS) или карусельной дисциплины диспетчеризации (Round Robin, RR), при которой ресурс делится поровну между всеми  $n$  пользователями  $1/n$  [3, 11–13], дисциплина DPS позволяет более гибко настроить разделение ресурса по весам  $w$ . Опишем функционирование системы при помощи случайного процесса  $(N_1(t), N_2(t))$  с состояниями вида  $(n_1, n_2)$ , где  $n_1$  и  $n_2$  — число передаваемых блоков данных 1-го и 2-го класса соответственно в некоторый момент времени. По дисциплине DPS доля объема ресурса, которую занимает  $i$ -блок, составляет  $w_i/(w_1n_1 + w_2n_2)$ ,  $i = 1, 2$ . Обозначим долю объема ресурса, занимаемого всеми  $i$ -блоками данных в состоянии  $(n_1, n_2)$ , как

$$\alpha_i(n_1, n_2) = \frac{w_i n_i}{w_1 n_1 + w_2 n_2}, \quad i = 1, 2, \quad \alpha_1(n_1, n_2) + \alpha_2(n_1, n_2) = 1.$$

Пространство состояний случайного процесса  $(N_1(t), N_2(t))$  имеет вид:

$$\mathcal{X} = \{(n_1, n_2) : n_1 \geq 0, n_2 \geq 0, b(n_1 + n_2) \geq C\},$$

а элементы матрицы интенсивностей переходов определяются соотношениями

$$\begin{aligned} q[(n_1, n_2), (n_1 + 1, n_2)] &= \lambda_1 p(n_1, n_2), & n_1 + n_2 < N; \\ q[(n_1, n_2), (n_1, n_2 + 1)] &= \lambda_2 p(n_1, n_2), & n_1 + n_2 < N; \\ q[(n_1, n_2), (n_1 - 1, n_2)] &= \frac{C}{\mu_1 \alpha_1(n_1, n_2)}, & n_1 > 0; \\ q[(n_1, n_2), (n_1, n_2 - 1)] &= \frac{C}{\mu_2 \alpha_2(n_1, n_2)}, & n_2 > 0, \end{aligned}$$

где  $p(n_1, n_2)$  — вероятность того, что в соте достаточно ресурсов для обеспечения минимальной скорости  $b$  передачи  $n_1 + n_2$  блоков данных.

Один из возможных вариантов задания весов  $w_i$ , чтобы в разделении ресурса учесть задержку передачи данных, — ориентироваться на остаточное время передачи блоков данных. В состоянии  $(n_1, n_2)$  это время для  $i$ -блока при условии, что состояние не будет изменяться, равно

$$T_i(n_1, n_2) = \frac{\mu_1^{-1}}{\alpha_i(n_1, n_2)C/n_i} = \frac{w_1 n_1 + w_2 n_2}{w_i \mu_i}, \quad i = 1, 2.$$

Таким образом, из того, что остаточное время передачи должно быть одинаковым для всех блоков данных  $T_1(n_1, n_2) = T_2(n_1, n_2)$ , следует, что веса можно взять равными  $w_1 = \mu_2$  и  $w_2 = \mu_1$ :

$$\alpha_1(n_1, n_2) = \frac{\mu_2 n_1}{\mu_2 n_1 + \mu_1 n_2}; \quad \alpha_2(n_1, n_2) = \frac{\mu_1 n_2}{\mu_2 n_1 + \mu_1 n_2}.$$

В качестве показателя эффективности разделения ресурса по задержке передачи данных можно воспользоваться индексом Джейна (Jain's Fairness Index) [14] относительно среднего времени  $W_i$  передачи блока данных:

$$J(W_1, W_2) = \frac{1}{2} \frac{(W_1 + W_2)^2}{W_1^2 + W_2^2}.$$

Здесь

$$W_1 = \frac{1}{\lambda_1(1-B)} \sum_{n_1=1}^{\lfloor C/b \rfloor} \sum_{n_2=0}^{\lfloor (C-bn_1)/b \rfloor} \pi(n_1, n_2);$$

$$W_2 = \frac{1}{\lambda_2(1-B)} \sum_{n_2=1}^{\lfloor C/b \rfloor} \sum_{n_1=0}^{\lfloor (C-bn_2)/b \rfloor} \pi(n_1, n_2),$$

где  $B$  — вероятность блокировки запроса на передачу блока данных:

$$B = \sum_{n_1=0}^{\lfloor C/b \rfloor} \sum_{n_2=0}^{\lfloor (C-bn_1)/b \rfloor} [1 - p(n_1, n_2)] \pi(n_1, n_2);$$

$\pi(n_1, n_2)$  — стационарное распределение.

## 5 Заключение

Разделение ресурса беспроводной сети на базе задержки передачи данных — один из подходов к эластичному управлению радиоресурсами. В данной работе построена модель обслуживания двух классов эластичного трафика в виде системы массового обслуживания с дисциплиной дискриминаторного разделения процессора в рамках решения задачи разделения ресурса. Для обеспечения справедливости относительно задержки передачи данных учитывается остаточное время передачи блоков данных, а в качестве показателя эффективности разделения ресурса предложено использовать индекса Джейна по среднему времени передачи блока. В дальнейшем планируется проведение численного анализа показателей эффективности и расширение модели на произвольное число классов.

## Литература

1. ITU TS. Network 2030 architecture framework, 2020.
2. 3GPP TS 23.501. System architecture for the 5G system (5GS). Release 18.4.0, 2023.
3. Кочеткова И. А., Власкина А. С., Ву Н. Н., Шоргин В. С. Система массового обслуживания с управляемым по сигналам перераспределением приборов для анализа нарезки ресурсов сети 5G // Информатика и её применения, 2021. Т. 15. Вып. 3. С. 91–97. doi: 10.14357/19922264210312.
4. Власкина А. С., Бурцева С. А., Кочеткова И. А., Шоргин С. Я. Управляемая система массового обслуживания с эластичным трафиком и сигналами для анализа нарезки ресурсов в сети радиодоступа // Информатика и её применения, 2022. Т. 16. Вып. 3. С. 90–96. doi: 10.14357/19922264220312.
5. Молчанов Д. А., Бегишев В. О., Самуйлов К. Е., Кучерявый Е. А. Сети 5G/6G: архитектура, технологии, методы анализа и расчета. — М.: РУДН, 2022. 516 с.
6. Lopez-Sanchez M., Villena-Rodriguez A., Gomez G., Martin-Vega F. J., Aguayo-Torres M. C. Latency fairness optimization on wireless networks through deep reinforcement learning // IEEE T. Veh. Technol., 2023. Vol. 72. No. 4. P. 5407–5412. doi: 10.1109/TVT.2022.3224368.
7. Kwon J., Ryu J., Lee J. H., Joung J. Improving end-to-end latency fairness using a reinforcement-learning-based network scheduler // Appl. Sci. — Basel, 2023. Vol. 13. No. 6. Art. 3397. doi: 10.3390/app13063397.
8. 3GPP TR 38.901. Study on channel model for frequencies from 0.5 to 100 GHz. Release 17.1.0, 2023.
9. Кочеткова И. А., Куцазли А. И., Харин П. А., Шоргин С. Я. Модель схемы приоритетного доступа трафика URLLC и eMBB в сети пятого поколения в виде ресурсной системы массового обслуживания // Информатика и её применения, 2021. Т. 15. Вып. 4. С. 87–92. doi: 10.14357/19922264210412.
10. Яшков С. Ф. Математические вопросы теории систем обслуживания с разделением процессора // Итоги науки и техники. Сер.: Теория вероятностей. Математическая статистика. Теоретическая кибернетика, 1990. Т. 29. С. 3–82.
11. Samouylov K. E., Gudkova I. A. Recursive computation for a multi-rate model with elastic traffic and minimum rate guarantees // Congress (International) on Ultra Modern Telecommunications and Control Systems and Workshops Proceedings. — Piscataway, NJ, USA: IEEE, 2010. P. 1065–1072. doi: 10.1109/ICUMT.2010.5676509.
12. Shorgin S. Y., Samouylov K. E., Gudkova I. A., Markova E. V., Sopin E. S. Approximating performance measures of radio admission control model for non real-time services with maximum bit rates in LTE // AIP Conf. Proc., 2015. Vol. 1648. Art. 250004. doi: 10.1063/1.4912508.
13. Kochetkova I., Leonteva K., Ghebrial I., Vlaskina A., Burtseva S., Kushchazli A., Samouylov K. Controllable queuing system with elastic traffic and signals for resource capacity planning in 5G network slicing // Future Internet, 2024. Vol. 16. No. 1. Art. 18. 23 p. doi: 10.3390/fi16010018.
14. Jain R., Chiu D. M., Hawe W. R. A quantitative measure of fairness and discrimination for resource allocation in shared computer systems: DEC Research Report TR-301. — Hudson, MA, USA: Eastern Research Laboratory, Digital Equipment Corp., 1984. 37 p.

Поступила в редакцию 15.01.24

## RADIO RESOURCE ALLOCATION FOR ELASTIC TRAFFIC TRANSMISSION WITH LATENCY FAIRNESS IN WIRELESS NETWORKS

I. A. Kochetkova<sup>1,2</sup>, A. S. Vlaskina<sup>1</sup>, P. L. E. Eseng Mangué<sup>1</sup>, and V. S. Shorgin<sup>2</sup>

<sup>1</sup>RUDN University, 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation

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

**Abstract:** Elastic resource management is anticipated to be prevalent in the forthcoming 6G, or 2030, networks. This approach entails not only adaptive allocation of resources but also the consideration of evolving service requirements, user behavior, and the dynamic state of the radio channel. One of the methods for distributing resources among network slices may not be equitable in terms of the required data transfer rates but aims for fairness concerning the data transmission delays. The paper formalizes the issue of resource allocation based on data transmission latency as a queuing system with a discriminatory processor-sharing discipline for classes of elastic traffic. The authors propose determining the discipline parameters — the weights for each class — according to the residual data transmission time. The fairness of resource allocation is evaluated using the Jain’s fairness index which is based on average transmission times.

**Keywords:** wireless network; network 2030; 6G; resource allocation; latency fairness; elastic traffic; queuing system; processor sharing

**DOI:** 10.14357/08696527240102

**EDN:** WXYPQV

### Acknowledgments

The publication has been supported by the RUDN University Scientific Projects Grant System, project No. 025319-2-000.

### References

1. ITU TS. Network 2030 architecture framework, 2020.
2. 3GPP TS 23.501. 2023. System architecture for the 5G system (5GS). Release 18.4.0.
3. Kochetkova, I. A., A. S. Vlaskina, N. N. Vu, and S. Ya. Shorgin. 2021. Queuing system with signals for dynamic resource allocation for analyzing network slicing in 5G networks. *Informatika i ee Primeneniya — Inform. Appl.* 15(3):91–97. doi: 10.14357/19922264220312.
4. Vlaskina, A. S., S. A. Burtseva, I. A. Kochetkova, and S. Ya. Shorgin. 2022. Controllable queuing system with elastic traffic and signals for analyzing network slicing. *Informatika i ee Primeneniya — Inform. Appl.* 16(3):90–96. doi: 10.14357/19922264220312.

5. Moltchanov, D. A., V. O. Begishev, K. E. Samouylov, and Y. A. Koucheryavy. 2022. *Seti 5G/6G: arkhitektura, tekhnologii, metody analiza i rascheta* [5G/6G networks: Architecture, technologies, analysis methods, and calculations]. Moscow: RUDN University. 516 p.
6. Lopez-Sanchez, M., A. Villena-Rodriguez, G. Gomez, F.J. Martin-Vega, and M.C. Aguayo-Torres. 2023. Latency fairness optimization on wireless networks through deep reinforcement learning. *IEEE T. Veh. Technol.* 72(4):5407–5412. doi: 10.1109/TVT.2022.3224368.
7. Kwon, J., J. Ryu, J.H. Lee, and J. Joung. 2023. Improving end-to-end latency fairness using a reinforcement-learning-based network scheduler. *Appl. Sci. — Basel* 13(6):3397. doi: 10.3390/app13063397.
8. 3GPP TR 38.901. 2023. Study on channel model for frequencies from 0.5 to 100 GHz. Release 17.1.0.
9. Kochetkova I. A., A. I. Kushchazli, P. A. Kharin, and S. Ya. Shorgin. 2021. Model' skhemy prioritetnogo dostupa trafika URLLC i eMBB v seti pyatogo pokoleniya v vide resursnoy sistemy massovogo obsluzhivaniya [Model for analyzing priority admission control of URLLC and eMBB communications in 5G networks as a resource queuing system]. *Informatika i ee Primeneniya — Inform. Appl.* 15(4):87–92. doi: 10.14357/19922264210412.
10. Yashkov, S. F. 1992. Mathematical problems in the theory of shared-processor systems. *J. Soviet Mathematics* 58(2):101–147. doi: 10.1007/BF01097426.
11. Samouylov, K. E. and I. A. Gudkova. 2010. Recursive computation for a multi-rate model with elastic traffic and minimum rate guarantees. *Congress (International) on Ultra Modern Telecommunications and Control Systems and Workshops Proceedings*. Piscataway, NJ: IEEE. 1065–1072. doi: 10.1109/ICUMT.2010.5676509.
12. Shorgin, S. Y., K. E. Samouylov, I. A. Gudkova, E. V. Markova, and E. S. Sopin. 2015. Approximating performance measures of radio admission control model for non real-time services with maximum bit rates in LTE. *AIP Conf. Proc.* 1648:250004. doi: 10.1063/1.4912508.
13. Kochetkova, I., K. Leonteva, I. Ghebrial, A. Vlaskina, S. Burtseva, A. Kushchazli, and K. Samouylov. 2024. Controllable queuing system with elastic traffic and signals for resource capacity planning in 5G network slicing. *Future Internet* 16(1):18. 23 p. doi: 10.3390/fi16010018.
14. Jain R., D. M. Chiu, W. R. Hawe. 1984. A quantitative measure of fairness and discrimination for resource allocation in shared computer systems. Hudson, MA: Eastern Research Laboratory, Digital Equipment Corp. DEC Research Report TR-301. 37 p.

Received January 15, 2024

## Contributors

**Kochetkova Irina A.** (b. 1985) — Candidate of Science (PhD) in physics and mathematics, associate professor, Department of Probability Theory and Cyber Security, RUDN University, 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; senior scientist, Federal Research Center “Computer Science and

Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; kochetkova-ia@rudn.ru

**Vlaskina Anastasiya S.** (b. 1995) — Candidate of Science (PhD) in physics and mathematics, teaching assistant, Department of Probability Theory and Cyber Security, RUDN University, 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; vlaskina-as@rudn.ru

**Eseng Mangué Pedro Luis Eyama** (b. 1992) — PhD student, Department of Probability Theory and Cyber Security, RUDN University, 6 Miklukho-Maklaya Str., Moscow 117198, Russian Federation; pedroluis\_eyamaeseng@yahoo.es

**Shorgin Vsevolod S.** (b. 1978) — Candidate of Science (PhD) in technology, senior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; vshorgin@ipiran.ru



## МЕТОДИКА ДЕСИНХРОНИЗАЦИИ ПРИ СИНТЕЗЕ САМОСИНХРОННЫХ СХЕМ\*

*Ю. А. Степченков<sup>1</sup>, Д. В. Хилько<sup>2</sup>, Ю. Г. Дьяченко<sup>3</sup>, Н. В. Морозов<sup>4</sup>,  
Д. Ю. Степченков<sup>5</sup>, Г. А. Орлов<sup>6</sup>*

**Аннотация:** Самосинхронные (СС) цифровые схемы обладают несомненными преимуществами в сравнении с синхронными аналогами. Однако синтез СС-схем в настоящее время требует от пользователя глубоких познаний в области СС-схемотехники, поскольку уровень его автоматизации еще далек от уровня, достигнутого в промышленных системах автоматизированного проектирования (САПР) синхронных схем, и ориентирован на задание поведения синтезируемой схемы в специфичных форматах. Проектирование СС-схем более трудоемко и специфично в сравнении с синхронными схемами из-за необходимости соблюдения строгих принципов их реализации. Важный этап синтеза СС-схемы из исходного Verilog-описания синхронной схемы — десинхронизация, которая обеспечивает абстрагирование от глобального тактового сигнала и подготовку к использованию асинхронного запрос-ответного взаимодействия. Статья рассматривает методику ее реализации и принципы формализации, обеспечивающие корректное построение СС-схемы на основе эвристических алгоритмов определения взаимосвязи функциональных блоков (ФБ) в составе синтезируемой схемы и организации их взаимодействия в строгом соответствии с дисциплиной работы СС-схемы.

**Ключевые слова:** синхронная схема; Verilog-описание; самосинхронная схема; автоматизированный синтез; десинхронизация; индикация; управление

**DOI:** 10.14357/08696527240103

**EDN:** XGZCWU

### 1 Введение

Принципы построения СС-схем были сформулированы Д. Маллером в 1950-х гг. [1]. Самосинхронные схемы, в отличие от синхронных аналогов, не

\* Исследование выполнено в рамках гранта Российского научного фонда (проект 22-19-00237).

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

<sup>2</sup> Федеральний исследовательский центр «Информатика и управление» Российской академии наук, DKhilko@ipiran.ru

<sup>3</sup> Федеральний исследовательский центр «Информатика и управление» Российской академии наук, diauga@mail.ru

<sup>4</sup> Федеральний исследовательский центр «Информатика и управление» Российской академии наук, NMorozov@ipiran.ru

<sup>5</sup> Федеральний исследовательский центр «Информатика и управление» Российской академии наук, Stepchenkov@mail.ru

<sup>6</sup> Федеральний исследовательский центр «Информатика и управление» Российской академии наук, davlubobrov@yandex.ru

нуждаются во внешней синхронизации в виде глобального тактового сигнала. Они функционируют на основе сигналов, разрешающих переключение и подтверждающих его успешное завершение [2–5]. В зависимости от предположений о временных задержках элементов и цепей традиционно выделяют четыре типа СС-схем, которые предполагают различные методы проектирования всей системы [2]:

- (1) нечувствительные к задержкам (delay-insensitive, DI);
- (2) квазинечувствительные к задержкам, (quasi-delay-insensitive, QDI);
- (3) не зависящие от скорости (speed-independent, SI);
- (4) со связанными данными (bundled-data, BD).

Поведение схем типа DI, QDI и SI не зависит от задержек сигналов в элементах. DI-схемы нечувствительны также и к задержкам в цепях, в то время как QDI-схемы предполагают, что задержки сигналов на отрезках цепи после разветвления одинаковые, а SI-схемы считают задержки любой цепи нулевыми. BD-схемы используют встроенные линейки согласованных задержек вместо индикации завершения переключений, поэтому они не обладают свойством самопроверяемости по отношению к константным неисправностям.

В зарубежной литературе известны несколько программных средств для разработки и верификации СС-схем: Balsa [6], Petrify [7], UNCLE [8], Workcraft [9]. Однако все они предполагают, что пользователь хорошо знаком с теорией СС-схем. Они используют специальные языки и графические средства для описания поведения СС-схем, но не всегда обеспечивают получение результата, действительно представляющего собой СС-схему [2].

Альтернативный подход к проектированию СС-схем базируется на формализованном преобразовании исходного поведенческого или аппаратного описания синхронной схемы на языке описания аппаратуры Verilog или VHDL в самосинхронное с помощью существующих САПР синхронных схем [10–12]. Такое преобразование получило название «десинхронизация». Процедура десинхронизации автоматически заменяет исходную тактовую сеть сетью локальных сигналов управления, обеспечивающих запрос-ответное взаимодействие функциональных устройств синтезируемой схемы. Однако лишь некоторые известные средства синтеза (например, [13]) направлены на получение СС-схем, отличных от схем BD-класса.

Данная статья посвящена разработке методики десинхронизации, гарантирующей синтез работоспособного варианта СС-схемы классов DI, QDI и SI пользователем, малознакомым с самосинхроникой, на основе принципов и формальных методов построения СС-схем, обеспечивающих получение СС-схемы с близкими к оптимальным параметрами.

## 2 Принципы десинхронизации

В маршруте синтеза СС-схем по исходному синхронному поведенческому Verilog-описанию [14] этап десинхронизации занимает важное место. Он

предшествует этапам СС-кодирования информационных сигналов схемы и построения индикаторной подсхемы. Назначение этапа десинхронизации — преобразовать управление схемой к виду, отвечающему СС-дисциплине.

Исходное поведенческое описание синтезируемой синхронной схемы предполагает использование глобального тактового сигнала. В частном случае могут использоваться и локальные синхросигналы, производные от глобального сигнала. Глобальный тактовый сигнал управляет процессом обработки данных в ФБ синтезируемой схемы и обменом результатами этой обработки между блоками. Частота глобального тактового сигнала подбирается такой, чтобы обеспечить корректную работу синхронной схемы с учетом возможного ухудшения быстродействия ее ФБ во время эксплуатации из-за неблагоприятного воздействия окружающей среды и изменения параметров активных элементов схем ФБ в результате процессов старения.

В отличие от синхронного аналога, СС-схема управляет обработкой данных и их обменом между ФБ на основе сигналов запрос-ответной дисциплины, подтверждающих готовность ФБ к соответствующей операции. Рисунок 1 демонстрирует пример типовой СС-схемы, ФБ которой организованы в конвейер. Каждая ступень СС-конвейера содержит комбинационную часть (КЧ) и выходной регистр (ВР), а также индикаторные подсхемы КЧ (ИКЧ) и ВР (ИВР) и гистерезисный триггер (Г-триггер, [15]), обозначенный символом «Н». Индикаторные подсхемы ИКЧ<sub>*k*</sub> и ИВР<sub>*k*</sub> вместе с Г-триггером *k*-й ступени формируют сигнал, подтверждающий завершение переключения *k*-й ступени в текущую фазу работы и разрешающий фазовые переключения ВР (*k* – 1)-й ступени.

В результате глобальный тактовый сигнал управления синхронной схемой замещается подмножеством локальных сигналов управления  $E_k$ . Число сигналов  $E_k$  не превышает число ФБ синтезируемой схемы, ассоциирующихся с модулями в исходном Verilog-описании ее поведения.

В общем случае синтезируемая схема представляет собой совокупность ФБ, взаимодействующих друг с другом с помощью потоков данных, регулируемых сигналами управления. В синхронной схеме сигналы управления формируются на основе глобального тактового сигнала и условий, разрешающих работу того

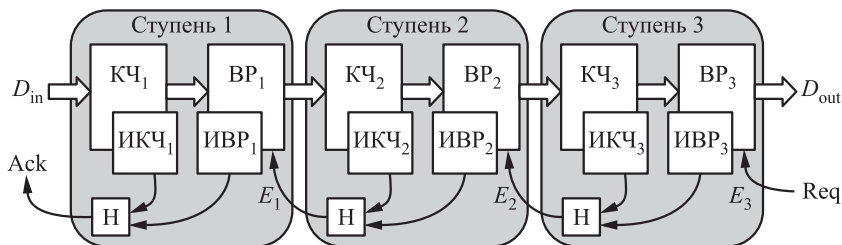


Рис. 1 Структура типовой СС-конвейера

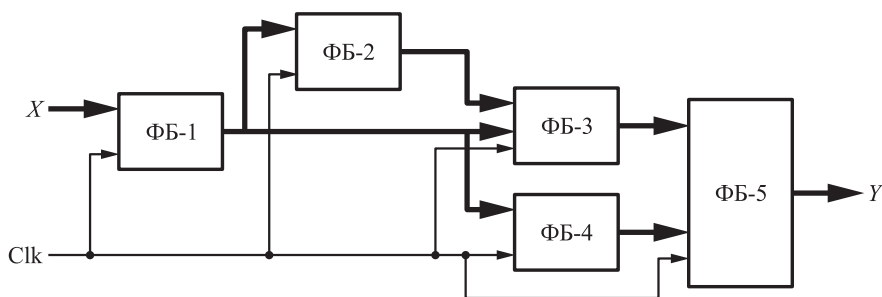


Рис. 2 Структурная синхронная схема

или иного ФБ в текущей операции. В СС-схеме сигналы управления формируются на основе индикаторных выходов ФБ. На рис. 2 показан пример схемы, в которой ФБ управляются глобальным тактовым сигналом  $Clk$ . Внутри ФБ могут использоваться локальные синхросигналы, производные от  $Clk$  и учитывающие те или иные логические условия. Поток данных обозначены широкими стрелками.

Базовый принцип функционирования СС-схемы заключается в следующем:  $k$ -й ФБ может начать переключение в спейсер (рабочую фазу) только в том случае, если ФБ, формирующие его входы, завершили свое переключение в спейсер (рабочую фазу), а ФБ — приемники выходов  $k$ -го ФБ — подтвердили свое переключение в рабочую фазу (спейсер). Успешное завершение переключения ФБ в каждую фазу подтверждается его индикаторным выходом. Индикаторные сигналы и формируемые на их основе сигналы управления должны играть роль диспетчеров процесса обработки потока данных без участия глобальных и локальных тактовых сигналов. Поэтому этап десинхронизации при синтезе СС-схемы должен выявить пары ФБ, обменивающихся между собой данными, и связать их друг с другом локальными сигналами управления, регулирующими фазовые переходы ФБ. Если у  $k$ -го ФБ несколько приемников (последователей), все они должны участвовать в формировании входа управления  $k$ -м ФБ.

Применение десинхронизации к схеме на рис. 2 приводит к СС-схеме, показанной на рис. 3. Выходы ФБ-1–ФБ-4 имеют парафазное кодирование. Выходы ФБ-5 могут быть бифазными, если их приемником служит синхронное окружение. ФБ-3 служит приемником для ФБ-1, однако его индикаторный выход  $E_2$  управляет ФБ-1 опосредованно через индикаторный выход ФБ-2, который вместе с индикаторным выходом ФБ-4 формирует сигнал управления  $E_1$ . Цель десинхронизации — выделить такого рода зависимости из исходного поведенческого описания синтезируемой схемы и реализовать их совокупностью локальных сигналов управления.

Входы  $ReqI$  и  $ReqO$  на рис. 3 служат сигналами запроса, формируемыми входным и выходным интерфейсом соответственно. Выходы  $AskI$  и  $AskO$

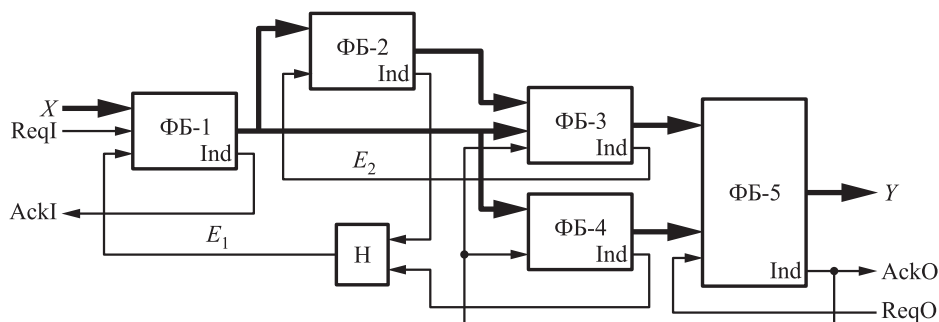


Рис. 3 Структурная СС-схема

служат сигналами ответа для входного и выходного интерфейсов соответственно. Вход ReqI может отсутствовать, если информационные входы  $X$  представлены в парафазном коде со спейсером.

Методика десинхронизации базируется на следующих принципах:

- синтезируемая схема рассматривается как совокупность ФБ, связанных друг с другом потоками данных и сигналами управления;
- тактовый вход ФБ замещается сигналом управления, формируемым подсхемой из Г-триггеров, объединяющей индикаторные выходы ФБ, представляющих собой прямых последователей данного ФБ;
- синхронная предустановка и запись заменяются СС-предустановкой с добавлением необходимой подсхемы окружения, реализующей процедуру СС-предустановки.

Таким образом, в процессе синтеза СС-схемы процедура десинхронизации должна заменить «деревья» глобального и локальных тактовых сигналов сетью сигналов фазового управления, производных от индикаторных выходов устройств синтезируемой схемы. Следует отметить, что приемниками сигналов управления в СС-схемах, как правило, служат регистры хранения результатов, регистры сдвига и счетчики.

### 3 Методика десинхронизации

Методика десинхронизации заключается в выполнении следующих процедур [14]:

- импорт исходного поведенческого Verilog-описания схемы и преобразование его в совокупность системы логических функций, иерархических модулей и блоков always;
- ввод явной иерархии — структурирование описания схемы путем выделения счетчиков, регистров, арифметических блоков, описываемых блоками always с многообразными переменными, в отдельные модули;

- разделение исходного описания схемы на ФБ на основе информации о парах «источник–приемник» для всех внутренних сигналов;
- определение особенностей работы ФБ (выполняемая функция, наличие динамической предустановки и т. д.);
- извлечение глобальных и локальных имен сигналов синхронизации, предустановки и производных от них сигналов;
- присвоение структурным элементам схемы, использующим сигналы синхронизации и предустановки, специальных атрибутов наследования, содержащих информацию об именах реальных сигналов — «родителей» портов синхровхода и предустановки;
- разделение совокупности структурных единиц и логических функций синтезируемой схемы на ступени конвейера;
- добавление в описание синтезируемой схемы функций подсхем управления и СС-предустановки.

Ввод иерархии заключается в выделении законченных функциональных устройств (арифметических блоков, счетчиков и регистров) в модули, образующие иерархию, что облегчает дальнейший анализ и синтез. Алгоритм такого преобразования включает анализ тела блоков *always* в составе исходного описания, извлечение типа функционального устройства, реализующего данный блок в схеме, ввод в описание синтезируемой схемы нового модуля, отображающего это устройство, и замену блоков *always* вызовами соответствующих модулей.

Как правило, сложные цифровые схемы имеют конвейерную структуру. Выявление структуры конвейера и привязка элементов схемы к его ступеням позволяет выявить дерево логических элементов внутри ступеней и построить оптимальную с точки зрения быстродействия индикаторную подсхему. Конвейер может иметь разветвленную структуру [16, 17].

Взаимодействие ступеней конвейера и ФБ в каждой ступени обеспечивается подсхемой управления. Она синтезируется с помощью эвристического формализованного подхода, основанного на анализе структуры конвейера и связности ФБ в его ступенях. Подсхема управления обеспечивает реализацию СС-дисциплины формирования и обработки промежуточных и выходных данных в конвейере, надлежащую индикацию синтезируемой схемы, реализацию ее запрос-ответного взаимодействия с входным и выходным окружением.

Сигналы управления формируются на основе совокупности индикаторных выходов ФБ и регистров, а также индикаторных сигналов комбинационных фрагментов путем их объединения подсхемой из Г-триггеров. Сигналы управления регулируют фазовые переходы регистров и преобразуют унарные и бифазные информационные сигналы в парафазные сигналы.

Использование при синтезе эвристических алгоритмов преобразования синхронной схемы в самосинхронную и шаблонов СС-реализации типовых цифровых устройств позволяет получить схемотехнические решения, близкие к результату ручного проектирования.

## 4 Заключение

Потребительские характеристики синтезируемых СС-схем зависят от эффективности методики десинхронизации. Она преобразует исходное синхронное описание синтезируемой схемы в устройство, работающее в соответствии с СС-дисциплиной взаимодействия его структурных частей и не содержащее глобальных и локальных цепей синхронизации.

Эвристические алгоритмы определения пар участников запрос-ответного взаимодействия, агрегирования фрагментов исходного описания синтезируемой схемы в ФБ и регистры хранения данных, разделения схемы на ступени конвейера обеспечивают формализацию и автоматизацию предложенной методики десинхронизации.

Предложенная методика десинхронизации позволяет интегрировать ее в промышленную САПР синхронных схем для придания ей нового качества — возможности синтеза СС-схем с приемлемыми потребительскими характеристиками силами разработчиков, не имеющих высокой квалификации в области СС-мотехники.

## Литература

1. *Muller D., Bartky W.* A theory of asynchronous circuits // Symposium (International) on the Theory of Switching Proceedings. — Harvard University Press, 1959. Vol. 1. P. 204–243.
2. *Sparsø J., Furber S.* Principles of asynchronous circuit design: A systems perspective. — Amsterdam: Kluwer Academic Pubs., 2001. 337 p. doi: 10.1007/978-1-4757-3385-3.
3. *Соколов И. А., Степченко Ю. А., Бобков С. Г., Захаров В. Н., Дьяченко Ю. Г., Рождественский Ю. В., Сурков А. В.* Базис реализации супер-ЭВМ эксафлопсного класса // Информатика и её применения, 2014. Т. 8. Вып. 1. С. 45–70. doi: 10.14357/1992264140106. EDN: RYYFGD.
4. *Yoshikawa S., Sannomiya S., Iwata M., Nishikawa H.* Pipeline stage level simulation method for self-timed data-driven processor on FPGA // 8th Electrical Engineering Congress (International). — Piscataway, NJ, USA: IEEE, 2020. Art. 229515. 5 p. doi: 10.1109/iEECON48109.2020.229515.
5. *Kushnerov A., Medina M., Yakovlev A.* Towards hazard-free multiplexer based implementation of self-timed circuits // 27th Symposium (International) on Asynchronous Circuits and Systems Proceedings. — Piscataway, NJ, USA: IEEE, 2021. P. 17–24. doi: 10.1109/ASYNC48570.2021.00011.
6. *Edwards D., Bardsley A., Jani L., Plana L., Toms W.* Balsa: A tutorial guide. — Manchester, U.K.: The University of Manchester, 2006. 157 p. <https://apt.cs.manchester.ac.uk/ftp/pub/apt/balsa/3.5/BalsaManual3.5.pdf>.
7. *Cortadella J., Kishinevsky M., Kondratyev A., Lavagno L., Yakovlev A.* Petri-fy: A tool for manipulating concurrent specifications and synthesis of asynchronous controllers // IEICE T. Inf. Syst., 1997. Vol. E80-D. No. 3. P. 315–325.
8. *Reese R. B., Smith S. C., Thornton M. A.* Uncle — an RTL approach to asynchronous design // 18th Symposium (International) on Asynchronous Circuits and



- Systems Proceedings. — Piscataway, NJ, USA: IEEE, 2012. P. 65–72. doi: 10.1109/ASYNC.2012.14.
9. *Poliakov I., Sokolov D., Mokhov A.* Workcraft: A static data flow structure editing, visualisation and analysis tool // Petri nets and other models of concurrency / Eds. J. Kleijn, A. Yakovlev. — Lecture notes in computer science ser. — Berlin, Heidelberg: Springer, 2007. Vol. 4546. P. 505–514. doi: 10.1007/978-3-540-73094-1\_30.
  10. *Cortadella J., Kondratyev A., Lavagno A., Christos P.* Desynchronization: Synthesis of asynchronous circuits from synchronous specifications // IEEE T. Comput. Aid., 2006. Vol. 25. Iss. 10. P. 1904–1921. doi: 10.1109/TCAD.2005.860958.
  11. *Andrikos N., Lavagno L., Pandini D., Sotiriou C. P.* A fully automated desynchronization flow for synchronous circuits // 44th ACM/IEEE Design Automation Conference Proceedings. — Piscataway, NJ, USA: IEEE, 2007. P. 982–985.
  12. *Zhou R., Chong K.-S., Gwee B.-H., Chang J. S.* Quasi-delay-insensitive compiler: Automatic synthesis of asynchronous circuits from verilog specification // 54th Midwest Symposium (International) on Circuits and Systems Proceedings. — Piscataway, NJ, USA: IEEE, 2011. Art. 6026577. 4 p. doi: 10.1109/MWSCAS.2011.6026577.
  13. *Zhou R., Chong K.-S., Gwee B.-H., Chang J. S., Ho W.-G.* Synthesis of asynchronous QDI circuits using synchronous coding specifications // Symposium (International) on Circuits and Systems Proceedings. — Piscataway, NJ, USA: IEEE, 2014. P. 153–156. doi: 10.1109/ISCAS.2014.6865088.
  14. *Зацаринный А. А., Степченко Ю. А., Дьяченко Ю. Г., Морозов Н. В., Степченко Д. Ю.* Автоматизация синтеза самосинхронных схем // Системы высокой доступности, 2023. Т. 19. № 3. С. 48–56. doi: 10.18127/j20729472-202303-04. EDN: IAPVXR.
  15. *Zakharov V., Stepchenkov Y., Diachenko Y., Rogdestvenski Y.* Self-timed circuitry retrospective // Conference (International) on Engineering Technologies and Computer Science Proceedings. — Piscataway, NJ, USA: IEEE, 2020. P. 58–64. doi: 10.1109/EnT48576.2020.00018.
  16. *Соколов И. А., Степченко Ю. А., Дьяченко Ю. Г., Морозов Н. В., Дьяченко Д. Ю.* Самосинхронный конвейер с переменным числом ступеней // Системы и средства информатики, 2023. Т. 33. № 1. С. 4–13. doi: 10.14357/08696527230101. EDN: FVYPZG.
  17. *Степченко Ю. А., Дьяченко Ю. Г., Степченко Д. Ю., Дьяченко Д. Ю., Орлов Г. А.* Мультиплексируемый самосинхронный конвейер // Системы и средства информатики, 2023. Т. 33. № 2. С. 4–12. doi: 10.14357/08696527230201.

Поступила в редакцию 29.12.23



## DESYNCHRONIZATION METHODOLOGY AT SELF-TIMED CIRCUIT SYNTHESIS

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

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation

**Abstract:** Self-timed (ST) digital circuits have undoubted advantages over synchronous counterparts. However, ST circuit synthesis currently requires the user to have deep knowledge in the ST circuit technology field, since its automation level is still far from the level achieved in industrial computer-aided design systems for synchronous circuits and is focused on assigning the synthesized circuit behavior in specific formats. Designing of ST circuits is more labor-intensive and specific in comparison with synchronous circuits because of the need to adhere to the strict principles for their implementation. Desynchronization is an important stage in the ST circuit synthesis based on the original Verilog description of the circuit operation algorithm. It provides circuit separation from the global clock and asynchronous request-acknowledge interaction usage preparation. The article considers the desynchronization implementation methodology and its formalization principles. The proposed method ensures an ST circuit correct construction based on heuristic algorithms determining the relationships between functional blocks in the synthesized circuit and organizing their interaction in strict accordance with the ST circuit operation discipline.

**Keywords:** synchronous circuit; Verilog description; self-timed circuit; automated synthesis; desynchronization; indication; control

**DOI:** 10.14357/08696527240103

**EDN:** XGZCWU

### Acknowledgments

The research was supported by the Russian Science Foundation (project No. 22-19-00237).

