

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

ТЕОРЕТИЧЕСКИЙ И ПРИКЛАДНОЙ НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

Том 27

2021

№ 4

САПР

КОМПЬЮТЕРНАЯ ГРАФИКА

МЕТОДЫ ПРОГРАММИРОВАНИЯ

ОПЕРАЦИОННЫЕ СИСТЕМЫ И СРЕДЫ

ТЕЛЕКОММУНИКАЦИИ
И ВЫЧИСЛИТЕЛЬНЫЕ СЕТИ

ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ

НЕЙРОСЕТИ И
НЕЙРОКОМПЬЮТЕРЫ

СТРУКТУРНЫЙ СИНТЕЗ

ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ

ПРИКЛАДНЫЕ ИНФОРМАЦИОННЫЕ
СИСТЕМЫ

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

ОПТИМИЗАЦИЯ И МОДЕЛИРОВАНИЕ

ИТ В ОБРАЗОВАНИИ

ГИС

Рисунки к статье С. В. Шайтуры, Д. А. Галкина
«ПРОТОКОЛ ЦИФРОВОЙ ПОДПИСИ НА ОСНОВЕ КРИПТОКОДОВЫХ КОНСТРУКЦИЙ»

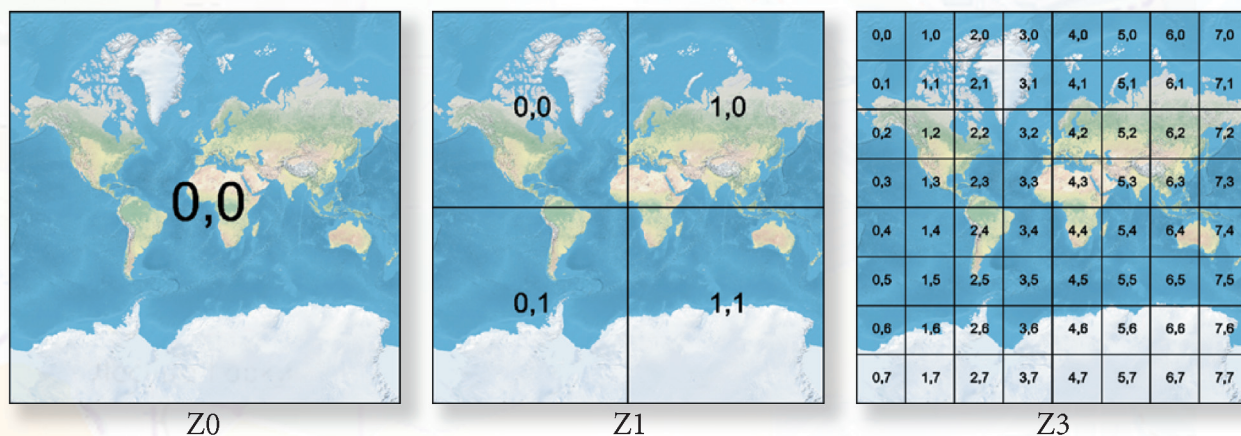


Рис. 2. Формирование тайлов:
 типовой размер тайла 256×256 пикселей; градус на пиксель при Z_1 : 0,703125

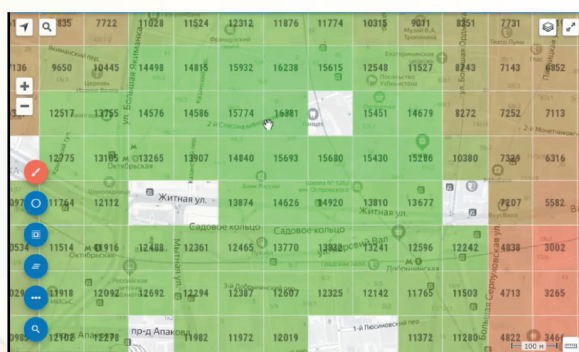


Рис. 3. Карта с прогнозом финансовых показателей премиальных магазинов одежды



Рис. 5. Агрегирующий слой «Население 1000»