### References

1. Muller, D. E., and W. C. Bartky. 1959. A theory of asynchronous circuits. *Symposium (International) on the Theory of Switching Proceedings*. Harvard University Press. 1:204–243.
2. Sparsø, J., and S. Furber. 2001. *Principles of asynchronous circuit design: A systems perspective*. Amsterdam: Kluwer Academic Pubs. 337 p. doi: 10.1007/978-1-4757-3385-3.
3. Sokolov, I. A., Yu. A. Stepchenkov, S. G. Bobkov, V. N. Zakharov, Yu. G. Diachenko, Yu. V. Rozhdestvenskiy, and A. V. Surkov. 2014. Bazis realizatsii super-EVM eksaflopsnogo klassa [Implementation basis of exaflops class supercomputer]. *Informatika*

- i ee Primeniya* — *Inform. Appl.* 8(1):45–70. doi: 10.14357/19922264140106. EDN: RYYFGD.
4. Yoshikawa, S., S. Sannomiya, M. Iwata, and H. Nishikawa. 2020. Pipeline stage level simulation method for self-timed data-driven processor on FPGA. *8th Electrical Engineering Congress (International)*. Piscataway, NJ: IEEE. Art. 229515. 5 p. doi: 10.1109/IEECON48109.2020.229515.
  5. Kushnerov, A., M. Medina, and A. Yakovlev. 2021. Towards hazard-free multiplexer based implementation of self-timed circuits. *27th Symposium (International) on Asynchronous Circuits and Systems Proceedings*. Piscataway, NJ: IEEE. 17–24. doi: 10.1109/ASYNC48570.2021.00011.
  6. Edwards, D., A. Bardsley, L. Jani, L. Plana, and W. Toms. 2006. *Balsa: A tutorial guide*. Manchester, U.K.: The University of Manchester. 157 p. Available at: <https://apt.cs.manchester.ac.uk/ftp/pub/apt/balsa/3.5/BalsaManual3.5.pdf> (accessed February 28, 2024).
  7. Cortadella, J., M. Kishinevsky, A. Kondratyev, L. Lavagno, and A. Yakovlev. 1997. Petrify: A tool for manipulating concurrent specifications and synthesis of asynchronous controllers. *IEICE T. Inf. Syst.* E80-D(3):315–325.
  8. Reese, R. B., S. C. Smith, and M. A. Thornton. 2012. Uncle — an RTL approach to asynchronous design. *18th Symposium (International) on Asynchronous Circuits and Systems Proceedings*. Piscataway, NJ: IEEE. 65–72. doi: 10.1109/ASYNC.2012.14.
  9. Poliakov, I., D. Sokolov, and A. Mokhov. 2007. Workcraft: A static data flow structure editing, visualisation and analysis tool. *Petri nets and other models of concurrency*. Eds. J. Kleijn and A. Yakovlev. Lecture notes in computer science ser. Berlin, Heidelberg: Springer. 4546:505–514. doi: 10.1007/978-3-540-73094-1\_30.
  10. Cortadella, J., A. Kondratyev, A. Lavagno, and P. Christos. 2006. Desynchronization: Synthesis of asynchronous circuits from synchronous specifications. *IEEE T. Comput. Aid. D.* 25(10):1904–1921. doi: 10.1109/TCAD.2005.860958.
  11. Andrikos, N., L. Lavagno, D. Pandini, and C. P. Sotiriou. 2007. A fully-automated desynchronization flow for synchronous circuits. *44th ACM/IEEE Design Automation Conference Proceedings*. Piscataway, NJ: IEEE. 982–985.
  12. Zhou, R., K.-S. Chong, B.-H. Gwee, and J. S. Chang. 2011. Quasi-delay-insensitive compiler: Automatic synthesis of asynchronous circuits from Verilog specifications. *54th Midwest Symposium (International) on Circuits and Systems Proceedings*. Piscataway, NJ: IEEE. Art. 6026577. 4 p. doi: 10.1109/MWSCAS.2011.6026577.
  13. Zhou, R., K.-S. Chong, B.-H. Gwee, J. S. Chang, and W.-G. Ho. 2014. Synthesis of asynchronous QDI circuits using synchronous coding specifications. *Symposium (International) on Circuits and Systems Proceedings*. Piscataway, NJ: IEEE. 153–156. doi: 10.1109/ISCAS.2014.6865088.
  14. Zatsarinnyy, A. A., Yu. A. Stepchenkov, Yu. G. D'yachenko, N. V. Morozov, and D. Yu. Stepchenkov. 2023. Avtomatizatsiya sinteza samosinkhronnykh skhem [Self-timed circuits design automation]. *Sistemy vysokoy dostupnosti* [Highly Available Systems] 19(3):48–56. doi: 10.18127/j20729472-202303-04. EDN: IAPVXR.
  15. Zakharov, V., Y. Stepchenkov, Y. Diachenko, and Y. Rogdestvenski. 2020. Self-timed circuitry retrospective. *Conference (International) on Engineering Technologies and Computer Science Proceedings*. Piscataway, NJ: IEEE. 58–64. doi: 10.1109/EnT48576.2020.00018.

16. Sokolov, I. A., Yu. A. Stepchenkov, Yu. G. Diachenko, N. V. Morozov, and D. Yu. Diachenko. 2023. Samosinkhronnyy konveyer s peremennym chislom stupeney [Self-timed pipeline with variable stage number]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 33(1):4–13. doi: 10.14357/08696527230101. EDN: FVYPZG.
17. Stepchenkov, Yu. A., Yu. G. Diachenko, D. Yu. Stepchenkov, D. Yu. Diachenko, and G. A. Orlov. 2023. Mul'tipleksiruemyy samosinkhronnyy konveyer [Multiplexed self-timed pipeline]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 33(2):4–12. doi: 10.14357/08696527230201.

*Received December 29, 2023*

## Contributors

**Stepchenkov Yuri A.** (b. 1951) — Candidate of Science (PhD) in technology, head of department, leading scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; YStepchenkov@ipiran.ru

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

**Diachenko Yuri G.** (b. 1958) — Candidate of Science (PhD) in technology, senior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; diaura@mail.ru

**Morozov Nikolai V.** (b. 1956) — senior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; NMorozov@ipiran.ru

**Stepchenkov Dmitri Y.** (b. 1973) — senior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; Stepchenkov@mail.ru

**Orlov Georgii A.** (b. 1994) — junior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; orlov.jaja@gmail.com

## РАЗРАБОТКА ИНФОРМАЦИОННОЙ СИСТЕМЫ ДЛЯ ХРАНЕНИЯ И ОБРАБОТКИ РЕЗУЛЬТАТОВ МОНИТОРИНГА СЕЛЬСКОХОЗЯЙСТВЕННЫХ УГОДИЙ ЮГА ДАЛЬНЕГО ВОСТОКА\*

*Д. В. Федосеев<sup>1</sup>, А. С. Степанов<sup>2</sup>, Л. В. Илларионова<sup>3</sup>, В. А. Елисеев<sup>4</sup>,  
И. О. Прохорец<sup>5</sup>, А. Н. Фролов<sup>6</sup>, А. И. Шамурина<sup>7</sup>*

**Аннотация:** Цифровизация сельского хозяйства стала одним из стратегических направлений развития экономики. Для хранения и обработки результатов спутникового мониторинга сельскохозяйственных земель предполагается развитие информационных систем (ИС) на федеральном и региональном уровне. Рассматриваются вопросы разработки соответствующей ИС для южной части Дальнего Востока. Описана двухуровневая система, включающая клиентский и серверный уровень. Реализованная система управления базами данных (СУБД) на основе PostgreSQL с расширением PostGIS имеет кластерную структуру, где в качестве отдельной базы данных (БД) кластера рассматривается субъект Российской Федерации. Элемент кластера логически разделен на 5 основных модулей, представляющих собой сгруппированные таблицы с данными разных типов. Для оптимизации работы ИС используются триггеры и триггерные функции, также предусмотрена возможность шардирования данных на уровне БД. Создание ИС для южной части Дальнего Востока обеспечит хранение и оперативный доступ к информации о данных дистанционного зондирования земель сельскохозяйственного назначения (ЗСН) макрорегиона.

**Ключевые слова:** информационная система; база данных; спутниковый мониторинг; сельское хозяйство

**DOI:** 10.14357/08696527240104

**EDN:** XPPVVV

---

\*Статья публикуется по представлению программного комитета VII Международной научно-практической конференции «Информационные технологии и высокопроизводительные вычисления» (ИТНПС-2023). Исследования выполнены при поддержке Российского научного фонда (проект 23-76-00007).

<sup>1</sup>Вычислительный центр Дальневосточного отделения Российской академии наук, d.fedoseev@mail.ru

<sup>2</sup>Дальневосточный научно-исследовательский институт сельского хозяйства, stepanfx@mail.ru

<sup>3</sup>Вычислительный центр Дальневосточного отделения Российской академии наук, illarionova\_l@list.ru

<sup>4</sup>Национальный исследовательский университет «Высшая школа экономики», vladeliseev05@mail.ru

<sup>5</sup>Вычислительный центр Дальневосточного отделения Российской академии наук, ivan0999@inbox.ru

<sup>6</sup>Дальневосточный научно-исследовательский институт сельского хозяйства, andriei-frolov-85@mail.ru

<sup>7</sup>Вычислительный центр Дальневосточного отделения Российской академии наук, sh-a-ig@mail.ru

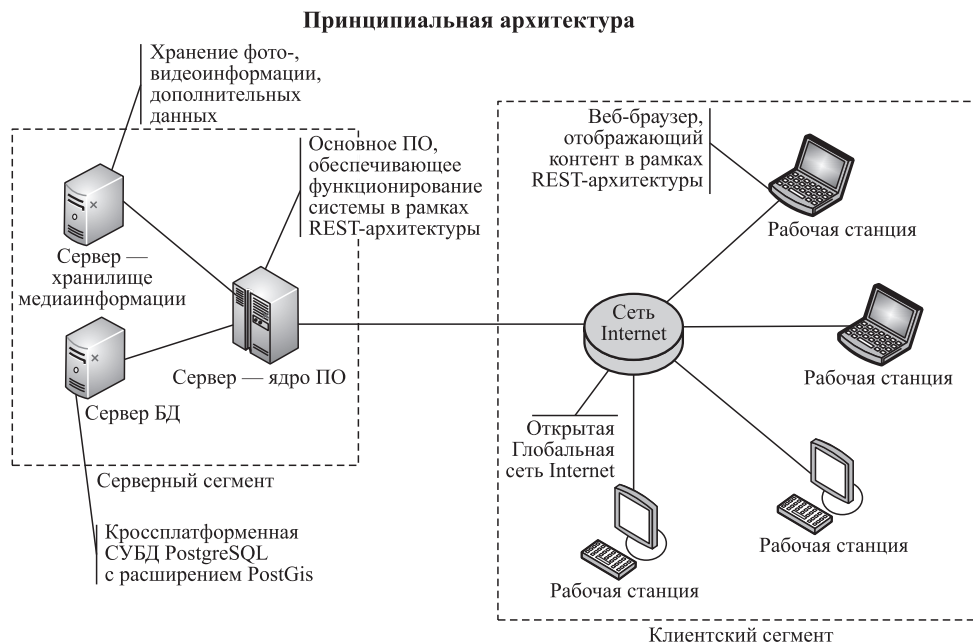
## 1 Введение

В соответствии с указом Президента РФ преобразование сельского хозяйства посредством внедрения цифровых технологий и платформенных решений относится к приоритетному стратегическому направлению развития отраслей экономики [1, 2]. Агропромышленный комплекс регионов в настоящее время должен развиваться с использованием методов машинного обучения и искусственного интеллекта, что предполагает построение ИС [3]. Для реализации Постановления Правительства РФ по эффективному вовлечению в оборот пахотных земель [4] необходимо также использовать современные цифровые технологии, в том числе ИС учета и мониторинга ЗСН [5, 6]. Для осуществления государственного мониторинга ЗСН Министерством сельского хозяйства ведется работа по развертыванию Единой ИС ЗХН с использованием спутниковых данных [7, 8]. При этом на текущий момент качество и полнота представленной информации о севообороте для некоторых регионов требует дополнительной верификации и внесения отсутствующих данных, в том числе для дальневосточных регионов. На современном этапе на рынке существуют много ИС для спутникового мониторинга ЗСН, которые различаются по своему функционалу, целевой группе, стране-разработчику и т. д.: Cropio, Ctrl2go, Агросигнал, IntTerraSkyScout, OneSoil, КосмосАгро, Вега-Science и многие другие [9–12]. Вместе с тем рассмотренные системы не обладают необходимым функционалом, обеспечивающим возможности хранения и обработки данных аэрокосмической съемки, атрибутивной информации, настройки пользовательского интерфейса, свободного доступа и прочих параметров для конкретного региона. Например, создание хранилища спутниковых снимков, а также результатов обработки изображений для субъектов южной части Дальневосточного федерального округа, обеспечение доступа к информации может существенно повысить эффективность реализации программы перехода к цифровому земледелию.

Таким образом, основная цель настоящего исследования — разработка ИС для обработки и хранения данных дистанционного зондирования Земли и атрибутивных данных, необходимых для решения задач цифрового земледелия для южных регионов Дальнего Востока.

## 2 Разработка архитектуры информационной системы

Архитектура разрабатываемой ИС схематически представлена на рис. 1 и имеет в основе структуру, состоящую из двух уровней: клиентского и серверного. Последний, в свою очередь, разбит на ряд подуровней. Так, серверный уровень делится на подуровень «СУБД», подуровень «хранение медиаданных» и подуровень, отвечающий за реализацию прикладной логики программного обеспечения (ПО). Клиентский уровень предполагает организацию процесса предоставления данных конечному пользователю и получения информации от него. Взаимодействие с конечным пользователем в этой схеме реализуется на классическом клиент-серверном шаблоне, в основе которого лежит архитектур-



**Рис. 1** Принципиальная схема ИС

ный стиль взаимодействия компонентов распределенного приложения в сети — REST (Representational State Transfer), в типичном случае представляющий собой обычный HTTP-запрос (GET или POST), в котором необходимые данные передаются в качестве параметров самого запроса [13].

Пользователи ИС посредством веб-браузера обращаются к серверу, на котором развернуто ядро ПО системы, отвечающее как за взаимодействие с пользователями, так и за работу с СУБД и внешним хранилищем медиаинформации. Кроме того, ПО обеспечивает выполнение ряда задач по обработке и анализу имеющихся данных. Разграничение доступа к ИС решается встроенным механизмом аутентификации — для каждого типа пользователей создается соответствующая роль в СУБД, которой выдаются необходимые права, так называемая RBAC (role-based access control) модель [14].

Доступ к системе возможен с любой рабочей станции, имеющей подключение к сети Интернет.

### **3 Принципы построения и структура системы управления базами данных**

Для построения модуля, отвечающего за хранение информации, были выбраны следующие инструменты: СУБД PostgreSQL v. 15 с дополнительным пространственным модулем-расширением PostGIS v. 3.3.2, операционная систе-

ма (ОС) Debian v. 11, а также ряд свободно распространяемых программных продуктов — Qgis, DBeaver и DBVisualiser.

В качестве основных факторов, обеспечивших выбор PostgreSQL, можно отметить следующие: свободное распространение и доступность; поддержка стандарта ANSI SQL 2011; поддержка подхода NoSQL (key-value и JSON); наличие расширяемой системы встроенных языков программирования; наличие георасширения PostGIS (возможность работы с географическими объектами, наличие необходимых геофункций и геотипов данных для расчетов); надежность и устойчивость; кроссплатформенность — поддержка Linux, MacOS, а также MS Windows; масштабируемость; полноценное использование современных архитектур многоядерных процессоров [15]. При этом расширение PostGIS обеспечивает возможности хранения пространственных данных, пространственное индексирование, обладает широким спектром пространственных функций, позволяет хранить и обрабатывать растровые данные, осуществлять геокодирование и обратное геокодирование, а также имеет возможности интеграции со сторонними геоинформационными системами, такими как QGIS, GeoServer, MapServer, ArcGIS, Tableau и др. В качестве ОС было предложено использовать Debian 11 Bullseye, отличительные особенности которой заключаются в возможности ее применения как для серверов, так и для рабочих станций, поддержке сообществом разработчиков в течение пяти лет с момента выпуска, наибольшем среди всех дистрибутивов хранилище программных пакетов — готовых к использованию программ и библиотек.

Структурно СУБД представляет собой кластер, где элементами кластера выступают отдельные БД, содержащие информацию о регионах исследования (на текущий момент — субъект РФ), сформированных по географически-территориальному, административному или иному принципам (рис. 2). Элементы кластера одинаковы по своей внутренней структуре, в том числе в части структуры и наполнения схем, таблиц, а также функционирования разработанных триггеров и триггерных функций. Такой подход обеспечивает не только логическое разделение информации на основе территориально-административного принципа, но и предусматривает возможность будущего шардирования БД [16]. Каждый структурный элемент кластера логически разделен на 5 основных модулей, представляющих собой сгруппированные таблицы. Модуль «Атрибутивная информация о полях» содержит данные о структуре субъекта РФ (перечень муниципалитетов), принадлежности ЗСН (список собственников полей), типе пахотных земель (в том числе пастбища и сенокосы), видах сельскохозяйственных культур со спецификацией по сортам. Модуль «Информация о вегетационных индексах», как следует из названия, содержит данные о рассчитанных различными способами индексах вегетации, в том числе NDVI (normalized difference vegetation index), EVI (enhanced vegetation index), GNDVI (green NDVI), CVI (chlorophyll vegetation index), а также других, в случае необходимости. Таким образом, полученная информация, рассчитанная по конкретному заданному индексу, хранится в отдельной обособленной таблице.





Рис. 2 Структура СУБД

На текущий момент в соответствующих полях таблицы хранятся рассчитанные по различным методикам еженедельные композитные значения индексов вегетации. Таким образом, для каждого вегетационного индекса может быть создано несколько таблиц в соответствии с числом использованных методик. Кроме того, геоданные, поступающие со спутников, имеют различное пространственное разрешение (пиксель), что также обуславливает потребность в отдельной таблице в БД. Центральным связующим блоком служит модуль «Геоинформация», содержащий геоданные о форме поля в виде мультиполигона, а также информацию о его площади, значениях агрохимических показателей, фактической и прогнозной урожайности; датах, дозах, кратности внесения пестицидов, гербицидов, удобрений; сроках посева и уборки и пр. Четвертый и пятый модули БД содержат соответственно медиаинформацию, относящуюся к конкретному полю с комментариями, и информацию о способах расчета, использованных моделях, полученных результатах.

#### 4 Структурная схема, встроенная логика и визуализация базы данных

В общем виде отдельная разработанная БД в кластере включает следующие схемы двух типов: Accounting, которая содержит информацию, не изменяемую во времени, а также схемы YEAR1, YEAR2, YEAR3 и т. д., основу которых составляют данные, актуальные для конкретного года (рис. 3).



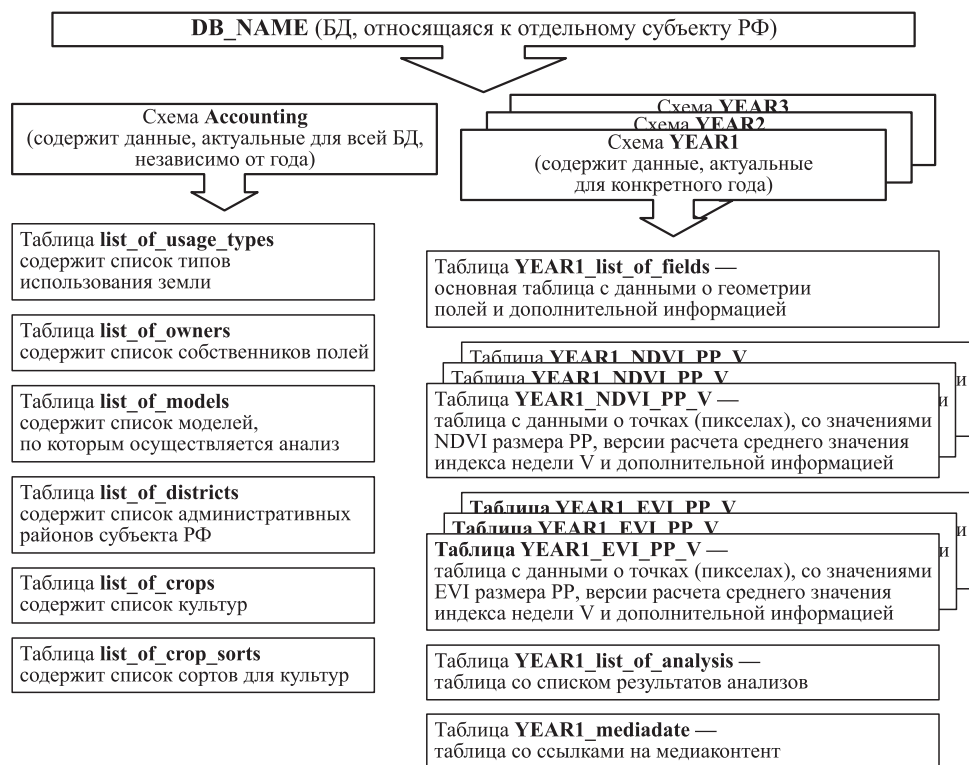


Рис. 3 Структурная схема БД

Данные в схеме Accounting собраны в таблицы: list\_of\_crops, list\_of\_crop\_sorts, list\_of\_districts, list\_of\_models, list\_of\_owners и list\_of\_usage\_types. Для каждого года исследуемого периода сформированы таблицы list\_of\_fields, NDVI\_PP\_V, EVI\_PP\_V, list\_of\_analysis и mediadata. Внутренняя структура таблиц и число полей в них остаются неизменными, при этом общее число таблиц, содержащих данные о вегетационных индексах, может варьироваться год от года в зависимости от потребности пользователей и имеющихся данных.

Работа с БД, схемами, таблицами и их атрибутами и отношениями проводится в свободно распространяемом кроссплатформенном ПО DBever.

Для оптимизации функционирования ИС часть преобразований данных и сопутствующих расчетов может быть перенесена из подуровня, отвечающего за реализацию прикладной логики ПО, внутрь самой СУБД посредством применения триггеров и триггерных функций.

На рис. 4 схематически представлена работа триггеров и триггерных функций. В качестве примера приведены триггер area\_insert для таблицы list\_of\_fields, вызывающий функцию area\_insert(), которая производит расчет

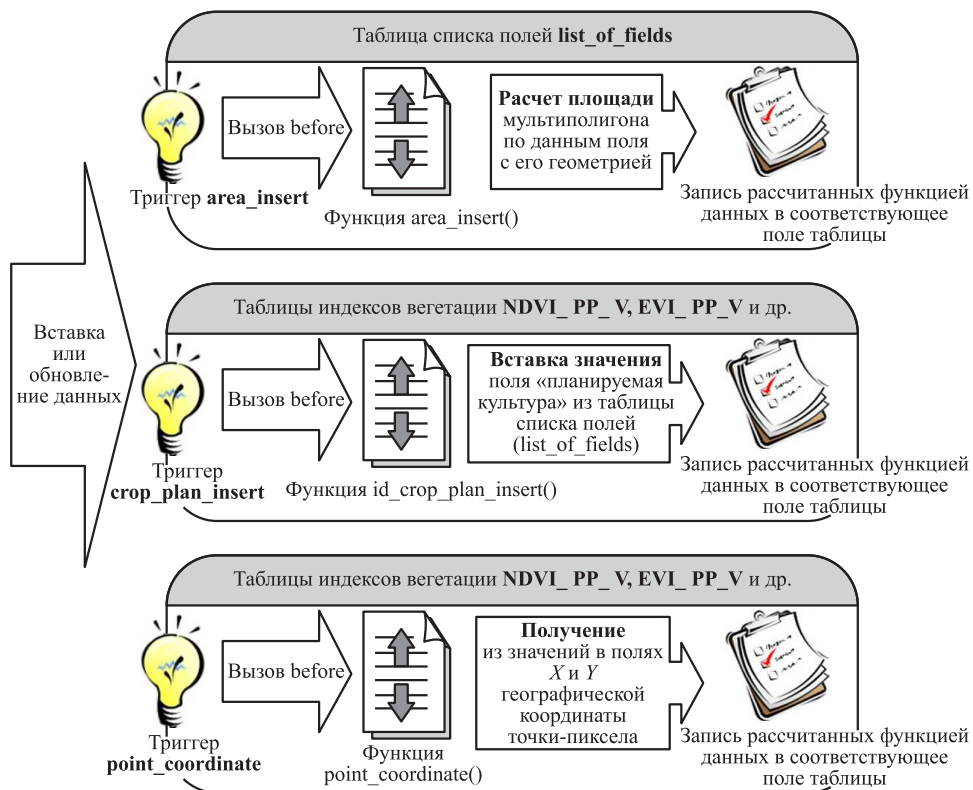


Рис. 4 Встроенная логика БД

площади мультиполигона по его геоданным. Похожим образом работает связка триггер `point_coordinate` для таблиц с вегетационными индексами и функция `point_coordinate()`, которая по имеющимся значениям `X` и `Y` рассчитывает географические координаты пикселя. В результате действия триггера `crop_plan_insert` и триггерной функции `id_crop_plan_insert()` формируется запись в текущую строку в столбец `id_crop_plan` id-значения планируемой к посадке культуры из таблицы полей `list_of_fields`.

Для выполнения условия уникальности каждого поля `geom` в таблице `list_of_fields` применяется связка триггера `hash_geom` и функции `hash_geom()`, результатом расчета которой выступает хеш-значение данных поля `geom`. Далее на уже полученное хеш-значение накладывается условие уникальности. Данный функционал обусловлен тем, что при прямой индексации поля `geom` размера индекса В-Дереа зачастую недостаточно.

Геоинформация, представленная в БД, может быть визуально отображена в свободно распространяемом ПО QGIS, в том числе с возможностью выбора

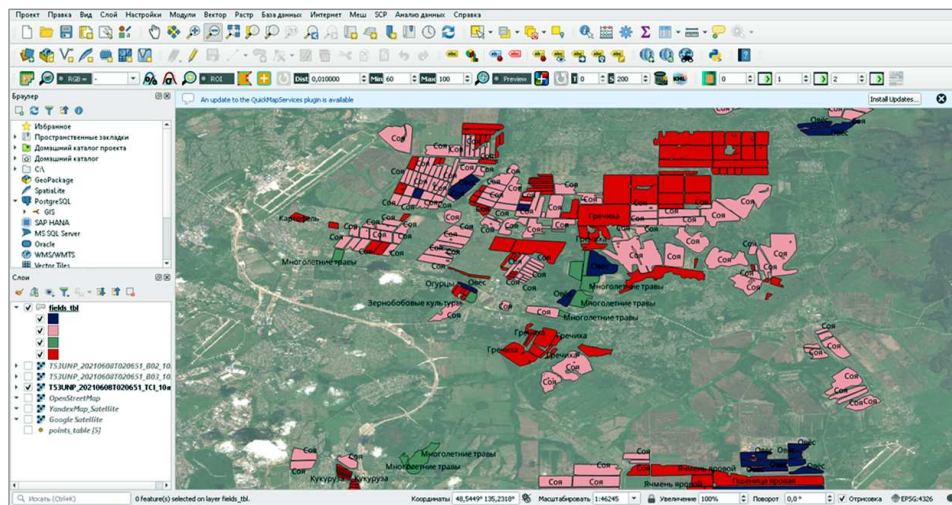


Рис. 5 Отображение фрагмента БД «Хабаровский край» в QGIS

отображения дополнительных слоев, представляющих как растровые изображения (спутниковые снимки), так и пространственные данные (карты) сервисов TMS, WMS, WMTS, ESRI ArcGISService или XYZ-тайлы. На рис. 5 представлен фрагмент геоданных из БД по Хабаровскому краю с графической информацией по севообороту в 2022 г.

## 5 Заключение

Разработанная ИС представляет собой двухуровневую систему для обработки и хранения данных для решения задач цифрового земледелия. Основные факторы, которые учитывались при построении СУБД, — это возможность работы с геоданными, доступность, надежность, устойчивость, кроссплатформенность и пр. Реализованная СУБД имеет кластерную структуру, где каждый структурный элемент (субъект РФ) логически разделен на 5 основных модулей, представляющих собой сгруппированные таблицы. Для оптимизации работы ИС используются триггеры и триггерные функции, а также предусмотрена возможность шардирования БД. На текущий момент проводится тестирование ИС и наполнение БД «Хабаровский край» и «Амурская область».

Структура БД позволит расширять ее путем внесения новой информации (например, дополнительных таблиц с физико-химической и микробиологической характеристикой почв). В настоящий момент в разработке модули фильтрации и визуализации, которые обеспечивают в интерактивном режиме построение графиков временных рядов вегетационных индексов, отобранных по дате, произрастающей культуре, номеру поля и т. п.

Создание региональной БД обеспечит хранение и оперативный доступ к информации о данных дистанционного зондирования сельскохозяйственных угодий региона, что, безусловно, будет способствовать решению задач цифрового земледелия для южных регионов Дальнего Востока. В дальнейшем доступ к БД получат научное сообщество, сельхозпроизводители, а также органы власти регионального и федерального уровня. Информационная система позволит решать вопросы контроля севооборота, оценки эффективности проведения сельскохозяйственных работ, мониторинга состояния посевов на уровне сельскохозяйственных предприятий, а также обеспеченности данными научного сообщества.

## **Литература**

1. *Кирюшин В. И., Иванов А. Л., Козубенко И. С., Савин И. Ю.* Цифровое земледелие // Вестник российской сельскохозяйственной науки, 2018. № 5. С. 4–9. doi: 10.30850/vrsn/2018/5/4-9. EDN: YLEQTR.
2. *Воронин Б. А., Митин А. Н., Пичугин О. А.* Управление процессами цифровизации сельского хозяйства России // Аграрный вестник Урала, 2019. Т. 83. № 4. С. 86–95. doi: 10.32417/article\_5cfa04a236d520.12761241. EDN: OTJXFT.
3. *Сёмкин А. Г., Воронин Е. А.* Концептуальная идентификация сельскохозяйственного макрорегиона как объекта управления пространственным развитием производства методами машинного обучения в информационном пространстве цифровой экономики // Экономика сельского хозяйства России, 2022. № 2. С. 62–68. doi: 10.32651/222-62. EDN: EREOEN.
4. О Государственной программе эффективного вовлечения в оборот земель сельскохозяйственного назначения: Постановление Правительства РФ от 14.05.2021 № 731. <http://static.government.ru/media/files/H0r3EQe7gpGEJvEtfACIXtnJ4gt6Xpr2.pdf>.
5. *Козубенко И. С.* О некоторых аспектах информационного обеспечения управления земельными ресурсами в АПК // Управление рисками в АПК, 2020. № 2(36). С. 18–26. doi: 10.53988/24136573-2020-02-03. EDN: DHVNWf.
6. *Студенкова Н. А., Добротворская Н. И., Аврунёв Е. И., Козина М. В., Пяткин В. П.* Актуальные вопросы инвентаризации и кадастрового учета земель сельскохозяйственного назначения // Вестник СГУГиТ, 2021. Т. 26. № 6. С. 140–149. doi: 10.33764/2411-1759-2021-26-6-140-149.
7. О вводе в эксплуатацию Единой федеральной информационной системы о землях сельскохозяйственного назначения и землях, используемых или предоставленных для ведения сельского хозяйства в составе земель иных категорий: Приказ Министерства сельского хозяйства РФ от 2 апреля 2018 г. № 130.
8. *Буланов К. А., Денисов П. В., Лупян Е. А., Мартыанов А. С., Середа И. И., Трошко К. А., Толпин В. А., Барталев С. А., Хвостиков С. А.* Блок работы с данными дистанционного зондирования Земли Единой федеральной информационной системы о землях сельскохозяйственного назначения // Современные проблемы дистанционного зондирования Земли из космоса, 2019. Т. 16. № 3. С. 171–182. doi: 10.21046/2070-7401-2019-16-3-171-182. EDN: SXGLGW.
9. *Лупян Е. А., Барталев С. А., Толпин В. А., Жарко В. О., Крашенинникова Ю. С., Оксюкевич А. Ю.* Использование спутникового сервиса ВЕГА в региональных системах дистанционного мониторинга // Современные проблемы ди-

- станционного зондирования Земли из космоса, 2014. Т. 11. № 3. С. 215–232. EDN: THXYDN.
10. Глебов И. П., Юрлова А. В., Павлова Е. Н., Ершов Ю. О. Применение программного обеспечения «Агросигнал» как фактора повышения инновационной активности сельскохозяйственных предприятий // Аграрный научный ж., 2018. № 12. С. 84–88. doi: 10.28983/asj.v0i12.666. EDN: YQDIST.
  11. Линьков Ю. В., Моторин О. А., Парфентьев М. В. Подходы к классификации цифровых сервисов для АПК и развитие интегрированных решений с учетом геосервисов // Управление рисками в АПК, 2021. Т. 39. № 1. С. 82–91. doi: 10.53988/24136573-2021-01-07. EDN: CWYVTU.
  12. Shaikh T. A., Rasool T., Lone F. R. Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming // Comput. Electron. Agr., 2022. Vol. 198. Art. 107119. 29 p. doi: 10.1016/j.compag.2022.107119.
  13. Гаврилюк В. И., Гаврилюк В. И. RESTful системы: основные принципы и применение // Молодой ученый, 2020. № 13(303). С. 4–6. EDN: LHFEDV.
  14. Liu D., Dong A., Yan B., Yu J. DF-RBAC: Dynamic and fine-grained role-based access control scheme with smart contract // Procedia Comput. Sci., 2021. Vol. 187. P. 359–364. doi: 10.1016/j.procs.2021.04.074.
  15. Ilba M. Parallel algorithm for improving the performance of spatial queries in SQL: The use cases of SQLite/SpatiaLite and PostgreSQL/PostGIS databases // Comput. Geosci., 2021. Vol. 155. Art. 104840. 10 p. doi: 10.1016/j.cageo.2021.104840.
  16. Mahajan D., Blakeney C., Zong Z. Improving the energy efficiency of relational and NoSQL databases via query optimizations // Sustain. Comput. — Infor., 2019. Vol. 22. P. 120–133. doi: 10.1016/j.suscom.2019.01.017.

Поступила в редакцию 01.11.23

---

---

## DEVELOPMENT OF INFORMATION SYSTEM FOR STORING AND PROCESSING THE RESULTS OF AGRICULTURAL LAND MONITORING IN THE SOUTH OF THE FAR EAST

*D. V. Fedoseev<sup>1</sup>, A. S. Stepanov<sup>2</sup>, L. V. Illarionova<sup>1</sup>, V. A. Eliseev<sup>3</sup>,  
I. O. Prokhorets<sup>1</sup>, A. N. Frolov<sup>2</sup>, and A. I. Shamurina<sup>1</sup>*

<sup>1</sup>Computing Center of the Far Eastern Branch of the Russian Academy of Sciences, 65 Kim Yu Chen Str., Khabarovsk 680000, Russian Federation

<sup>2</sup>Far Eastern Research Institute of Agriculture, 13 Klubnaya Str., Vostochnoe 680521, Khabarovsk Territory, Russian Federation

<sup>3</sup>National Research University Higher School of Economics, 20 Myasnitskaya Str., Moscow 101000, Russian Federation

**Abstract:** Digitalization of agriculture is one of the strategic directions of economic development. To store and process the results of satellite monitoring of agricultural lands, it is planned to develop information systems (IS) at the federal

and regional levels. The article discusses the issues of developing an appropriate IS for the southern part of the Far East. A two-level system including client and server levels is described. The implemented database (DB) management system based on PostgreSQL with the PostGIS extension has a cluster structure where the subject of the Russian Federation is considered as a cluster. The cluster is logically divided into 5 main modules which are grouped tables with data of different types. Triggers and trigger functions are used to optimize the operation of the IS and the possibility of sharing DB is also provided. The DB is being tested and filled for the Khabarovsk and the Amur Regions.

**Keywords:** information systems; database; satellite monitoring; agriculture

**DOI:** 10.14357/08696527240104

**EDN:** XPPVVV

## Acknowledgments

The paper is published on the proposal of the Program Committee of the 7th International Scientific and Practical Conference “Information Technologies and High Performance Computing” (ITHPC-2023). The research was supported by the Russian Science Foundation (project No. 23-76-00007).

## References

1. Kiryushin, V. I., A. L. Ivanov, I. S. Kozubenko, and I. Yu. Savin. 2018. Tsifrovoye zemledelie [Digital farming]. *Vestnik rossiyskoy sel'skokhozyaystvennoy nauki* [Vestnik of the Russian Agricultural Science] 5:4–9. doi: 10.30850/vrsn/2018/5/4-9. EDN: YLEQTR.
2. Voronin, B. A., A. N. Mitin, and O. A. Pichugin. 2019. Upravlenie protsessami tsifrovizatsii sel'skogo khozyaystva Rossii [Management of digitalization processes in agriculture of Russia]. *Agrarnyy vestnik Urala* [Agrarian Bulletin of the Urals] 183(4):86–95. doi: 10.32417/article\_5cfa04a236d520.12761241. EDN: OTJXFT.
3. Semkin, A. G., and E. A. Voronin. 2022. Kontseptual'naya identifikatsiya sel'skokhozyaystvennogo makroregiona kak ob'ekta upravleniya prostranstvennym razvitiem proizvodstva metodami mashinnogo obucheniya v informatsionnom prostranstve tsifrovoy ekonomiki [Conceptual identification of the agricultural macroregion as an object of control of spatial development of production by machine learning methods in the information space of digital economy]. *Ekonomika sel'skogo khozyaystva Rossii* [Economics of Agriculture of Russia] 2:62–68. doi: 10.32651/222-62. EDN: EREOEH.
4. O Gosudarstvennoy programme effektivnogo vovlecheniya v oborot zemel' sel'skokhozyaystvennogo naznacheniya i razvitiya meliorativnogo kompleksa Rossiyskoy Federatsii: Postanovlenie Pravitel'stva Rossiyskoy Federatsii ot 14.05.2021 No. 731 [On the state program for the effective involvement of agricultural lands in circulation and the development of the reclamation complex of the Russian Federation: Decree of the Government of the Russian Federation No. 731 dated 14.05.2021]. Available at: <http://static.government.ru/media/files/H0r3EQe7gpGEJvEtfACIXtnJ4gt6Xpr2.pdf> (accessed February 27, 2024).



5. Kozubenko, I. S. 2020. O nekotorykh aspektakh informatsionnogo obespecheniya upravleniya zemel'nymi resursami v APK [On some aspects of information supply of agricultural land resources management]. *Upravlenie riskami v APK* [Agricultural Risk Management] 2(36):18–26. doi: 10.53988/24136573-2020-02-03. EDN: DHVNWF.
6. Studenkova, N. A., N. I. Dobrotvorskaya, E. I. Avrunev, M. V. Kozina, and V. P. Pyatkin. 2021. Aktual'nye voprosy inventarizatsii i kadaastrovogo ucheta zemel' sel'skokhozyaystvennogo naznacheniya [Current issues of inventory and cadastral registration of agricultural land]. *Vestnik SGUGiT* [Vestnik of the Siberian State University of Geosystems and Technologies] 26(6):140–149. doi: 10.33764/2411-1759-2021-26-6-140-149.
7. O vvode v ekspluatatsiyu Edinoy federal'noy informatsionnoy sistemy o zemlyakh sel'skokhozyaystvennogo naznacheniya i zemlyakh, ispol'zuemykh ili predostavlennykh dlya vedeniya sel'skogo khozyaystva v sostave zemel' inykh kategoriy: Prikaz Ministerstva sel'skogo khozyaystva RF ot 2 aprelya 2018 g. No. 130 [On the commissioning of the Unified Federal Information System on agricultural lands and lands used or provided for agriculture as part of lands of other categories: Order of the Ministry of Agriculture of the Russian Federation No. 130 dated 02.04.2018].
8. Bulanov, K. A., P. V. Denisov, E. A. Loupian, A. S. Martyanov, I. I. Sereda, K. A. Troshko, V. A. Tolpin, S. A. Bartalev, and S. A. Khvostikov. 2019. Blok raboty s dannymi distantsionnogo zondirovaniya Zemli Edinoy federal'noy informatsionnoy sistemy o zemlyakh sel'skokhozyaystvennogo naznacheniya [Remote sensing unit of Integrated federal information system on agricultural lands]. *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa* [Current Problems in Remote Sensing of the Earth from Space] 16(3):171–182. doi: 10.21046/2070-7401-2019-16-3-171-182. EDN: SXGLGW.
9. Loupian, E. A., S. A. Bartalev, V. A. Tolpin, V. O. Zharko, Yu. S. Krashennikova, and A. Yu. Oksyukevich. 2014. Ispol'zovanie sputnikovogo servisa VEGA v regional'nykh sistemakh distantsionnogo monitoringa [VEGA satellite service applications in regional remote monitoring systems]. *Sovremennye problemy distantsionnogo zondirovaniya Zemli iz kosmosa* [Current Problems in Remote Sensing of the Earth from Space] 11(3):215–232. EDN: THXYDN.
10. Glebov, I. P., A. V. Yurlova, E. N. Pavlova, and Yu. O. Ershov. 2018. Primenenie programmnoogo obespecheniya “Agrosignal” kak faktora povysheniya innovatsionnoy aktivnosti sel'skokhozyaystvennykh predpriyatiy [The use of software “Agrosignal” as a factor in increasing the innovative activity of agricultural enterprises]. *Agrarnyy nauchnyy zh.* [Agrarian Scientific J.] 12:84–88. doi: 10.28983/asj.v0i12.666. EDN: YQDIST.
11. Linkov, Yu. V., O. A. Motorin, and M. V. Parfent'ev. 2021. Podkhody k klassifikatsii tsifrovyykh servisov dlya APK i razvitie integrirrovannykh resheniy s uchetom geoservisov [Approaches to the classification of digital services for APC and the development of integrated solutions taking into account geoservices]. *Upravlenie riskami v APK* [Agricultural Risk Management] 39(1):82–91. doi: 10.53988/24136573-2021-01-07. EDN: CWYVTU.
12. Shaikh, T. A., T. Rasool, and F. R. Lone. 2022. Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming. *Comput. Electron. Agr.* 198:107119. 29 p. doi: 10.1016/j.compag.2022.107119.

13. Gavriilyuk, V. I., and V. I. Gavriilyuk. 2020. RESTful sistemy: osnovnye printsipy i primeneniye [RESTful systems: Basic principles and application]. *Molodoy uchenyy* [Young Scientist] 13(303):4–6. EDN: LHFEDV.
14. Liu, D., A. Dong, B. Yan, and J. Yu. 2021. DF-RBAC: Dynamic and fine-grained role-based access control scheme with smart contract. *Procedia Comput. Sci.* 187:359–364. doi: 10.1016/j.procs.2021.04.074.
15. Ilba, M. 2021. Parallel algorithm for improving the performance of spatial queries in SQL: The use cases of SQLite/Spatialite and PostgreSQL/PostGIS databases. *Comput. Geosci.* 155:104840. 10 p. doi: 10.1016/j.cageo.2021.104840.
16. Mahajan, D., C. Blakeney, and Z. Zong. 2019. Improving the energy efficiency of relational and NoSQL databases via query optimizations. *Sustain. Comput. — Infor.* 22:120–133. doi: 10.1016/j.suscom.2019.01.017.

*Received November 1, 2023*

## **Contributors**

**Fedoseev Denis V.** (b. 1994) — PhD student, Computing Center of the Far Eastern Branch of the Russian Academy of Sciences, 65 Kim Yu Chen Str., Khabarovsk 680000, Russian Federation; d.fedoseev@mail.ru

**Stepanov Alexey S.** (b. 1974) — Doctor of Science in pharmacy, leading scientist, Far Eastern Research Institute of Agriculture, 13 Klubnaya Str., Vostochnoe 680521, Khabarovsk Territory, Russian Federation; stepanfx@mail.ru

**Illarionova Lyubov V.** (b. 1977) — Candidate of Science (PhD) in physics and mathematics, senior scientist, Computing Center of the Far Eastern Branch of the Russian Academy of Sciences, 65 Kim Yu Chen Str., Khabarovsk 680000, Russian Federation; illarionova\_l@list.ru

**Eliseev Vladislav A.** (b. 2005) — student, National Research University Higher School of Economics, 20 Myasnitskaya Str., Moscow 101000, Russian Federation; vladeliseev05@mail.ru

**Prokhorets Ivan O.** (b. 1999) — PhD student, engineer, Computing Center of the Far Eastern Branch of the Russian Academy of Sciences, 65 Kim Yu Chen Str., Khabarovsk 680000, Russian Federation; ivan0999@inbox.ru

**Frolov Andrey N.** (b. 1985) — junior scientist, Far Eastern Research Institute of Agriculture, 13 Klubnaya Str., Vostochnoe 680521, Khabarovsk Territory, Russian Federation; andriei-frolov-85@mail.ru

**Shamurina Anna I.** (b. 1988) — Candidate of Science (PhD) in technology, senior scientist, Computing Center of the Far Eastern Branch of the Russian Academy of Sciences, 65 Kim Yu Chen Str., Khabarovsk 680000, Russian Federation; sh-a-ig@mail.ru



# КАРТОГРАФИРОВАНИЕ ЗЕМЕЛЬ СЕЛЬСКОХОЗЯЙСТВЕННОГО НАЗНАЧЕНИЯ ХАБАРОВСКОГО КРАЯ МЕТОДАМИ МАШИННОГО ОБУЧЕНИЯ С ИСПОЛЬЗОВАНИЕМ ИЗОБРАЖЕНИЙ SENTINEL-2\*

*И. О. Прохорец<sup>1</sup>, А. С. Степанов<sup>2</sup>*

**Аннотация:** Проведение автоматизированной классификации сельскохозяйственных земель с использованием методов машинного обучения (machine learning, ML) стало одной из важнейших задач при переходе к цифровому сельскому хозяйству. Классификация пахотных земель Хабаровского района проводилась методами случайного леса (random forest, RF), минимального расстояния (minimum distance, MD), кластеризацией K-средних (K-means) по снимкам спутников Sentinel-2 за июль, август, сентябрь и октябрь 2022 г. В качестве входных данных рассматривались значения спектральных каналов, EVI (Enhanced Vegetation Index) и NDVI (Normalized Difference Vegetation Index). По результатам статистической обработки было установлено, что при изменении даты съемки и типа входных данных наибольшую стабильность продемонстрировал метод RF. Точность распознавания пахотных земель Хабаровского района в 2022 г. составила 92,5% при использовании в классификаторе значений NDVI, рассчитанных по сентябрьскому снимку Sentinel-2. Предложенный подход может применяться для автоматизированного классифицирования и последующего картографирования с экспертной корректировкой пахотных земель южной части Дальнего Востока.

**Ключевые слова:** картографирование; машинное обучение; спутниковый мониторинг; пахотные земли; классификация; Хабаровский край

**DOI:** 10.14357/08696527240105

**EDN:** YEZNXD

## 1 Введение

В последнее время в Российской Федерации осуществляется цифровая трансформация сельского хозяйства, заключающаяся во внедрении цифровых технологий для оптимизации производительности, рентабельности и устойчивости сельскохозяйственных предприятий [1, 2]. В аграрной области внедрение платформы цифрового сельского хозяйства базируется на принципах цифрового

---

\*Статья публикуется по представлению программного комитета VII Международной научно-практической конференции «Информационные технологии и высокопроизводительные вычисления» (ИТНПС-2023).

<sup>1</sup>Вычислительный центр Дальневосточного отделения Российской академии наук, ivan0999@inbox.ru

<sup>2</sup>Дальневосточный научно-исследовательский институт сельского хозяйства, stepanfx@mail.ru

земледелия [3]. Картографирование земель сельскохозяйственного назначения, поддержание баз данных в актуальном состоянии, контроль севооборота и перманентное уточнение границ отдельных полей стали важными задачами земледелия, при этом в качестве эффективных подходов к их решению используются данные дистанционного зондирования Земли (ДЗЗ) [4, 5]. Необходимо отметить, что реальные контуры пахотных земель зачастую не совпадают с границами, которые представлены в разных федеральных и коммерческих базах данных, в том числе в Единой федеральной информационной системе о землях сельскохозяйственного назначения (ЕФИС ЗСН) [6]. Это связано, во-первых, с ошибками передачи информации, во-вторых, с низкой заинтересованностью сельскохозяйственных производителей в достоверности, в-третьих, с нерегулярностью корректировки границ. Вместе с тем непрерывно происходящее воздействие природных и техногенных факторов, разбалансированность структуры посевных площадей, нарушение севооборотов, рост доли залежных земель обуславливают необходимость постоянного уточнения границ полей, что особенно актуально для южной части российского Дальнего Востока [7]. В целом для картографирования пахотных земель используются данные ДЗЗ: спутниковой и аэрофотосъемки. Применение беспилотных летательных аппаратов гарантирует высокую точность определения пашни и границ отдельных полей, но при этом относится к достаточно трудозатратным методам, а использование данных ДЗЗ из космоса обычно требует экспертной корректировки для повышения точности картографирования.

Картографирование пахотных земель РФ проводилось с использованием алгоритма локально-адаптивной классификации (Locally Adaptive Global Mapping Algorithm, LAGMA), разработанного в Институте космических исследований Российской академии наук: особое внимание уделялось методам и алгоритмам предварительной обработки спутниковых данных с целью максимального улучшения входных данных для классификации [8]. Общая точность распознавания пахотных земель в различных регионах России с использованием LAGMA варьировалась в диапазоне от 80% до 98%, при этом снижалась для дальневосточных регионов. Работы по картографированию пахотных земель на национальном уровне выполнялись на основе временных рядов, построенных для 7-дневных композитов NDVI, полученных с помощью спектрорадиометра MODIS (Moderate Resolution Imaging Spectroradiometer) [9]. Метод продемонстрировал достаточно хорошие результаты для полей с площадью более 20 га: общая точность метода составила от 85% до 95% для разных территорий. В ряде работ проводился сравнительный анализ целесообразности использования временных рядов индексов NDVI и EVI при классификации наземных покровов и определения типа растительности [10, 11]. В работе по оценке растительных покровов в Бангладеш было установлено, что для разных методов машинного обучения ошибка при применении NDVI составила от 1,5% до 5%, а при использовании EVI — от 4% до 7% [11]. Китайские ученые показали, что временные ряды NDVI и EVI, рассчитанные по изображениям со спутников Sentinel и Landsat, могут быть применены для классифицирования пахотных земель и распознавания культур

в провинции Хэбэй с высокой степенью точности [10]. При этом NDVI превзошел EVI в классификации культур на ранних стадиях созревания. Исследователями из Индии описан подход к картографированию земель Западного Годавари: были выбраны 4 класса (водные объекты, инфраструктура, залежи и пашня) в качестве параметров для ML методом опорных векторов и RF, рассматривались значения четырех каналов временных серий снимков со спутников Sentinel и Landsat [12]. Было установлено, что общая точность метода с использованием Landsat составила около 90%, с использованием Sentinel — 94%.

Вместе с тем использование временных рядов значений каналов или рассчитанных индексов вегетации для ML требует большего количества вычислительных часов и, соответственно, способствует росту вычислительных затрат в сравнении с картографированием, базирующемся на использовании одного спутникового изображения. В исследовании, посвященном картографированию лесов Бразилии, была проведена оценка точности классификации на основе временных рядов ежемесячных значений NDVI в сравнении с классификацией по одному снимку [13]. Как было установлено, точность метода с использованием временной серии была существенно выше, чем при применении метода классификации по одиночному изображению (соответственно 89% и 81,4%).

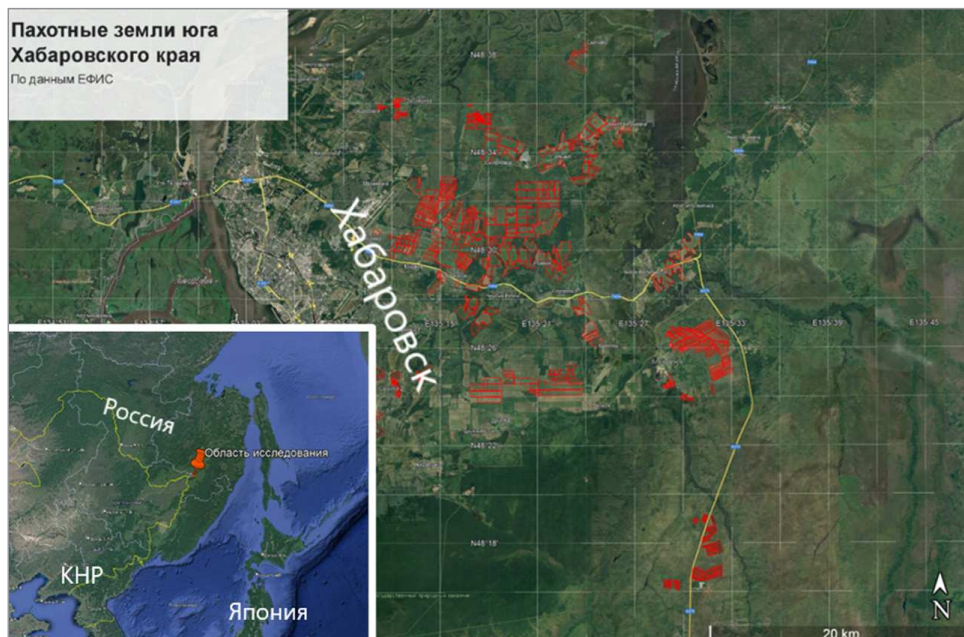
В связи с вышеизложенным для российского Дальнего Востока представляет интерес выполнение исследования для оценки возможности идентификации пахотных земель на основе спутникового изображения, поиск метода и параметров ML, а также даты съемки. Цель работы — разработка подходов к автоматизированному определению границ пахотных земель с использованием ДЗЗ на примере Хабаровского района. Для достижения цели были поставлены следующие задачи:

- получение и обработка мультиспектральных спутниковых изображений для Хабаровского района, расчет значений каналов и индексов вегетации;
- классификация изображений с использованием разных методов и типов входных данных ML;
- оценка точности разных методов.

## 2 Материалы и методы

В качестве области исследования рассматривался Хабаровский муниципальный район, представляющий собой структурную единицу Хабаровского края (рис. 1).

Хабаровский район состоит из двух частей: северо-западной (общей площадью более 24 000 км<sup>2</sup>), которая преимущественно относится к горной местности, и юго-восточной (общей площадью около 6000 км<sup>2</sup>), где и осуществляется хозяйственная деятельность. Земли Хабаровского района располагаются в юго-восточной части, относящейся к Среднеамурской низменности, в пойме реки Амур. Основные типы земель — это пахотные земли, луга, болота, мари, гари, смешанные леса, водные объекты и кустарниковые ассоциации [14].



**Рис. 1** Область исследования (пахотные земли Хабаровского района)

Для исследования были получены 4 безоблачных мультиспектральных спутниковых снимка Sentinel-2 уровня обработки 1С с разрешением 10 м за 13.07.23, 24.08.23, 11.09.23 и 08.10.23. Использовались значения каналов В2 (BLUE, 490 нм), В3 (560 нм), В4 (RED, 665 нм), В5 (705 нм), В6 (740 нм), В7 (783 нм), В8 (NIR, 842 нм), В8а (865 нм), В11 (1610 нм), В12 (2190 нм), а также рассчитанные индексы NDVI и EVI. Индексы NDVI и EVI вычисляются по следующим формулам:

$$NDVI = \frac{(B8) - (B4)}{(B8) + (B4)}; \quad (1)$$

$$EVI = 2,5 \frac{(B8) - (B4)}{(B8) + 6(B4) - 7,5(B2) + 1}. \quad (2)$$

Рассматривались 5 классов подстилающей поверхности: пахотные земли с высокой степенью развития зеленой биомассы (класс 1), пахотные земли с низкой степенью развития зеленой биомассы (класс 2), лесостариковая растительность (класс 3), вода (класс 4) и инфраструктура (класс 5).

Для классификации были выбраны следующие алгоритмы: случайного леса (RF), минимального расстояния (MD) и кластеризация методом k-средних (K-means) [14].

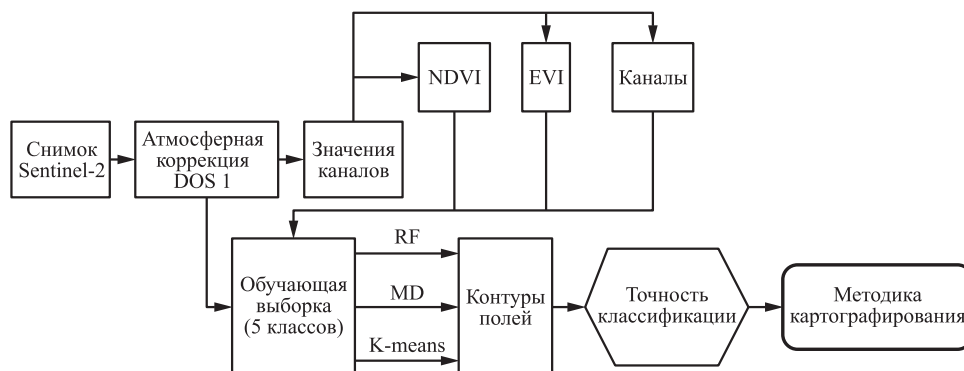


Рис. 2 Схема исследования

Схема исследования представлена на рис. 2.

Алгоритм исследования выглядел следующим образом: на первом этапе происходила обработка спутникового снимка Sentinel-2 уровня 1С с атмосферной коррекцией, затем по формулам (1) и (2) вычислялись индексы NDVI и EVI для каждого пикселя. Для каждого из классов формировалась обучающая выборка: класс 1 — 642 га; класс 2 — 350 га; класс 3 — 12 600 га; класс 4 — 8600 га; класс 5 — 15 га.

Далее тремя методами ML, где в качестве параметров рассматривались либо значения NDVI, либо EVI, либо набор значений всех каналов (за исключением B1, B9 и B10) проводилась классификация всех пикселей снимка.

Для оценки точности классификации рассчитывалась доля верно отнесенных пикселей классов 1 и 2 в класс пахотных земель (шейп-файл сформирован путем экспертной обработки полученных данных из ЕФИС ЗСН для Хабаровского района). Точность классификации  $A$ , а также число пикселей каждого класса, отнесенных к классу пахотных земель  $M_{cl}$ , определялись по формулам:

$$A = \frac{TP_1 + TP_2}{N_{shp}} \cdot 100\%;$$

$$M_{cl} = \frac{N_{cl}}{N_{shp}} \cdot 100\%,$$

где  $TP_1$  и  $TP_2$  — число пикселей классов 1 и 2;  $N_{shp}$  — число всех пикселей, принадлежащих полигонам шейп-файла;  $N_{cl}$  — число пикселей каждого класса. При этом пиксели классов 3–5, отнесенные к пахотным землям, очевидно, идентифицированы неверно, поэтому значение  $M_{cl}$  в данном случае служило ошибкой распознавания —  $E_{cl}$ .

Сравнительная оценка достоверности различий в точности алгоритмов классификации (при использовании разных методов ML, параметров и дат съемки)

**Таблица 1** Доля идентифицированных пикселей каждого класса в классе пахотные земли  $M_{cl}$ , %

Метод ML	Тип данных ML	Класс 1	Класс 2	Класс 3	Класс 4	Класс 5
13.07.22						
RF	Каналы	51,2	29,3	11,0	0,1	8,4
	EVI	52,1	20,0	21,3	0,1	6,5
	NDVI	50,2	24,6	16,9	0,1	8,2
MD	Каналы	54,5	32,2	10,9	0,1	2,3
	EVI	21,3	26,7	13,4	37,3	1,2
	NDVI	27,3	26,7	15,4	30,5	0,0
K-means	Каналы	59,3	29,7	6,4	0,0	4,5
	EVI	19,2	20,0	5,0	55,6	0,0
	NDVI	14,3	24,8	11,1	49,8	0,0
24.08.22						
RF	Каналы	56,8	8,2	17,6	0,1	17,1
	EVI	26,7	22,8	42,8	0,0	7,4
	NDVI	28,2	16,6	38,2	0,0	17,7
MD	Каналы	47,9	22,7	29,0	0,2	0,1
	EVI	41,3	24,6	32,9	0,0	1,0
	NDVI	35,2	22,5	41,7	0,0	0,3
K-means	Каналы	45,6	6,4	28,8	0,1	18,8
	EVI	45,6	18,1	20,1	0,1	16,0
	NDVI	58,9	16,7	24,1	0,0	0,2
11.09.22						
RF	Каналы	84,7	5,6	2,2	0,0	7,3
	EVI	57,9	14,5	25,8	0,0	1,7
	NDVI	65,6	26,8	5,2	0,0	2,1
MD	Каналы	60,5	34,9	4,4	0,1	0,0
	EVI	49,5	26,2	23,3	0,0	0,7
	NDVI	61,9	21,3	16,3	0,0	0,3
K-means	Каналы	44,9	6,3	10,1	0,0	38,5
	EVI	47,2	28,8	16,3	0,0	7,4
	NDVI	59,8	26,1	13,6	0,0	0,3
08.10.22						
RF	Каналы	52,7	4,3	40,1	0,0	2,6
	EVI	22,7	38,3	34,4	0,0	4,3
	NDVI	27,9	41,1	26,0	0,0	4,9
MD	Каналы	20,6	39,1	39,6	0,2	0,3
	EVI	47,9	22,2	27,0	0,0	2,6
	NDVI	39,1	33,0	26,1	0,0	1,6
K-means	Каналы	39,6	19,8	20,5	0,0	19,9
	EVI	57,0	27,8	9,7	0,0	5,3
	NDVI	49,2	37,0	13,6	0,0	0,0



проводилась с использованием дисперсионного анализа. Парное сравнение проводилось по критерию Фишера ( $\alpha = 0,05$ ).

### 3 Результаты

В табл. 1 представлена оценка доли идентифицированных пикселей каждого класса для снимков за июль, август, сентябрь и октябрь 2022 г. В целом достаточно ожидаемо, что  $E_{cl}$  в классах 3 и 5 обычно существенно выше, чем в классе 4, что связано с тем, что объекты инфраструктуры предполагают соседство с деревьями и кустарниками, обладающими вегетационной активностью, а пахотные земли зачастую переходят в залежные, покрытые не только травянистой, но и кустарниковой растительностью.

При классифицировании с использованием июльского снимка MD и K-means демонстрируют низкую эффективность, что связано с высокими значениями  $E_{cl}$  в классе «вода».

Необходимо отметить, что в августе, в сравнении с июлем, наблюдается снижение доли класса 2, так как в южной части Хабаровского края начинается активная вегетация тех культур, которые были посеяны в июле (в частности, гречиха), а также зарастание кормовыми травами убранных полей с ранними зерновыми. Вместе с тем возрастала ошибка определения классов 3 и 5, что также связано с наступлением максимума вегетации. В октябре наблюдалось перераспределение доли класса 1 в пользу класса 2, что обусловлено уборкой и распашкой полей. При этом в процессе использования октябрьских снимков затрудняется поиск границ между полями, что связано со снижением вегетационной активности деревьев и кустарников в лесополосе.

Как следует из табл. 2, при использовании MD и K-means дата съемки не оказывала значимого влияния на результаты классификации. Связано это с тем, что в зависимости от комбинации даты съемки, типа входных данных и метода ML точность определения пахотных земель могла находиться в диапазоне от 40% до 95%. Таким образом, для проведения классификации этими методами необходимо подготовить серию снимков в период с июля по сентябрь, оценить точность на обучающей выборке для разных дат и впоследствии распознать классы, что увеличит трудоемкость процесса картографирования.

Достоверно установлено ( $p < 0,05$ ), что дата снимка оказывает значимое влияние на среднюю точность при применении RF. Наиболее высокая средняя точность 85,1% была достигнута для сентябрьского снимка. Это обусловлено в том числе и тем, что для Хабаровского района характерны заморозки в начале сентября, что вызвало снижение вегетационной активности и способствовало более высокому качеству разделения классов 1 и 3. Точность алгоритма при использовании в качестве входных данных значений NDVI превысила 92%.

На рис. 3 представлен результат работы классификатора RF по значениям NDVI за 11 сентября 2022 г. Были идентифицированы распаханное поле;

**Таблица 2** Точность распознавания пахотных земель Хабаровского района с использованием разных методов MLA, %

Метод ML	Тип данных ML	13.07.2022	24.08.2022	11.09.2022	08.10.2022	<i>p</i>
RF	Каналы	80,5	65,1	90,4	57,1	—
	EVI	72,1	49,6	72,4	61,1	
	NDVI	74,7	44,9	92,5	69,0	
	$X_{ср}$	<b>75,8</b>	<b>53,2</b>	<b>85,1</b>	<b>62,4</b>	
MD	Каналы	86,6	70,6	95,4	59,7	—
	EVI	47,9	66,0	75,8	70,2	
	NDVI	54,0	57,8	83,2	72,1	
	$X_{ср}$	<b>62,8</b>	<b>64,8</b>	<b>84,8</b>	<b>67,4</b>	
K-means	Каналы	89,0	52,1	51,3	59,4	—
	EVI	39,2	63,7	76,1	84,8	
	NDVI	39,0	75,7	86,0	86,3	
	$X_{ср}$	<b>55,7</b>	<b>63,8</b>	<b>71,1</b>	<b>76,8</b>	



**Рис. 3** Результаты классификации подстилающей поверхности Хабаровского района (RF, NDVI, 11.09.2022): (а) окрестности с. Восточное; (б) подтопление поймы р. Амур



часть полей темно-зеленого цвета, очевидно, представляла собой поля, перешедшие в состояние залежи; также достаточно хорошо определялись лесополосы и границы отдельных полей (рис. 3, *a*). На рис. 3, *б* хорошо виден разлив Петропавловского озера, находящегося в пойме Амура, и подтопление близлежащих сельскохозяйственных угодий.

## 4 Заключение

Таким образом, для картографирования подстилающей поверхности предложен следующий подход: на основе снимка спутника Sentinel-2, полученного в первой половине сентября, рассчитываются значения NDVI для каждого пикселя, проводится обучение и классификация изображения методом RF. Точность распознавания пахотных земель в 2022 г. на примере Хабаровского района составила 92,5%. Разработанная методика, базирующаяся на классификации одного мультиспектрального снимка, может быть применена для картографирования сельскохозяйственных угодий Хабаровского края, в частности для оценки нерационального использования земель, выявления залежей и подтоплений, уточнения границ отдельных полей. Использование методики предполагает экспертную ручную корректировку границ отдельных полей после проведения автоматизированной классификации. В дальнейшем планируется применение описанного подхода для картографирования пахотных земель муниципальных образований южной части российского Дальнего Востока.

## Литература

1. Ведомственный проект «Цифровое сельское хозяйство». — М.: Росинформгротех, 2019. 48 с. <https://rosinformagrotech.ru/data/elektronnye-kopii-izdaniy/normativnye-dokumenty-spravochniki-katalogi-i-dr/send/66-normativnye-dokumenty-spravochniki-katalogi/1346-vedomstvennyj-proekt-tsifrovoe-selskoe-khozyajstvo-2019>.
2. Будзко В. И., Меденников В. И. Условия результативного применения технологий искусственного интеллекта в агропромышленном комплексе ЕАЭС // Труды Института системного анализа Российской академии наук, 2023. Т. 73. № 1. С. 148–158. doi: 10.14357/20790279230117. EDN: YDJZZP.
3. Кирюшин В. И., Иванов А. Л., Козубенко И. С., Савин И. Ю. Цифровое земледелие // Вестник российской сельскохозяйственной науки, 2018. № 5. С. 4–9. doi: 10.30850/vrsn/2018/5/4-9. EDN: YLEQTR.
4. Villa-Henriksen A., Edwards G., Pesonen L., Green O., Sørensen C. Internet of Things in arable farming: Implementation, applications, challenges and potential // Biosyst. Eng., 2020. Vol. 191. P. 60–84. doi:10.1016/j.biosystemseng.2019.12.013.
5. Елтошкина Н. В. Геоинформационное картографирование земель сельскохозяйственного назначения // Московский экономический ж., 2022. № 3. С. 31–45. doi: 10.55186/2413046X.2022.7.3.142. EDN: ZXMGSO.

6. Студенкова Н. А., Добротворская Н. И., Аврунев Е. И., Козина М. В., Пяткин В. П. Актуальные вопросы инвентаризации и кадастрового учета земель сельскохозяйственного назначения // Вестник СГУГиТ, 2021. № 6. С. 140–149. doi: 10.33764/2411-1759-2021-26-6-140-149.
7. Трофимов И. А., Трофимова Л. С., Яковлева Е. П., Шпедт А. А., Асеева Т. А. Агрландшафтно-экологическое районирование — основа устойчивого развития сельского хозяйства Восточной Сибири и Дальнего Востока // Биосфера, 2022. № 3. С. 193–199. doi: 10.24855/biosfera.v14i3.695. EDN: BFDPAХ.
8. Барталев С. А., Егоров В. А., Жарко В. О., Луян Е. А., Плотников Д. Е., Хвостиков С. А., Шабанов Н. В. Спутниковое картографирование растительного покрова России. — М.: ИКИ РАН, 2016. 208 с. EDN: YOBJAV.
9. Waldner F., De Abellejra D., Santiago V., Zhang M., Wu B., Plotnikov D., Bartalev S., Lavreniuk M., Skakun S., Kussul N., Le Maire G., Dupuy S., Jarvis I., Defourny P. Towards a set of agrosystem-specific cropland mapping methods to address the global cropland diversity // Int. J. Remote Sens., 2016. Vol. 37. Iss. 14. P. 3196–3231. doi: 10.1080/01431161.2016.1194545.
10. Hao P., Tang H., Chen Z., Meng Q., Kang Y. Early-season crop type mapping using 30-m reference time series // J. Integr. Agr., 2020. Vol. 19. Iss. 7. P. 1897–1911. doi: 10.1016/S2095-3119(19)62812-1.
11. Roy B. Optimum machine learning algorithm selection for forecasting vegetation indices: MODIS NDVI & EVI // Remote Sensing Applications Society Environment, 2021. Vol. 23. Art. 100582. 4 p. doi: 10.1016/j.rsase.2021.100582.
12. Savitha C., Talari R. Mapping cropland extent using Sentinel-2 datasets and machine learning algorithms for an agriculture watershed // Smart Agricultural Technology, 2023. Vol. 4. Art. 100193. 10 p. doi: 10.1016/j.atech.2023.100193.
13. Andrade J., Cunha J., Silva J., Rufino I., Galvã C. Evaluating single and multi-date Landsat classifications of land-cover in a seasonally dry tropical forest // Remote Sensing Applications Society Environment, 2021. Vol. 22. Art. 100515. 13 p. doi: 10.1016/j.rsase.2021.100515.
14. Ганзей С. С., Ермошин В. В., Мишина Н. В. Динамика использования земель в бассейне Амура в XX веке // География и природные ресурсы, 2010. № 1. С. 30–38.
15. Shaikh T. A., Rasool T., Lone F. R. Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming // Comput. Electron. Agr., 2022. Vol. 198. Art. 107119. doi: 10.1016/j.compag.2022.107119.

Поступила в редакцию 31.01.24

## MAPPING OF THE KHABAROVSK REGION ARABLE LANDS BY MACHINE LEARNING USING SENTINEL-2 IMAGES

*I. O. Prokhorets<sup>1</sup> and A. S. Stepanov<sup>2</sup>*

<sup>1</sup>Computing Center of the Far Eastern Branch of the Russian Academy of Sciences, 65 Kim Yu Chen Str., Khabarovsk 680000, Russian Federation

<sup>2</sup>Far Eastern Research Institute of Agriculture, 13 Klubnaya Str., Vostochnoe 680521, Khabarovsk Territory, Russian Federation

**Abstract:** Automated classification of arable lands using machine learning methods is one of the most important tasks in the transition to digital agriculture. The classification of arable lands in the Khabarovsk Region was carried out using random forest (RF), minimum distance (MD), and K-means clustering methods based on Sentinel-2 images for July, August, September, and October 2022. The values of spectral bands, EVI (Enhanced Vegetation Index), and NDVI (Normalized Difference Vegetation Index) were considered as input data. Based on the results of statistical processing, it was found that the RF method demonstrated the greatest stability when changing the date of shooting and the type of input data. The accuracy of recognition of arable lands in the Khabarovsk Region in 2022 was 92.5% when using NDVI values calculated from the September Sentinel-2 image in the classifier. The proposed approach can be used for automated classification and subsequent mapping with expert correction of arable lands in the southern part of the Far East.

**Keywords:** mapping; machine learning; satellite monitoring; arable lands; classification; Khabarovsk Region

**DOI:** 10.14357/08696527240105

**EDN:** YEZNXD

### Acknowledgments

The paper is published on the proposal of the Program Committee of the 7th International Scientific and Practical Conference “Information Technologies and High Performance Computing” (ITHPC-2023).

### References

1. *Vedomstvennyy proekt “Tsifrovoe sel’skoe khozyaystvo”* [Departmental project “Digital Agriculture”]. 2019. Moscow: Rosinformagrotekh. 48 p. Available at: <https://rosinformagrotech.ru/data/elektronnye-kopii-izdanij/normativnye-dokumenty-spravochniki-katalogi-i-dr/send/66-normativnye-dokumenty-spravochniki-katalogi/1346-vedomstvennyj-proekt-tsifrovoe-selskoe-khozyajstvo-2019> (accessed February 29, 2024).

2. Budzko, V. I., and V. I. Medennikov. 2023. Usloviya rezul'tativnogo primeneniya tekhnologiy iskusstvennogo intellekta v agropromyshlennom komplekse EAES [Conditions for the effective application of artificial intelligence technologies in the agro-industrial complex of the EAEU]. *Trudy Instituta sistemnogo analiza Rossiyskoy akademii nauk* [Proceedings of the Institute for Systems Analysis of the Russian Academy of Sciences] 73(1):148–158. doi: 10.14357/20790279230117. EDN: YDJZZP.
3. Kiryushin, V. I., A. L. Ivanov, I. S. Kozubenko, and I. Yu. Savin. 2018. Tsifrovoye zemledelie [Digital farming]. *Vestnik rossiyskoy sel'skokhozyaystvennoy nauki* [Vestnik of the Russian Agricultural Science] 5:4–9. doi: 10.30850/vrsn/2018/5/4-9. EDN: YLEQTR.
4. Villa-Henriksen, A., G. Edwards, L. Pesonen, O. Green, and C. Sørensen. 2020. Internet of Things in arable farming: Implementation, applications, challenges and potential. *Biosyst. Eng.* 191:60–84. doi: 10.1016/j.biosystemseng.2019.12.013.
5. Eltoshkina, N. V. 2022. Geoinformatsionnoe kartografirovaniye zemel' sel'skokhozyaystvennogo naznacheniya [Geoinformation mapping of agricultural land]. *Moskovskiy ekonomicheskiy zh.* [Moscow Economic J.] 7(3):31–45. doi: 10.55186/2413046X.2022.7.3.142. EDN: ZXMGSO.
6. Studenkova, N. A., N. I. Dobrotvorskaya, E. I. Avrunev, M. V. Kozina, and V. P. Pyatkin. 2021. Aktual'nye voprosy inventarizatsii i kadastravnogo ucheta zemel' sel'skokhozyaystvennogo naznacheniya [Current issues of inventory and cadastral registration of agricultural land]. *Vestnik SGUGiT* [Vestnik of SSUGT] 26(6):140–149. doi: 10.33764/2411-1759-2021-26-6-140-149.
7. Trofimov, I. A., L. S. Trofimova, E. P. Yakovleva, A. A. Shpedt, and T. A. Aseeva. 2022. Agrolandshaftno-ekologicheskoye rayonirovaniye — osnova ustoychivogo razvitiya sel'skogo khozyaystva Vostochnoy Sibiri i Dal'nego Vostoka [Agro-landscape and ecological zoning is the basis for sustainable development of agriculture in Eastern Siberia and the Far East]. *Biosfera* [Biosfera] 14(3):193–199. doi: 10.24855/biosfera.v14i3.695. EDN: BFDPAx.
8. Bartalev, S. A., V. A. Egorov, V. O. Zharko, E. A. Loupian, D. E. Plotnikov, S. A. Khvostikov, and N. V. Shabanov. 2016. *Sputnikovoye kartografirovaniye rastitel'nogo pokrova Rossii* [Land cover mapping over Russia using Earth observation data]. Moscow: Russian Academy of Sciences' Space Research Institute. 208 p. EDN: YOBJAV.
9. Waldner, F., D. De Aballeyra, V. Santiago, M. Zhang, B. Wu, D. Plotnikov, S. Bartalev, M. Lavreniuk, S. Skakun, N. Kussul, G. Le Maire, S. Dupuy, I. Jarvis, and P. Defourny. 2016. Towards a set of agrosystem-specific cropland mapping methods to address the global cropland diversity. *Int. J. Remote Sens.* 37(14):3196–3231. doi: 10.1080/01431161.2016.1194545.
10. Hao, P., H. Tang, Z. Chen, Q. Meng, and Y. Kang. 2020. Early-season crop type mapping using 30-m reference time series. *J. Integr. Agr.* 19(7):1897–1911. doi: 10.1016/S2095-3119(19)62812-1.
11. Roy, B. 2021. Optimum machine learning algorithm selection for forecasting vegetation indices: MODIS NDVI & EVI. *Remote Sensing Applications Society Environment* 23:100582. 4 p. doi: 10.1016/j.rsase.2021.100582.
12. Savitha, C., and R. Talari. 2023. Mapping cropland extent using Sentinel-2 datasets and machine learning algorithms for an agriculture watershed. *Smart Agricultural Technology* 4:100193. 10 p. doi: 10.1016/j.atech.2023.100193.

13. Andrade, J., J. Cunha, J. Silva, I. Rufino, and C. Galvã. 2021. Evaluating single and multi-date Landsat classifications of land-cover in a seasonally dry tropical forest. *Remote Sensing Applications Society Environment* 22:100515. 13 p. doi: 10.1016/j.rsase.2021.100515.
14. Ganzei, S. S., V. V. Ermoshin, and N. V. Mishina. 2010. The dynamics of land use within the Amur basin in the 20th century. *Geography Natural Resources* 31(1):18–24. doi: 10.1016/j.gnr.2010.03.004.
15. Shaikh, T. A., T. Rasool, and F. R. Lone. 2022. Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming. *Comput. Electron. Agr.* 198:107119. 29 p. doi: 10.1016/j.compag.2022.107119.

*Received January 31, 2024*

### **Contributors**

**Prokhorets Ivan O.** (b. 1999) — PhD student, engineer, Computing Center of the Far Eastern Branch of the Russian Academy of Sciences, 65 Kim Yu Chen Str., Khabarovsk 680000, Russian Federation; ivan0999@inbox.ru

**Stepanov Alexey S.** (b. 1974) — Doctor of Science in pharmacy, leading scientist, Far Eastern Research Institute of Agriculture, 13 Klubnaya Str., Vostochnoe 680521, Khabarovsk Territory, Russian Federation; stepanfx@mail.ru

# NEURAL NETWORK ARCHITECTURE FOR ARTIFACTS DETECTION IN ZTF SURVEY

*T. A. Semenikhin*<sup>1</sup>

**Abstract:** Today, astronomers are faced with a challenge of handling vast volume of data as modern astronomical surveys are capable of generating terabytes of data in a single night. One of such survey is the Zwicky Transient Facility (ZTF), an automated sky survey that provides approximately a million alerts per observational night. However, a significant part of the detected objects turn out to be artifacts, i. e., phenomena of a nonastrophysical origin. Therefore, specialists must invest time in manually classifying these objects as there is currently no efficient method that can perform this task without human intervention. The goal of the work is the development of an algorithm to predict whether a light curve from the ZTF data releases (DRs) has a bogus nature or not, based on the sequence of frames. A labeled dataset provided by experts from SNAD team was utilized, comprising 2230 frames series. Due to substantial size of the frame sequences, the application of a variational autoencoder (VAE) was deemed necessary for mapping the images into lower-dimensional vectors. For the task of binary classification based on sequences of compressed frame vectors, a recurrent neural network (RNN) was employed. Several neural network models were considered and the quality metrics were assessed using k-fold cross-validation. The final performance metrics, including ROC-AUC =  $0.86 \pm 0.01$  and Accuracy =  $0.80 \pm 0.02$ , suggest that the model has practical utility. The code implementing the algorithm is available on [GitHub](#).

**Keywords:** neural network; data analysis; real-bogus classification

**DOI:** 10.14357/08696527240106

**EDN:** LRMTGD

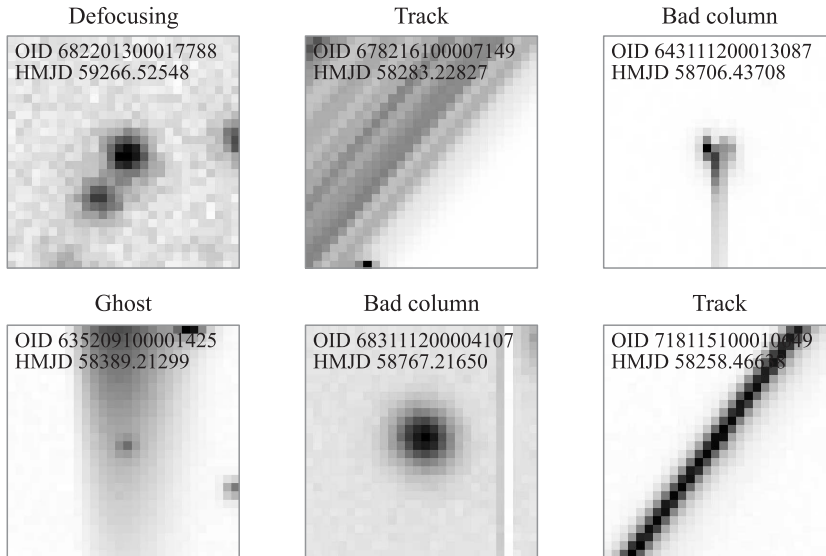
## 1 Introduction

Zwicky Transient Facility [1, 2] is an automated sky survey. The telescope generates a large amount of data every night. The survey data consist of images capturing specific parts of the sky which are later used to create DRs.

The purpose of creating DRs is to investigate any variable sources in the sky. The survey frames are processed using the widely-used algorithm SExtractor [3] which detects sources, determines their coordinates, and performs photometry on them. Subsequently, all variable sources are assigned a unique object identifier (OID) which is added to the so-called ZTF DRs. It is assumed that there should be no nonperiodic variable objects (such as supernovae, red dwarf flares, etc.) in the

---

<sup>1</sup>Sternberg Astronomical Institute, M. V. Lomonosov Moscow State University, 13 Universitetskyy Prospekt, Moscow 119234, Russian Federation



**Figure 1** Examples of artifacts found within ZTF data. Images size is  $28 \times 28$  pixels. Frame intensity inverted for better readability

DRs. However, due to the fact that the applied algorithms are not perfect, unusual objects are encountered in these data [4–6]. This motivates researchers, even those whose scientific interests are not related to variable stars, to work with DRs.

Unfortunately, among the ZTF data, so-called artifacts are often encountered (Fig. 1). An artifact is commonly referred to as any phenomenon of nonastrophysical nature. These can be effects related to instrument malfunction (defocusing, saturated columns in the CCD array, etc.) or effects associated with external conditions (atmospheric turbulence, clouds, bright satellites, etc.). Under certain circumstances, these phenomena can be classified as astrophysical objects by machine learning methods which significantly complicates working with ZTF data.

Classifying astrophysical objects is one of the primary and most challenging tasks when working with DR data because they lack class labels. As a result, instead of directly using a sequence of object images, photometric observations are often employed. Since the observational properties of many astrophysical classes have been sufficiently studied, a set of features can be derived from the light curve (temporal sequence of photometric observations) that allows for their differentiation. Subsequently, classification methods from classical machine learning can be applied to the constructed feature sets to solve the classification task. However, this approach can lead to situations where astrophysical phenomena and artifacts closely resemble each other in their photometric representation. For example, a passing

bright satellite may temporarily illuminate a specific region of the detector which can be mistaken for a brief flare of a red dwarf appearing as a short-lived peak in the light curve. Therefore, when tackling such problems, specialists need to spend time validating the obtained results. Artifacts can account for 68% of the total number of objects labeled as astrophysical objects by machine learning methods [4].