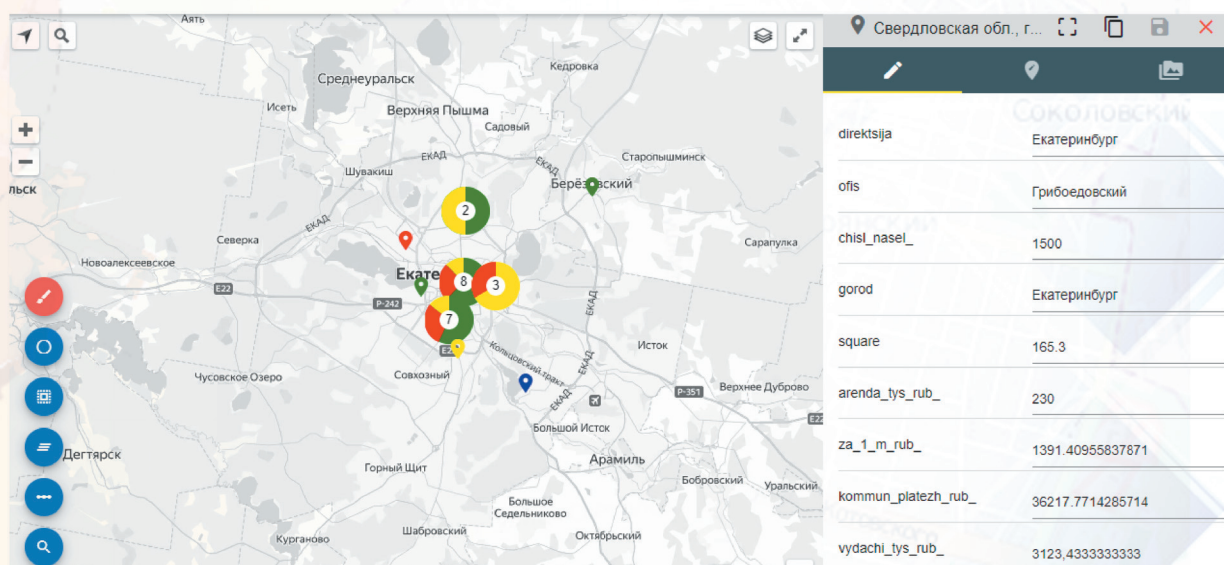


Рис. 12. Потенциальные локации для отделений банка

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Том 27
2021
№ 4

ТЕОРЕТИЧЕСКИЙ И ПРИКЛАДНОЙ НАУЧНО-ТЕХНИЧЕСКИЙ ЖУРНАЛ

Издается с ноября 1995 г.

DOI 10.17587/issn.1684-6400

УЧРЕДИТЕЛЬ

Издательство "Новые технологии"

СОДЕРЖАНИЕ

ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ И СЕТИ

- Ляхов П. А., Ионисян А. С., Валуева М. В., Ларикова А. С. Высокопроизводительная цифровая фильтрация на модифицированных умножителях с накоплением в системе остаточных классов с модулями специального вида 171

БАЗЫ ДАННЫХ

- Шайтура С. В., Галкин Д. А. Геомаркетинговый анализ больших данных . . . 180
- Малаховецкий Д. В., Разумовский А. И. Синтаксический анализ символьных массивов методом рекурсивного охвата и структурирования на примере данных VRML-формата 188

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ И ТЕХНОЛОГИИ

- Нгуен Т. В., Кравец А. Г. Оценка и прогнозирование тенденций развития научных исследований на основе библиометрического анализа публикаций . . . 195
- Поляков А. Н., Позевалкин В. В. Применение нейронной сети прямого пространства к прогнозированию тепловых характеристик станков . . . 202

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ЭКОНОМИКЕ, УПРАВЛЕНИИ И ПРОИЗВОДСТВЕ

- Коковин В. А., Евсиков А. А., Леонов А. П. Особенности организации и взаимодействия функциональных сетевых компонентов в распределенных управляющих системах 212

Главный редактор:

СТЕМПКОВСКИЙ А. Л.,
акад. РАН, д. т. н., проф.

Зам. главного редактора:

ИВАННИКОВ А. Д., д. т. н., проф.
ФИЛИМОНОВ Н. Б., д. т. н., с.н.с.