This work is dedicated to implementing an algorithm that would avoid the involvement of specialists in classifying artifact/nonartifact objects from DR. Currently, no attempts have been found in the literature to solve this problem.

## 2 Data

The labeled object dataset has been provided by SNAD team (<https://snad.space/>) and does not publicly available. Each object has a unique index, OID, and a set of labels indicating whether it is an artifact or not, along with additional details about the nature of the object (e. g., AGN — active galactic nucleus, variable, etc. Similarly, for artifacts: ghost, defocusing, etc.). For each object, there is a series of telescope images captured at different times. To obtain a list of links to all available images for a given OID, an API service ([http://finder.fits.ztf.snad.space/api/docs#operation/Get\\_URLs\\_for\\_all\\_exposures\\_by\\_object\\_ID\\_api\\_v1\\_urls\\_by\\_oid\\_get](http://finder.fits.ztf.snad.space/api/docs#operation/Get_URLs_for_all_exposures_by_object_ID_api_v1_urls_by_oid_get)) is used.

A telescope frame is an image of size  $3072 \times 3080$  depicting a portion of the sky. However, for this task, only a small area within this frame associated with a specific object is of interest. Using OID in the ZTF DR, coordinates are determined for the corresponding object (source). In turn, each telescope frame is tied to a coordinate grid. Thus, knowing the OID of the source allows determining its center on the telescope frame. To download a specific frame via a link, IRSA service ([https://irsa.ipac.caltech.edu/docs/program\\_interface/ztf\\_api.html](https://irsa.ipac.caltech.edu/docs/program_interface/ztf_api.html)) is used where, upon request, one can specify the coordinates of the source's center and the dimensions of the rectangular area. According to these specifications, the desired region will be cropped from the original image. This avoids downloading the entire large-sized image allowing the entire dataset to be downloaded in approximately one day.

The size of the cropped area was chosen to be  $28 \times 28$  pixels since the angular size of the sources of interest does not exceed  $\sim 10$  arcsec (1 pixel in ZTF frames corresponds to 1.012 arcsec). Thus, the region is generously sized and the input image size matches that of the popular MNIST [7] dataset. If object was close to the image border, the corresponding missing areas were filled with zeros (this is necessary for the correct operation of the PyTorch library). The frames have their background intensity values; so, each image needs to be normalized. Normalization corresponding to equation

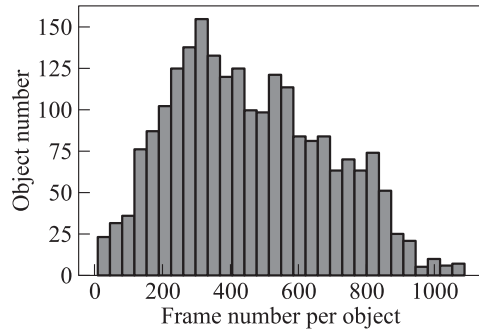
$$x_{\text{norm}} = \frac{x - \min(x)}{\max(x) - \min(x)}$$



where  $x$  is the frame;  $\mu$  is the mean intensity value of the frame; and  $\sigma$  is the standard deviation of the frame intensity was applied. Figure 1 shows examples of pre-processed frames.

The dataset contains 2230 objects including 1150 artifacts and 1080 astrophysical objects, totaling 1,015,177 images.

One of the main challenges of the task is that objects have a large number of observations (Fig. 2) and it is unknown at which specific time points anomalies were observed. Therefore, an algorithm capable of working with sequences of images rather than individual anomalous images is necessary.



**Figure 2** Distribution of images number per object

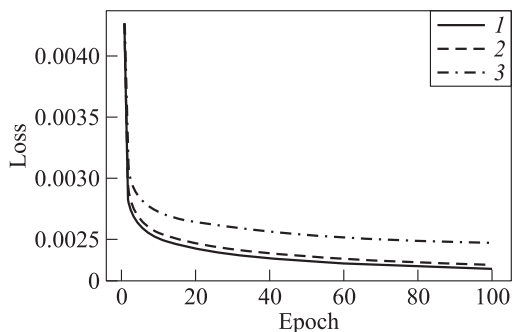
### 3 Methods

The implemented approach involves two stages. First, VAE is trained on all images in the dataset to obtain informative compressed representations of frames. Then, an RNN is trained on sequences of these compressed representations to solve the classification task. Combining these two stages into a single RNN with convolutional layers at the beginning is not feasible due to computational resource limitations. All neural networks are implemented using the Python library PyTorch.

#### 3.1 Variational autoencoder

The implemented encoder consists of 5 convolutional layers (with image channel dimensions: 32, 64, 128, 256, and 512, respectively) with the LeakyReLU activation function and batch normalization after each layer. Each output convolutional layer returns half the size of the input image; so, the encoder output is a vector whose size is a model parameter. The architecture of the decoder is symmetrical to that described above, except that transposed convolutional layers are used instead of convolutional layers.

The VAE was trained on all images from dataset. When selecting the parameters for the VAE, different levels of image compression were considered. In machine learning literature, the vector obtained at the output of the encoder is commonly referred to as the latent state. In Fig. 3, curves 1 and 2 represent the loss curves for latent state sizes of 36 and 78, respectively. It can be observed that increasing the latent state size requires more training epochs. In this work, a latent state size



**Figure 3** Loss functions during training a VAE: 1 — latent size = 36; 2 — latent size = 78; and 3 — random rotate

combine adaptive moment estimation with adaptive learning rates for efficient model optimization in machine learning. The loss function for the VAE takes the following form:

$$\begin{aligned}
 L &= \text{LogCosh}(x, \hat{x}) + \omega_{\text{kld}} \text{KL}(N(\mu, \sigma) || N(0, 1)) \\
 &= \ln \cosh(x - \hat{x}) + \omega_{\text{kld}} \left( \ln \frac{1}{\sigma} + \frac{\sigma^2 + \mu^2}{2} - \frac{1}{2} \right).
 \end{aligned}$$

Here,  $x$  and  $\hat{x}$  represent the original image and its reconstruction by the decoder, respectively;  $\mu$  and  $\sigma$  are vectors of means and variances obtained from the encoder's output; and  $\omega_{\text{kld}} = 8 \cdot 10^{-5}$  serves as the weight coefficient for the Kullback–Leibler divergence.

By using the compressed representations of the images, it is expected that the compressed representations of anomalous frames are outliers from the overall distribution of normal images. After the VAE was trained, the obtained compressed representations of frames were saved. The next stage of the work involved training an RNN for binary classification.

### 3.2 Recurrent neural network

The fundamental idea behind RNNs is the utilization of hidden states that capture information from previous time steps allowing them to maintain a form of memory and learn dependencies within sequential data. By using recurrent connections, RNNs can model sequences of arbitrary length making them powerful tools for tasks involving sequential information. However, traditional RNNs have issues with vanishing and exploding gradients leading to difficulties in learning long-range

of 36 proved to be optimal. Image augmentation through random rotation was also explored (curve 3 in Fig. 3); however, it was found that this significantly increased the number of required training epochs without a visible improvement in the model. Therefore, the final model does not include any augmentations. For model training, the Adam optimizer [8] was used with a learning rate =  $5 \cdot 10^{-5}$  for 100 epochs. The core idea of the Adam optimizer is to combine

dependencies. To address this, various RNN architectures like LSTM (Long Short-Term Memory) [9] and GRU (Gated Recurrent Unit) [10] have been developed which introduce gating mechanisms to better control the flow of information and mitigate gradient-related problems.

The model takes inputs of size (batch size, sequence length, and latent\_dim) where latent\_dim represents the dimensionality of the compressed representation. Embeddings for object frames along with their corresponding class labels are fed into an RNN. The model parameters are optimized using the Adam optimizer with a learning rate =  $10^{-4}$  for 500 epochs.

The baseline model considered was a GRU layer with a hidden state size of 128 and the loss function used during training the RNN was cross-entropy. In addition to the baseline model, the following variations were explored: a bidirectional GRU cell with  $L_2$  weight regularization (weight decay =  $10^{-5}$ ); a GRU layer with a modified loss function (the Tversky Index [11]) term was added to the loss function); and an LSTM layer with the same parameters as the baseline model.

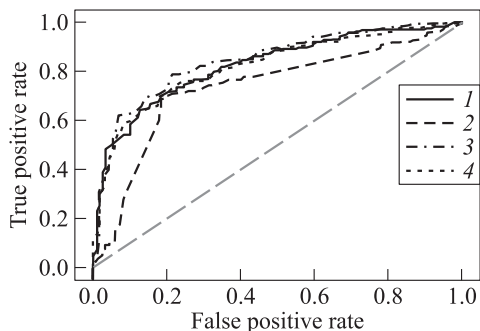
During the RNN training, we employed k-fold cross-validation ( $k = 5$ ). The dataset was randomly divided into 5 folds and then 5 models were trained, each having a specific fold as the test set and the union of the remaining 4 folds as the training set. Thus, the data split into training/test sets was unique for each model. The final quality metrics for each considered model were calculated as the average values across 5 models from the k-fold cross-validation split.

Applying the obtained model to the entire ZTF DR dataset is impractical as downloading all images for all objects in the ZTF DR would take months or even years. In practice, it is expected that the classifier's decision function will serve as an additional source of information for the active (employing partial teacher involvement through feedback loops) anomaly detection algorithm and the classifier will be applied only to a small subset of objects — candidates for anomalies [12]. In this case, determining the specific threshold of the decision function is the prerogative of the active anomaly detection algorithm based on expert feedback. Moreover, technically, such a threshold may vary for different feature space regions. Therefore, ROC-AUC was chosen as the main metric for evaluating model performance.

Bidirectional GRU is different from the regular GRU in that it processes input sequences in both forward and backward directions which enables it to capture contextual information from past and future time steps. However, when using

Model results. The metric values are averaged over 5 test folds as well as the standard deviation

Model name	ROC-AUC	Accuracy	F1-score
GRU	$0.84 \pm 0.01$	$0.78 \pm 0.02$	$0.79 \pm 0.01$
Bidirectional GRU + $L_2$	$0.86 \pm 0.01$	$0.80 \pm 0.02$	$0.80 \pm 0.01$
GRU + Tversky loss	$0.84 \pm 0.02$	$0.79 \pm 0.02$	$0.79 \pm 0.02$
LSTM	$0.76 \pm 0.04$	$0.76 \pm 0.03$	$0.77 \pm 0.02$



**Figure 4** The ROC curves examples of trained models for one of the cross-validation splits: 1 — GRU; 2 — LSTM; 3 — bidirectional GRU +  $L_2$ ; and 4 — GRU + Tversky loss

such an approach in this task, the model experienced overfitting; so,  $L_2$  weight regularization was added to it. The table as well as Fig. 4 show the main results for the considered models. The table indicates that LSTM performs worse than the other models. The three remaining models exhibit approximately the same level of quality based on the performance metrics. However, the bidirectional GRU +  $L_2$  appears slightly more stable across different data splits. Therefore, this model was selected for further work.

## 4 Concluding Remarks

Throughout the course of this study, sequences of ZTF object frames from a labeled dataset were downloaded and preprocessed. Using the `PyTorch` library, a VAE was implemented and trained enabling the compression of images into informative lower-dimensional vectors. Several models of RNNs were considered for the task of binary classification based on sequences of compressed image vectors. The quality metrics of the models were evaluated using  $k$ -fold cross-validation, with the best-performing model achieving an ROC-AUC =  $0.86 \pm 0.01$  and Accuracy =  $0.80 \pm 0.02$ . All the code used in the study is available on `GitHub` ([https://github.com/semtim/RB\\_ZTF](https://github.com/semtim/RB_ZTF)).

Thus, an algorithm has been developed that aims to classify objects from ZTF DR in a manner similar to that of a specialist who annotated the training dataset. This algorithm stands out from others by utilizing sequences of observation frames rather than photometric time series. While computationally more challenging, images contain more information about the object's radiation than its corresponding photometry.

The next stage of the work will involve: searching for the optimal architecture for both the autoencoder and the recurrent network; selecting the most effective loss function for the classifier in terms of our task; analyzing the dependence of the first and second type errors on the chosen threshold value; and conducting computational experiments using different optimizers. This algorithm can be employed within the scope of the SNAD project as the obtained quality metrics indicate its practical value. Since the training dataset used in this work was annotated by specialists from the SNAD project, it is expected that the implementation of this algorithm into the general pipeline will reduce the number of artifacts among anomalies. This will enable specialists to detect more new astrophysical objects.

## Acknowledgments

The paper is published on the proposal of the Program Committee of the 25th International Conference on Data Analytics and Management in Data Intensive Domains (DAMDID/RCDL 2023). The author thanks the SNAD team for the provided comments and support. The author acknowledges the support of M. V. Lomonosov Moscow State University Program of Development. The work was supported by Nonprofit Foundation for the Development of Science and Education “Intellect.”

## References

1. Bellm, E. C., S. R. Kulkarni, M. J. Graham, *et al.* 2019. The Zwicky Transient Facility: System overview, performance, and first results. *Publ. Astron. Soc. Pac.* 131(995):018002. 19 p. doi: 10.1088/1538-3873/aaecbe.
2. Bellm, E. C., S. R. Kulkarni, T. Barlow, *et al.* 2019. The Zwicky Transient Facility: Surveys and scheduler. *Publ. Astron. Soc. Pac.* 131(1000):068003. 13 p. doi: 10.1088/1538-3873/ab0c2a.
3. Bertin, E., and S. Arnouts. 1996. SExtractor: Software for source extraction. *Astron. Astrophys. Sup.* 117(2):393–404. doi: 10.1051/aas:1996164.
4. Malanchev, K. L., M. V. Pruzhinskaya, V. S. Korolev, *et al.* 2021. Anomaly detection in the Zwicky transient facility DR3. *Mon. Not. R. Astron. Soc.* 502(4):5147–5175. doi: 10.1093/mnras/stab316.
5. Ho-Sang, C., V. A. Villar, S. Cheung, *et al.* 2022. Searching for anomalies in the ZTF catalog of periodic variable stars. *Astrophys. J.* 932(2):118. 21 p. doi: 10.3847/1538-4357/ac69d4.
6. Pruzhinskaya, M. V., E. E. O. Ishida, A. K. Novinskaya, *et al.* 2023. Supernova search with active learning in ZTF DR3. *Astron. Astrophys.* 672:A111. 22 p. doi: 10.1051/0004-6361/202245172.
7. Deng, L. 2012. The MNIST database of handwritten digit images for machine learning research. *IEEE Signal Proc. Mag.* 29(6):141–142. doi: 10.1109/MSP.2012.2211477.
8. Kingma, D. P., and J. Ba. 2017. Adam: A method for stochastic optimization. *arXiv.org*. 15 p. Available at: <https://arxiv.org/abs/1412.6980> (accessed March 11, 2024).
9. Hochreiter, S., and J. Schmidhuber. 1997. Long short-term memory. *Neural Comput.* 9(8):1735–1780. doi: 10.1162/neco.1997.9.8.1735.
10. Cho, K., B. Merriënboer, C. Gulcehre, *et al.* 2014. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv.org*. 15 p. Available at: <https://arxiv.org/abs/1406.1078> (accessed March 11, 2024).
11. Salehi, S., D. Erdogmus, and A. Gholipour. 2017. Tversky loss function for image segmentation using 3D fully convolutional deep networks. *arXiv.org*. 9 p. Available at: <https://arxiv.org/abs/1706.05721> (accessed March 11, 2024).
12. Ishida, E. E. O., M. V. Kornilov, K. L. Malanchev, *et al.* 2021. Active anomaly detection for time-domain discoveries. *Astron. Astrophys.* 650:A195. 9 p. doi: 10.1051/0004-6361/202037709.

*Received December 6, 2023*

## Contributor

**Semenikhin Timofey A.** (b. 1999) — engineer, Sternberg Astronomical Institute, M. V. Lomonosov Moscow State University, 13 Universitetsky Prosp., Moscow 119234, Russian Federation; ofmafowo@gmail.com

---

---

# ПОИСК АРТЕФАКТОВ НА ИЗОБРАЖЕНИЯХ ОБЗОРА ZTF ПРИ ПОМОЩИ НЕЙРОННЫХ СЕТЕЙ\*

*Т. А. Семенихин*

Государственный астрономический институт имени П. К. Штернберга Московского государственного университета имени М. В. Ломоносова

**Аннотация:** Сегодня астрономам приходится работать с большими объемами данных, поскольку современные инструменты способны генерировать терабайты информации за одну ночь. Одним из таких инструментов стал автоматизированный обзор неба Zwicky Transient Facility (ZTF), который за одну ночь детектирует порядка миллиона новых вспышек. Однако значительная доля найденных объектов оказывается артефактами, т. е. явлениями, имеющими неастрофизическую природу. Поэтому специалистам приходится тратить время на классификацию объектов вручную, так как на текущий момент не существует эффективного метода, который делал бы это без участия человека. Целью данной работы ставилась реализация эффективного алгоритма, который по последовательности кадров объекта из обзора ZTF предсказывал бы, относится он к артефактам или нет. Для реализации алгоритма использована выборка, размеченная специалистами коллаборации SNAD и содержащая 2230 серий кадров объектов. Так как последовательности кадров достаточно велики, использован вариационный автоэнкодер, который позволяет отобразить изображение в вектор меньшей длины. Для решения задачи бинарной классификации по последовательности сжатых в векторы кадров применялась рекуррентная нейронная сеть. Были рассмотрены несколько моделей нейронных сетей, для оценки метрик качества использовалась k-fold кросс-валидация. Итоговые метрики качества составляют ROC-AUC = 0,86 ± 0,01 и точность 0,80 ± 0,02 и позволяют говорить, что модель имеет практическую ценность. Код с реализацией алгоритма доступен на GitHub.

---

\*Статья публикуется по представлению программного комитета XXV Международной конференции «Аналитика и управление данными в областях с интенсивным использованием данных» (DAMDID/RCDL 2023). Автор выражает благодарность команде SNAD за комментарии и поддержку. Работа выполнена с использованием оборудования, приобретенного за счет средств Программы развития Московского университета. Исследование выполнено при финансовой поддержке Некоммерческого фонда развития науки и образования «Интеллект».

**Ключевые слова:** нейронные сети; анализ данных; классификация реальных/ложных объектов

**DOI:** 10.14357/08696527240106

**EDN:** LRMTGD

## Литература

1. *Bellm E. C., Kulkarni S. R., Graham M. J., et al.* The Zwicky Transient Facility: System overview, performance, and first results // *Publ. Astron. Soc. Pac.*, 2019. Vol. 131. Iss. 995. Art. 018002. 19 p. doi: 10.1088/1538-3873/aacbbe.
2. *Bellm E. C., Kulkarni S. R., Barlow T., et al.* The Zwicky Transient Facility: Surveys and scheduler // *Publ. Astron. Soc. Pac.*, 2019. Vol. 131. Iss. 1000. Art. 068003. 13 p. doi: 10.1088/1538-3873/ab0c2a.
3. *Bertin E., Arnouts S.* SExtractor: Software for source extraction // *Astron. Astrophys. Sup.*, 1996. Vol. 117. No. 2. P. 393–404. doi: 10.1051/aas:1996164.
4. *Malanchev K. L., Pruzhinskaya M. V., Korolev V. S., et al.* Anomaly detection in the Zwicky Transient Facility DR3 // *Mon. Not. R. Astron. Soc.*, 2021. Vol. 502. No. 4. P. 5147–5175. doi: 10.1093/mnras/stab316.
5. *Ho-Sang C., Villar V. A., Cheung S., et al.* Searching for anomalies in the ZTF catalog of periodic variable stars // *Astrophys. J.*, 2022. Vol. 932. No. 2. Art. 118. 21 p. doi: 10.3847/1538-4357/ac69d4.
6. *Pruzhinskaya M. V., Ishida E. E. O., Novinskaya A. K., et al.* Supernova search with active learning in ZTF DR3 // *Astron. Astrophys.*, 2023. Vol. 672. Art. A111. 22 p. doi: 10.1051/0004-6361/202245172.
7. *Deng Li.* The MNIST database of handwritten digit images for machine learning research // *IEEE Signal Proc. Mag.*, 2012. Vol. 29. No. 6. P. 141–142. doi: 10.1109/MSP.2012.2211477.
8. *Kingma D. P., Ba J.* Adam: A method for stochastic optimization. — Cornell University, 2017. arXiv:1412.6980 [cs.LG]. 15 p.
9. *Sepp H., Jurgen S.* Long short-term memory // *Neural Comput.*, 1997. Vol. 9. No. 8. P. 1735–1780. doi: 10.1162/neco.1997.9.8.1735.
10. *Cho K., Merriënboer B., Gulcehre C., et al.* Learning phrase representations using RNN encoder-decoder for statistical machine translation. — Cornell University, 2014. arXiv:1406.1078 [cs.CL]. 15 p.
11. *Salehi S., Erdogmus D., Gholipour A.* Tversky loss function for image segmentation using 3D fully convolutional deep networks. — Cornell University, 2017. arXiv:1706.05721 [cs.CV]. 9 p.
12. *Ishida E. E. O., Kornilov M. V., Malanchev K. L., et al.* Active anomaly detection for time-domain discoveries // *Astron. Astrophys.*, 2021. Vol. 650. Art. A195. 9 p. doi: 10.1051/0004-6361/202037709.

Поступила в редакцию 06.12.2023



## ЖИЗНЕННЫЙ ЦИКЛ КИБЕРАТАК НА УЯЗВИМОСТИ СИСТЕМ МАШИННОГО ОБУЧЕНИЯ

*А. П. Сучков*<sup>1</sup>

**Аннотация:** Рассматривается проблематика обеспечения необходимого уровня информационной безопасности (ИБ) систем искусственного интеллекта (ИИ), что становится одним из ключевых факторов их широкого использования. На основе анализа данных о совершенных кибератаках на известные системы машинного обучения (СМО) проведено формирование полного набора стадий жизненного цикла (ЖЦ) компьютерных атак на СМО, определены целевые уязвимости СМО для каждой стадии ЖЦ, систематизированы способы и методы атак на уязвимости СМО, обозначены подходы к обеспечению ИБ СМО.

**Ключевые слова:** информационная безопасность; системы искусственного интеллекта; машинное обучение; виды защиты информации

**DOI:** 10.14357/08696527240107

**EDN:** QLGHEH

### 1 Введение

Системы, основанные на методах ИИ, переживают в настоящий момент стадию бурного роста как по широте приложений, так и по реальным достижениям по всем направлениям. Такие системы применяются в области финансов, государственных услуг, в медицине, транспорте и многих других. Большой прогресс достигнут в автоматизации распознавания образов, речи, в развитии робототехники во многом благодаря существенному прогрессу СМО. Превратившись в системы высокой доступности и массового использования системы ИИ все больше становятся объектом киберпреступлений, преследующих различные экономические, политические и социальные цели и опирающихся на высокотехнологичные методы, техники и тактики.

С целью обобщения опыта борьбы с киберпреступлениями против СМО создана и постоянно обновляется глобально доступная база знаний MITRE ATT&CK<sup>®</sup> [1] о тактике и приемах противника, основанная на реальных наблюдениях. База знаний ATT&CK используется в качестве основы для разработки конкретных моделей угроз и методов защиты информации в государственном и корпоративном секторе для создания эффективных продуктов и услуг в области кибербезопасности. База знаний ATT&CK открыта и доступна для бесплатного использования любым лицом или организацией.

---

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



Российский аналог MITRE ATT&CK — MITRE ATLAS™ (Adversarial Threat Landscape for Artificial-Intelligence Systems) [2], созданный факультетом безопасности информационных технологий Университета ИТМО [3], — это база знаний о кейсах, тактиках и техниках, применяемых злоумышленниками для атак на СМО. Она оформлена в виде интерактивной матрицы, основанной на реальных наблюдениях, исследованиях комплексной имитации реальных атак с целью оценки кибербезопасности систем и групп по безопасности. База знаний ATLAS создана по образцу фреймворка MITRE ATT&CK®<sup>®</sup>, ее тактики и техники дополняют матрицу ATT&CK.

База знаний ATLAS позволяет исследователям ориентироваться в ландшафте угроз для СМО. Число уязвимостей в них растет, и их использование увеличивает фронт атак на существующие системы. Матрица ATLAS была разработана для повышения осведомленности об угрозах ИБ СМО и представления их в виде, удобном для исследователей безопасности. Базы знаний этих систем содержат более сотни примеров реальных и имитированных кибератак на известные СМО, например:

- сбой служб машинного перевода — Google Translate, Bing Translator и Systran Translate;
- копирование модели GPT-2;
- сбой службы Microsoft Azure;
- обход системы ИИ в классификаторе изображений Microsoft (on-edge);
- выполнение вредоносного кода в MathGPT с помощью быстрого внедрения и др.

Техники, методы и тактики атак в базе знаний организованы в виде матрицы по этапам их разработки и выполнения. Интересно было бы упорядочить эти этапы по стадиям ЖЦ кибератак с привязкой к уязвимостям СМО и методам их защиты. Итак, задачами исследования ставятся:

- (1) формирование полного набора стадий ЖЦ компьютерных атак на СМО;
- (2) определение целевых уязвимостей СМО для каждой стадии ЖЦ;
- (3) систематизация способов и методов атак на уязвимости СМО;
- (4) подходы к обеспечению ИБ СМО.

## 2 Стадии жизненного цикла кибератак

Как всякая высокотехнологичная разработка, компьютерная атака на СМО проходит определенные стадии своего ЖЦ — от замысла, подготовки и реализации до организованного отхода. Естественно, в реальности не все эти стадии обязательны к реализации: например, атака может носить спонтанный характер или иметь минимальную подготовку с использованием готовых решений

без тщательного мониторинга системы. Однако рассмотрим наиболее полный набор стадий ЖЦ для описания всех возможных тактик, способов и техник, применяемых в кибератаках.

Итак, на основе анализа данных [1, 2] выделим 5 стадий ЖЦ кибератак с указанием решаемых на этих стадиях задач.

1. Целеполагание:
  - (а) определение объекта атаки;
  - (б) формирование замысла атаки;
  - (в) планирование подготовки и осуществления атаки.
2. Мониторинг объекта атаки:
  - (а) сбор доступной информации о СМО;
  - (б) исследование или сканирование системы жертвы;
  - (в) выявление уязвимостей объекта атаки;
  - (г) изучение существующего опыта атак уязвимостей известных моделей.
3. Подготовка ресурсов для атаки:
  - (а) получение артефактов СМО;
  - (б) подготовка технической инфраструктуры для атаки;
  - (в) подготовка программных средств для атаки;
  - (г) публикация вредоносных данных для обучения;
  - (д) создание учетных записей к различным сервисам.
4. Осуществление атаки:
  - (а) первоначальный доступ к СМО;
  - (б) кража сведений о системе, имен учетных записей, паролей и пр.;
  - (в) запуск вредоносного кода, встроенного в артефакты или программное обеспечение (ПО);
  - (г) закрепление в СМО;
  - (д) попытка получения доступа более высокого уровня;
  - (е) уклонение от защиты;
  - (ж) кража данных;
  - (з) манипулирование СМО и данными, прерывание их работы или их уничтожение.
5. Предотвращение обнаружения и идентификации субъекта атаки.

Упорядочим данные баз знаний [1, 2] в виде таблицы. Следует отметить, что в этих базах приведено около двухсот техник и тактик атак, поэтому в таблице в ряде случаев указано их зарегистрированное число и перечислены наиболее представительные типы.

Уязвимости СМО и меры противодействия атакам на них

ЖЦ	Уязвимости	Мероприятия по защите
1а, 1б, 1в	Популярность СМО. Средний и высокий возможный ущерб	Учет степени повышенных рисков при планировании мероприятий по ИБ*
2а	Доступность информации о научных исследованиях организации и персоналий: <ul style="list-style-type: none"> <li>– репозитории препринтов;</li> <li>– технические блоги;</li> <li>– журналы и материалы конференций;</li> <li>– информации с сайта организации;</li> <li>– открытые репозитории аналогичных приложений (Google Play, iOS App Store, macOS App Store и Microsoft Store)</li> </ul>	Анализ доступности критической информации о СМО и учет при планировании системы защиты. Ограничение размещения в открытом доступе критических данных*
2б	Доступность данных о СМО: <ul style="list-style-type: none"> <li>– учетные данные;</li> <li>– адреса электронной почты;</li> <li>– имена сотрудников;</li> <li>– информация о сети;</li> <li>– информация об организации;</li> <li>– фишинг с целью получения информации;</li> <li>– поиск в закрытых источниках;</li> <li>– поиск в открытых технических базах данных;</li> <li>– поиск открытых веб-сайтов/доменов</li> </ul>	Минимизация объема и чувствительности данных, доступных внешним пользователям*
2в	Уязвимости инфраструктуры: <ul style="list-style-type: none"> <li>– уязвимость серверов;</li> <li>– уязвимость целевых сред и служб для эксплойта;</li> <li>– уязвимость программных средств;</li> <li>– уязвимость моделей</li> </ul>	Анализ и учет уязвимостей инфраструктуры. Выявление аномалий в потоках данных*
2г	Наличие в открытом доступе: <ul style="list-style-type: none"> <li>– исследования об уязвимостях распространённых моделей;</li> <li>– ранее полученные результаты, касающиеся нужного класса моделей, включая: <ul style="list-style-type: none"> <li>○ публикации статей с подробностями реализации успешной атаки;</li> <li>○ ранее созданные реализации этих атак в виде готовых программных средств</li> </ul> </li> </ul>	Поиск в открытом доступе данных об успешных атаках на выбранную модель. Учет степени повышенных рисков при планировании мероприятий по ИБ*

*Продолжение таблицы на с. 84*

Уязвимости СМО и меры противодействия атакам на них (продолжение)

ЖЦ	Уязвимости	Мероприятия по защите
За	Наличие в общедоступных источниках, включая облачные хранилища, общедоступных сервисов, а также ПО репозиториях данных, артефактов машинного обучения: <ul style="list-style-type: none"> <li>– программный стек, используемый для обучения и развертывания моделей;</li> <li>– данные обучения и тестирования;</li> <li>– конфигурации и параметры моделей</li> </ul>	Поиск в общедоступных источниках данных об артефактах СМО. Учет степени повышенных рисков при планировании мероприятий по ИБ
Зб	Уязвимости доступной инфраструктуры с использованием: <ul style="list-style-type: none"> <li>– доменных имен (похожие, устаревшие и др.);</li> <li>– DNS-серверов;</li> <li>– виртуальных частных серверов;</li> <li>– своих серверов;</li> <li>– ботнетов;</li> <li>– популярных веб-сервисов;</li> <li>– бессерверного подхода;</li> <li>– вредоносной рекламы</li> </ul>	Регистрация доменов с похожими именами. Сканирование интернета на предмет выявления серверов, приобретенных злоумышленниками. Возможен поиск уникальных характеристик, связанных с ПО противника, если они известны. Средства блокировки рекламы для предотвращения выполнения вредоносного кода*
Зв	Возможности получения ПО: <ul style="list-style-type: none"> <li>– бесплатное скачивание ПО с открытым кодом;</li> <li>– покупка ПО;</li> <li>– кража ПО и/или лицензии на ПО у сторонних организаций;</li> <li>– взлом пробных версий;</li> <li>– разработка собственных версий ПО</li> </ul>	Радикально — отказ от использования общедоступного недоверенного ПО, иначе см.*
Зг	Внедренные уязвимости в модели машинного обучения: <ul style="list-style-type: none"> <li>– подготовка и опубликование вредоносных данных для обучения;</li> <li>– отравление существующих данных для обучения;</li> <li>– создание возможности активации уязвимости образцами данных с бэкдор-триггером</li> </ul>	Радикально — отказ от использования общедоступных недоверенных обучающих данных. Мероприятия по нормализации обучающих наборов данных и по повышению устойчивости модели СМО

Продолжение таблицы на с. 85

Уязвимости СМО и меры противодействия атакам на них (продолжение)

ЖЦ	Уязвимости	Мероприятия по защите
Зд	<p>Уязвимости, связанные с доступом к сервисам для использования в целевых атаках, чтобы получить доступ к необходимым ресурсам для постановки атак или для выдачи себя за жертву:</p> <ul style="list-style-type: none"> <li>– создание и поддержание учетных записей в сервисах, которые могут быть использованы во время целевых атак — социальные сети, электронная почта, сайты и т. п.;</li> <li>– создание и использование онлайн-персонажа</li> </ul>	<p>Анализ содержания сетевого трафика. Мониторинг активности в социальных сетях, связанных с организацией СМО*</p>
4а	<p>Уязвимости в системе доступа по различным векторам входа, используемые для закрепления в системе</p> <ul style="list-style-type: none"> <li>– компрометация цепочек поставок аппаратно-программных средств, ПО, данных, моделей;</li> <li>– доступ к модели СМО через API, доступ к данным из физической среды, полный доступ к модели, косвенно через взаимодействие с продуктом или сервисом, использующим машинное обучение;</li> <li>– использование существующих учетных записей;</li> <li>– обход систем обнаружения вредоносного ПО или сетевого сканирования, основанных на машинном обучении;</li> <li>– репликация через съемные носители;</li> <li>– фишинг</li> </ul>	<p>Внедрение процесса управления исправлениями для проверки неиспользуемых зависимостей. Внедрение непрерывного мониторинга источников уязвимостей. Использование многофакторной идентификации. Сегментация сети. Надлежащее управление учетными записями и разрешениями, используемыми сторонами в доверительных отношениях. Ограничение использования USB-устройств и съемных носителей в сети</p>
4б	<p>Уязвимости в политике ведения учетных записей:</p> <ul style="list-style-type: none"> <li>– устаревший порядок использования учетных записей;</li> <li>– недостатки в политике ведения паролей</li> </ul>	<p>Использование политики условного доступа для блокировки входов с устройств, не соответствующих требованиям, или с IP-адресов за пределами определенных организацией диапазонов. Ключи должны периодически обновляться и надлежащим образом защищаться. Отслеживание необычного поведения учетных записей</p>

*Продолжение таблицы на с. 86*

Уязвимости СМО и меры противодействия атакам на них (*продолжение*)

ЖЦ	Уязвимости	Мероприятия по защите
4в	Использование уязвимостей с возможностями запуска вредоносного кода: <ul style="list-style-type: none"> <li>– запуск пользователем с умыслом или по невнимательности;</li> <li>– запуск с помощью артефактов машинного обучения;</li> <li>– использование инструментов удаленного доступа для запуска сценария</li> </ul>	Существует более 40 видов возможностей запуска вредоносного кода и техник защиты от них [1], в частности администрирование: <ul style="list-style-type: none"> <li>– облачных сервисов;</li> <li>– интерпретаторов команд и сценариев;</li> <li>– служб контейнеров;</li> <li>– клиентских приложений;</li> <li>– механизмов межпроцессной связи;</li> <li>– создания собственных API;</li> <li>– планирования задач;</li> <li>– функций пользователей</li> </ul>
4г	Уязвимости системы доступа при перезапусках, изменении учетных данных и других прерываниях, которые могут привести к прекращению доступа: <ul style="list-style-type: none"> <li>– манипулирование сроками действия паролей;</li> <li>– использование фоновых задач;</li> <li>– системы автозапуска;</li> <li>– сценариев инициализации загрузки и входа в систему;</li> <li>– внешних удаленных сервисов;</li> <li>– процессов аутентификации</li> </ul>	Существует более 20 видов возможностей закрепления и техник защиты от них [1], в частности администрирование: <ul style="list-style-type: none"> <li>– сроков действия паролей;</li> <li>– фоновых задач;</li> <li>– системы автозапуска;</li> <li>– сценариев инициализации загрузки и входа в систему;</li> <li>– внешних удаленных сервисов;</li> <li>– процессов аутентификации</li> </ul>
4д	Уязвимости в разграничении прав доступа на уровне: <ul style="list-style-type: none"> <li>– системном — корневой уровень;</li> <li>– локального администратора;</li> <li>– учетной записи пользователя с правами администратора;</li> <li>– учетных записей пользователей, имеющих доступ к определенной системе или выполняющих определенную функцию</li> </ul>	Существует более 10 видов возможностей атак и техник защиты от них [1], в частности администрирование: <ul style="list-style-type: none"> <li>– аудита операционной системы;</li> <li>– выполнения приложений;</li> <li>– конфигурации операционной системы;</li> <li>– управления привилегированными учетными записями;</li> <li>– прав доступа к файлам и каталогам</li> </ul>
4е	Методы, используемые для уклонения от защиты, включают: <ul style="list-style-type: none"> <li>– удаление/отключение ПО безопасности;</li> <li>– запутывание/шифрование данных и сценариев;</li> <li>– использование доверенных процессов и злоупотребление ими</li> </ul>	Существует более 40 видов возможностей атак и техник защиты от них [1], в частности администрирование: <ul style="list-style-type: none"> <li>– системы ПО ИБ;</li> <li>– системы архивирования и восстановления;</li> <li>– доверенных процессов</li> </ul>

*Продолжение таблицы на с. 87*

Уязвимости СМО и меры противодействия атакам на них (продолжение)

ЖЦ	Уязвимости	Мероприятия по защите
4ж	<p>Уязвимости в системе доступа к модели машинного обучения, дающие возможности:</p> <ul style="list-style-type: none"> <li>– получения данных по своему каналу управления или альтернативному каналу;</li> <li>– установления ограничений на размер передачи;</li> <li>– кражи артефактов машинного обучения;</li> <li>– вывода о принадлежности образца данных к обучающему набору, что считается нарушением конфиденциальности;</li> <li>– извлечение копии модели</li> </ul>	<p>Этот тип атак не может быть легко устранен с помощью превентивных мер контроля, поскольку он основан на злоупотреблении системными функциями. Использование системы обнаружения и предотвращения сетевых вторжений. Настройка сетевого брандмауэра, чтобы разрешать вход в сеть и выход из нее только необходимым портам. Строгий контроль идентификации и управления доступом</p>
4з	<p>Уязвимости СМО и данных, позволяющие прерывать работу или их уничтожить путем:</p> <ul style="list-style-type: none"> <li>– создания состязательных данных, которые нарушают корректную идентификацию содержимого данных моделью;</li> <li>– создания потока запросов с целью ухудшить или остановить работу всего сервиса;</li> <li>– ухудшения производительности целевой модели, направляя бесполезные запросы или вычислительно сложные входные данные на различные сервисы машинного обучения, чтобы увеличить затраты на работу этих сервисов;</li> <li>– извлечения артефактов машинного обучения, чтобы украсть интеллектуальную собственность и причинить экономический вред;</li> <li>– засорение мусорными данными, что приводит к увеличению числа ложных обнаружений</li> </ul>	<p>Существует 14 видов атак и техник защиты от них [1], в частности:</p> <ul style="list-style-type: none"> <li>– внедрение планов аварийного восстановления информационных технологий, содержащих процедуры регулярного создания резервных копий данных, которые могут использоваться для восстановления данных организации;</li> <li>– фильтрация пограничного трафика, блокируя адреса источников, на которые направлена атака, блокируя порты, на которые нацелены, или блокируя протоколы, используемые для транспортировки;</li> <li>– обучение пользователей распознавать методы социальной инженерии, используемые для совершения финансовых краж;</li> <li>– отслеживание выполняемых команд и аргументов двоичных файлов, участвующих в выключении или перезагрузке систем</li> </ul>

Окончание таблицы на с. 88

Уязвимости СМО и меры противодействия атакам на них (окончание)

ЖЦ	Уязвимости	Мероприятия по защите
5	<p>Уязвимости в системах безопасности, в том числе, основанных на методах машинного обучения для предотвращения отслеживания [4]:</p> <ul style="list-style-type: none"> <li>– использование обычных сервисов, таких как Google или Twitter, что облегчает злоумышленникам скрытие в ожидаемом шуме;</li> <li>– маскировка путем переименования инструментов и двоичных файлов в соответствии с файлами и программами на взломанном устройстве;</li> <li>– отключение регистрации событий с помощью AUDITPOL перед практически действиями с клавиатуры и включение обратно после них;</li> <li>– создание правил брандмауэра для минимизации исходящих пакетов для определенных протоколов перед запуском действий по перечислению сетей, которые потом удалялись;</li> <li>– отключение служб безопасности на целевых хостах, использование временных меток для изменения временных меток артефактов, а также процедур и инструментов очистки, чтобы препятствовать обнаружению вредоносных внедрений DLL в уязвимых средах</li> </ul>	<p>Необходимо реализовывать комплексные меры обеспечения ИБ, к одной из первоочередных можно отнести процессы мониторинга и реагирования на киберинциденты. В случае недостатка собственных сил и средств на рынке РФ существует достаточное число провайдеров сервисов безопасности [4]</p>

\*Кибератаки на эти уязвимости не могут быть легко предотвращены с помощью превентивных мер защиты, поскольку они основаны на подходах, выходящих за рамки средств защиты и контроля предприятия.