Редакционный совет:

БЫЧКОВ И. В., акад. РАН, д. т. н.
ЖУРАВЛЕВ Ю. И.,
акад. РАН, д. ф.-м. н., проф.
КУЛЕШОВ А. П.,
акад. РАН, д. т. н., проф.
ПОПКОВ Ю. С.,
акад. РАН, д. т. н., проф.
РУСАКОВ С. Г.,
чл.-корр. РАН, д. т. н., проф.
РЯБОВ Г. Г.,
чл.-корр. РАН, д. т. н., проф.
СОЙФЕР В. А.,
акад. РАН, д. т. н., проф.
СОКОЛОВ И. А.,
акад. РАН, д. т. н., проф.
СУЕТИН Н. В., д. ф.-м. н., проф.
ЧАПЛЫГИН Ю. А.,
акад. РАН, д. т. н., проф.
ШАХНОВ В. А.,
чл.-корр. РАН, д. т. н., проф.
ШОКИН Ю. И.,
акад. РАН, д. т. н., проф.
ЮСУПОВ Р. М.,
чл.-корр. РАН, д. т. н., проф.

Редакционная коллегия:

АВДОШИН С. М., к. т. н., доц.
АНТОНОВ Б. И.
БАРСКИЙ А. Б., д. т. н., проф.
ВАСЕНИН В. А., д. ф.-м. н., проф.
ВАСИЛЬЕВ В. И., д. т. н., проф.
ВИШНЕКОВ А. В., д. т. н., проф.
ДИМИТРИЕНКО Ю. И., д. ф.-м. н., проф.
ДОМРАЧЕВ В. Г., д. т. н., проф.
ЗАБОРОВСКИЙ В. С., д. т. н., проф.
ЗАРУБИН В. С., д. т. н., проф.
КАРПЕНКО А. П., д. ф.-м. н., проф.
КОЛИН К. К., д. т. н., проф.
КУЛАГИН В. П., д. т. н., проф.
КУРЕЙЧИК В. В., д. т. н., проф.
ЛЬВОВИЧ Я. Е., д. т. н., проф.
МАРТЫНОВ В. В., д. т. н., проф.
МИХАЙЛОВ Б. М., д. т. н., проф.
НЕЧАЕВ В. В., к. т. н., проф.
ПОЛЕШУК О. М., д. т. н., проф.
ПРОХОРОВ С. А., д. т. н., проф.
САКСОНОВ Е. А., д. т. н., проф.
СОКОЛОВ Б. В., д. т. н., проф.
СОЛОВЬЕВ Р. А., д. т. н., в. н. с.
ТИМОНИНА Е. Е., д. т. н., проф.
УСКОВ В. Л., к. т. н. (США)
ФОМИЧЕВ В. А., д. т. н., проф.
ШИЛОВ В. В., к. т. н., доц.

Редакция:

БЕЗМЕНОВА М. Ю.

Информация о журнале доступна по сети Internet по адресу <http://novtex.ru/IT>.
Журнал включен в систему Российского индекса научного цитирования и базу данных RSCI на платформе Web of Science.
Журнал входит в Перечень научных журналов, в которых по рекомендации ВАК РФ должны быть опубликованы научные результаты диссертаций на соискание ученой степени доктора и кандидата наук.

INFORMATION TECHNOLOGIES INFORMACIONNYYE TEHNOLOGII

Vol. 27
2021
No. 4

THEORETICAL AND APPLIED SCIENTIFIC AND TECHNICAL JOURNAL

Published since November 1995

DOI 10.17587/issn.1684-6400

ISSN 1684-6400

CONTENTS

COMPUTING SYSTEMS AND NETWORKS

- Lyakhov P. A., Ionisyan A. S., Valueva M. V., Larikova A. S.** High Performance Digital Filtering on Modified Multiply and Accumulate Unit in Residue Number System with Special Type of Moduli 171

DATABASE

- Shaytura S. V., Galkin D. A.** Geomarketing Big Data Analysis 180

- Malakhovetsky D. V., Razumowsky A. I.** Parsing Character Arrays by Recursive Scoping and Structuring Using the Example of VRML Data 188

INTELLIGENT SYSTEMS AND TECHNOLOGIES

- Nguyen T. V., Kravets A. G.** Evaluation and Prediction of Trends in the Development of Scientific Research Based on Bibliometric Analysis of Publications 195

- Polyakov A. N., Pozevalkin V. V.** Application of a Feedforward Neural Network to Predicting the Thermal Characteristics of Machine Tools 202

INFORMATION TECHNOLOGIES IN ECONOMY, MANAGEMENT AND PRODUCTION

- Kokovin V. A., Evsikov A. A., Leonov A. P.** Features of the Organization and Interaction of Functional Network Components in Distributed Control Systems 212

Editor-in-Chief:

Stempkovsky A. L., Member of RAS,
Dr. Sci. (Tech.), Prof.