### 3 Подходы к защите информации для систем машинного обучения

К основным видам защиты информации для СМО можно отнести (ГОСТ Р 50922-2006):

- техническую защиту информации;
- криптографическую защиту информации, включая обучающие данные;
- физическую защиту информации;
- правовую защиту информации;

К основным видам защиты СМО необходимо добавить [5, 6]:

- обеспечение системной и функциональной безопасности;
- повышение устойчивости СМО.



Техническая и физическая защита, обеспечивающая конфиденциальность, целостность и доступность информации, осуществляется на основе существующих нормативно-методических документов и не отличается от комплекса мероприятий по обеспечению безопасности информации для любых автоматизированных информационных систем. При этом следует обратить внимание на то, что на начальных этапах ЖЦ кибератак эти виды защиты затруднены, так как выходят за пределы компетенции предприятия и могут быть реализованы только путем ограничения взаимодействия СМО с внешними системами и сервисами и выполнения персоналом определенных обязательств. Другая особенность обеспечения ИБ СМО кроется в их специфических свойствах. Одна из основных проблем обеспечения ИБ СМО на основе технологии машинного обучения заключается в том, что создание и функционирование таких систем невозможно без разнообразных процедур обучения, дообучения и самообучения с использованием зачастую огромных массивов данных из различных источников — баз данных, баз знаний, экспертных данных, коллекций изображений и аудиозаписей, сред сенсорного и имитационного моделирования, «цифровых двойников». Очень часто источники этих данных не входят в число доверенных и формируются в открытом доступе с целью тестирования новых подходов к созданию СМО самыми разными авторами и научным сообществом в целом.

Еще одна специфическая особенность — фактор применения интеллектуальных автономных и мобильных СМО, что предъявляет специфические требования к сохранению конфиденциальности информации и информационных процессов в смысле расширения спектра угроз ИБ (возможность извлечения и инверсии моделей и т. п.).

С точки зрения информационных процессов, влияющих на ИБ СМО, к одним из основных специфических угроз следует отнести процедуры обучения систем ИИ, в ходе которых настраивается и может меняться не только информационная база, но и алгоритмы интеллектуальных выводов. Эта проблема устойчивости СМО состоит в том, что даже небольшие изменения в обучающих наборах данных могут привести к кардинальным изменениям интеллектуальных способностей таких систем, вплоть до их существенного нарушения. Важно также очень трудно проверяемое качество обучающих данных (репрезентативность, статистическая несмещенность, достаточный объем).

Таким образом, с точки зрения ИБ СМО данные свойства таких систем предоставляют злоумышленникам более широкий спектр уязвимостей для вредоносного воздействия. Специфика защиты информации и СМО может потребовать дополнительных мер на стадии их проектирования и разработки, таких как изменение гиперпараметров моделей, применения процедур верификации обучающих данных, создание интеллектуальных подсистем тестирования на предмет устойчивости к вредоносным воздействиям. Это осуществляется в рамках обеспечения *системной* безопасности путем определения целевых областей применения СМО и ее критических элементов, формирования реализуемых системных требований на основе анализа эксплуатационных и технологических рисков.

## 4 Заключение

1. Системы машинного обучения относятся к одному из видов автоматизированных информационных систем, и к ним должны применяться все существующие виды защиты информации (правовая, техническая, криптографическая, физическая, функциональная, системная), регламентированные существующей нормативно-методической базой.
2. Основные специфические уязвимости СМО несут процедуры обучения систем ИИ, в ходе которых настраивается и может меняться не только информационная база, но и алгоритмы интеллектуальных выводов.
3. Предложена классификация уязвимостей СМО и мер защиты информации по стадиям ЖЦ кибератак.
4. Обсуждены специфические вопросы защиты информации СМО и ряд мероприятий по обеспечению ИБ этих систем.

## Литература

1. MITRE ATT&CK<sup>®</sup>. <https://attack.mitre.org>.
2. MITRE ATLAS<sup>™</sup> (Adversarial Threat Landscape for Artificial-Intelligence Systems). <https://atlas.securityhub.ru>.
3. Университет ИТМО. <https://itmo.ru/?ysclid=lqbyvy76bx557068225>.
4. Microsoft рассказала, как хакеры избежали обнаружения при атаке SolarWinds. [https://habr.com/ru/search/?q=уклонение%20от%20обнаружения&target\\_type=posts&order=relevance](https://habr.com/ru/search/?q=уклонение%20от%20обнаружения&target_type=posts&order=relevance).
5. Гаврилов В. Е., Зацаринный А. А. Исследование проблем нормативно-методического регулирования в области информационной безопасности процессов создания и внедрения информационных технологий, разрабатываемых в рамках программы «Цифровая экономика» // Вестник Воронежского института ФСИН России, 2020. № 3. С. 30–37. EDN: HNRNCW.
6. Зацаринный А. А., Сучков А. П. Некоторые подходы к анализу факторов, влияющих на информационную безопасность систем искусственного интеллекта // Системы и средства информатики, 2023. Т. 33. № 3. С. 95–107. doi: 10.14357/08696527230308. EDN: QNXZBN.

*Поступила в редакцию 20.12.23*

---

---

## THE LIFECYCLE OF CYBERATTACKS ON MACHINE LEARNING SYSTEM VULNERABILITIES

*A. P. Suchkov*

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation

**Abstract:** The article discusses the problems of ensuring the necessary level of information security of artificial intelligence systems which is becoming one of the

key factors in their widespread use. Based on the analysis of data on committed cyberattacks on well-known machine learning systems (MLS), a complete set of stages of the life cycle of computer attacks on MLS was formed, target MLS vulnerabilities for each stage of the life cycle were determined, methods of attacks on MLS vulnerabilities were systematized, and approaches to ensuring information security of the system were outlined.

**Keywords:** information security; artificial intelligence systems; machine learning; types of information protection

**DOI:** 10.14357/08696527240107

**EDN:** QLGHEH

## References

1. MITRE ATT&CK<sup>®</sup>. Available at: <https://attack.mitre.org> (accessed February 20, 2024).
2. MITRE ATLAS<sup>™</sup> (Adversarial Threat Landscape for Artificial-Intelligence Systems). Available at: <https://atlas.securityhub.ru> (accessed February 20, 2024).
3. Universitet ITMO [ITMO University]. Available at: <https://itmo.ru/?ysclid=lqbyvy76bx557068225> (accessed February 20, 2024).
4. Microsoft rasskazala, kak khakery izbezhali obnaruzheniya pri atake SolarWinds [Microsoft told how hackers evaded detection in the SolarWinds attack]. Available at: <https://habr.com/ru/news/538476/> (accessed February 20, 2024).
5. Zatsarinnyy, A. A., and A. P. Suchkov. 2023. Nekotorye podkhody k analizu faktorov, vliyayushchikh na informatsionnyuyu bezopasnost' sistem iskusstvennogo intellekta [Some approaches to the analysis of factors affecting the information security of artificial intelligence systems]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 33(3):95–107. doi: 10.14357/08696527230308. EDN: QNXZBN.
6. Gavrilov, V. E., and A. A. Zatsarinnyy. 2020. Issledovanie problem normativno-metodicheskogo regulirovaniya v oblasti informatsionnoy bezopasnosti protsessov sozdaniya i vnedreniya informatsionnykh tekhnologiy, razrabatyvaemykh v ramkakh programmy “Tsifrovaya ekonomika” [Study of the problems of normative-methodical management in the field of information security processes for the creation and implementation of information technologies developed in the framework of the “Digital economy”]. *Vestnik Voronezhskogo instituta FSIN Rossii* [Bulletin of the Voronezh Institute of the Federal Penitentiary Service of Russia] 3:30–37. EDN: HNRNCW.

*Received December 20, 2023*

## Contributor

**Suchkov Alexander P.** (b. 1954) — Doctor of Science in technology, leading scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; ASuchkov@frccsc.ru

## ВЫРАБОТКА ИСКУССТВЕННЫМИ ИНТЕЛЛЕКТУАЛЬНЫМИ РЕФЛЕКСИВНЫМИ АГЕНТАМИ ТАКТИКИ ВЕДЕНИЯ ПЕРЕГОВОРОВ\*

*С. В. Листонад*<sup>1</sup>

**Аннотация:** Предлагается метод выработки переговорной тактики искусственными гетерогенными интеллектуальными агентами рефлексивно-активных систем, предназначенных для компьютерного моделирования рассуждений в длительно существующих коллективах специалистов по решению практических проблем. Агенты таких систем рефлексивно моделируют рассуждения друг друга, в ходе решения проблемы вырабатывая стратегии и тактики ведения переговоров в соответствии с моделями контрагентов, что обеспечивает самоорганизацию агентов в сильном смысле без явного централизованного управления. Благодаря этим особенностям рефлексивно-активные системы искусственных гетерогенных интеллектуальных агентов (РАСИГИА) способны учитывать такие свойства практических проблем, как неоднородность, динамичность и комплексность, аналогично моделируемым коллективам специалистов.

**Ключевые слова:** коллектив специалистов; рефлексия; переговоры; гибридная интеллектуальная многоагентная система

**DOI:** 10.14357/08696527240108

**EDN:** RHXQRD

### 1 Введение

Переговоры — один из основных механизмов децентрализованного согласования действий в РАСИГИА [1]. Данные системы предназначены для компьютерного моделирования процессов и эффектов, которые присутствуют в длительно существующих коллективах специалистов, решающих практические проблемы. В отличие от гибридных интеллектуальных многоагентных систем [2, 3], применявшихся для этих же целей при решении проблем в области логистики и энергоснабжения, в РАСИГИА дополнительно моделируются рефлексивные механизмы [4, 5]. Благодаря этим механизмам реальные коллективы специалистов способны работать в условиях отсутствия непосредственного взаимодействия между их членами [4], а также проявлять высокую продуктивность при совместном поиске решения творческих задач [6]. В случае РАСИГИА рефлексивные механизмы определения стратегии и тактики переговоров облегчают согласование

---

\* Исследование выполнено за счет гранта Российского научного фонда № 23-21-00218, <https://rscf.ru/project/23-21-00218/>.

<sup>1</sup>Федеральный исследовательский центр «Информатика и управление» Российской академии наук, ser-list-post@yandex.ru

целей, моделей предметной области и норм поведения, обеспечивая эволюционную стадию самоорганизации системы и формирование коллективного субъекта.

Переговоры агентов — процесс их взаимодействия для достижения собственных целей, в рамках которого преимущества одного из них в контроле над ситуацией, инициативой при выработке и принятии решений не предопределены разработчиками системы [7]. Под переговорной стратегией агента понимается обобщенный подход по достижению желаемого распределения выгоды в переговорах между ним самим и его оппонентом [8]. Переговорная тактика агента — детализированный метод по реализации переговорной стратегии, в том числе с использованием манипулятивных приемов и рефлексивного управления. Разработке метода выработки стратегии переговоров агентами РАСИГИА посвящена работа [1]. В настоящей работе предлагается метод выработки переговорной тактики рефлексивным агентом в рамках выбранной стратегии.

## 2 Переговорные тактики агентов

При разработке метода выработки стратегии переговоров агентов [1] в рамках бихевиористского подхода [9–12] были выделены следующие стратегии переговоров в коллективах специалистов: избегание (уклонение), приспособление, конкуренция, сотрудничество и компромисс. Экспериментальные исследования подходов к выбору переговорных стратегий в реальных коллективах специалистов [9] позволяют предполагать, что с точки зрения максимизации результата работы РАСИГИА предпочтительным выглядит выбор стратегии сотрудничества всеми ее агентами. Выбор стратегии агентом обусловлен соотношением показателей заботы о своей выгоде и выгоде других агентов, которые устанавливаются его разработчиком или пользователем РАСИГИА, а также длительностью применения стратегии [1]. Выбранная стратегия реализуется с использованием различных характерных для нее тактических приемов [13].

Стратегия избегания применяется агентом, если его устраивает текущее положение дел и он не заинтересован в получении выгоды другим агентом. Данной стратегии соответствует тактика выжидания, когда оглашение собственной позиции насколько возможно оттягивается, а предложения другого агента не принимаются к рассмотрению [13].

Стратегия приспособления не требует для своей реализации специальных тактик ведения переговоров, так как агент, избравший ее, соглашается с любым исходом, не заботясь о достижении собственной цели.

При реализации стратегии конкуренции агент ориентируется на получение краткосрочной выгоды в результате переговоров, а не на построение долгосрочных отношений с другими агентами [14] путем учета их интересов. Для реализации данной стратегии могут быть использованы следующие тактики:

- тактика салями [13, 14], при которой агент предоставляет партнеру по переговорам информацию о своих приоритетах или делает уступки небольшими порциями, чтобы сохранить конфиденциальность и получить преимущество;

- тактика первоначального завышения требований [13, 14], в рамках которой рефлексивный агент, проанализировав соотношение собственной цели и предполагаемой цели агента-оппонента в соответствии с построенной ранее моделью последнего, выдвигает предложения, превышающие его потребности, чтобы делать уступки в ходе переговоров с минимальным ущербом для достижения своей цели;
- тактика принуждения [13, 14], которая актуальна, если агент имеет преимущество в контроле над ресурсами либо обладает возможностями, необходимыми оппоненту для его работы. В этом случае он может использовать угрозы, включая ложные (блеф), ограничения доступа оппонента к таким ресурсам и возможностям для получения преимущества в переговорах;
- тактика дискредитации оппонента [13], подразумевающая, что агент вместо обсуждения и оценки предложенной оппонентом альтернативы пытается поставить под сомнение его способность предлагать качественные решения, например на основе анализа характеристик используемых им моделей рассуждений или предыдущего опыта совместной работы;
- тактика расстановки ложных акцентов [13], которая схожа с тактикой первоначального завышения требований, но актуальна при переговорах по нескольким вопросам. В рамках данной тактики агент делает акцент на важности решения вопроса, который в действительности для него второстепенен, и в обмен на уступку по нему требует от оппонента смены позиции по ключевому вопросу;
- тактика изменения условий в последний момент [13], при использовании которой агент-оппонент, не имея возможности из-за временных ограничений проанализировать предложенные изменения и предложить свою альтернативу, вынужден согласиться на менее выгодные для него условия;
- тактика повышения требований [13], в рамках которой агент выдвигает новые требования и условия для достижения соглашения с каждой уступкой оппонента;
- тактика пакетного предложения [13], в соответствии с которой предложения по обсуждаемым вопросам группируются агентом в пакет, включающий требования и уступки оппоненту. Данный пакет может быть принят или отклонен оппонентом только целиком, т. е., будучи заинтересованным в каких-то из предложенных уступок агента, оппонент принимает на себя и обязательства по выполнению всех требований из пакета;
- тактика двойного смысла [13], подразумевающая, что агент закладывает в формируемое соглашение положения с двойным смыслом или расплывчатые формулировки с выгодой для себя, например используя концепты более высокого уровня таксономической иерархии, чем хотел бы оппонент;
- тактика загадки [13], в рамках которой агент отправляет оппоненту сообщения, которые противоречат друг другу и затрудняют моделирование агента оппонентом, если последний обладает рефлексией первого ранга [4].

Стратегия сотрудничества (совместного решения проблемы) предполагает, что агент в рамках переговоров стремится максимизировать не только свой выигрыш, но и выигрыш оппонента для выстраивания долгосрочных отношений с ним. Данной стратегии релевантны следующие тактики:

- тактика вынесения спорных вопросов за скобки [13], в рамках которой агенты в первую очередь договариваются по вопросам с минимальными разногласиями, откладывая переговоры по поводу остальных на более поздний срок;
- тактика пробного шара [13], которая состоит в первоначальном формулировании предложения агентом в наиболее общем виде, что соответствует некоторой идее в переговорах реальных специалистов. В случае согласия оппонента с предложенной агентом «идеей» она конкретизируется, и агенты приступают к согласованию деталей;
- тактика обсуждения агентами своих целей, а не требований [14], чтобы при генерации новых решений учитывать базовые потребности друг друга, а не запросы, сформированные в рамках частной модели предметной области каждого из них;
- тактика моделирования мозгового штурма с генерацией множества новых возможных решений, отвечающих потребностям обоих агентов, без предвзятого отношения к какому-либо решению [14]. Для этого агент при генерации собственных решений может использовать решения, предложенные партнером, в качестве основы, дорабатывая их с учетом собственной цели;
- тактика прозрачности, предполагающая передачу агентом партнеру как можно большего объема информации о своей цели и модели предметной области для снижения конфликтов и выработки партнером решений, учитывающих точку зрения агента на проблему [14].

Стратегия компромисса направлена на достижение в результате переговоров справедливого с точки зрения обоих агентов соглашения, для чего каждый из них должен пойти на определенные уступки, сократив свой выигрыш от желаемого или максимально возможного. Эта стратегия может быть реализована следующими тактическими приемами:

- определение точек соприкосновения [14], т. е. областей, в которых оба агента могут прийти к согласию;
- упорядочение целей [14], в рамках которого агент устанавливает приоритетные требования и цели, а также вторичные, от которых можно отказаться;
- торг [14], при котором агент предлагает уступки и требует взамен уступок от агента-оппонента.

### **3 Метод выработки переговорной тактики рефлексивным агентом**

В рамках реализации переговорной стратегии агент может использовать одну или несколько реализуемых им тактик в соответствии с выбранной стратегией,



сложившейся ситуации и моделью агента-оппонента по переговорам. Предлагаемый метод основан на подходе рассуждений по прецедентам (*англ.* case-based reasoning) [15]. Метод выработки тактики включает в себя две функции (действия): «выбор тактики и ее настройка» и «сохранение результатов применения тактики».

Входные параметры для работы функции «выбор тактики и ее настройка»: «выбранная стратегия»  $ngstr \in \{\text{«избегание»}, \text{«приспособление»}, \text{«конкуренция»}, \text{«компромисс»}, \text{«сотрудничество»}\}$  [1], «модель агента-оппонента»  $agm$ , описываемая формулой (1), «критерий оптимальности переговоров»  $opc$  — целевая функция переговоров для данного агента, «множество тактик агента»  $ACT_{tct}^{ag}$  из рассмотренных в предыдущем разделе, «множество тактик агента, уже использованных при реализации стратегии»  $ACT_{tctud}^{ag}$ , «вероятность случайного выбора тактики»  $trp \in [0, 1]$ , «число выбираемых прецедентов»  $k \in \mathbb{N}$ . Результат работы функции «выбор тактики и ее настройка» — тактика  $act_{tct}^{ag} \in ACT_{tct}^{ag}$  со значениями параметров ее настройки, если они предусмотрены.

Параметр «модель агента-оппонента» представляет собой следующий кортеж:

$$agm = \langle id, gl, LANG, ont, prot, ACT, AGM^{ag} \rangle, \quad (1)$$

где  $id$  — идентификатор агента;  $gl$  — цель агента;  $LANG$  — множество языков, сообщения на которых могут быть записаны или прочитаны агентом;  $ont$  — онтология (модель предметной области) агента;  $prot$  — модель протокола решения проблемы, разработанная агентом;  $ACT$  — множество действий, реализуемых агентом;  $AGM^{ag}$  — множество моделей других агентов, также описываемых кортежем (1).

Функция «выбор тактики и ее настройка» может быть представлена следующей последовательностью шагов:

- (1) инициализировать входные переменные полученными значениями;
- (2) сгенерировать случайное число  $p \in [0, 1]$ , если  $p > trp$ , перейти к п. 5;
- (3) случайным образом выбрать тактику  $act_{tct}^{ag}$  из множества  $ACT_{tct}^{ag}$ , релевантную выбранной стратегии, а также установить значения параметров ее настройки;
- (4) перейти к п. 12;
- (5) извлечь из библиотеки прецедентов  $k$  наиболее похожих на сложившуюся ситуацию (выбранной стратегии  $ngstr$ , модели агента-оппонента  $agm$ , критерий оптимальности переговоров  $opc$ ) прецедентов следующего вида:

$$prcdt = \langle ngstr, agm, opc, act_{tct}^{ag}, mtcd, mdopc, nprcdt \rangle,$$

где  $act_{tct}^{ag}$  — тактика агента для использования в ходе переговоров с ее настройками;  $mtcd$  — средняя длительность использования тактики;  $mdopc$  — среднее изменение критерия оптимальности переговоров;  $nprcdt$  — число случаев возникновения прецедента;



- (6) из сформированного в п. 5 множества прецедентов  $PRCDT$  удалить прецеденты, содержащие использованные тактики из множества  $ACT_{tctud}^{ag}$

$$PRCDT^- = PRCDT \setminus \{prcdt \setminus prcdt \in PRCDT \wedge \text{Пр}_4(prcdt) \in ACT_{tctud}^{ag}\},$$

где  $\text{Пр}_i$  — проекция вектора на  $i$ -ю компоненту;

- (7) если  $PRCDT^- = \emptyset$ , перейти к п. 3;
- (8) из множества  $PRCDT^-$  сформировать список  $PRCDTL$ , отсортированный сначала по убыванию степени сходства с текущей ситуацией, а при равных значениях степени сходства по убыванию среднего изменения критерия оптимальности в результате переговоров  $mdopc$  и возрастанию средней длительности использования тактики  $mtcd$ ;
- (9) из списка  $PRCDTL$  выбрать первый элемент  $prcdt$ ;
- (10) если сходство  $prcdt$  с текущей ситуацией равно единице, т. е. ситуация в точности повторяет прецедент, выбрать из прецедента тактику с ее настройками  $act_{tct}^{ag} = \text{Пр}_4(prcdt)$  и перейти к п. 12;
- (11) если несколько прецедентов из  $PRCDTL$  содержат ту же тактику, что и  $prcdt$  с разными параметрами и значениями сходства с текущей ситуацией, выбрать тактику  $act_{tct}^{ag} = \text{Пр}_4(prcdt)$ , а параметры установить, экстраполировав зависимость между параметрами тактики и сходством с ситуацией, соответствующие сходству, равному единице;
- (12) вернуть в качестве результата выбранную тактику  $act_{tct}^{ag}$  со значениями параметров ее настройки и завершить работу функции.

Функция «сохранение результатов применения тактики» вызывается после того, как агент в течение некоторого числа итераций переговоров применяет выбранную тактику и либо достигает поставленной цели переговоров, либо принимает решение о бесперспективности ее дальнейшего использования. Входными параметрами функции служат «выбранная стратегия»  $ngstr^*$ , «модель агента-оппонента»  $agm^*$ , «критерий оптимальности переговоров»  $opc^*$ , «тактика агента»  $act_{tct}^{ag*}$ , «длительность использования тактики»  $tcd^*$ , «изменение критерия оптимальности переговоров»  $dopc^*$ . Выполняется поиск в базе прецедентов на предмет наличия прецедента с такой же выбранной стратегией  $ngstr^*$ , моделью агента-оппонента  $agm^*$ , критерием оптимальности переговоров  $opc^*$  и тактикой  $act_{tct}^{ag*}$  со значениями параметров ее настройки. Если прецедент отсутствует в базе прецедентов, создается прецедент следующего вида:

$$prcdt^* = \langle ngstr^*, agm^*, opc^*, act_{tct}^{ag*}, tcd^*, dopc^*, 1 \rangle.$$

В противном случае у найденного прецедента  $\text{prcdt}^f$  корректируются средняя длительность использования тактики  $\text{tcd}^f$ , среднее изменение критерия оптимальности переговоров  $\text{dopc}^f$  и число случаев возникновения прецедента  $\text{nprcdt}^f$  с учетом информации из нового прецедента  $\text{prcdt}^*$  следующим образом:

$$\begin{aligned}\text{tcd}^f &= \left( \text{tcd}^f \cdot \text{nprcdt}^f + \text{tcd}^* \right) \left( \text{nprcdt}^f + 1 \right)^{-1}; \\ \text{dopc}^f &= \left( \text{dopc}^f \cdot \text{nprcdt}^f + \text{dopc}^* \right) \left( \text{nprcdt}^f + 1 \right)^{-1}; \\ \text{nprcdt}^f &= \text{nprcdt}^f + 1,\end{aligned}$$

после чего функция «сохранение результатов применения тактики» завершает свою работу.

## 4 Заключение

В статье разработан метод выработки переговорной тактики рефлексивными агентами с использованием сформированных ими моделей других агентов системы. Предложенный метод позволяет выбирать и менять тактику ведения переговоров, релевантную выбранной стратегии и сложившейся ситуации в ходе переговоров с учетом модели агента-оппонента и предыдущего опыта ведения переговоров. В случае отсутствия релевантных прецедентов в базе, а также с некоторой заданной долей вероятности тактика и значения параметров ее настройки выбираются случайным образом, что позволяет вносить разнообразие в базу прецедентов. В качестве недостатка предложенного подхода следует отметить необходимость длительного накопления прецедентов для обеспечения выбора релевантной ситуации и модели агента-оппонента тактики ведения переговоров, что может быть компенсировано накоплением прецедентов на этапе тестирования РАСИГИА.

## Литература

1. Листопад С. В., Луцко А. С. Метод выработки стратегии поведения при переговорах искусственных гетерогенных интеллектуальных агентов рефлексивно-активных систем // *Нейроинформатика, ее приложения и анализ данных: Мат-лы XXXI Всеросс. семинара.* — Красноярск: Институт вычислительного моделирования СО РАН, 2023. С. 83–90.
2. Колесников А. В., Кириков И. А., Листопад С. В. Гибридные интеллектуальные системы с самоорганизацией: координация, согласованность, спор. — М.: ИПИ РАН, 2014. 189 с.
3. Колесников А. В., Листопад С. В. Функциональная структура гибридной интеллектуальной многоагентной системы гетерогенного мышления для решения проблемы восстановления распределительной электросети // *Системы и средства информатики*, 2019. Т. 29. № 1. С. 41–52. doi: 10.14357/08696527190104. EDN: QDLMBW.

4. *Левфевр В. А.* Рефлексия. — М.: Когито-Центр, 2003. 496 с.
5. *Новиков Д. А., Чартишвили А. Г.* Рефлексия и управление: математические модели. — М.: Физматлит, 2012. 412 с. EDN: PFGVWP.
6. *Семёнов И. Н., Степанов С. Ю., Найдёнов М. И., Найдёнова Л. А.* Рефлексия в организации мышления при совместном решении задач // Новые исследования в психологии и возрастной физиологии, 1989. № 1. С. 4–10.
7. *Рыбкин А. Г., Эмих О. К.* Стратегия сложных переговоров. — М.: ИНФРА-М, 2019. 260 с.
8. Negotiation strategy. <https://www.negotiations.com/definition/negotiation-strategy>.
9. *Pruitt D.* Strategic choice in negotiation // Negotiation theory and practice / Eds. J. W. Breslin, J. Z. Rubin. — Cambridge: The Program on Negotiation at Harvard Law School, 1991. P. 27–46.
10. *Carnevale P., Pruitt D.* Negotiation and mediation // Annu. Rev. Psychol., 2003. Vol. 43. P. 531–582. doi: 10.1146/annurev.ps.43.020192.002531.
11. *Lewicki R. J., Hiam A.* Mastering business negotiation: A working guide to making deals and resolving conflict. — San Francisco, CA, USA: Jossey-Bass, 2006. 320 p.
12. *Стремовская А. Л.* Возможные стратегии ведения переговорного процесса // Российский внешнеэкономический вестник, 2012. № 8. С. 79–88. EDN: PJCGJX.
13. *Шеретов С. Г.* Ведение переговоров. — Алматы: Юрист, 2008. 92 с.
14. Negotiation strategy: Types, techniques & examples. <https://thestrategystory.com/blog/negotiation-strategy-types-techniques-examples>.
15. *Варшавский П. Р., Еремеев А. П.* Моделирование рассуждений на основе прецедентов в интеллектуальных системах поддержки принятия решений // Искусственный интеллект и принятие решений, 2009. № 2. С. 45–57.

*Поступила в редакцию 30.01.24*

---

---

## DEVELOPMENT OF NEGOTIATION TACTICS BY ARTIFICIAL INTELLIGENT REFLEXIVE AGENTS

*S. V. Listopad*

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation

**Abstract:** The paper proposes a method for developing negotiation tactics by artificial heterogeneous intelligent agents of reflexive-active systems designed for computer modeling of reasoning in long-term teams of specialists in solving practical problems. Agents of such systems reflexively model each other’s reasoning developing strategies and tactics for negotiating, while solving a problem, in accordance with the models of counterparties which ensures self-organization of agents in a strong sense without explicit centralized control. Thanks to these features, reflexive-active systems of artificial heterogeneous intelligent agents are able to take into account such properties of practical problems as heterogeneity, dynamism, and complexity similar to simulated teams of specialists.

**Keywords:** team of specialists; reflection; negotiations; hybrid intelligent multiagent system

**DOI:** 10.14357/08696527240108

**EDN:** RHXQRD

## Acknowledgments

The work was supported by the Russian Science Foundation, project No. 23-21-00218.

## References

1. Listopad, S. V., and A. S. Luchko. 2023. Metod vyrabotki strategii povedeniya pri peregovorakh iskusstvennykh geterogennykh intellektual'nykh agentov refleksivno-aktivnykh sistem [Method for developing a behavioral strategy during negotiations of artificial heterogeneous intelligent agents of reflexive-active systems]. *Neyroinformatika, ee prilozheniya i analiz dannykh: Mat-ly XXXI Vseross. seminar* [Neuroinformatics, its applications and data analysis: Proceedings of the 31st All-Russian Seminar]. Krasnoyarsk: Institute of Computational Modelling SB RAS. 83–90.
2. Kolesnikov, A. V., I. A. Kirikov, and S. V. Listopad. 2014. *Gibridnye intellektual'nye sistemy s samoorganizatsiyey: koordinatsiya, soglasovannost', spor* [Hybrid intelligent systems with self-organization: Coordination, consistency, and dispute]. Moscow: IPI RAN. 189 p.
3. Kolesnikov, A. V., and S. V. Listopad. 2019. Funktsional'naya struktura gibridnoy intellektual'noy mnogoagentnoy sistemy geterogenogo myshleniya dlya resheniya problemy vosstanovleniya raspredelitel'noy elektroseti [Functional structure of the hybrid intelligent multiagent system of heterogeneous thinking for solving the problem of restoring the distribution power grid]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(1):41–52. doi: 10.14357/08696527190104. EDN: QDLMBW.
4. Lefevr, V. A. 2003. *Refleksiya* [Reflection]. Moscow: Kogito-Tsentr. 496 p.
5. Novikov, D. A., and A. G. Chkhartishvili. 2012. *Refleksiya i upravlenie: matematicheskie modeli* [Reflection and control: Mathematical models]. Moscow: Fizmatlit. 412 p. EDN: PFGVWP.
6. Semenov, I. N., S. Yu. Stepanov, M. I. Naydenov, and L. A. Naydenova. 1989. Refleksiya v organizatsii myshleniya pri sovmestnom reshenii zadach [Reflection in the organization of thinking during joint problem solving]. *Novye issledovaniya v psikhologii i vozrastnoy fiziologii* [New Research in Psychology and Age-Related Physiology] 1:4–10.
7. Rybkin, A. G., and O. K. Emikh. 2019. *Strategiya slozhnykh peregovorov* [Strategy for difficult negotiations]. Moscow: INFRA-M. 260 p.
8. Negotiation strategy. Available at: <https://www.negotiations.com/definition/negotiation-strategy> (accessed February 26, 2024).
9. Pruitt, D. 1991. Strategic choice in negotiation. *Negotiation theory and practice*. Eds. J. W. Breslin and J. Z. Rubin. Cambridge: The Program on Negotiation at Harvard Law School. 27–46.

10. Carnevale, P., and D. Pruitt. 2003. Negotiation and mediation. *Annu. Rev. Psychol.* 43:531–582. doi: 10.1146/annurev.ps.43.020192.002531.
11. Lewicki, R. J., and A. Hiam. 2006. *Mastering business negotiation: A working guide to making deals and resolving conflict*. San Francisco, CA: Jossey-Bass. 320 p.
12. Stremovskaya, A. L. 2012. Vozmozhnye strategii vedeniya peregovornogo protsessa [Potential negotiation strategies]. *Rossiyskiy vneshneekonomicheskiy vestnik* [Russian Foreign Economic J.] 8:79–88. EDN: PJCGJX.
13. Sheretov, S. G. 2008. *Vedenie peregovorov* [Negotiating]. Almaty: Yurist. 92 p.
14. Negotiation strategy: Types, techniques & examples. Available at: <https://thestrategystory.com/blog/negotiation-strategy-types-techniques-examples> (accessed February 26, 2024).
15. Varshavskii, P. R., and A. P. Ereemeev. 2010. Modeling of case-based reasoning in intelligent decision support systems. *Scientific Technical Information Processing* 37:336–345. doi: 10.3103/S0147688210050096.

*Received January 30, 2024*

## **Contributor**

**Listopad Sergey V.** (b. 1984) — Candidate of Science (PhD) in technology, senior scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; ser-list-post@yandex.ru

## НЕКОТОРЫЕ СЕТЕВЫЕ АСПЕКТЫ ПРОГРАММНОГО ОПРЕДЕЛЕНИЯ «ОБЛАЧНЫХ» ПЛАТФОРМ

*В. Б. Егоров*<sup>1</sup>

**Аннотация:** Термин «облачные вычисления» прочно вошел в сферу информационных технологий (ИТ). Требования к современным «облачным» платформам в гибкой ориентированной на приложения инфраструктуре наиболее эффективно реализуются их программным определением. Важнейшим инфраструктурным компонентом программно-определяемой «облачной» платформы выступает сеть, поскольку она не только консолидирует все ее инфраструктурные компоненты, но и обеспечивает клиентский доступ к сервисам «облака». В программно-определяемой «облачной» платформе естественно смотрится сеть, организованная также на принципах программного определения, которое способно придать «облачной» инфраструктуре необходимую гибкость и управляемость. Настоящая статья рассматривает некоторые вопросы программного определения «облачных» платформ, уделяя особое внимание ряду сетевых аспектов: от имеющейся терминологической неопределенности в отношении сетевой инфраструктуры центров обработки данных (ЦОД) до некоторых особенностей сети в ориентированной на приложения инфраструктуре и вопросов управления программно-определяемыми «облачными» сетями.

**Ключевые слова:** «облачная» платформа; программно-определяемое «облако»; программно-определяемые сети

**DOI:** 10.14357/08696527240109

**EDN:** CGZXIX

### 1 Введение

Термин «облачные вычисления», или просто «облака», прочно вошел в сферу ИТ, а в последнее время и в повседневный быт миллионов людей. Согласно Рекомендации SP 800-145 Национального института по стандартам и технологиям США (National Institute of Standards and Technology), «“облачные вычисления” — это модель обеспечения сетевого доступа по требованию к некоторому общему фонду конфигурируемых вычислительных ресурсов (сетям передачи данных, серверам, устройствам хранения данных, приложениям и сервисам — как вместе, так и по отдельности), которые могут быть быстро предоставлены и освобождены с минимальными эксплуатационными затратами или обращениями к провайдеру». Поскольку в основе модели «облачных вычислений» лежит сетевой доступ к ресурсам, особое место в «облачных» платформах занимает

---

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

предоставляющая этот доступ и консолидирующая «облачную» инфраструктуру сеть.

К современным «облачным» платформам в качестве важнейших предъявляются требования гибкости, эластичности и масштабируемости вычислений, а также сервисной ориентированности и динамичной адаптируемости к рабочим нагрузкам (workloads) [1]. Эти требования наиболее эффективно реализуются программно-определяемыми «облачными» платформами, в которых, в свою очередь, центральное место занимает программное определение сети (software-defined networking, SDN) [2]. Настоящая статья рассматривает некоторые вопросы организации программно-определяемых «облачных» платформ, уделяя особое внимание ряду сетевых аспектов.

## 2 Программно-определяемая «облачная» платформа

Современная виртуализированная инфраструктура ЦОД обзавелась программно-определяемой сетью и программно-определяемыми системами хранения данных (СХД), на подходе программно-определяемая флеш-память [3]. Чтобы получить современную «облачную» платформу, к такой виртуализированной инфраструктуре необходимо добавить автоматизацию всех выполняемых действий [2]. К этому хорошему рецепту остается только добавить, что одной из самых распространенных и действенных форм «автоматизации всех выполняемых действий» для ИТ-платформ становится их программное определение. Именно программное определение дает возможность реализовать требования к современной «облачной» платформе, включая принципиальную для «облака» сервисную ориентированность [4]. Поэтому «облачные» платформы развиваются в направлении программно-определяемых платформ, главная особенность которых в соответствии с общей концепцией программного определения заключается в централизованном управлении, реализуемом интеллектуальным программным обеспечением (ПО) в логически едином контроллере, в данном случае — контроллере «облака». Вряд ли в настоящее время программно-определяемое «облако» (software-defined cloud, SDC) можно считать общепринятой практикой, но в той или иной мере отвечающие требованиям SDC продукты постепенно внедряются поставщиками «облачных» услуг, в числе которых Amazon, Google и Open Compute Project [1].

Важнейшим инфраструктурным компонентом любой ИТ-системы выступает сеть, поскольку она связывает воедино для совместной работы прочие компоненты: «именно сетевые сервисы превращают некое множество серверов и СХД в законченную ИТ-систему» [2]. Даже из самых общих соображений в SDC естественно смотрится сеть, организованная на принципах SDN, а следующее краткое, но выразительное определение SDC прямо увязывает его с SDN: «SDC — это подход, при котором... SDN и “облачные вычисления” идут рука об руку» [5]. Главная цель SDN в SDC — придание «облачной» инфраструктуре требуемой динамической управляемости [6].

Но как заставить подружиться и «идти рука об руку» SDN и «облаку»? Самое естественное решение — переложить функции контроллера SDN-сети на контроллер «облака», что вполне соответствует централизации управления в парадигме программного определения. Тогда универсальный контроллер, управляя всеми физическими и виртуальными ресурсами SDC, будет также контролировать и SDN-сеть [5]. Автоматизированное с помощью интеллигентного ПО управление ИТ-системами, обеспечивающее согласование бизнес-запросов с приложениями, данными и инфраструктурой, нередко называют «оркестровкой» (orchestration). В таком контексте контроллер «облака» оказывается своеобразным «дирижером» (orchestrator) «оркеструемой» ИТ-системы. В функции «дирижера» SDC входит «обеспечение гибкости последовательного распространения политик» [7], а также контроль взаимосвязей и взаимодействия между рабочими нагрузками [5].

Преобразование традиционного «облака» в SDC в конечном счете предполагает переход на программно-определяемую «облачную» платформу. На пути обретения такой платформы помимо объективных трудностей создания «дирижеров» [8] могут возникнуть дополнительные искусственные проблемы из-за имеющейся терминологической неопределенности в отношении инфраструктуры ЦОД и использования этой неопределенности недобросовестной рекламой. В этом плане показателен пример компании SoftLine. С одной стороны, она агитирует за переход к программно-определяемой инфраструктуре ЦОД, ставя целью отвязать оборудование от определенного поставщика: «Так как программно-определяемая инфраструктура строится на базе стандартных серверов, то она не требует покупки СХД и оборудования для сети передачи. Такая концепция позволяет гибко подходить к выбору поставщика оборудования, не привязываясь к определенному вендору». Однако тут же в отношении ПО для программно-определяемого ЦОД компания рекомендует комплексное решение от одного единственного вендора: «...внедрение комплексного решения для построения программно-определяемого ЦОД с отдельной консолью для управления всеми компонентами, которые требуются для его построения... позволяет построить по-настоящему простую в эксплуатации современную инфраструктуру, подходящую для работы любых приложений. При этом пропадают проблемы, вызванные совместимостью отдельных компонентов, а также снижается совокупная стоимость владения и риски при эксплуатации инфраструктуры, так как администрировать стек продуктов одного производителя значительно проще, чем «зоопарк» систем от разных вендоров» [9], т. е. для оборудования, которое SoftLine не поставляет, пользователь волен выбирать поставщика, но когда дело касается ПО, ему, по существу, навязывается единственный комплексный поставщик, а именно: SoftLine. Однако настоящая проблема не в рекламных уловках. Хуже то, что неопределенность терминологии компания использует для откровенного подлога, так как вопреки ее рекламе «строится на базе стандартных серверов» вовсе не программно-определяемая, а (гипер)конвергентная инфраструктура. А программно-определяемыми могут быть и «нестандартные»



серверы в традиционной инфраструктуре. Более того, программно определяться могут не только серверы, но также СХД и сети. Единственный же поставщик в случае программно-определяемого «облака» обретает смысл, если он предоставляет готовую под ключ платформу SDC, в которой оборудование (даже если это «зоопарк» систем от разных вендоров») и ПО комплексно увязаны, сбалансированы и «оркестрованы» продавцом платформы для эффективной совместной работы [10].

### 3 «Облачная» инфраструктура, ориентированная на приложения

Важным достоинством технологии виртуализации стала возможность миграции виртуальных машин (ВМ). С проникновением виртуализации в «облачные» сферы пространство миграций распространилось за рамки не только сервера, но и целого ЦОД. Для переноса ВМ в другой ЦОД «облака» возможны разные причины [2]:

- рост рабочей нагрузки на приложения ВМ при ограниченных ресурсах инфраструктуры ЦОД;
- катастрофическая утеря работоспособности ЦОД с невозможностью быстрого ее восстановления (disaster recovery);
- поддержка методологии автоматизации процессов сборки, настройки и развертывания ПО (development & operations) для новых клиентских сервисов;
- пространственное перемещение источника или потребителя сервиса, связанного с работающим на ВМ приложением.

Мигрирующая в другой ЦОД ВМ должна «прихватить» с собой свои сетевые компоненты и настройки, причем с точки зрения пользователей порядок работы с приложениями ВМ не должен меняться. Конечно, эти задачи в том или ином виде решаются и в традиционных ЦОД, но полноценно решить задачу в автоматизированном режиме в масштабах SDC способна только программно-определяемая ориентированная на приложения платформа. Поэтому «внедрение программно-определяемого и ориентированного на приложения ЦОД с SDN имеет жизненно важное значение для перехода в “облако»» [11].

Примером такого ориентированного на приложения подхода к «облачному» ЦОД может служить программный продукт компании Cisco Systems для управления сетевой инфраструктурой, ориентированной на приложения (application-centric infrastructure, ACI). «Архитектура ACI применяет модель политик для хранения сетевых настроек приложения, что позволяет автоматически распространять изменения на все устройства сети и, соответственно, сокращать время и ресурсы, необходимые для разворачивания приложений, более того, автоматизировать этот процесс» [2]. Сквозные политики ACI делают возможной автоматизацию как развертывания, так и миграции ВМ в масштабах всего «облака» и, как следствие, становятся важным шагом к SDC. В отношении сети

существенно также то, что в АСІ вычислительный и сетевой компоненты инфраструктуры настолько взаимосвязаны, что разворачиваются, настраиваются и управляются в тесном взаимодействии единым «дирижером».

#### **4 Особенности управления «облачными» сетями**

Очевидным преимуществом использования SDN в «облаках» выглядит возможность непосредственной гибко управляемой коммутации в виртуальных оверлейных сетях. Действительно, задаваемые SDN-контроллером узлам сети политики поведения могут включать указания для коммутации непосредственно по виртуальным адресам, например идентификаторам виртуальной сети (virtual network identifiers) в случае расширяемой локальной сети (virtual extensible local-area network, VxLAN). Кроме того, поскольку в SDN-сети правила коммутации формулируются и задаются контроллером, сетевые устройства плоскости данных, на первый взгляд, избавляются от необходимости следить за состоянием сети и, как следствие, поддерживать протоколы маршрутизации. Но, к сожалению, на практике SDN-сети все равно остаются традиционными IP-сетями, сохраняющими во всех своих узлах маршрутизаторы, которые оказываются незаменимыми, например, в процессе развертывания и конфигурирования SDN-сетей [12]. Кроме того, в случае виртуальных, в том числе «облачных», сетей их IP-функциональность обуславливается наиболее распространенными технологиями виртуализации: маршрутизирующая инкапсуляция (generic routing encapsulation) использует непосредственно сетевой протокол IP, а VxLAN — транспортный протокол UDP, работающий поверх IP.

Как правило, модель оверлейных сетей предполагает, что несущая физическая сеть идеальна, т. е. имеет полносвязную топологию и неограниченную пропускную способность. Однако физические сети имеют ограничения и по связности, и по пропускной способности, к тому же в них могут случаться отказы. Виртуализация сети, при всех ее достоинствах, фактически означает, что сетевой администратор вынужден иметь дело с двумя сетями вместо одной. И хотя наложенные сети работают поверх физических, это разные сети с совершенно разными инструментами администрирования и устранения возникающих проблем [2]. В результате SDN-контроллер, даже имея возможность непосредственно управлять виртуальными сетями, вынужден рутинно коррелировать все свои действия в виртуальном пространстве с реальными возможностями физических сетей и, главное, согласовывать трафик во всех сетях, физических и виртуальных, с многочисленными соглашениями об уровне обслуживания (service level agreements), которые заключаются и действуют без всяких ссылок и скидок на сетевую виртуализацию и программное определение.

Концепция SDN изначально была нацелена на физические сети. Для управления физической сетью у SDN-контроллера есть и возможности, и средства, предоставляемые, в частности, протоколом OpenFlow. Но оверлейные виртуальные сети объективно оказываются вне компетенции SDN-контроллера. Они

создаются, управляются и ликвидируются средствами виртуализации в связи с другими виртуальными сущностями, ВМ и виртуализированными сетевыми функциями (virtualized network functions). Чтобы в «облачной» SDN-сети потенциальная возможность прямого управления виртуальными сетями стала реальной, необходимо обеспечить тесное взаимодействие SDN-контроллера с контроллером «облака». Для SDC это фактически означает, как отмечалось выше, что первый должен фактически «раствориться» во втором. В согласии с этим тезисом в [6] предлагается платформа для интегрированного управления вычислительными и сетевыми ресурсами в «облачной» инфраструктуре, разработанная на OpenStack и OpenDaylight. Платформа способна выполнять размещение и миграцию ВМ, планирование сетевых потоков с распределением полос пропускания, мониторинг вычислительных и сетевых ресурсов в режиме реального времени, а также контроль энергопотребления.

## 5 Заключение

«Облачные» платформы развиваются в направлении программно-определяемых платформ, которые характеризуются централизованным управлением с помощью интеллектуального ПО. Программное определение придает им гибкость, масштабируемость, сервисную ориентированность и динамичную адаптируемость к рабочим нагрузкам. Важнейшим инфраструктурным компонентом «облачной» платформы выступает сеть, поскольку она связывает воедино прочие компоненты и предоставляет клиентский доступ к сервисам. В SDC естественной выглядит организация сети на принципах SDN, причем SDN-контроллер в этом случае должен стать составной частью SDC-контроллера.

При выборе поставщика «облачной» платформы помимо объективных трудностей, связанных с ее «оркестровкой», дополнительные искусственные проблемы могут возникнуть из-за имеющейся недобросовестной рекламы и терминологической неопределенности. В этих условиях предпочтение может быть отдано предложению готовой под ключ платформы SDC, в которой оборудование и ПО комплексно увязаны, сбалансированы и «оркестрованы» поставщиком для эффективной совместной работы.

Существенным шагом к SDC может стать инфраструктура на принципах ACI, которая упрощает автоматизацию развертывания приложений и миграции ВМ в масштабах всего «облака».

## Литература

1. *Kranz G.* SDDC (software-defined data center) // TechTarget, 2017. [https://www.techtarget.com/searchdatacenter/definition/SDDC-software-defined-data-center#:~:text=An%20SDDC%20\(software%2Ddefined%20data,configuration%20are%20abstracted%20from%20hardware.](https://www.techtarget.com/searchdatacenter/definition/SDDC-software-defined-data-center#:~:text=An%20SDDC%20(software%2Ddefined%20data,configuration%20are%20abstracted%20from%20hardware.)

2. *Xopouhux Д.* SDDC — основа новой облачной инфраструктуры // Jet Info, 2014. No. 11. P. 16–21. [https://www.jetinfo.ru/wp-content/uploads/2014/12/jet-info\\_11\\_2014.pdf](https://www.jetinfo.ru/wp-content/uploads/2014/12/jet-info_11_2014.pdf).
3. *Robb D.* Top software-defined data center (SDDC) trends // Datamation, 2022. <https://www.datamation.com/data-center/software-defined-data-center-trends>.
4. *Crockett E.* Top software-defined data center providers (SDDC) // Datamation, 2022. <https://www.datamation.com/data-center/software-defined-data-center-providers>.
5. *Abbasi A. A., Abbasi A., Shamshirband S., Chronopoulos A. T., Persico V., Pescapé A.* Software-defined cloud computing: A systematic review on latest trends and developments // IEEE Access, 2019. Vol. 7. P. 93294–93314. doi: 10.1109/ACCESS.2019.2927822.
6. *Son J., Buyya R.* SDCon: Integrated control platform for software-defined clouds // IEEE T. Parallel. Distrib. Syst., 2019. Vol. 30. P. 230–244. doi: 10.1109/TPDS.2018.2855119.
7. *Kameshwar L.* Software defined networking: Connecting hybrid clouds seamlessly // Wipro, 2023. <https://www.wipro.com/infrastructure/software-defined-networking-connecting-hybrid-clouds-seamlessly>.
8. *Егоров В. Б.* Некоторые вопросы программного определения центров обработки данных // Системы и средства информатики, 2020. Т. 30. № 2. С. 103–112. doi: 10.14357/08696527200210. EDN: FXXGJY.
9. Как построить полноценное гибридное облако и снизить стоимость владения ИТ-инфраструктурой // С-News, 2021. [https://softline.cnews.ru/articles/2021-07-14\\_virtualnyj-tsod\\_i\\_gibridnoe\\_oblako](https://softline.cnews.ru/articles/2021-07-14_virtualnyj-tsod_i_gibridnoe_oblako).
10. *Егоров В. Б.* Некоторые вопросы деагрегации и компонентности инфраструктуры центра обработки данных // Системы и средства информатики, 2023. Т. 33. № 2. С. 101–110. doi: 10.14357/08696527230210. EDN: OEXVDB.
11. *Steen T.* Why software-defined solutions are key to a cloud-ready architecture // SHI, 2020. <https://blog.shi.com/next-generation-infrastructure/why-software-defined-solutions-are-key-to-a-cloud-ready-architecture>.
12. *Егоров В. Б.* К вопросу о соотношении программно определяемых и традиционных IP-сетей // Системы и средства информатики, 2022. Т. 32. № 1. С. 73–82. doi: 10.14357/08696527220107. EDN: OGWOXL.

*Поступила в редакцию 05.10.23*

---

## SOME NETWORKING ASPECTS OF SOFTWARE-DEFINED CLOUD PLATFORMS

*V. B. Egorov*

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation

**Abstract:** The term “cloud computing” has firmly entered the area of information technologies. The requirements to modern cloud platforms in a flexible application-centric infrastructure are most effectively implemented with

software-defined cloud platforms. The most important infrastructure component of the software-defined cloud is the network which not only consolidates all infrastructure components of the platform but also provides the client access to cloud services. In the software-defined cloud, a software-defined network looks most naturally being able to provide the cloud infrastructure with the necessary flexibility and manageability. The article examines some issues of the software-defined cloud platforms paying special attention to a number of network aspects from the existing terminological uncertainty regarding the data center infrastructures to some management features of the software-defined cloud networks and the role of the networking in an application-centric infrastructure.

**Keywords:** cloud platform; software-defined cloud; software-defined networking

**DOI:** 10.14357/08696527240109

**EDN:** CGZXIX

## References

1. Kranz, G. 2017. SDDC (software-defined data center). *TechTarget*. Available at: [https://www.techtarget.com/searchdatacenter/definition/SDDC-software-defined-data-center#:~:text=An%20SDDC%20\(software%2Ddefined%20data,configuration%20are%20abstracted%20from%20hardware](https://www.techtarget.com/searchdatacenter/definition/SDDC-software-defined-data-center#:~:text=An%20SDDC%20(software%2Ddefined%20data,configuration%20are%20abstracted%20from%20hardware) (accessed February 20, 2024).
2. Khoroshikh, D. 2014. SDDC — osnova novoy oblachnoy infrastruktury [SDDC is the basis of the new cloud infrastructure]. *Jet Info* 11:16–21. Available at: [https://www.jetinfo.ru/wp-content/uploads/2014/12/jetinfo\\_11\\_2014.pdf](https://www.jetinfo.ru/wp-content/uploads/2014/12/jetinfo_11_2014.pdf) (accessed February 20, 2024).
3. Robb, D. 2022. Top software-defined data center (SDDC) trends. *Datamation*. Available at: <https://www.datamation.com/data-center/software-defined-data-center-trends/> (accessed February 20, 2024).
4. Crockett, E. 2022. Top software-defined data center providers (SDDC). *Datamation*. Available at: <https://www.datamation.com/data-center/software-defined-data-center-providers/> (accessed February 20, 2024).
5. Abbasi, A. A., A. Abbasi, S. Shamshirband, A. T. Chronopoulos, V. Persico, and A. Pescapé. 2019. Software-defined cloud computing: A systematic review on latest trends and developments. *IEEE Access* 7:93294–93314. doi: 10.1109/ACCESS.2019.2927822.
6. Son, J., and R. Buyya. 2019. SDCon: Integrated control platform for software-defined clouds. *IEEE T. Parall. Distr.* 30(1):230–244. doi: 10.1109/TPDS.2018.2855119.
7. Kameshwar, L. 2023. Software defined networking: Connecting hybrid clouds seamlessly. *Wipro*. Available at: <https://www.wipro.com/infrastructure/software-defined-networking-connecting-hybrid-clouds-seamlessly> (accessed February 20, 2024).
8. Egorov, V. B. 2020. Nekotorye voprosy programmnoy opredeleniya tse ntrov obrabotki dannykh [Some issues of software-defined data centers]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 30(2):103–112. doi: 10.14357/08696527200210. EDN: FXXGJY.

9. Kak postroit' polnotsennoe gibridnoe oblako i snizit' stoimost' vladeniya IT-infrastrukturoy [How to build a full-fledged hybrid cloud and reduce the ownership cost of the IT-infrastructure]. *C·News*. Available at: [https://softline.cnews.ru/articles/2021-07-14\\_virtualnyj\\_tsod\\_i\\_gibridnoe\\_oblako](https://softline.cnews.ru/articles/2021-07-14_virtualnyj_tsod_i_gibridnoe_oblako) (accessed February 20, 2024).
10. Egorov, V. B. 2023. Nekotorye voprosy dezagregatsii i komponuemosti infrastruktury tsentra obrabotki dannykh [Some issues of disaggregation and composability of the data center infrastructure]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 33(2): 101–110. doi: 10.14357/08696527230210. EDN: OEXVDB.
11. Steen, T. 2020. Why software-defined solutions are key to a cloud-ready architecture. *SHI*. Available at: <https://blog.shi.com/next-generation-infrastructure/why-software-defined-solutions-are-key-to-a-cloud-ready-architecture> (accessed February 20, 2024).
12. Egorov, V. B. 2022. K voprosu o sootnoshenii programmno opredelyaemykh i traditsionnykh IP-setey [Interrelation between the software-defined and conventional IP-networks]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 32(1):73–82. doi: 10.14357/08696527220107. EDN: OGWOXL.

*Received October 5, 2023*

## **Contributor**

**Egorov Vladimir B.** (b. 1948)—Candidate of Science (PhD) in technology, leading scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; VEgorov@ipiran.ru

## МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ОЦЕНКИ ВЛИЯНИЯ ГЛОБАЛЬНЫХ ИЗМЕНЕНИЙ В МИРЕ НА ЦИФРОВОЙ СЛЕД РОССИЙСКИХ УНИВЕРСИТЕТОВ НА ПРИМЕРЕ АГРАРНЫХ\*

*В. И. Будзко<sup>1</sup>, В. И. Меденников<sup>2</sup>*

**Аннотация:** Рассматриваются вопросы определения эффективности российских университетов на примере аграрных по подготовке специалистов на базе математических моделей оценки их цифровых следов, оставленных на различных интернет-площадках. Дан анализ все более возрастающего значения этих следов на современном этапе развития цифровизации высшего образования. Показано, что при грамотной организации на основе интеграционных механизмов интернет-представительства учебных заведений как преподаватели, так и студенты будут обладать качественно новыми средствами для обмена образовательными и научными знаниями и идеями, а также будет обеспечиваться эффективный трансфер их в экономику, что приведет к росту эффективности использования интеллектуального потенциала страны. Продемонстрировано влияние в этом случае цифрового следа на конкурентоспособность образовательных организаций. Для расчета рейтингов на основе цифровых следов были разработаны математические модели, формализующие выделенные цифровые следы, сформированы на их основе методики, при помощи которых проведено сравнение эффективности использования университетами образовательных и научных информационных ресурсов. Исходной информацией данных расчетов послужили результаты проведенного мониторинга и анализа объемов и состояния научно-образовательных ресурсов, имеющихся на сайтах аграрных вузов. Показаны результаты проведенной оценки влияния глобальных изменений в мире в виде введенных санкций против России на цифровой след посредством сравнения рейтингов, полученных в 2022 г., с рейтингами, рассчитанными в 2016 г., и динамики объемов и качества рассмотренных ресурсов на указанных сайтах.

**Ключевые слова:** аграрные университеты; цифровой след; рейтинги; методика оценки

**DOI:** 10.14357/08696527240110

**EDN:** DLPCFL

---

\*Статья публикуется по представлению программного комитета XXV Международной конференции «Аналитика и управление данными в областях с интенсивным использованием данных» (DAMDID/RCDL 2023). Работа выполнена при финансовой поддержке гранта Министерства науки и высшего образования Российской Федерации, внутренний номер 00600/2020/51896, соглашение от 21.04.2022 № 075-15-2022-319.

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

<sup>2</sup>Федеральный исследовательский центр «Информатика и управление» Российской академии наук, dommed@mail.ru



## **1 Введение**

Ориентация российского научного и образовательного пространства на передовые западные разработки отрицательно повлияла на успешное развитие инновационной деятельности. Поэтому требуется соответствующее реформирование образования и науки, в том числе в аграрной инновационной сфере России. Для оценки эффективности деятельности организаций, представляющих это направление, необходимы научно-обоснованные методики. В условиях стремительного роста объемов доступных научных и образовательных ресурсов стали создаваться цифровые инструменты, которые обеспечивают эффективное извлечение необходимых знаний из этих ресурсов, прежде всего текстового формата, а также всесторонний обмен результатами исследований между учеными на основе унификации представлений этих знаний. В качестве примера можно привести работу [1].

Так, необходимость активного и глубокого анализа текущих достижений при цифровизации аграрных областей, в том числе в научно-образовательных учреждениях, рассмотрена при обосновании методики максимально эффективной организации процессов севооборота с применением адекватных математических моделей [2] и результатов исследований в области обеспечения необходимого уровня информационной безопасности применения средств цифровизации [3].

Цифровизация рыночных механизмов в развитых странах идет высокими темпами, например в виде проекта RePec [4]. Этому способствуют отработанные на протяжении длительного периода механизмы трансфера знаний в экономику и совершенствования образования.

В современных условиях при ограничении возможности применения западных достижений необходимо развивать собственные технологии с ускоренным внедрением их в экономику. Для оценки достигнутого уровня должны быть разработаны метрики оценки результативности деятельности научно-образовательных учреждений. В настоящее время в этой сфере в области образования страна ориентируется на распорядительные документы Министерства образования и науки, в частности приказ № 662 от 5.08.2013 «Об осуществлении мониторинга системы образования» [5]. Однобокость определенных в нем требований связана с учетом лишь образовательной деятельности. Но решение о выделении ресурсов на инновационную деятельность учреждениям в зависимости от уровня публикационной активности снижает эффективность вложений по сравнению с процедурой, принятой в развитых странах, где в качестве показателей принимаются результаты научных исследований.

В работе рассматривается подход к формированию метрики оценки научно-образовательных учреждений путем определения их рейтингов с помощью цифровых следов, позволяющих рассчитать и интегральные оценки эффективности использования ими информационных научно-образовательных ресурсов (ИНОР).



В период активной фазы пандемии в условиях самоизоляции на имидж университетов основное влияние стало оказывать грамотное и эффективное представительство университета в интернете — цифровой след. В настоящее время широко распространено мнение, что цифровой след может быть активным (АЦС) или пассивным (ПЦС). Первый содержит информацию, публикуемую на сайте ее владельцем или уполномоченным им лицом. Второй — информацию, собираемую без ведома ее владельца. Но по мере совершенствования интернет-технологий значительное развитие получили так называемые агрегаторы информации, освещающие многие стороны деятельности организаций с использованием независимых от владельца сайтов информационных источников. Будем называть такой цифровой след полуактивным (ПАЦС).

Необходимость введения понятия ПАЦС диктуется требованием формирования единого информационного пространства (цифровой платформы) ИНОР отрасли в соответствии с мировыми тенденциями цифровой трансформации, в частности заключающимися в интеграции информационных ресурсов, алгоритмов их обработки и инструментальной составляющей в виде программного обеспечения и электронного оборудования. Однако в настоящее время технологии разработки сайтов научно-исследовательских институтов и вузов аграрного профиля вступают в противоречия с данными тенденциями. Например, число видов программного обеспечения разработки сайтов вузов за 5 лет выросло на 66%. Такой неконтролируемый рост программного обеспечения с одновременным увеличением и его разнообразия без внятных рекомендаций со стороны Минобрнауки по его применению привел к тому, что исчезла возможность автоматически считывать данные с сайтов с последующим использованием их в других, растущих по экспоненте информационных системах, в частности и в самом Минобрнауки России, что лишает перспектив формирование единой цифровой платформы ИНОР.

Активный цифровой след университетов России опирается на информацию, содержащуюся на их сайтах и отражающую следующие виды ИНОР: нормативно-правовая информация, разработки, консультационная деятельность, публикации, пакеты прикладных программ, дистанционное обучение, тематические базы данных, электронная биржа труда, электронная торговая площадка [6]. Пассивный цифровой след содержит значимые выходные данные сервиса Site-Auditor — анализа и аудита сайтов [7], который разрабатывался для оценки имиджа организаций по различным рейтингам. Информация ПАЦС содержит данные из отчетов университетов о самообследовании, собираемую Минобрнауки [5].

Недооценка АЦС и ПЦС привела к тому, что не был исполнен указ Президента РФ от 2012 г., по которому планировалось, что более пяти университетов страны к 2020 г. должны были войти в первую сотню (топ-100) рейтингов ведущих университетов мира. Счетная палата в конце 2020 г. проанализировала результаты реализации указа, используя рейтинги QS [8], THE [9], ARWU [10], и заявила, что указ не был исполнен, поскольку ни одного российского универ-

ситета, которые финансировались по соответствующей программе, не оказалось в топ-100. При этом на данную программу были потрачены из бюджета 80 млрд руб. [11]. По указанным выше причинам такой результат можно было предсказать в самом начале реализации указанной программы [6]. Поэтому необходимо направить усилия на исследования по формированию рейтингов университетов, опираясь на комплекс их цифровых следов, интегральный показатель которых равносителен комплексной оценке эффективности использования ИНОР образовательными организациями. Полученная методика формирования рейтингов при таком подходе дает шанс сблизить российский подход в этом отношении с западным, что очень актуально в свете неизбежности в ближайшие годы перехода на единую цифровую платформу ИНОР, дополняющую требования Минобрнауки к оценке деятельности образовательных учреждений. Разработанные математические модели формирования цифровых следов образовательных учреждений позволяют рациональнее перераспределять ресурсы в целях повышения эффективности образования и науки в университетах.

В статье проводится анализ математических моделей, которые могут составить основу методик расчета рейтингов университетов на базе перечисленных выше цифровых следов, для чего потребовалось:

- (1) осуществить мониторинг объемов и состояния ИНОР на сайтах аграрных университетов в силу причин, указанных ниже;
- (2) конкретизировать понятие цифрового следа в отношении этих организаций с их классификацией;
- (3) разработать математические модели, формализующие выделенные цифровые следы;
- (4) на основе модели разработать методику формирования рейтингов для оценки эффективности использования университетами ИНОР;
- (5) провести оценку влияния глобальных изменений в мире на цифровой след посредством сравнения полученных рейтингов.

Выбор аграрных университетов объясняется несколькими причинами. Во-первых, представительством в большинстве регионов (53 вуза), что важно для применения статистических методов; во-вторых, в сельском хозяйстве развитых стран идет активная индустриализация производства продукции на основе цифровой трансформации, которая позволила значительно увеличить востребованность производством научных исследований, носивших до этого в большинстве случаев чисто теоретический характер. При этом данный процесс сопровождается требованием такого же ускоренного трансфера результатов исследований в бизнес, как и само развитие цифровой экономики. Чтобы поддержать данные темпы трансфера, в развитых в аграрном отношении странах пошли по пути создания соответствующих инновационных аграрных центров как интеграторов науки и производства. В-третьих, сельское хозяйство, имеющее более четверти всего

биологического разнообразия на нашей планете, представляется наиболее ярким представителем экосистем с огромным разнообразием природных факторов и биологических видов, что требует пристального внимания государства и научного сообщества из-за нарастающих угроз деградации земель и всей экосистемы Земли.

## 2 Методика формирования рейтингов университетов на основе пассивного цифрового следа

Для учета ПЦС сельскохозяйственных вузов при разработке методики формирования их рейтингов лучше всего подходит технология сайтометрии в виде наиболее известного сервиса Site-Auditor, предназначенного как раз для анализа и аудита сайтов [7]. Выходом сервиса служат показатели, представленные в табл. 1 с весами, полученными на основе работ [6, 12–14].

Значения показателей в классах были получены следующим образом:

- «Каталоги» — рассчитывались суммированием индексов присутствия (1 — есть; 0 — нет) сайтов в каталогах соответствующих сервисов: Mail.ru, Яндекс, DMOZ, Рамблер ТОП;
- «Индексация» — рассчитывались по данным поисковых сервисов Яндекс, Google, Bing и Seznam;
- «Проблемы» — рассчитывались суммированием индексов присутствия (1 — есть; 0 — нет) сайтов Яндекс.АГС и Spamhaus (IP);
- «Рейтинговые классы» — рассчитывались по данным сервисов (сайтов) Alexa (local и global), Яндекс и Google PR;
- «Социальные сервисы» — рассчитывались суммированием индикаторов в соцсетях Mail.ru (Мой мир), Facebook и GooglePlus;

**Таблица 1** Выходные показатели сервиса Site-Auditor

№	Перечень классов показателей	Число в классе	Вес
1	Каталоги	4	8%
2	Индексация	4	8%
3	Проблемы	2	5%
4	Рейтинг Alexa (local)	1	4%
5	Рейтинг Alexa (global)	1	4%
6	Рейтинг Яндекс (тИЦ)	1	8%
7	Рейтинг Google PR	1	8%
8	Социальные сервисы	3	5%
9	Ссылки с сайта	2	10%
10	Ссылки на сайт	4	40%
	<b>ИТОГО</b>	<b>23</b>	<b>100%</b>

- «Ссылки с сайта» — рассчитывались в виде средней математической величины показателей с сайтов Linkpad и Bung;
- «Ссылки на сайт» — рассчитывались аналогично ссылкам с сайта по данным сервисов Google, Alexa, Majestic и Linkpad.

Детальные расчеты содержатся в [6].

Целый ряд показателей из различных классов в количестве 17, выдаваемых сервисом Site-Auditor, не был принят во внимание в силу их незначительности и малой актуальности.

Формализуем учет ПЦС сельскохозяйственных вузов при разработке методики формирования их рейтингов:

- $P_2^m$  — индикатор оценки  $m$ -го вуза;
- $d_{rm}^2$  — величина  $r$ -го показателя оценки сайта  $m$ -го вуза;
- $q_{rm}^2$  — индикатор  $r$ -го показателя оценки сайта  $m$ -го вуза:

$$q_{rm}^2 = \frac{d_{rm}^2}{\max_m d_{rm}^2};$$

- $\omega_r^2$  — вес индикатора  $r$ -го показателя оценки сайта.

Определим  $P_2^m$ :

$$P_2^m = \sum_k \omega_k^2 q_{km}^2. \quad (1)$$

Выражение (1) позволило рассчитать 10 частичных рейтингов по данным сервиса Site-Auditor, которые в дальнейшем были просуммированы с весами из табл. 1. Результаты расчетов рейтингов вузов показаны далее в общей табл. 10.

### 3 Методика формирования рейтингов университетов на основе полуактивного цифрового следа

Для разработки методики формирования рейтингов аграрных вузов на основе ПАЦС воспользуемся данными из отчетов о самообследовании, представленными в [5]. Выборка показателей, относящихся к поставленной цели, приведена в табл. 2. Попытки определить веса показателей методами математической статистики не увенчались успехом, хотя был использован значительный набор наиболее известных из них: расчет коэффициентов конкордации Кендалла; корреляционный анализ; расчет матриц компетентности; вероятностная модель оценивания. Поэтому веса были взяты почти равными с небольшими колебаниями, связанными с округлением значений.

По аналогии с разд. 2 формализуем учет ПАЦС сельскохозяйственных вузов при разработке методики формирования их рейтингов:

- $P_5^m$  — индикатор оценки  $m$ -го вуза;
- $d_{hm}^5$  — величина  $h$ -го показателя оценки сайта  $m$ -го вуза;

**Таблица 2** Показатели ПАЦС

№	Показатель	Вес
1	Средний балл абитуриентов, принятых в вуз по результатам единого государственного экзамена на очное обучение	5,89%
2	Число цитирований в Web of Science, исходя из расчета на 100 НПП	5,87%
3	Число цитирований в Scopus, исходя из расчета на 100 НПП	5,89%
4	Число цитирований в РИНЦ, исходя из расчета на 100 НПП	5,87%
5	Число статей, индексируемых в Web of Science, исходя из расчета на 100 НПП	5,89%
6	Число статей, индексируемых в Scopus, исходя из расчета на 100 НПП	5,87%
7	Число публикаций в РИНЦ, исходя из расчета на 100 НПП	5,89%
8	Объем НИОКР, исходя из расчета на одного НПП	5,87%
9	Удельный вес доходов от НИОКР в суммарных доходах организации	5,89%
10	Удельный вес численности НПП с ученой степенью кандидата наук в общей численности НПП организации	5,87%
11	Удельный вес численности НПП с ученой степенью доктора наук в общей численности НПП организации	5,89%
12	Доходы организации по всем видам финансовой деятельности, исходя из расчета на одного НПП	5,87%
13	Доходы университета из средств от доходной деятельности в расчете на одного НПП	5,89%
14	Отношение средней зарплаты НПП в организации (по всем видам финансового обеспечения (деятельности)) к средней зарплате по экономике региона	5,87%
15	Общая площадь помещений, в которых осуществляется образовательная деятельность, исходя из расчета на одного студента	5,89%
16	Количество вычислительной техники, приходящейся на одного студента	5,87%
17	Удельный вес численности студентов, живущих в общежитиях, к общей численности нуждающихся в них	5,89%

**Примечания:** НПП — научно-педагогический работник; НИОКР — научно-исследовательские и опытно-конструкторские работы.

–  $q_{hm}^5$  — индикатор  $h$ -го показателя оценки сайта  $m$ -го вуза:

$$q_{hm}^5 = \frac{d_{hm}^5}{\max_m d_{hm}^5};$$

–  $\omega_{hm}^5$  — вес индикатора  $h$ -го показателя оценки сайта  $m$ -го вуза.

Тогда получим

$$P_5^m = \sum_h \omega_h^5 q_{hm}^5.$$

Результаты расчетов рейтингов вузов показаны далее в общей табл. 10.

#### 4 Методика формирования рейтингов университетов на основе активного цифрового следа

Роль АЦС в условиях возрастающего влияния интернета на научно-образовательную деятельность приобретает с каждым годом все более существенное значение, чем просто имидж университета. При грамотном следовании интеграционным тенденциям цифровой трансформации экономики в интернете АЦС обеспечивает качественно новые возможности как для широкого обмена образовательными и научными знаниями и идеями, так и для эффективного трансфера их в экономику и роста интеллектуального потенциала общества за счет совершенствования системы образования [6]. Такое научно-обоснованное представительство университетов в интернете наиболее полно будет реализовано при создании и внедрении, как уже отмечалось выше, единой цифровой платформы ИНОР, интегрирующей все ее компоненты, представленные на сайтах научных и образовательных организаций. Наиболее важные виды ИНОР, такие как публикации, разработки, материалы консультационной деятельности, дистанционное обучение, нормативно-правовая информация, базы данных, пакеты прикладных программ, наиболее востребованы в аграрной сфере [6, 15].

Кроме ИНОР на сайтах образовательных организаций по мере совершенствования цифровой экономики начинают появляться зачатки электронной торговой площадки и электронной биржи труда, следуя современным веяниям оказания в онлайн-режиме услуг в виде подобных сервисов. Поэтому помимо ИНОР к АЦС отнесем также и сервисы электронной торговой площадки и электронной биржи труда. А поскольку все информационные ресурсы на сайтах вузов размещаются в неоднородном виде (например, электронная торговая площадка и электронная биржа труда представлены в широком диапазоне от простейших досок объявлений до полнофункциональных площадок и бирж), то их требовалось привести к определенному однородному виду средствами онтологического моделирования, единого словаря параметров, единых форматов представления данных и единой системы классификаторов. Это обеспечило разработку типовой методики оценки как АЦС, так и результатов научно-образовательной деятельности аграрных вузов с формированием их рейтингов на основе данного цифрового следа. Для получения объективных данных о качестве и объемах информационных ресурсов на соответствующих сайтах был проведен их мониторинг на основе сформированной анкеты, включающей свыше двухсот показателей образовательного и научного характера деятельности вузов (свыше 120 из них относятся к оценке самой организации; 40 — к оценке факультетов; 46 — к оценке деятельности кафедр; 6 — к общей оценке непосредственно сайта).

Веса показателей оценки  $m$ -го вуза получены экспертным путем посредством анализа результатов исследований специалистов в области образования [6, 12, 13], анализа существующих методик формирования различных рейтингов образовательных учреждений, путем анкетирования научно-педагогических работников

Тимирязевской академии, а также применением соответствующих статистических подходов.

Однородные информационные ресурсы приобретали на основе учета возможностей хранения контента сайтов в базах данных под управлением некоторой системы управления базами данных, предоставляемых провайдерами, в которых информационные ресурсы хранятся в форме каталогов и полноформатном представлении в упорядоченном виде, когда они дают возможность навигации по организациям, авторам, тематической рубрикации, ключевым словам и т. п. При отсутствии же баз данных обычно и каталоги, и полноформатное представление на сайтах хранятся в неупорядоченном виде.

Интегральная оценка университета по АЦС формулируется как сумма взвешенных классов оценок информационных ресурсов, сумма весов которых равна 1: видов представления ИНОР, электронной торговой площадки, электронной биржи труда. На этом основании опишем параметры методики (табл. 3–8).

Формализуем учет АЦС сельскохозяйственных вузов при разработке методики формирования их рейтингов:

**Таблица 3** Индикаторы степени интеграции информационных ресурсов

№	Индикатор	Вес
1	Неупорядоченное представление (список)	0,1
2	Упорядоченное представление	0,9
	ИТОГО	1,0

**Таблица 4** Индикаторы форм хранения информационных ресурсов

№	Индикатор	Вес
1	Каталожная форма	0,3
2	Полноформатная форма	0,7
	ИТОГО	1,0

**Таблица 5** Индикаторы видов ИНОР

№	Индикатор	Вес
1	Разработки	0,3
2	Публикации	0,2
3	Базы данных	0,05
4	Пакеты прикладных программ	0,05
5	Дистанционное обучение	0,05
6	Консультационная деятельность	0,3
7	Нормативно-правовая информация	0,05
	ИТОГО	1,0

**Таблица 6** Индикаторы оценки электронной торговой площадки

№	Индикатор	Вес
1	Доска объявлений в неструктурированном формате	0,05
2	Доска объявлений в структурированном формате	0,1
3	Поиск торгового партнера по заданному критерию	0,2
4	Интеграция информационных процессов всех торговых операций	0,25
5	Полная автоматизация электронной торговли	0,4
	ИТОГО	1,0

**Таблица 7** Индикаторы оценки электронной биржи труда

№	Индикатор	Вес
1	Доска объявлений в неструктурированном формате	0,1
2	Доска объявлений в структурированном формате	0,2
3	Автоматизированная электронная биржа труда (автоматический поиск)	0,6
4	Наличие ссылок на другие электронные биржи труда	0,1
	ИТОГО	1,0

**Таблица 8** Частные оценки АЦС

№	Индикатор	Вес
1	Индикаторы видов ИНОР	0,7
2	Индикаторы оценки электронной торговой площадки	0,15
3	Индикаторы оценки электронной биржи труда	0,15
	ИТОГО	1,0

- $i$  — индикатор степени интеграции информационных ресурсов,  $i \in M$  (см. табл. 3);
- $l$  — индикатор форм хранения информационных ресурсов,  $l \in L$  (см. табл. 4);
- $n$  — индикатор видов ИНОР,  $n \in N$  (см. табл. 5);
- $m$  — номер вуза,  $m \in M$ ;
- $P_j^m$  — частная  $j$ -я оценка АЦС  $m$ -го вуза,  $j \in J$  (см. табл. 8);
- $P^m$  — интегральная оценка АЦС  $m$ -го вуза;
- $\alpha_i^1$  — вес  $i$ -го индикатора степени интеграции информационных ресурсов;
- $\alpha_l^2$  — вес  $l$ -го индикатора форм хранения информационных ресурсов;
- $\alpha_n^3$  — вес  $n$ -го индикатора видов ИНОР;
- $\beta_j$  — вес  $j$ -й частной оценки АЦС (см. табл. 8);
- $v_{ilm}^m$  — объем ИНОР  $i$ -го индикатора степени интеграции информационных ресурсов,  $l$ -го индикатора форм хранения информационных ресурсов,  $n$ -го индикатора видов ИНОР  $m$ -го вуза;



- $\lambda_{iln}^m$  — соответственно оценка ИНОР  $i$ -го индикатора степени интеграции информационных ресурсов,  $l$ -го индикатора форм хранения информационных ресурсов,  $n$ -го индикатора видов ИНОР  $m$ -го вуза:

$$\lambda_{iln}^m = \frac{v_{iln0}^m}{\max_m v_{iln0}^m};$$

- $d_{sm}^3$  — оценка сайта по электронной торговой площадке  $s$ -го индикатора  $m$ -го вуза (см. табл. 6);
- $\omega_s^3$  — вес  $s$ -го индикатора оценки сайта по электронной торговой площадке (см. табл. 6);
- $d_{gm}^4$  — оценка сайта по электронной бирже труда  $g$ -го индикатора  $m$ -го вуза (см. табл. 6);
- $\omega_s^4$  — вес  $g$ -го индикатора оценки сайта по электронной бирже труда (см. табл. 7).

В результате получим:

$$P^m = \sum_j P_j^m,$$

где

$$P_1^m = \sum_{i,l,n} \lambda_{iln}^m \alpha_i^1 \alpha_l^2 \alpha_n^3; \quad P_3^m = \sum_s \omega_s^3 d_{sm}^3.$$

## 5 Анализ цифровых следов и результаты расчетов рейтингов аграрных вузов

Анализ результатов мониторинга сайтов аграрных вузов подтвердил упомянутый в начале статьи тезис о том, что в нашей стране имеется недооценка АЦС и ПЦС, что нашло отражение в табл. 9, в которой приведены объемные и качественные характеристики ИНОР, свидетельствующие о довольно низкой

**Таблица 9** Объем и состояние ИНОР на сайтах аграрных вузов

Виды ИНОР	И1	И2	И3	И4	И5
Разработки	85%	3685/2791	392/0	338/566	249/0
Публикации	89%	18650/14033	409/55081	345/344	1/21114
Базы данных	11%	531/0	46/0	0/0	0/0
Пакеты прикладных программ	2%	829/20	3/0	26/0	1/0
Дистанционное обучение	12%	1196/11	0/0	0/0	4/0
Консультации	25%	217/75	44/0	8/0	1/0
Нормативно-правовая информация	89%	66/1971	409/0	329/91	1/0

информативности исследуемых ресурсов на сайтах. В табл. 9 индексы столбцов несут следующую информацию: И1 — процент сайтов, на которых имелись данные ИНОР в 2022 г.; И2 — объем ресурсов в виде неупорядоченного представления; И3 — объем ресурсов в виде упорядоченного каталога; И4 — объем ресурсов в виде неупорядоченного полноформатного представления; И5 — объем ресурсов в виде упорядоченного полноформатного представления. Объемные и качественные характеристики ИНОР, имеющиеся на уровне вуза, приведены через косую черту за 2016 и 2022 гг. Рост числа публикаций на фоне резкого снижения качества и количества остальных ресурсов связан с влиянием пандемии на структуру форм хранения учебной литературы. В период пандемии на рынке было предложено большое число электронно-библиотечных систем, которые стали востребованы образовательными организациями: Руслан, ИРБИС64+, Сетевая электронная библиотека аграрных вузов, Znanium, ИРБИС64, IPRbooks, МегаПро, Марк Web, Буки, Юрайт, AgriLib и еще целая линейка менее значимых. Соответственно, это отразилось на резком росте числа публикаций в виде упорядоченного каталога (до 55 082 экз.) и в форме полноформатного представления их (до 21 114). Такая большая разница в числах объясняется фактом поддержки многими электронно-библиотечными системами лишь упорядоченных каталогов. Заметим, что, поскольку в представленных электронно-библиотечных системах собраны коллекции публикаций многих организаций, данные о числе этого вида ИНОР на сайте не могут характеризовать в должной мере АЦС конкретного вуза. При этом наблюдается снижение количества и качества публикаций в доминирующих до пандемии формах с появлением ограничений доступа к ним в силу появления паролей.