Deputy Editor-in-Chief:

Ivannikov A. D., Dr. Sci. (Tech.), Prof.
Filimonov N. B., Dr. Sci. (Tech.), Prof.

Chairman:

Bychkov I. V., Member of RAS,
Dr. Sci. (Tech.), Prof.
Zhuravljov Yu. I., Member of RAS,
Dr. Sci. (Phys.-Math.), Prof.
Kuleshov A. P., Member of RAS,
Dr. Sci. (Tech.), Prof.
Popkov Yu. S., Member of RAS,
Dr. Sci. (Tech.), Prof.
Rusakov S. G., Corresp. Member of RAS,
Dr. Sci. (Tech.), Prof.
Ryabov G. G., Corresp. Member of RAS,
Dr. Sci. (Tech.), Prof.
Soifer V. A., Member of RAS,
Dr. Sci. (Tech.), Prof.
Sokolov I. A., Member of RAS,
Dr. Sci. (Phys.-Math.), Prof.
Suetin N. V.,
Dr. Sci. (Phys.-Math.), Prof.
Chaplygin Yu. A., Member of RAS,
Dr. Sci. (Tech.), Prof.
Shakhnov V. A., Corresp. Member of RAS,
Dr. Sci. (Tech.), Prof.
Shokin Yu. I., Member of RAS,
Dr. Sci. (Tech.), Prof.
Yusupov R. M., Corresp. Member of RAS,
Dr. Sci. (Tech.), Prof.

Editorial Board Members:

Avdoshin S. M., Cand. Sci. (Tech.), Ass. Prof.
Antonov B. I.
Barsky A. B., Dr. Sci. (Tech.), Prof.
Vasenin V. A., Dr. Sci. (Phys.-Math.), Prof.
Vasiliev V. I., Dr. Sci. (Tech.), Prof.
Vishnekov A. V., Dr. Sci. (Tech.), Prof.
Dimitrienko Yu. I., Dr. Sci. (Phys.-Math.), Prof.
Domrachev V. G., Dr. Sci. (Tech.), Prof.
Zaborovsky V. S., Dr. Sci. (Tech.), Prof.
Zarubin V. S., Dr. Sci. (Tech.), Prof.
Karpenko A. P., Dr. Sci. (Phys.-Math.), Prof.
Kolin K. K., Dr. Sci. (Tech.)
Kulagin V. P., Dr. Sci. (Tech.), Prof.
Kureichik V. V., Dr. Sci. (Tech.), Prof.
Ljvovich Ya. E., Dr. Sci. (Tech.), Prof.
Martynov V. V., Dr. Sci. (Tech.), Prof.
Mikhailov B. M., Dr. Sci. (Tech.), Prof.
Nechaev V. V., Cand. Sci. (Tech.), Ass. Prof.
Poleschuk O. M., Dr. Sci. (Tech.), Prof.
Prokhorov S. A., Dr. Sci. (Tech.), Prof.
Saksonov E. A., Dr. Sci. (Tech.), Prof.
Sokolov B. V., Dr. Sci. (Tech.)
Solovyev R. A., Dr. Sci. (Tech.)
Timonina E. E., Dr. Sci. (Tech.), Prof.
Uskov V. L. (USA), Dr. Sci. (Tech.)
Fomichev V. A., Dr. Sci. (Tech.), Prof.
Shilov V. V., Cand. Sci. (Tech.), Ass. Prof.

Editors:

Bezmenova M. Yu.

Complete Internet version of the journal at site: <http://novtex.ru/IT>.

According to the decision of the Higher Certifying Commission of the Ministry of Education of Russian Federation, the journal is inscribed in "The List of the Leading Scientific Journals and Editions wherein Main Scientific Results of Theses for Doctor's or Candidate's Degrees Should Be Published"

П. А. Ляхов, канд. физ.-мат. наук, доц., e-mail: ljahov@mail.ru,
А. С. Ионисян, канд. физ.-мат. наук, доц., e-mail: asion@mail.ru,
М. В. Валуева, аспирант, e-mail: mriya.valueva@mail.ru,
А. С. Ларикова, аспирант, e-mail: larikova@gmail.com,
Северо-Кавказский федеральный университет

Высокопроизводительная цифровая фильтрация на модифицированных умножителях с накоплением в системе остаточных классов с модулями специального вида*

Предложен способ выполнения цифровой фильтрации в системе остаточных классов с использованием модифицированных умножителей с накоплением. Проведен анализ цифровых фильтров, использующих арифметику системы остаточных классов, и представлены результаты аппаратного моделирования на FPGA. Показано, что использование системы остаточных классов позволяет увеличить частоту цифровых фильтров примерно в 4 раза, а аппаратные затраты уменьшить в 3 раза по сравнению с использованием традиционной позиционной системы счисления. Полученные результаты открывают возможность для эффективной аппаратной реализации цифровых фильтров на современных устройствах (FPGA, ASIC и др.) для решения практических задач, таких как шумоподавление, усиление и подавление частот, интерполяция, децимация, эквализация и многих других.

Ключевые слова: цифровая обработка сигналов, цифровой фильтр, модулярные умножители с накоплением, система остаточных классов

Введение

Цифровая фильтрация является ядром цифровой обработки сигналов, так как она лежит в основе решения большинства практических задач этой области: шумоподавления [1], усиления и подавления частот [2], интерполяции [3], децимации [4], эквализации [5] и многих других. Инструментом реализации цифровой фильтрации являются цифровые фильтры (ЦФ), которые принято делить на фильтры с конечной импульсной характеристикой (КИХ-ЦФ) и фильтры с бесконечной импульсной характеристикой (БИХ-ЦФ).

В практике цифровой обработки сигналов существует потребность в увеличении производительности устройств. Одним из методов удовлетворения данной потребности является переход к параллельной архитектуре вычислений. Система остаточных классов (СОК) благодаря

свойству отсутствия межразрядных переносов и параллельному выполнению операции сложения чисел может быть эффективно использована в приложениях с преобладающей долей операций сложения, вычитания и умножения [6, 7]. Выбор набора модулей СОК оказывает большое влияние на производительность цифрового устройства [8, 9].

В данной работе для увеличения производительности и сокращения аппаратных затрат предлагается использовать модули специального вида 2^k , $k \in N$ и $2^k - 1$, $k \in N$, $k > 1$. В статье проведен анализ и представлены результаты аппаратного моделирования на FPGA КИХ-ЦФ, использующих модифицированные умножители с накоплением, в СОК и в традиционной позиционной системе счисления (ПСС).

Система остаточных классов

В СОК числа представляются в базисе попарно взаимно простых чисел, называемых модулями, $\beta = \{m_1, \dots, m_n\}$, $\text{НОД}(m_i, m_j) = 1$ для $i \neq j$.

*Работа выполнена при поддержке Российского фонда фундаментальных исследований (№ 19-07-00130 А), совета по грантам Президента Российской Федерации (проекты СП-126.2019.5 и МК-3918.2021.1.6).

Произведение всех модулей СОК $M = \prod_{i=1}^n m_i$ называется динамическим диапазоном системы. Любое целое число $0 \leq X \leq M$ может быть единственным образом представлено в СОК в виде вектора $\{x_1, x_2, \dots, x_n\}$, где $x_i = |X|_{m_i}$ — это наименьший неотрицательный вычет от целого неотрицательного X по модулю m_i [10].

Динамический диапазон СОК обычно делится на две примерно равные части таким образом, чтобы около половины диапазона представляли положительные числа, а остальную часть диапазона — отрицательные. Таким образом, любое целое число, удовлетворяющее одному из двух соотношений:

$$-\frac{M-1}{2} \leq X \leq \frac{M-1}{2}, \text{ для нечетных } M, \quad (1)$$

$$-\frac{M}{2} \leq X \leq \frac{M}{2} - 1, \text{ для четных } M, \quad (2)$$

может быть представлено в СОК.

Операции сложения, вычитания и умножения в СОК определяются формулами

$$A \pm B = (|a_1 \pm b_1|_{m_1}, \dots, |a_n \pm b_n|_{m_n}); \quad (3)$$

$$A \times B = (|a_1 \times b_1|_{m_1}, \dots, |a_n \times b_n|_{m_n}). \quad (4)$$

Равенства (3)—(4) демонстрируют параллельную природу СОК, свободную от поразрядных переносов.