Для оценки влияния глобальных изменений в мире на цифровой след аграрных университетов в 2022 г., как уже отмечалось, был проведен еще один мониторинг их сайтов. В табл. 10 приводятся их рейтинги по АЦС, ПЦС, ПАЦСи интегральный рейтинг (IR) за 2016 и 2022 гг.

**Таблица 10** Рейтинги аграрных вузов (2016/2022 гг.)

Вуз	Статус	АЦС	ПЦС	ПАЦС	IR
Алтайский	ГАУ	22,5/25	8/2	27,5/35	24/22
Арктический	ГАТУ	25/18	44/54	52/49	33/50
Башкирский	ГАУ	33/14	19/19	7/7	21/10
Белгородский	ГАУ	7/8	41/8	10/13	4/7
Брянский	ГАУ	20/13	21/16	2/18	10/9
Великолукская	ГСХА	10/20	40/37	27,5/44	15/30
Волгоградский	ГАУ	13/5	35/39	23/48	9/16
Вологодская	ГМХА	11/22	18/11	38/12	12/19
Воронежский	ГАУ	54/42	9/20	39/23	49,5/46
Вятский	ГАТУ	6/31	33/33	30/24	14/39

Окончание табл. 10 на с. 123

**Таблица 10** (окончание) Рейтинги аграрных вузов (2016/2022 гг.)

Вуз	Статус	АЦС	ПЦС	ПАЦС	IR
Северного Зауралья	ГАУ	50/54	31/27	49/15	53/48
Горский	ГАУ	31/7	52/52	37/45	42/20
ГУЗ	ГАУ	35/50	29/13	21/5	35/36
Дагестанский	ГАУ	46/46	54/51	35/27	49,5/53
Дальневосточный	ГАУ	27/9	27/43	44/26	32/23
Донской	ГАУ	32/28	11/21	4/19	27/27
Ивановская	ГСХА	39/38	43/47	12/37	40/47
Иркутский	ГАУ	18/24	32/39	50/46,5	29/45
Казанская	ГАВМ	36/12	53/55	17/52	34/42
Казанский	ГАУ	8/10	28/30	15/17	6/28
Кабардино-Балкарский	ГАУ	48/23	49/46	26/29	47/34
Костромская	ГСХА	30/27	47/35	51/39	39/41
Красноярский	ГАУ	2/21	22/23	32/6	3/12
Кубанский	ГАУ	1/40	10/18	47/46,5	5/49
Кузбасская	ГСХА	3/35	12/17	22/14	7/35
Курганская	ГСХА	22,5/39	48/41	31/42	26/38
Курская	ГСХА	19/29	34/12	48/20	25/11
Мичуринский	ГАУ	16/45	7/29	29/10	13/43
Московская	ГАВМБ	43/49	30/5	1/3	31/26
Нижегородская	ГСХА	17/52	24/22	20/34	22/51
Новосибирский	ГАУ	28/32	4/10	18/21	17/6
Омский	ГАУ	24/51	5/9	8/8	19/44
Оренбургский	ГАУ	26/30	13/4	45/36	30/21
Орловский	ГАУ	5/3	20/40	14/52	2/18
Пензенский	ГАУ	14/26	46/45	6/41	11/33
Пермский	ГАТУ	12/15	26/24	33/30	16/17
Приморская	ГСХА	37/33	25/34	34/25	37/24
РГАЗУ	ГАУ	52/41	2/26	53/33	51/37
РГАУ-МСХА	ГАУ	4/16	1/6	16/2	1/2
Рязанский	ГАТУ	29/1	37/31	36/4	36/4
Самарский	ГАУ	40/44	15/7	46/28	41/32
Санкт-Петербургская	ГАВМ	53/47	45/32	25/52	52/52
Санкт-Петербургский	ГАУ	34/11	3/3	19/31	25/5
Саратовский	ГАУ	9/2	6/1	11/16	8/3
Смоленская	ГСХА	51/53	50/48	43/52	54/54
Ставропольский	ГАУ	38/4	14/15	9/1	28/1
Тверская	ГСХА	44/36	51/50	54/32	48/40
Удмуртский	ГАУ	45/6	42/28	13/38	45/14
Ульяновский	ГАУ	42/48	23/49	5/11	38/31
Уральский	ГАУ	47/34	36/44	3/9	44/8
Чувашский	ГАУ	21/43	38/36	24/22	20/15
Южно-Уральский	ГАУ	49/37	16/42	42/52	46/29
Ярославская	ГСХА	41/17	39/25	41/40	43/13

Представленные в табл. 10 данные показывают, что Ставропольский ГАУ, Рязанский ГАТУ и Удмуртский ГАУ значительно продвинулись в интегральном рейтинге. А Ставропольский ГАУ потеснил даже Тимирязевскую академию с первого места и при незначительном снижении общего числа разработок увеличил число публикаций с 594 до 6327, нормативно-правовой информации — с 0 до 222, пакетов прикладных программ — с 0 до 50, баз данных — с 0 до 5, а Кубанский ГАУ, входивший в лидирующую пятерку рейтинга 2016 г., снизил число разработок с 2196 до 44, публикаций — с 8475 до 4249, нормативно-правовой информации — с 37 до 0, пакетов прикладных программ — с 874 до 0, баз данных — с 538 до 0.

## 6 Заключение

Из материалов мониторинга сайтов и из табл. 10 следует, что в настоящее время вузы развивают свои сайты, ориентируясь лишь на нормативные указания Минобрнауки, включая в их контент в основном лишь ресурсы, требуемые регулирующими организациями и оставляющие в интернет-пространстве в значительной степени ППС. Процедуры совершенствования цифровых инструментов, влияющих на иные следы, отданы на усмотрение вузов. Финансирование лишь публикационной активности привело к тому, что социальный заказ на цифровые инструменты трансфера ИНОР в экономику, связанный с санкциями против России, реализуется слабо.

## Литература

1. *Девяткин Д. А.* Система распределенного построения случайных лесов деревьев решений с линейными и нелинейными разделителями // Системы высокой доступности, 2022. Т. 18. № 3. С. 59–68. doi: 10.18127/j20729472-202203-05. EDN: QSTZWW.
2. *Будзко В. И., Меденников В. И.* Математическая модель оптимизации структуры севооборотов на основе единой цифровой платформы управления сельскохозяйственным производством // Системы высокой доступности, 2022. Т. 18. № 4. С. 5–13. doi: 10.18127/j20729472-202204-01. EDN: RPXYGA.
3. *Будзко В. И.* Проблемы цифровой трансформации агротехнологических процессов // Системы высокой доступности, 2021. Т. 17. № 4. С. 5–23. doi: 10.18127/j20729472-202104-01. EDN: LEDOZI.
4. Research papers in economics. <http://repec.org>.
5. Об утверждении Порядка проведения самообследования образовательной организацией: Приказ Министерства образования и науки РФ от 14 июня 2013 г. № 462. <https://www.edu.ru/file/docs/2013/06/56145.pdf>.
6. *Меденников В. И., Сальников С. Г., Личман А. А., Муратова Л. Г., Горбачев М. И., Тухина Н. Ю.* Методика оценки эффективности использования информационных научно-образовательных ресурсов. — М.: Аналитик, 2017. 250 с. EDN: XUVTNP.

7. Site-Auditor. <https://freesoft.ru/windows/siteauditor>.
8. QS world university & business school rankings. <https://www.qs.com/rankings>.
9. World university rankings. <https://www.timeshighereducation.com/world-university-rankings>.
10. Academic ranking of world universities. <https://www.shanghairanking.com/rankings/arwu/2023>.
11. Вузы из проекта «5–100» так и не вошли в топ-100 международных рейтингов. [https://www.rbc.ru/society/18/02/2021/602cbdf9a7947765cbb58e5?from=from\\_main\\_10](https://www.rbc.ru/society/18/02/2021/602cbdf9a7947765cbb58e5?from=from_main_10).
12. *Сироткин Г. В.* Системный анализ факторов качества образования в вузе // Прикаспийский журнал: управление и высокие технологии, 2013. № 2(22). С. 109–118. EDN: QJENET.
13. *Гузаева М. Ю.* Использование информационных ресурсов науки и образования для повышения эффективности реализации новых форм обучения. <http://pedsovet.su/publ/164-1-0-1048>.
14. *Антюхов А. В., Фомин Н. В.* Разработка фонда оценочных средств в контексте ФГОС ВПО // Вестник Брянского государственного университета, 2024. № 1. С. 112–117. EDN: TLFARH.
15. *Medennikov V., Raikov A.* Creating the requirements to the national platform “Digital Agriculture” // CEUR Workshop Procee., 2020. Vol. 2763. P. 13–18. doi: 10.30987/conferencearticle\_5fce27715a3742.47428784. EDN: NXHNTW.

*Поступила в редакцию 28.12.23*

---

## MATHEMATICAL MODELING FOR ASSESSING THE IMPACT OF GLOBAL CHANGES IN THE WORLD ON THE DIGITAL FOOTPRINT OF RUSSIAN AGRICULTURAL UNIVERSITIES

*V. I. Budzko and V. I. Medennikov*

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation

**Abstract:** Issues of determining the effectiveness of training specialists at agricultural Russian universities using mathematical models for assessing their digital traces left on various Internet sites are considered. The increasing importance of these traces at the present stage of development of higher education digitalization has been determined. It is shown that teachers and students will have qualitatively new means for exchanging educational and scientific knowledge and ideas and their effective transfer to the economy will be ensured based on the integration mechanisms of the Internet representation of educational institutions with their competent organization. This will increase the efficiency of using the country’s intellectual potential. The impact of the digital footprint on the competitiveness of educational organizations is shown. To calculate ratings

based on digital traces, mathematical models that formalize the identified digital footprints were developed and methods were formed on their basis which were used to compare the effectiveness of the use of educational and scientific information resources by universities. The impact of the imposed sanctions against Russia on the digital footprint was assessed by comparing the ratings obtained in 2022 with the ratings calculated in 2016. The dynamics of changes in the volume and quality of resources presented on these sites are given.

**Keywords:** agricultural universities; digital footprint; ratings; assessment methodology

**DOI:** 10.14357/08696527240110

**EDN:** DLPCFL

## Acknowledgments

The paper is published on the proposal of the Program Committee of the 25th International Conference on Data Analytics and Management in Data Intensive Domains (DAMDID/RCDL 2023). This work was supported by the grant from the Ministry of Science and Higher Education of the Russian Federation, internal number 00600/2020/51896, Agreement dated 21.04.2022 No. 075-15-2022-319.

## References

1. Devyatkin, D. A. 2022. Sistema raspredelennogo postroeniya sluchaynykh lesov derev'ev resheniy s lineynymi i nelineynymi razdelitelyami [System for distributed training of Kernel forests]. *Sistemy vysokoy dostupnosti* [Highly Available Systems] 18(3):59–68. doi: 10.18127/j20729472-202203-05. EDN: QCTZWW.
2. Budzko, V. I., and V. I. Medennikov. 2022. Matematicheskaya model' optimizatsii struktury sevooborotov na osnove edinoy tsifrovoy platformy upravleniya sel'skokhozyaystvennym proizvodstvom [Mathematical model for optimizing the structure of cropping rotations on the basis of a unified digital platform for agricultural production management]. *Sistemy vysokoy dostupnosti* [Highly Available Systems] 18(4):5–15. doi: 10.18127/j20729472-202204-01. EDN: RPXYGA.
3. Budzko, V. I. 2021. Problemy tsifrovoy transformatsii agrotekhnologicheskikh protsessov [Problems of digital transformation of agrotechnological processes]. *Sistemy vysokoy dostupnosti* [Highly Available Systems] 17(4):5–23. doi: 10.18127/j20729472-202104-01. EDN: LEDOZI.
4. Research papers in economics. Available at: <http://repec.org> (accessed March 5, 2024).
5. Ob utverzhdenii Poryadka provedeniya samoobsledovaniya obrazovatel'noy organizatsiy: Prikaz Ministerstva obrazovaniya i nauki RF ot 14 iyunya 2013 g. No. 462 [On approval of the Procedure for conducting self-examination by an educational organization: Order of the Ministry of Education and Science of the Russian Federation No. 462 dated 14.06.2013]. Available at: <https://www.edu.ru/file/docs/2013/06/56145.pdf> (accessed March 5, 2024).
6. Medennikov, V. I., S. G. Salnikov, A. A. Lichman, L. G. Muratova, M. I. Gorbachev, and N. Yu. Tukhina. 2017. Metodika otsenki effektivnosti ispol'zovaniya informatsionnykh nauchno-obrazovatel'nykh resursov [Methodology for assessing the effectiveness

- of using information scientific and educational resources]. Moscow: Analitik. 250 p. EDN: XUYPNP.
7. Site-Auditor. Available at: <https://freesoft.ru/windows/siteauditor> (accessed March 5, 2024).
  8. QS World university & business school rankings. Available at: <https://www.qs.com/rankings> (accessed March 5, 2024).
  9. World university rankings. Available at: <https://www.timeshighereducation.com/world-university-rankings> (accessed March 5, 2024).
  10. Academic ranking of world universities. Available at: <https://www.shanghairanking.com/rankings/arwu/2023> (accessed March 5, 2024).
  11. Vuzy iz proyekta “5–100” tak i ne voshli v top-100 mezhdunarodnykh reytingov [Universities from the “5–100” project were not included in the top 100 international rankings]. Available at: [https://www.rbc.ru/society/18/02/2021/602cbdff9a7947765cbb58e5?from=from\\_main\\_10](https://www.rbc.ru/society/18/02/2021/602cbdff9a7947765cbb58e5?from=from_main_10) (accessed March 5, 2024).
  12. Sirotkin, G. V. 2013. Sistemnyy analiz faktorov kachestva obrazovaniya v vuze [System analysis of factors of quality education in institutes of higher education]. *Prikaspiyskiy zhurnal: upravlenie i vysokie tekhnologii* [Caspian J.: Management and High Technologies] 2(22):109–118. EDN: QJEHET.
  13. Guzaeva, M. Yu. 2010. Ispol’zovanie informatsionnykh resursov nauki i obrazovaniya dlya povysheniya effektivnosti realizatsii novykh form obucheniya [Using information resources of science and education to improve the efficiency of implementing new forms of education]. Available at: <http://pedsosvet.su/publ/164-1-0-1048> (accessed March 5, 2024).
  14. Antyukhov, A. V., and N. V. Fomin. 2014. Razrabotka fonda otsenochnykh sredstv v kontekste FGOS VPO [Elaboration of funds of appraisal means in the context of federal state educational standards of higher professional education]. *Vestnik Bryanskogo gosudarstvennogo universiteta* [The Bryansk state university herald] 1:112–117. EDN: TLFARH.
  15. Medennikov, V. I., and A. N. Raikov. 2020. Creating the requirements to the national platform “Digital Agriculture.” *CEUR Workshop Procee.* 2763:13–18. doi: 10.30987/conferencearticle.5fce27715a3742.47428784. EDN: NXHNTW.

*Received December 28, 2023*

## Contributors

**Budzko Vladimir I.** (b. 1944) — Doctor of Science in technology, Academician of the Academy of Cryptography of the Russian Federation, principal scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; vbudzko@ipiran.ru

**Medennikov Viktor I.** (b. 1949) — Doctor of Science in technology, leading scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; dommed@mail.ru

## КОЛЛЕКТИВНОЕ РАЗРЕШЕНИЕ СУЩНОСТЕЙ В ТЕХНОЛОГИИ ПОДДЕРЖКИ КОНКРЕТНО-ИСТОРИЧЕСКИХ ИССЛЕДОВАНИЙ

*И. М. Адамович<sup>1</sup>, О. И. Волков<sup>2</sup>*

**Аннотация:** Статья посвящена дальнейшему развитию распределенной технологии поддержки конкретно-исторических исследований (ПКИИ), основанной на принципах краудсорсинга и ориентированной на широкий круг не относящихся к профессиональным историкам и биографам пользователей. Развитие осуществляется за счет включения в технологию алгоритма разрешения сущностей при обработке номинативных документов, осуществляющего коллективное разрешение, при котором сущности для совпадающих ссылок определяются совместно, а не независимо, и представляющего собой модификацию жадного алгоритма агломеративной кластеризации. Приведены подробное описание подхода, лежащего в основе алгоритма, и его высокоуровневый псевдокод. Представлен анализ его эффективности на данных с различной степенью неоднозначности имен, оценена степень неоднозначности имен конкретно-исторических данных. Сделан вывод о целесообразности включения алгоритма в технологию. Намечены направления дальнейших исследований по определению настраиваемых параметров алгоритма.

**Ключевые слова:** конкретно-историческое исследование; распределенная технология; разрешение сущностей; жадный алгоритм; мера реляционного сходства

DOI: 10.14357/08696527240111

EDN: GQPGXQ

### 1 Введение

Информатизация общества охватывает все сферы общественной жизни. Через внедрение информационных технологий претерпевает изменения и научная деятельность. Данная статья посвящена развитию программных средств информационно-аналитической поддержки такого направления научной деятельности, как конкретно-исторические исследования. Значимость такой поддержки на современном этапе обусловлена вовлечением в исследовательский процесс не только членов профессионального исторического сообщества, но и самых широких слоев непрофессионалов в связи со все возрастающим интересом к частной, семейной истории.

---

<sup>1</sup>Федеральный исследовательский центр «Информатика и управление» Российской академии наук, Adam@amsd.com

<sup>2</sup>Федеральный исследовательский центр «Информатика и управление» Российской академии наук, Volkov@amsd.com



В основе любого конкретно-исторического исследования лежит извлечение информации из источников. В [1] приведена классификация исторических источников, используемых в историческом исследовании. Наибольшую ценность для конкретно-исторического исследования представляют источники, содержащие персональную информацию, включая имя персоны (так называемые номинативные источники). В свою очередь, они подразделяются на источники персональной структуры и источники семейной структуры [2].

Одна из наиболее сложных задач, возникающих при обработке номинативных источников, — связывание записей (Record Linkage), также называемое задачей сопоставления данных, связывания данных и задачей разрешения сущностей. Эта задача состоит в поиске в различных источниках записей, которые относятся к одной персоне или одной сущности. Сложность этой задачи объясняется рядом негативных особенностей исторических документов, таких как пропуски, неточности, различные способы написания имен, отсутствие фамилий у крестьян вплоть до второй половины XIX в., частое неуказание персональной информации для женщин и детей и т. п. Но, несмотря на эти особенности, существуют и используются подходы к решению задачи автоматизированного связывания записей в номинативных источниках при проведении конкретно-исторических исследований.

Первый подход заключается в применении метода восстановления истории семей (ВИС), разработанного для задач исторической демографии [1]. Данный метод заключается в поатрибутном сравнении записей с помощью вероятностных алгоритмов. Наиболее успешно метод ВИС применяется при обработке метрических записей и подобных персональных источников демографических событий. Но при обработке источников семейной структуры метод ВИС сталкивается с серьезными проблемами [3].

Второй подход опирается на наличие в записях исторического регистра дополнительной реляционной информации, что повышает точность связывания записей за счет совместного, а не независимого определения сущностей для совпадающих ссылок. Для нужд конкретно-исторических исследований данный подход был реализован в рамках технологии ПККИИ [4, 5].

В [6] для технологии ПККИИ была предложена надстройка для методов ВИС, опирающаяся помимо сходства атрибутов также и на сходство семейных структур. Недостаток данной реализации подхода заключается в опоре не на весь объем реляционной информации, а на сравнение двух семейных структур, относящихся к двум номинативным документам семейного типа. Отсюда вытекает потребность в развитии этого подхода.

В [7] предложен более универсальный подход к коллективному разрешению сущностей, основанный на новом алгоритме реляционной кластеризации, который использует как атрибутивную, так и реляционную информацию. Целью данной статьи ставится исследование соответствия предложенного алгоритма специфике конкретно-исторических данных.

## 2 Описание подхода

В [7] вводятся следующие обозначения: в задаче разрешения сущностей дан набор ссылок  $R = \{r_i\}$ , где каждая ссылка  $r$  имеет атрибуты  $r.A_1, r.A_2, \dots, r.A_k$ . Ссылки соответствуют некоторому набору неизвестных сущностей  $E = \{e_i\}$ .  $r.E$  обозначает объект, которому соответствует ссылка  $r$ . Задача состоит в том, чтобы восстановить скрытый набор сущностей  $E = \{e_i\}$  и метки сущностей  $r.E$  для отдельных ссылок, учитывая наблюдаемые атрибуты ссылок. Ссылки наблюдаются не независимо, а объединенными в группы. Можно рассмотреть гиперграф, вершины которого представляют собой ссылки. Тогда эти группы ссылок соответствуют его гиперребрам. Набор гиперребер обозначается как  $H = \{h_i\}$ . Каждое гиперребро  $h$  также может иметь атрибуты, которые обозначаются  $h.A_1, h.A_2, \dots, h.A_l$ , а  $h.R$  обозначает набор ссылок, которые оно соединяет. Ссылка  $r$  может принадлежать нескольким гиперребрам, и  $r.H$  используется для обозначения набора гиперребер, в котором участвует  $r$ .

Применительно к конкретно-историческим данным гиперребро соответствует одной записи в источнике номинативного типа. Упоминаемые в записи персоны соответствуют ссылкам. Сущности представляют собой набор исторических личностей, упоминаемых в исследуемых источниках.

Суть подхода к коллективному разрешению сущностей состоит в том, что решения не принимаются независимо, но вместо этого одно решение по разрешению сущностей оказывает влияние на другие решения за счет связанности сущностей посредством гиперребер.

В основе подхода лежит разбиение ссылок на кластеры по признаку соответствия этих ссылок одной и той же сущности. Подход использует жадный алгоритм (greedy algorithm) агломеративной кластеризации, где на любом этапе текущий набор  $C = \{c_i\}$  кластеров ссылок отражает текущее представление о соответствии ссылок сущностям. Метка кластера для ссылки  $r$  обозначается как  $r.C$ .

Мера сходства  $\text{sim}$  двух кластеров  $c_i$  и  $c_j$  определяется следующим образом:

$$\text{sim}(c_i, c_j) = (1 - \alpha) \text{sim}_A(c_i, c_j) + \alpha \text{sim}_R(c_i, c_j), \quad 0 \leq \alpha \leq 1, \quad (1)$$

где  $\text{sim}_A()$  — мера сходства на основе атрибутов;  $\text{sim}_R()$  — мера реляционного сходства.

Мера сходства двух ссылок на основе атрибутов  $\text{sim}_A(r_i, r_j)$  широко используется в рамках метода восстановления истории семей [1]. В рамках традиционного подхода разработаны различные меры для разных типов атрибутов. Общая мера вычисляется как их взвешенная сумма. Мера сходства двух кластеров  $c_i$  и  $c_j$  на основе атрибутов служит обобщением для меры сходства двух ссылок. Определяем сходство  $\text{sim}_A(c_i, c_j)$  между двумя кластерами  $c_i$  и  $c_j$  как наилучшее попарное соответствие атрибутов между их ссылками. Поскольку ссылки в любом кластере не упорядочены, каждая ссылка  $r \in c_i$  может быть сопоставлена любой

ссылке  $r' \in c_j$ . Итак, для каждой ссылки  $r \in c_i$  находим наилучшее соответствие в  $c_j$ :

$$\text{sim}_A(r, c_j) = \max_{r' \in c_j} \text{sim}_A(r, r').$$

Для симметрии также вычисляется наилучшее соответствие кластеру  $c_i$  для каждой ссылки в  $c_j$ , а затем берется среднее значение по всем ссылкам в двух кластерах, чтобы получить  $\text{sim}_A(c_i, c_j)$ .

Не относится к традиционным и требует определения мера реляционного сходства между двумя кластерами ссылок  $\text{sim}_R$ .

### 3 Меры реляционного сходства кластеров

Для определения меры реляционного сходства для каждого кластера следует учитывать ссылки, связанные с помощью гиперребер со ссылками этого кластера. Определим набор гиперребер  $c.H$  кластера  $c$  как

$$c.H = \bigcup_{r \in R \wedge r.C=c} \{h | h \in H \wedge r \in h.R\}.$$

Для любого кластера  $c$  множество других кластеров, с которыми  $c$  связан через свой набор гиперребер  $c.H$ , образуют его окрестность  $\text{Nbr}(c)$ :

$$\text{Nbr}(c) = \bigcup_{h \in c.H, r \in h.R} \{c_j | c_j = r.C\}.$$

Если кластеры рассматривать как мультимножества, то арифметическая сумма кластеров, связанных с кластером  $c$  через набор его гиперребер, дает альтернативное определение окрестности  $\text{Nbr}_B(c)$ , учитывающее множественность вхождения различных соседних кластеров.

В литературе предложено и оценено множество различных показателей для измерения общности между наборами данных [8]. Для меры  $\text{sim}_R$  могут быть адаптированы некоторые из них. Мера определяется через понятие окрестности  $\text{Nbr}$ . Это определение игнорирует множественность вхождения кластера в окрестность. Поэтому для каждой меры определяется также вариант, учитывающий многократное появление общих кластеров в окрестностях, с опорой на оценку  $\text{Nbr}_B$  вместо  $\text{Nbr}$ .

#### 3.1 Оценка общих соседей

Для двух кластеров  $c_i$  и  $c_j$  оценка общих соседей определяется как

$$\text{CommonNbr}(c_i, c_j) = \frac{1}{K} |\text{Nbr}(c_i) \cap \text{Nbr}(c_j)|,$$

где  $K$  — нормализующая константа, подобранная таким образом, чтобы мера была меньше 1 для любой пары кластеров.

### 3.2 Коэффициент Жаккарда

Коэффициент Жаккарда используется для нормализации размера окрестности:

$$\text{Jaccard}(c_i, c_j) = \frac{|\text{Nbr}(c_i) \cap \text{Nbr}(c_j)|}{|\text{Nbr}(c_i) \cup \text{Nbr}(c_j)|}.$$

### 3.3 Оценка Адамик–Адар

Если кластер часто связан с множеством различных кластеров, то его присутствие в общей окрестности может быть не столь значительно, как для кластера, который встречается реже. Это учитывает оценка Адамик–Адар:

$$\text{AdarNbr}(c_i, c_j) = \frac{\sum_{c \in \text{Nbr}(c_i) \cap \text{Nbr}(c_j)} u(c)}{\sum_{c \in \text{Nbr}(c_i) \cup \text{Nbr}(c_j)} u(c)}, \quad (2)$$

где уникальность кластера

$$u(c) = \frac{1}{\log(|\text{Nbr}(c)|)}.$$

### 3.4 Оценка неоднозначности имен Адар

Для разрешения именованных сущностей предлагается оценка  $\text{AdarName}$ , основанная на степени уникальности имен. Определим неоднозначность  $\text{Amb}(r.\text{Name})$  ссылочного имени как вероятность того, что две случайно выбранные ссылки с  $\text{Name} = n$  соответствуют разным сущностям. Для ссылочного атрибута  $A_1$ , обозначаемого  $R.A_1$ , оценка неоднозначности значения  $n$  для атрибута равна:

$$\text{Amb}(r.A_1) = \frac{|\sigma_{R.A_1=r.A_1}(R)|}{|R|},$$

где  $|\sigma_{R.A_1=r.A_1}(R)|$  обозначает число ссылок со значением  $r.A_1$  для  $A_1$ . Когда для ссылок доступны несколько некоррелированных атрибутов  $A_i$ , этот подход может быть обобщен для получения более точных оценок неоднозначности.

Уникальность кластера  $c$  определяется как величина, обратно пропорциональная средней неоднозначности имен его ссылок:

$$u(c) = \frac{1}{\text{Avg}_{r \in c}(\text{Amb}(r.\text{Name}))},$$

что позволяет определить  $\text{AdarName}$  аналогично  $\text{AdarNbr}$  с помощью формулы (2).

### 3.5 Оценка через окрестности более высокого порядка

Окрестность второго порядка  $\text{Nbr}^2(c)$  для кластера  $c$  определяется как объединение множества окрестностей всех соседних кластеров:

$$\text{Nbr}^2(c) = \bigcup_{c' \in \text{Nbr}(c)} \text{Nbr}(c').$$

Оценка  $\text{Path3Sim}$  для  $c_i$  и  $c_j$  определяется как среднее значение оценок Жаккарда для окрестностей первого и второго порядка:

$$\begin{aligned} \text{Path3Sim}(c_i, c_j) = & \frac{1}{3} [\text{Jaccard}(\text{Nbr}(c_i), \text{Nbr}(c_j)) + \\ & + \text{Jaccard}(\text{Nbr}^2(c_i), \text{Nbr}(c_j)) + \text{Jaccard}(\text{Nbr}(c_i), \text{Nbr}^2(c_j))]. \end{aligned}$$

### 3.6 Негативные ограничения, связанные с отношениями

Могут существовать и негативные ограничения для разрешения сущностей, возникающие из взаимосвязей. Например, во многих случаях две ссылки, появляющиеся в одной и той же гиперссылке, не могут ссылаться на один и тот же объект. Поэтому сходство между двумя парами кластеров в формуле (1) установим равным нулю, если их объединение нарушает какие-либо реляционные ограничения.

## 4 Алгоритм реляционной кластеризации

После выбора меры сходства  $\text{sim}_R$  для пары кластеров для коллективного разрешения сущностей можно использовать жадный алгоритм агломеративной кластеризации [9], который находит ближайшую пару кластеров на каждом шаге и объединяет их. Высокоуровневый псевдокод для алгоритма представлен на рисунке. Описание отдельных механизмов алгоритма приведено ниже.

### 4.1 Механизм блокировки

Цель блокировки состоит в том, чтобы разделить ссылки на возможно перекрывающиеся сегменты и для уменьшения объема вычислений рассматривать как потенциальные совпадения только пары ссылок внутри каждого сегмента.

Алгоритм блокировки выполняет один проход по списку ссылок и распределяет их по сегментам, используя меру сходства на основе атрибутов  $\text{sim}_A$ . На каждом шаге прохода алгоритма каждый сегмент  $S_i$  имеет репрезентативную ссылку  $rs_i$ , для которой  $\sum_{r \in S_i} \text{sim}_A(r, rs_i)$  максимальна. Каждая новая ссылка сравнивается с репрезентативной для каждого сегмента. Ссылка присваивается всем сегментам, для которых сходство превышает пороговое значение. Если подходящий сегмент не найден, для этой ссылки создается новый сегмент.

1. Группировка ссылок в сегменты  $\{S_i\}$  с помощью механизма блокировки
2. Инициализация кластеров с помощью механизма начальной загрузки
3. Инициализация очереди с приоритетом (priority queue) PQ
4. Для всех пар кластеров  $c_i, c_j$  таких, что  $\exists l : c_i \in S_l, c_j \in S_l$
5. Вставка  $\langle \text{sim}(c_i, c_j), c_j, c_j \rangle$  в PQ
6. Пока PQ не пуста
7. Извлечение  $\langle \text{sim}(c_i, c_j), c_j, c_j \rangle$  из PQ
8. Если  $\text{sim}(c_i, c_j)$  меньше порогового значения, то остановка алгоритма
9. Объединение  $c_i$  и  $c_j$  в новый кластер  $c_{ij}$
10. Удаление записей, содержащих  $c_i$  и  $c_j$ , из PQ
11. Для каждого кластера  $c_k$  такого, что  $\exists l : c_{ij} \in S_l, c_k \in S_l$
12. Вставка  $\langle \text{sim}(c_{ij}, c_k), c_{jj}, c_k \rangle$  в PQ
13. Для каждого кластера  $c_n$  такого, что  $c_n \in \text{Nbr}(c_{ij})$
14. Для каждого кластера  $c_k$  такого, что  $\exists l : c_k \in S_l, c_n \in S_l$
15. Обновить значение  $\text{sim}(c_k, c_n)$  в PQ

Высокоуровневое описание алгоритма реляционной кластеризации

## 4.2 Механизм начальной загрузки

Если исходно каждой ссылке будет соответствовать отдельный кластер, то все слияния кластеров на начальных шагах алгоритма будут выполняться только на основе сравнения атрибутов. Многие из таких первоначальных слияний могут оказаться неточными, особенно для ссылок с неоднозначными значениями атрибутов. Механизм начальной загрузки просматривает каждую ссылочную пару, которая потенциально относится к одной сущности (как определено блокировкой), и определяет, станет ли она кандидатом для отнесения к одному кластеру. Для этого механизм сначала проверяет, точно ли совпадают их атрибуты. При совпадении он определяет, какие значения атрибутов неоднозначны, а какие нет, используя оценку неоднозначности  $\text{Amb}()$ . Пара неоднозначных ссылок формирует кандидата на загрузку, если их гиперребра совпадают в следующем смысле: два гиперребра  $h_1$  и  $h_2$  имеют  $k$ -точное совпадение, если существует по крайней мере  $k$  пар ссылок  $(r_i, r_j)$ ,  $r_i \in h_1.R$ ,  $r_j \in h_2.R$  с точно совпадающими атрибутами, т. е.  $r_i.A = r_j.A$ . Две ссылки  $r_1$  и  $r_2$  считаются кандидатами, если любая пара их гиперссылок имеет  $k$ -точное совпадение. Значение  $k$  для проверки гиперребер зависит от неоднозначности предметной области. Более высокое значение  $k$  следует использовать для областей с высокой неоднозначностью. После определения кандидатов исходные кластеры создаются с использованием популярного алгоритма union-find [10].

## 4.3 Объединение кластеров и обновление мер сходства

После выполнения процедуры начальной загрузки алгоритм итеративно объединяет наиболее похожую пару кластеров и обновляет меры сходства до тех пор, пока мера сходства не упадет ниже некоторого заданного порога. Это по-

казано в строках 6–15 рисунка. Шаги обновления меры сходства для связанных кластеров в строках 13–15 принципиально отличают коллективную реляционную кластеризацию от традиционного алгоритма агломеративной кластеризации.

## 5 Результаты экспериментальных проверок

В [7] приведены результаты испытаний алгоритма на тестовых наборах данных со степенью неоднозначности ссылок от 0,25% до 84% и со средним числом сущностей на одно неоднозначное имя от 2,33 до 32. Эффективность алгоритма оценивалась по F1-мере (среднее гармоническое точности и полноты). Проверки показали прирост эффективности по сравнению с традиционным алгоритмом разрешения сущностей на базе атрибутов с 0,98 до 0,995 (1,5%) на данных с низкой степенью неоднозначности и с 0,568 до 0,819 (44,2%) на данных с высокой степенью неоднозначности.

Опыт проведения конкретно-исторических исследований показывает, что степень неоднозначности данных в таких исследованиях может варьироваться в широких пределах и зависит от глубины исследования, изучаемого периода, сословной принадлежности изучаемых персон и ряда других факторов. Поскольку алгоритм обеспечивает прирост эффективности по сравнению с традиционным подходом во всем диапазоне степени неоднозначности данных, его применение для обработки конкретно-исторических данных, несомненно, оправданно. Но во многих случаях следует ожидать существенного повышения качества разрешения сущностей от применения рассматриваемого алгоритма в исследовании. Например, крестьяне в Российской Империи часто не имели фамилий вплоть до отмены крепостного права в 1861 г., что стало причиной высокой неоднозначности ссылок в таких номинативных исторических документах, как Метрические книги, Ревизские сказки, Исповедные ведомости и т. п.

Вычислительная сложность алгоритма оценивается в [7] как  $O(n\gamma \log n)$  для  $n$  входных ссылок, где  $\gamma$  представляет степень связности между ссылками. Тесты показывают более высокое время выполнения для алгоритма коллективной реляционной кластеризации по сравнению с традиционным подходом, основанным на атрибутах. Но, поскольку время выполнения остается в пределах десятков секунд даже на тестовых данных, содержащих сотни тысяч записей, в конкретно-историческом исследовании, где данные, как правило, содержат сотни и тысячи записей, этим параметром можно пренебречь.

## 6 Выводы

Рассмотренный в статье алгоритм разрешения сущностей показал свою высокую эффективность на данных с высоким уровнем неоднозначности ссылок, что соответствует специфике конкретно-исторических данных. Из этого следует принципиальная возможность применения нового подхода к решению задач

автоматизированного связывания записей в номинативных источниках при осуществлении конкретно-исторических исследований. Для этого при включении алгоритма в состав технологии ПКИИ необходимо осуществить выбор меры реляционного сходства, наиболее соответствующую специфике конкретно-исторических данных, а также конкретизировать на этой основе настраиваемые параметры алгоритма.

## **Литература**

1. *Антонов Д. Н.* Восстановление истории семей: метод, источники, анализ: Дис. . . . канд. ист. наук. — М., 2000. 290 с. EDN: QDBKMR.
2. *Торвальдсен Г.* Номинативные источники в контексте всемирной истории переписей: Россия и Запад // Известия Уральского федерального университета. Сер. 2: Гуманитарные науки, 2016. Т. 18. № 3. С. 9–28. doi: 10.15826/izv2.2016.18.3.041. EDN: WYDBXL.
3. *Юрченко Н. Л.* Некоторые проблемы использования ревизских сказок как источника по исторической демографии // Вспомогательные исторические дисциплины, 1993. Т. XXIV. С. 183–189.
4. *Адамович И. М., Волков О. И.* Технология распределенного автоматизированного анализа исторических текстов // Системы и средства информатики, 2016. Т. 26. № 3. С. 148–161. doi: 10.14357/08696527160311. EDN: WWSZIJ.
5. *Адамович И. М., Волков О. И.* Единая технология поддержки конкретно-исторических исследований // Системы и средства информатики, 2019. Т. 29. № 1. С. 194–205. doi: 10.14357/08696527190116. EDN: MZLQZG.
6. *Адамович И. М., Волков О. И.* Структурный подход к связыванию записей в технологии поддержки конкретно-исторических исследований // Системы и средства информатики, 2022. Т. 32. № 1. С. 94–103. doi: 10.14357/08696527220109. EDN: RLZRXH.
7. *Bhattacharya I., Getoor L.* Collective entity resolution in relational data // ACM T. Knowl. Discov. D., 2007. Vol. 1. Iss. 1. Art. 5. 36 p. doi: 10.1145/1217299.1217304.
8. *Вовченко А. Е., Калиниченко Л. А., Ковалев Д. Ю.* Методы разрешения сущностей и слияния данных в ETL-процессе и их реализация в среде Hadoop // Информатика и её применения, 2014. Т. 8. Вып. 4. С. 94–109. doi: 10.14357/19922264140412. EDN: PJYYLB.
9. *Roughgarden T.* Algorithms illuminated. Part 3: Greedy algorithms and dynamic programming. — New York, NY, USA: Soundlikeyourself Pubs., 2019. 217 p.
10. *Konger S.* What does the union-find algorithm actually do? // Scientific World, 2022. <https://scientificworld.org/union-find-algorithm>.

*Поступила в редакцию 09.01.24*



## COLLECTIVE ENTITY RESOLUTION IN TECHNOLOGY OF CONCRETE HISTORICAL INVESTIGATION SUPPORT

*I. M. Adamovich and O. I. Volkov*

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation

**Abstract:** The article is devoted to the further development of a distributed technology of concrete historical investigation support based on the principles of crowdsourcing and focused on a wide range of users who are not professional historians and biographers. Development is carried out by including in the technology an entity resolution algorithm for nominative documents processing that performs collective resolution in which entities for matching links are determined jointly. This algorithm is a modification of the greedy agglomerative clustering algorithm. The article provides a detailed description of the approach underlying the algorithm and provides its high-level pseudocode. The analysis of its effectiveness on data with varying degrees of ambiguity of names is given and the degree of ambiguity of names of concrete historical data is estimated. The conclusion about the expediency of including the algorithm in the technology is made. The directions of further research on determining the configurable parameters of the algorithm are outlined.

**Keywords:** concrete historical investigation; distributed technology; entity resolution; greedy algorithm; relational similarity measure

**DOI:** 10.14357/08696527240111

**EDN:** GQPGXQ

### References

1. Antonov, D. N. 2000. Vosstanovlenie istorii semey: metod, istochniki, analiz [Restoring family history: Method, sources, and analysis]. Moscow. PhD Diss. 290 p. EDN: QDBKMR.
2. Thorvaldsen, G. 2016. Nominativnye istochniki v kontekste vseмирnoy istorii perepisey: Rossiya i Zapad [Nominative data and global census history: Russia and the West]. *Izvestiya Ural'skogo federal'nogo universiteta. Ser. 2: Gumanitarnye nauki* [Izvestia. Ural Federal University J. Ser. 2: Humanities and Arts] 18(3):9–28. doi: 10.15826/izv2.2016.18.3.041. EDN: WYDBXL.
3. Yurchenko, N. L. 1993. Nekotorye problemy ispol'zovaniya revizskikh skazok kak istochnika po istoricheskoy demografii [Some problems of using census lists as a source for historical demography]. *Vspomogatel'nye istoricheskie distsipliny* [Auxiliary Historical Disciplines] XXIV:183–189.
4. Adamovich, I. M., and O. I. Volkov. 2016. Tekhnologiya raspredelenного avtomatizirovannogo analiza istoricheskikh tekstov [The distributed automated technology of historical texts analysis]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 26(3):148–161. doi: 10.14357/08696527160311. EDN: WWSZIJ.

5. Adamovich, I. M., and O. I. Volkov. 2019. Edinaya tekhnologiya podderzhki konkretno-istoricheskikh issledovaniy [Unified technology of concrete historical research support]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 29(1):194–205. doi: 10.14357/08696527190116. EDN: MZLOGZ.
6. Adamovich, I. M., and O. I. Volkov. 2022. Strukturnyy podkhod k svyazyvaniyu zapisey v tekhnologii podderzhki konkretno-istoricheskikh issledovaniy [Structural approach to record linking in technology of concrete historical investigation support]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 32(1):94–103. doi: 10.14357/08696527220109. EDN: RLZRXH.
7. Bhattacharya, I., and L. Getoor. 2007. Collective entity resolution in relational data. *ACM T. Knowl. Discov. D.* 1(1):5. 36 p. doi: 10.1145/1217299.1217304.
8. Vovchenko, A. E., L. A. Kalinichenko, and D. Yu. Kovalev. 2014. Metody razresheniya sushchnostey i sliyaniya dannykh v ETL-protsesse i ikh realizatsiya v srede Hadoop [Methods of entity resolution and data fusion in the ETL-process and their implementation in the Hadoop environment]. *Informatika i ee primeneniya — Inform. Appl.* 8(4):94–109. doi: 10.14357/19922264140412. EDN: PJYYLB.
9. Roughgarden, T. 2019. *Algorithms illuminated. Part 3: Greedy algorithms and dynamic programming*. New York, NY: Soundlikeyourself Publs. 217 p.
10. Konger, S. 2022. What does the union-find algorithm actually do? *Scientific World*. Available at: <https://scientificworld.org/union-find-algorithm/> (accessed March 4, 2024).

*Received January 9, 2024*

## **Contributors**

**Adamovich Igor M.** (b. 1934)— Candidate of Science (PhD) in technology, leading scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; adam@amsd.com

**Volkov Oleg I.** (b. 1964)— leading programmer, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation; volkov@amsd.com

## О ПАРАЛЛЕЛИЗМЕ В РАЗВИТИИ БАЗИСНЫХ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, ПРОСТРАНСТВЕННЫХ ТИПОВ ЦИВИЛИЗАЦИЙ И ТИПОЛОГИИ ЗНАКОВЫХ СИСТЕМ

*С. Н. Гринченко*<sup>1</sup>

**Аннотация:** Представлены основные структурные и пространственно-временные характеристики авторской информатико-кибернетической модели (ИКМ) самоуправляющейся иерархо-сетевой системы Человечества. Этапы ее системной глобальной эволюции сопоставлены с классификацией типов знаковых систем (ТЗС) А. Б. Соломоника. Установлена высокая степень соответствия содержания этапов усложнения базисных информационных технологий (БИТ) и пространственных типов цивилизаций (ПТЦ) в ходе системной глобальной эволюции с элементами классификации ТЗС в тройках: (1) БИТ «Сигнальные позы/звуки/движения» – ПТЦ «пред-пред-цивилизации» – ТЗС «Естественные знаковые системы»; (2) БИТ «Мимика/жесты» – ПТЦ «пред-цивилизации» – ТЗС «Образные системы»; (3) БИТ «Речь/язык» – ПТЦ «протоцивилизации» – ТЗС «Языковые системы»; (4) БИТ «Письменность/чтение» – ПТЦ «локальные цивилизации» – ТЗС «Системы записи»; (5) БИТ «Тиражирование текстов/книгопечатание» – ПТЦ «региональные и субконтинентальные цивилизации» – ТЗС «Формализованные знаковые системы первого порядка» (ФЗС-1); (6) БИТ «Локальный компьютер» – ПТЦ «Планетарная цивилизация» – ТЗС «Формализованные знаковые системы второго порядка» (ФЗС-2). Отмечена возможность дополнения классификации ТЗС путем введения понятий «Формализованные знаковые системы третьего порядка» (ФЗС-3) и «Формализованные знаковые системы четвертого порядка» (ФЗС-4), как аналога следующим в глобальной эволюции этапам БИТ «Телекоммуникации/сети» – ПТЦ «Цивилизация околоземного космоса» и «нано-БИТ (ИИ-БИТ)» – ПТЦ «Цивилизация промежуточного космоса» соответственно. Предлагается расширить содержательную трактовку элементов классификации ТЗС посредством введения в нее соответствующих пространственно-временных параметров ИКМ и рассматривать ТЗС как определяющие культурные признаки соответствующих ПТЦ. Сделан вывод о том, что последовательность возникновения и повышения уровня абстракции ТЗС можно интерпретировать как еще одно отражение процессов системной глобальной эволюции Человечества.

**Ключевые слова:** самоуправляющаяся система Человечества; системная глобальная эволюция; информатико-кибернетическая модель; базисные информационные технологии; пространственные типы цивилизаций; типы знаковых систем; семиотическая классификация А. Б. Соломоника

DOI: 10.14357/08696527240112

EDN: IAPYGP

---

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

## 1 О моделировании системной глобальной эволюции Человечества

Процессы глобальной эволюции иерархо-сетевой самоуправляющейся системы Человечества предлагается рассматривать, опираясь на ее ИКМ [1]. В качестве маркера этих системных процессов удобно использовать последовательность создания и усложнения БИТ [2] (рис. 1). На рис. 1 *восходящие стрелки*, имеющие структуру «многие – к одному», отражают **поисковую активность** представителей соответствующих ярусов в иерархии; *нисходящие сплошные стрелки*, имеющие структуру «один – ко многим», отражают **целевые критерии** поисковой оптимизации системной энергетики; *нисходящие пунктирные стрелки*, имеющие структуру «один – ко многим», отражают **системную память** личностно-производственно-социального — результат адаптивных влияний представителей вышележащих иерархических ярусов на структуру и поведение вложенных в них нижележащих.

Количественные (пространственно-временные) параметры ИКМ основаны на некоторой модификации геометрической прогрессии со знаменателем  $e^e = 15,15426\dots$ , выявленной А. В. Жирмунским и В. И. Кузьминым при исследовании критических уровней в развитии биосистем [3].

Данные пространственные параметры (линейные — радиусы кругов той же площади), монотонно увеличивающиеся в ходе глобальной эволюции, относятся к ориентировочным размерам ареалов устойчиво самоуправляющихся сообществ/социумов, ПТЦ и др., а монотонно уменьшающиеся — к величинам точности антропогенных воздействий в системе и др. Временные параметры ИКМ также носят ориентировочный, реперный характер.

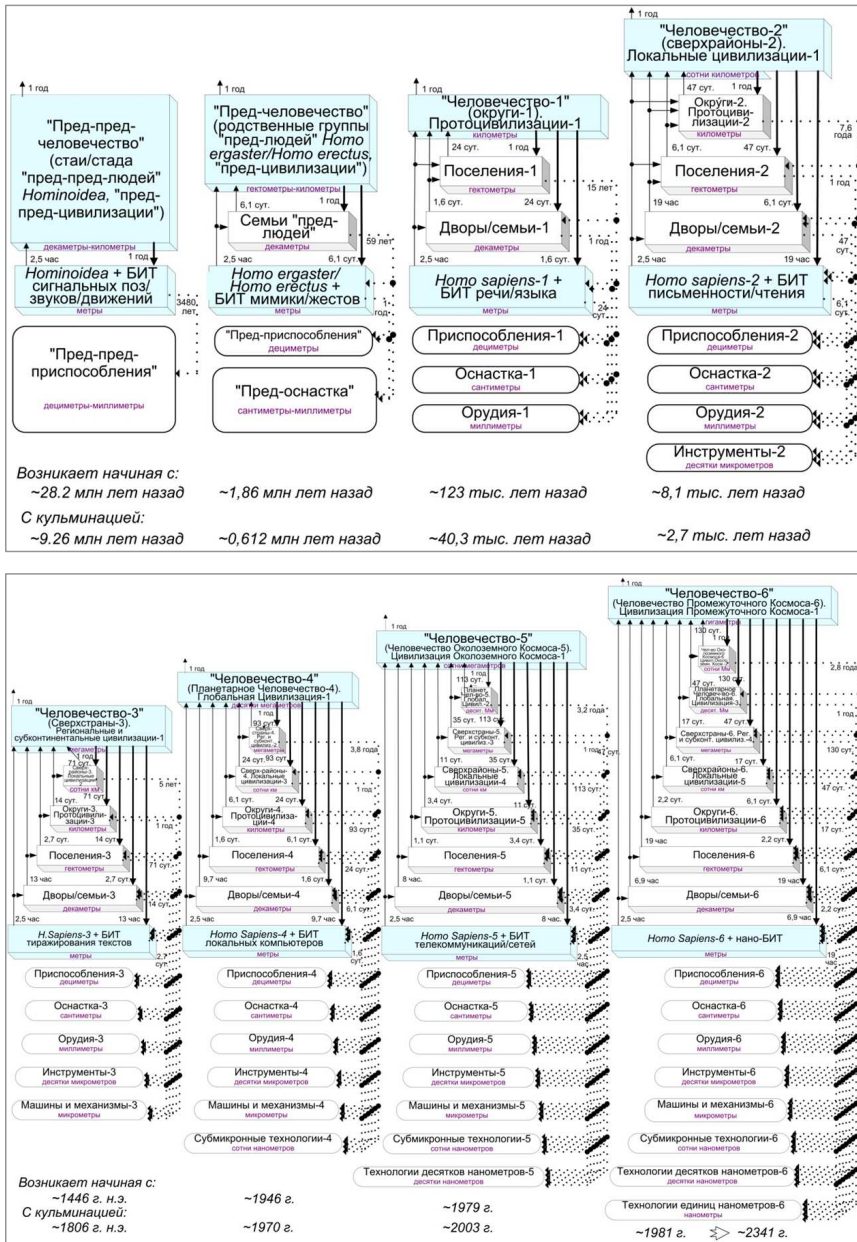
Практика применения ИКМ (в совокупности с фибоначчиевой моделью археологической эпохи Ю. Л. Щаповой) к описанию исторического процесса приведена в публикациях [4, 5].

В ходе описываемого системного глобального эволюционного процесса выполняются:

- (1) принцип *системной кумуляции* — возникновение в системе Человечества новых подсистем не означает элиминации ранее возникших — все они сосуществуют, активно взаимодействуют и коэволюционируют;
- (2) принцип *системной согласованности* — возникновение новых подсистем сопровождается кардинальными переменами в структуре и приспособительном поведении ранее возникших при уменьшении их вклада в общий ход приспособительного поведения [6].

Одним из потоков глобальной эволюции самоуправляющейся системы Человечества выступает последовательность усложняющихся ПТЦ [7, 8].

Как представляется, среди множества определений понятия «цивилизация» наиболее универсально и точно представление Н. Н. Моисеева, по мнению которого цивилизацией «целесообразно называть некоторую общность людей, связанную не только культурными традициями, технологическими навыками,



**Рис. 1** Этапы глобальной эволюции самоуправляющейся системы Человечества. Приведены типичные пространственно-временные характеристики структурных элементов подсистем и процессов приспособительного поведения в них

особенностями политического и общественного уклада, но прежде всего структурой ценностей, особенностями духовного мира. Если угодно — общностью миропонимания и близостью мировоззрений. Итак, цивилизация — это прежде всего некоторая общность людей!» [9, с. 88].

Исходя из того факта, что в центре ИКМ находится Человек, и опираясь на него, логика данной публикации намечает переход к анализу семиотической реальности, которая «собирает в себе основные знаковые свершения *человеческой цивилизации* (курсив мой. — С. Г.), и к ней обращается каждый, кто в своих исследованиях природы, общества или индивидуума использует опыт предыдущих поколений ученых» [10, с. 2].

## 2 О классификации типов семиотических знаковых систем

Известный теоретик семиотики А. Б. Соломоник отмечает: «под типом знаковых систем я понимаю набор знаковых систем с одинаковыми характеристиками знака, на котором они базируются. Например, слово — базисный знак языка, отличается по своим характеристикам и свойствам от математического символа, и я объединяю все языковые знаковые системы в один тип, а математико-формализованные знаковые системы — в другой» [10, с. 30].

Более того, согласно его представлению о *классификации* ТЗС, «естественные знаки являются первыми в истории знакового взаимодействия между природой и человеком. Они отличаются от всех прочих знаков своим естественным происхождением; все другие знаки придуманы человеком... Следующим типом знаковых систем, пришедшим вслед за естественными и надстраивающимися над ними как в филогенезе человечества, так и в онтогенезе каждого из нас, являются системы образные, имеющие образ (имидж) в качестве базисного знака... Язык — это следующая стадия в процессе развития человека и человечества, которая приходит после овладения началами образного представления... Вслед



**Рис. 2** Типы знаковых систем и их базисные знаки согласно классификации А. Б. Соломоника [11, с. 76]



за языковыми знаковыми системами приходит очередь систем записи... самые высокие слои в нашей иерархии... формализованные системы первого и второго порядка» [11, с. 78–81] (рис. 2).

### 3 О сопоставлении этапов глобальной системной эволюции и типов знаковых систем

Опираясь на положение А. Б. Соломоника о надстраивании ТЗС как в филогенезе человечества, так и в онтогенезе каждого из индивидов, сопоставим последовательность повышения степени абстракции таких типов в процессах их генезиса с этапами глобальной эволюции БИТ в системе Человечества [1] (см. таблицу).

Сопоставление 2-й, 4-й и 5-й колонок таблицы позволяет сделать следующие выводы о существенном *параллелизме по своему содержанию* этапов глобальной эволюции Человечества (на примере БИТ и ПТЦ) и типов семиотических знаковых систем.

1. Составляющие пары «БИТ сигнальных поз/звуков/движений» (формируемая представителями Биосферы *Hominoidea*, которым еще очень долго развиваться до уровня Человека) — «Естественные знаковые системы» выглядят вполне подобными, при этом им не противоречит соответствующая составляющая в последовательности ПТЦ — «пред-пред-цивилизации».
2. Составляющие пары «БИТ мимики/жестов» (формируемая еще весьма далекими предками человека *Homo erectus*) — «Образные системы» также позволяют соотнести их как вполне подобных между собой, при этом им также не противоречит соответствующий ПТЦ — «пред-цивилизации».
3. Составляющие пары «БИТ речи/языка» — «Языковые системы» оказываются совпадающими, что называется, «по определению». При этом они вполне корреспондируют с соответствующим представителем последовательности ПТЦ — «протоцивилизациями». Как представляется, определяющим их культурным признаком служит возникновение «семиотических языковых систем».
4. Составляющие пары «БИТ письменности/чтения» — «Системы записи» также оказываются полностью совпадающими. При этом они также вполне корреспондируют с соответствующими представителями последовательностей ПТЦ — «локальными цивилизациями». Предположительно, определяющим их культурным признаком служит возникновение «семиотических систем записи».
5. Возможная интерпретация ФЗС-1 достаточно широка и вполне позволяет соотнести ее с двойкой «БИТ тиражирования текстов/книгопечатания» — «ПТЦ региональные и субконтинентальные цивилизации». Можно также выдвинуть предположение, что именно возникновение ФЗС-1 стало определяющим культурным признаком последних.

Сопоставление этапов системной глобальной эволюции БИТ, сообществ/социумов, ПТЦ и ТЗС

Этапы системной глобальной эволюции [1]					
	Базисная информационная технология. Субъект	Времена старта; кульминации скорости развития	Пространственный ареал (радиус эквивалентного круга/шара). Сообщество/социум, ПТЦ	П	Типы семиотических знаковых систем [11]
1	2	3	4	А	5
1	Сигнальные позы/звуки/движения. <i>Hominoides</i>	~ 28,2 млн лет назад; ~ 9,26 млн лет назад	~ 4,2–64 м. «Двор»/семья, «пред-пред-цивилизации»	Р	Естественные знаковые системы
2	+ Мимика/жесты. <i>Homo erectus</i>	~ 1,86 млн лет назад; ~ 0,612 млн лет назад	~ 64 м – 1 км. + «Поселение»/род; + «пред-цивилизации»	А	+ Образные системы
3	+ Речь/язык. <i>Homo sapiens-1</i>	~ 123 тыс. лет назад; 40,3 тыс. лет назад	~ 1–15 км. + «Округа»/племя; + «протоцивилизации»	Л	+ Языковые системы
4	+ Письменность/чтение. <i>Homo sapiens-2</i>	~ 8,1 тыс. лет назад; 2,7 тыс. лет назад	~ 15–223 км. + «Сверхрайон» (национальное государство); + локальные цивилизации	Л	+ Системы записи
5	+ Тиражирование текстов/книгопечатание. <i>Homo sapiens-3</i>	~ 1446 г.; ~ 1806 г.	~ 223–3370 км. + «Сверхстрана» (наднациональная империя, держава и т. п.); + региональные и субконтинентальные цивилизации	Л Е	+ Формализованные системы первого порядка
6	+ Локальный компьютер. <i>Homo sapiens-4</i>	~ 1946 г.; 1970 г.	~ 3,37–51 тыс. км. + Планета Земля; + Планетарная (глобальная) цивилизация	Л	+ Формализованные системы второго порядка
7	+ Телекоммуникации/сети. <i>Homo sapiens-5</i>	~ 1979 г.; ~ 2003 г.	~ 51–773 тыс. км; + Околосредний космос (пространственный объем); + Цивилизация околосреднего космоса	И З	+ ???
8	+ Нано-БИТ (ИИ-БИТ). <i>Homo sapiens-6</i>	~ 1981 г.; 2341 г. (прогноз)	~ 0,773–11,7 млн км. + Промежуточный космос (пространственный объем); + Цивилизация промежуточного космоса	М	+ ???
...	...	...	...	...	...

**Примечание:** знаки «плюс» во 2-й, 4-й и 5-й колонках указывают на дополнительную новизну фактора по отношению к ранее возникшему («принцип системной кумуляции» [6]).



6. Аналогично возможная интерпретация ФЗС-2 достаточно широка и вполне позволяет соотнести ее с двойкой «БИТ локальных компьютеров» – «ПТЦ планетарная (глобальная) цивилизация». По аналогии предположим, что определяющим ее культурным признаком служит именно возникновение ФЗС-2<sup>1</sup>.
7. В свою очередь, отсутствующие в классификации ТЗС позиции — вероятные седьмой и восьмой по порядку возникновения, которым естественно присвоить рабочие наименования ФЗС-3 и ФЗС-4, естественным образом следует поставить в соответствие следующим двойкам: «БИТ коммуникаций/сетей (актуальной на сегодня)» – «ПТЦ Цивилизация околоземного космоса» (с определяющим ее культурным признаком ФЗС-3) и «нано-БИТ (ИИ-БИТ) (перспективной)» – «ПТЦ Цивилизация промежуточного космоса» (с определяющим ее культурным признаком ФЗС-4) соответственно. Как представляется, базисными знаками ФЗС-3, отражающими знаковую систему БИТ коммуникаций/сетей, и ФЗС-4, отражающими знаковую систему нано-БИТ (ИИ-БИТ), могут выступать *синтетические конструкции различной степени интеграции*, включающие те или иные — ранее возникшие в глобальной эволюции — элементы ТЗС: образных, языковых, систем записи, ФЗС-1 и ФЗС-2 (дальнейшее прояснение этих конкретных вопросов требует участия специалистов-семиотиков).

#### 4 Заключение

Последовательность возникновения и усложнения БИТ служит маркером *системной глобальной эволюции* Человечества, в ходе которой возникают и усложняются не только ПТЦ [8], но и общественно-экономические формации [12], феномен «многомерный иерархический территориальный суверенитет» [13, 14] и др. По аналогии и последовательность возникновения и повышения уровня абстракции/интеграции ТЗС (в классификации А. Б. Соломоника) можно интерпретировать как отражение *системной глобальной эволюции*. И если это предположение окажется адекватным, становится возможным ввести понятия ФЗС-3 и ФЗС-4 в параллели с седьмым и восьмым этапами системной глобальной эволюции Человечества соответственно, опираясь на пространственно-временные параметры ее соответствующих этапов [15].

В свою очередь, выявленный параллелизм в развитии ПТЦ и классификации ТЗС позволяет рассматривать последние — в контексте системной глобальной эволюции Человечества — именно как существенные эволюционно-цивилизационные культурные характеристики.

<sup>1</sup>При этом следует отметить, что либо работы Лейбница, Ньютона и др., использующие базисный знак «символ с переменным значением», заметно опередили свое (модельно прогнозируемое) время, либо интерпретация базисного знака ФЗС-2 нуждается в пересмотре — отражении знаковой системы БИТ локальных компьютеров.

## Литература

1. Гринченко С. Н. Метаэволюция (систем неживой, живой и социально-технологической природы). — М.: ИПИ РАН, 2007. 456 с.
2. Гринченко С. Н., Шапова Ю. Л. Генезис информационных технологий как маркер генезиса иерархий в системе Человечества: модельное представление // Современные информационные технологии и ИТ-образование, 2019. Т. 15. № 2. С. 421–430. doi: 10.25559/SITITO.15.201902.421-430. EDN: GKXOKI.
3. Жирмунский А. В., Кузьмин В. И. Критические уровни в процессах развития биологических систем. — М.: Наука, 1982. 179 с.
4. Шапова Ю. Л., Гринченко С. Н. Введение в теорию археологической эпохи: числовое моделирование и логарифмические шкалы пространственно-временных координат. — М.: ФИЦ ИУ РАН, 2017. 236 с.
5. Шапова Ю. Л., Гринченко С. Н., Кокорина Ю. Г. Информатико-кибернетическое и математическое моделирование археологической эпохи: логико-понятийный аппарат. — М.: ФИЦ ИУ РАН, 2019. 136 с.
6. Гринченко С. Н. Общение людей и информационные технологии: принципы системной кумуляции и системной согласованности // Мир психологии, 2020. № 3(103). С. 235–245. doi: 10.51944/2073-8528-2020\_3\_235. EDN: UICFKK.
7. Гринченко С. Н. Системные закономерности развития цивилизации. — Красноярск: СФУ, 2011. 34 с.
8. Гринченко С. Н. О глобальной коэволюции информационных технологий, общественно-экономических формаций и цивилизаций: кибернетический взгляд // Моисеевские чтения. Гуманитарные вызовы и угрозы XXI века. — М.: МосГУ, 2023. Т. I. С. 65–70. EDN: EANPHT.
9. Моисеев Н. Н. Универсум. Информация. Общество. — М.: Устойчивый мир, 2001. 200 с.
10. Соломоник А. Б. Язык как знаковая система. — М.: Наука, 1992. 223 с.
11. Соломоник А. Б. Очерк общей семиотики. — Минск: МЕТ, 2009. 191 с.
12. Гринченко С. Н. О сравнительном генезисе базисных информационных технологий и общественно-экономических формаций Человечества // Системы и средства информатики, 2022. Т. 32. № 3. С. 147–155. doi: 10.14357/08696527220314. EDN: RWDREG.
13. Гринченко С. Н. Феномен «территориального суверенитета» с кибернетических позиций: многомерность, иерархическая структура, глобальная эволюция // Ученый совет, 2022. Т. XIX. № 5. С. 323–329. doi: 10.33920/nik-02-2205-04. EDN: JWVUFV.
14. Гринченко С. Н. Базисные информационные технологии, информационная война и многомерный иерархический территориальный суверенитет: этапы глобально-космической коэволюции // Системы и средства информатики, 2023. Т. 33. № 1. С. 147–154. doi: 10.14357/08696527230114. EDN: SNVMZT.
15. Гринченко С. Н. О параллелизме этапов глобальной эволюции базисных информационных технологий и типов семиотических знаковых систем // Современные информационные технологии и ИТ-образование, 2023. Т. 19. № 2. С. 454–462.

*Поступила в редакцию 03.02.24*

## ON THE COMPARISON OF STAGES OF SYSTEMIC GLOBAL EVOLUTION OF HUMANKIND AND TYPES OF SEMIOTIC SIGN SYSTEMS

*S. N. Grinchenko*

Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119133, Russian Federation

**Abstract:** The main structural and spatiotemporal characteristics of the author’s informatics-cybernetic hierarchical-network model of the self-controlling system of Humankind are presented. The stages of its systemic global evolution are compared with the classification of types of sign systems (TSS) by A. B. Solomonick. A high degree of correspondence has been established between the content of the stages of complication of basic information technologies (BIT) and spatial types of civilizations (STC) in the course of global evolution with the elements of the TSS classification in threes: (i) BIT “Signal poses/sounds/movements”–STC “pre-pre-civilizations”–TSS “Natural sign systems;” (ii) BIT “Mimics/gestures”–STC “pre-civilizations”–TSS “Image systems;” (iii) BIT “Speech/language”–STC “protocivilizations”–TSS “Language systems;” (iv) BIT “Writing/reading”–STC “Local civilizations”–TSS “Writing systems;” (v) BIT “Replication of texts/book printing”–STC “regional and subcontinental civilizations”–TSS “Formalized sign systems of the first order;” and (vi) BIT “Local computer”–STC “Planetary Civilization”–TSS “Formalized sign systems of the second order.” The possibility of supplementing the classification of TSS by introducing the concepts of “Formalized sign systems of the third order” and “Formalized sign systems of the fourth order” as an analogue to the following stages in global evolution BIT “Telecommunications/networks”–STC “Civilization of Near-Earth Space” and “nano-BIT (AI-BIT)”–STC “Civilization of Intermediate Cosmos,” respectively. It is also proposed to expand the meaningful interpretation of the elements of the TSS classification by introducing into it the corresponding spatiotemporal parameters of informatics-cybernetic model and also consider TSS as defining cultural characteristics of the corresponding STC. It is concluded that the sequence of emergence and increasing the level of abstraction of the TSC can be interpreted as another reflection of the processes of the systemic global evolution of Humanity.

**Keywords:** self-controlling system of Humankind; systemic global evolution; informatics-cybernetic model; basic information technologies; spatial types of civilizations; types of sign systems; semiotic classification by A. B. Solomonick

**DOI:** 10.14357/08696527240112

**EDN:** IAPYGP

### References

1. Grinchenko, S.N. 2007. *Metaevolyutsiya (sistem nezhivoy, zhivoy i sotsial’no-tekhnologicheskoy prirody)* [Metaevolution (of inanimate, animate, and socio-technological nature systems)]. Moscow: IPIRAN. 456 p.

2. Grinchenko, S. N., and Yu. L. Shchapova. 2019. Genezis informatsionnykh tekhnologiy kak marker genezisa ierarkhiy v sisteme Chelovechestva: model' noe predstavlenie [Genesis of information technologies as a marker of the genesis of hierarchies in the humankind's system: A model representation]. *Sovremennye informatsionnye tekhnologii i IT-obrazovanie* [Modern Information Technologies and IT-Education] 15(2):421–430. doi: 10.25559/SITITO.15.201902.421-430. EDN: GKXOKI.
3. Zhirmunskiy, A. V., and V. I. Kuz'min. 1982. *Kriticheskie urovni v protsessakh razvitiya biologicheskikh sistem* [Critical levels in the development of biological systems]. Moscow: Nauka. 179 p.
4. Shchapova, Y. L., and S. N. Grinchenko. 2017. *Vvedenie v teoriyu arkheologicheskoy epokhi: chislovoe modelirovanie i logarifmicheskie shkaly prostranstvenno-vremennykh koordinat* [Introduction to the theory of the archaeological epoch: Numerical modeling and logarithmic scales of space-time coordinates]. Moscow: FRC CSC RAS. 236 p.
5. Shchapova, Yu. L., S. N. Grinchenko, and Yu. G. Kokorina. 2019. *Informatiko-kiberneticheskoe i matematicheskoe modelirovanie arkheologicheskoy epokhi: logiko-ponyatiynyy apparat* [Informatics-cybernetic and mathematical modeling of the archaeological epoch: Logical-conceptual apparatus]. Moscow: FRC CSC RAS. 136 p.
6. Grinchenko, S. N. 2020. Obshchenie lyudey i informatsionnye tekhnologii: printsipy sistemnoy kumulyatsii i sistemnoy soglasovannosti [Communication of people and information technology: Principles of systemic cumulation and systemic coordination]. *Mir psikhologii* [World of Psychology] 3(103):235–245. doi: 10.51944/2073-8528\_2020\_3\_235. EDN: UICFKK.
7. Grinchenko, S. N. 2011. *Sistemnye zakonomernosti razvitiya tsivilizatsii* [Systemic patterns of civilization development]. Krasnoyarsk: SFU. 34 p.
8. Grinchenko, S. N. 2023. O global'noy koevolutsii informatsionnykh tekhnologiy, obshchestvenno-ekonomicheskikh formatsiy i tsivilizatsiy: kiberneticheskiy vzglyad [On the global coevolution of information technologies, socioeconomic formations and civilizations: A cybernetic view]. *Moiseevskie chteniya. Gumanitarnye vyzovy i ugrozy XXI veka* [Moiseev's readings. Humanitarian challenges and threats of the 21st century]. Moscow: MosGU. 1:65–70. EDN: EANPHT.
9. Moiseev, N. N. 2001. *Universum. Informatsiya. Obshchestvo* [Universum. Information. Society]. Moscow: Ustoychivyy mir. 200 p.
10. Solomonick, A. B. 1992. *Yazyk kak znakovaya sistema* [Language as a sign system]. Moscow: Nauka. 223 p.
11. Solomonick, A. B. 2009. *Ocherk obshchey semiotiki* [Essay on general semiotics]. Minsk: MET. 191 p.
12. Grinchenko, S. N. 2022. O sravnitel'nom genezise bazisnykh informatsionnykh tekhnologiy i obshchestvenno-ekonomicheskikh formatsiy Chelovechestva [On the comparative genesis of basic information technologies and socioeconomic formations of the Humankind]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 32(3):147–155. doi: 10.14357/08696527220314. EDN: RWDREG.
13. Grinchenko, S. N. 2022. Fenomen "territorial'nogo suvereniteta" s kiberneticheskikh pozitsiy: mnogomernost', ierarkhicheskaya struktura, global'naya evolyutsiya [The phenomenon of "territorial sovereignty" from the cybernetic perspective: Multidimensionality, hierarchical structure, and global evolution]. *Uchenyy sovet* [Academic Council] XIX(5):323–329. doi: 10.33920/nik-02-2205-04. EDN: JWVUFV.

14. Grinchenko, S.N. 2023. Bazisnye informatsionnye tekhnologii, informatsionnaya voyna i mnogomernyy ierarkhicheskiy territorial'nyy suverenitet: etapy global'no-kosmicheskoy koevolyutsii [Basic information technologies, information warfare, and multidimensional hierarchical territorial sovereignty: Stages of global-space coevolution]. *Sistemy i Sredstva Informatiki — Systems and Means of Informatics* 33(1):146–153. doi: 10.14357/08696527230114. EDN: SNVMZT.
15. Grinchenko, S. N. 2023. O parallelizme etapov global'noy evolyutsii bazisnykh informatsionnykh tekhnologiy i tipov semioticheskikh znakovykh sistem [On the parallelism of the stages of the global evolution of basic information technologies and types of semi-otic sign systems]. *Sovremennye informatsionnye tekhnologii i IT-obrazovanie* [Modern Information Technologies and IT-education]. 19(2):454–462.

*Received February 3, 2024*

### **Contributor**

**Grinchenko Sergey N.** (b. 1946) — Doctor of Science in technology, professor, principal scientist, Federal Research Center “Computer Science and Control” of the Russian Academy of Sciences, 44-2 Vavilov Str., Moscow 119333, Russian Federation; sgrinchenko@ipiran.ru

---

## ОБ АВТОРАХ

---

**Адамович Игорь Михайлович** (р. 1934) — кандидат технических наук, ведущий научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Будзко Владимир Игоревич** (р. 1944) — доктор технических наук, академик Академии криптографии Российской Федерации, главный научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Власкина Анастасия Сергеевна** (р. 1995) — кандидат физико-математических наук, ассистент кафедры теории вероятностей и кибербезопасности Российского университета дружбы народов имени Патриса Лумумбы

**Волков Олег Игоревич** (р. 1964) — ведущий программист Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Гринченко Сергей Николаевич** (р. 1946) — доктор технических наук, профессор, главный научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Дьяченко Юрий Георгиевич** (р. 1958) — кандидат технических наук, старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Егоров Владимир Борисович** (р. 1948) — кандидат технических наук, ведущий научный сотрудник Федерального исследовательского центра «Информатика и управление»

**Елисеев Владислав Антонович** (р. 2005) — студент Национального исследовательского университета «Высшая школа экономики»

**Илларионова Любовь Викторовна** (р. 1977) — кандидат физико-математических наук, старший научный сотрудник Вычислительного центра Дальневосточного отделения Российской академии наук

**Кочеткова Ирина Андреевна** (р. 1985) — кандидат физико-математических наук, доцент кафедры теории вероятностей и кибербезопасности Российского университета дружбы народов имени Патриса Лумумбы; старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Листопад Сергей Викторович** (р. 1984) — кандидат технических наук, старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Меденников Виктор Иванович** (р. 1949) — доктор технических наук, ведущий научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Морозов Николай Викторович** (р. 1956) — старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Орлов Георгий Александрович** (р. 1994) — младший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Прохорец Иван Олегович** (р. 1999) — инженер, аспирант Вычислительного центра Дальневосточного отделения Российской академии наук

**Семенihin Тимофей Александрович** (р. 1999) — инженер Государственного астрономического института имени П. К. Штернберга Московского государственного университета имени М. В. Ломоносова

**Синицын Игорь Николаевич** (р. 1940) — доктор технических наук, профессор, заслуженный деятель науки РФ, главный научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук; профессор кафедры «Мехатроника и теоретическая механика» Московского авиационного института (национального исследовательского университета)

**Степанов Алексей Сергеевич** (р. 1974) — доктор фармацевтических наук, ведущий научный сотрудник Дальневосточного научно-исследовательского института сельского хозяйства

**Степченко Дмитрий Юрьевич** (р. 1973) — старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Степченко Юрий Афанасьевич** (р. 1951) — кандидат технических наук, заведующий отделом, ведущий научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Сучков Александр Павлович** (р. 1954) — доктор технических наук, ведущий научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Федосеев Денис Владимирович** (р. 1994) — аспирант Вычислительного центра Дальневосточного отделения Российской академии наук

**Фролов Андрей Николаевич** (р. 1985) — младший научный сотрудник Дальневосточного научно-исследовательского института сельского хозяйства

**Хилько Дмитрий Владимирович** (р. 1987) — старший научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Шамурина Анна Игоревна** (р. 1988) — кандидат технических наук, старший научный сотрудник Вычислительного центра Дальневосточного отделения Российской академии наук

**Шоргин Всеволод Сергеевич** (р. 1978) — кандидат технических наук, старший научный сотрудник Института проблем информатики Федерального исследовательского центра «Информатика и управление» Российской академии наук

**Эсенг Манге Педро Луис Эяма** (р. 1992) — аспирант кафедры теории вероятностей и кибербезопасности Российского университета дружбы народов имени Патриса Лумумбы



---

## 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 предоставленной автором(ами) рукописи не должна нарушать положения глав 69, 70 раздела VII части IV Гражданского кодекса, которые определяют права на результаты интеллектуальной деятельности и средства индивидуализации, в том числе авторские права, в РФ.

Ответственность за нарушение авторских прав, в случае предъявления претензий к редакции журнала, несут авторы статей.

Направляя рукопись в редакцию, авторы сохраняют свои права на данную рукопись и при этом передают учредителям и редколлегии журнала неисключительные права на издание статьи на русском языке (или на языке статьи, если он отличен от русского) и на перевод ее на английский язык, а также на ее распространение в России и за рубежом. Каждый автор должен представить в редакцию подписанный с его стороны «Лицензионный договор о передаче неисключительных прав на использование произведения», текст которого размещен по адресу <http://www.ipiran.ru/publications/licence.doc>. Этот договор может быть представлен в бумажном (в 2-х экз.) или в электронном виде (отсканированная копия заполненного и подписанного документа).

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 10 страниц указанного формата. При превышении указанного объема редколлегия вправе потребовать от автора сокращения объема рукописи.

Сокращения слов, помимо стандартных, не допускаются. Допускается минимальное количество аббревиатур.

Все страницы рукописи нумеруются.

Шаблоны оформления представлены в интернете:

[http://www.ipiran.ru/journal/template\\_iiep\\_ssi\\_2024.zip](http://www.ipiran.ru/journal/template_iiep_ssi_2024.zip)

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/2019\\_29\\_03\\_rus/authors.asp](http://www.ipiran.ru/journal/collected/2019_29_03_rus/authors.asp) и  
[http://www.ipiran.ru/journal/collected/2019\\_29\\_03\\_eng/authors.asp](http://www.ipiran.ru/journal/collected/2019_29_03_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 (ссылка 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.net/ru/bgn/> бесплатной программой транслитерации русского текста в латиницу.

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. *Russ. 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 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 resursoberegayushchie 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. — New York, NY, USA: Wiley, 1974. 521 p.)

2. В англоязычной части:

Timoshenko, S. P., D. H. Young, and W. Weaver. 1974. *Vibration problems in engineering*. 4th ed. New York, NY: 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]. Moscow. D.Sc. Diss. 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;
  - указывать в тексте письма название статьи, авторов и журнал, в который направляется статья;
  - 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: [ssi@frccsc.ru](mailto:ssi@frccsc.ru) (Стригина Светлана Николаевна)

<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 10 pages of the specified format. If the specified size exceeded, the editorial board is entitled to require the author to reduce the manuscript.

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/journal/template\\_jiep\\_ssi\\_2024.zip](http://www.ipiran.ru/journal/template_jiep_ssi_2024.zip)

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/2019\\_29\\_03\\_rus/authors.asp](http://www.ipiran.ru/journal/collected/2019_29_03_rus/authors.asp) and  
[http://www.ipiran.ru/journal/collected/2019\\_29\\_03\\_eng/authors.asp](http://www.ipiran.ru/journal/collected/2019_29_03_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.net/ru/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. *Russ. 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 Publ. 352 p.

**Dissertation and Thesis:**

Kozhunova, O. S. 2009. Tekhnologiya razrabotki semanticheskogo slovaryia informatsionnogo monitoringa [Technology of development of semantic dictionary of information monitoring system]. Moscow: IPI RAN. PhD Thesis. 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 Publ. 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;
  - the article title, authors' names, and the journal title, whereto the paper is being submitted, in the text of the e-mail;
  - 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: [ssi@frccsc.ru](mailto:ssi@frccsc.ru) (to Svetlana Strigina)

[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 34 No.1 Year 2024

Editor-in-Chief and Chair of Editorial Council  
Academician I. A. Sokolov

## IN THIS ISSUE:

- ANALYTICAL MODELING OF STOCHASTIC SYSTEMS WITH RANDOM PARAMETERS  
AND UNSOLVED DERIVATIVES  
*I. N. Sinitsyn* 4
- RADIO RESOURCE ALLOCATION FOR ELASTIC TRAFFIC TRANSMISSION  
WITH LATENCY FAIRNESS IN WIRELESS NETWORKS  
*I. A. Kochetkova, A. S. Vlaskina, P. L. E. Eseng Mangué,  
and V. S. Shorgin* 23
- DESYNCHRONIZATION METHODOLOGY AT SELF-TIMED CIRCUIT SYNTHESIS  
*Yu. A. Stepchenkov, D. V. Khilko, Yu. G. Diachenko, N. V. Morozov,  
D. Yu. Stepchenkov, and G. A. Orlov* 33
- DEVELOPMENT OF INFORMATION SYSTEM FOR STORING AND PROCESSING  
THE RESULTS OF AGRICULTURAL LAND MONITORING IN THE SOUTH  
OF THE FAR EAST  
*D. V. Fedoseev, A. S. Stepanov, L. V. Illarionova, V. A. Eliseev,  
I. O. Prokhorets, A. N. Frolov, and A. I. Shamurina* 44
- MAPPING OF THE Khabarovsk REGION ARABLE LANDS  
BY MACHINE LEARNING USING SENTINEL-2 IMAGES  
*I. O. Prokhorets and A. S. Stepanov* 57
- NEURAL NETWORK ARCHITECTURE FOR ARTIFACTS DETECTION  
IN ZTF SURVEY  
*T. A. Semenikhin* 70
- THE LIFECYCLE OF CYBERATTACKS ON MACHINE LEARNING  
SYSTEM VULNERABILITIES  
*A. P. Suchkov* 80
- DEVELOPMENT OF NEGOTIATION TACTICS BY ARTIFICIAL INTELLIGENT  
REFLEXIVE AGENTS  
*S. V. Listopad* 92