Восстановление числа X по остаткам $\{x_1, x_2, \dots, x_n\}$ основано на китайской теореме об остатках (КТО) [10]:

$$X = \left| \sum_{i=1}^n \left| M_i^{-1} \right|_{m_i} x_i \right|_M, \quad (5)$$

где $M_i = \frac{M}{m_i}$. Элемент $|M_i^{-1}|_{m_i}$ означает мультипликативный обратный элемент для M_i .

Преимущества представления чисел в СОК могут быть сформулированы следующим образом [8].

1. В СОК отсутствует распространение переноса между арифметическими блоками, и числа большой размерности представляются в виде небольших остатков, что приводит к ускорению в обработке данных.

2. При представлении данных с использованием СОК числа большой размерности кодируются в набор небольших остатков, соответственно, уменьшается сложность арифметических устройств в каждом канале модуля, что облегчает и упрощает работу вычислительной системы.

3. В СОК отсутствует зависимость между арифметическими блоками при выполнении модульных операций, следовательно, ошибка

в одном канале не распространяется на другие. Таким образом, упрощается процесс обнаружения и исправления ошибок.

Однако, несмотря на преимущества, СОК имеет ряд недостатков. Основным недостатком СОК является сложность выполнения операций определения знака, сравнения двух чисел, деления и некоторых других.

Обоснование выбора набора модулей системы остаточных классов

Важной задачей при разработке прикладной системы, использующей вычисления в СОК, является выбор набора модулей $\{m_1, \dots, m_n\}$. В большинстве современных работ, посвященных прикладному применению СОК, используется модуль, равный степени двойки, т. е. 2^k , $k \in N$, где N — множество натуральных чисел. Это объясняется тем, что вычисления по модулю 2^k являются наиболее простыми с точки зрения аппаратной реализации, так как они могут быть реализованы в виде обычных арифметических устройств двоичной системы счисления с шириной k бит. Использование модуля 2^k , а также требование попарной взаимной простоты всех модулей СОК, приводит к тому, что все остальные модули системы должны быть нечетными. Для реализации вычислений, требующих суммирования большого числа слагаемых, необходимо использовать наиболее эффективные техники суммирования по модулю, аналогичные используемым в традиционной двоичной системе счисления. Такие техники суммирования разработаны лишь для модулей вида 2^k , $2^k \pm 1$, $k \in N$ [11].

Практическая реализация вычислительного канала по модулю $2^k + 1$, $k \in N$, требует введения дополнительной логики по методу "diminished-1" для отслеживания нулевой кодовой комбинации, что является нежелательным явлением при разработке системы с минимальными аппаратными и временными затратами [12, 13]. Таким образом, целесообразно использовать лишь модули вида $2^k - 1$, $k \in N$, в качестве нечетных модулей системы.

Модулярные сумматоры

Для модулей вида 2^k эффективными подходами к сложению являются сумматоры с сохранением переноса (carry save adder, CSA) [10] и параллельно-префиксные сумматоры Когге—Стоуна (Kogge—Stone adder, KSA) [14].

Базовым устройством при выполнении арифметических операций в ПСС и СОК явля-

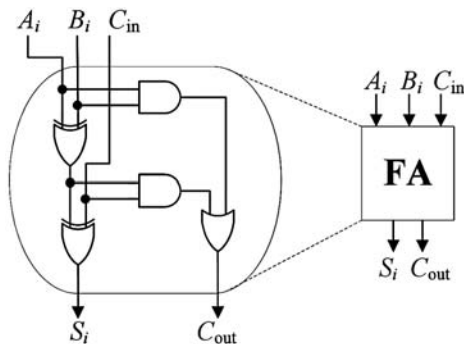


Рис. 1. Логическая схема полного сумматора

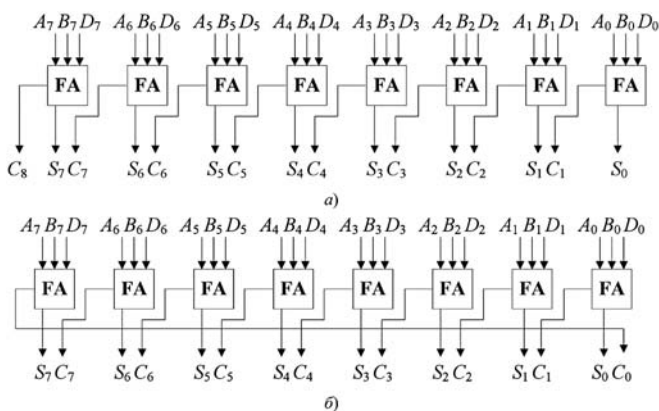


Рис. 2. Логическая схема 8-битного сумматора с сохранением переноса:

a — в ПСС; *б* — по модулю $2^8 - 1$

ется полный сумматор (full adder, FA), изображенный на рис. 1. На вход устройства поступают биты A_i , B_i и C_{in} , которые преобразуются в выходные сигналы S_i и C_{out} по формулам

$$\begin{aligned} S_i &= A_i \oplus B_i \oplus C_{in}, \\ C_{out} &= (A_i \& B_i) \vee (C_{in} \& (A_i \oplus B_i)). \end{aligned} \quad (6)$$

Выходной сигнал S_i является суммой, а выходной сигнал C_{out} — переносом, полученным в полном сумматоре.

На рис. 2 представлен сумматор CSA. Данное устройство преобразует три вектора входных данных A , B и D в два выходных вектора: сумму S и перенос C . При этом количество

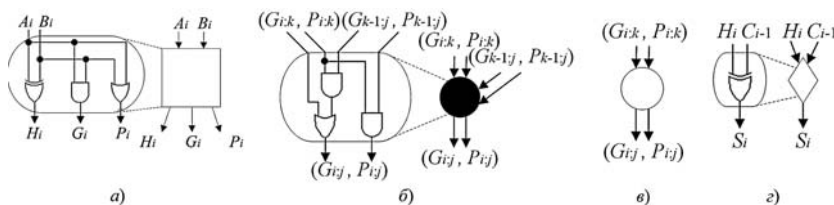


Рис. 3. Устройство блоков сложения параллельно-префиксного сумматора:

a — блок предварительных вычислений; *б, в* — блоки вычисления переносов; *г* — блок вычисления суммы на заключительном этапе

информации для обработки на последующем шаге сокращается в 1,5 раза. На рис. 2, *a* представлена логическая схема сумматора CSA в ПСС, сложение по модулю 2^k проводится аналогично, отличие заключается лишь в обрезке старшего бита переноса. Для сложения по модулю $2^k - 1$ требуется подавать старший бит переноса в первый разряд (рис. 2, *б*), такой сумматор CSA называется сумматором CSA с циклическим переносом (end around carry, EAC) и обозначается EAC-CSA.

Принцип функционирования сумматора KSA для сложения n -битных чисел $A = \sum_{i=0}^{n-1} 2^i A_i$ и $B = \sum_{i=0}^{n-1} 2^i B_i$ заключается в следующем. Сначала осуществляется предварительное вычисление битов G_i , генерирующих перенос, битов P_i , передающих перенос, и полусумм H_i для любого i , $0 \leq i \leq n - 1$ (рис. 3, *a*):

$$G_i = A_i \& B_i, P_i = A_i \vee B_i, H_i = A_i \oplus B_i. \quad (7)$$

Вторая стадия сложения, называемая параллельно-префиксной сетью, вычисляет сигналы переноса C_i для $0 \leq i \leq k - 1$ с использованием G_i и P_i (рис. 3, *б, в*). Для этого используется оператор " $\circ$ ", который связывает пары генерирующих и передающих битов и определен как

$$(G, P) \circ (G', P') = (G \vee (P \& G'), P \& P'). \quad (8)$$

Последовательное вычисление пар генерирующих и передающих битов (G, P) будем обозначать как $(G_{i:j}, P_{i:j})$, $i > j$, где соответствующая пара вычислена на основе битов $i, i - 1, \dots, j$ следующим образом:

$$(G_{i:j}, P_{i:j}) = (G_i, P_i) \circ (G_{i-1}, P_{i-1}) \circ \dots \circ (G_j, P_j). \quad (9)$$

Обозначение " $i : j$ " в формуле (9) показывает вычисление пары (G, P) для i -го бита с учетом всех генерирующих и передающих битов до j -го бита включительно. Так как перенос $C_i = G_{i:0}$ для всех $i > 0$, то все переносы могут быть вычислены с использованием только оператора " $\circ$ " [14].

На третьей стадии вычисляется сумма (рис. 3, *г*)

$$\begin{aligned} S_0 &= H_0 \oplus C_{in}, S_i = H_i \oplus C_{i-1} \\ \text{для } 1 \leq i \leq k - 1, S_k &= C_{k-1}. \end{aligned} \quad (10)$$

На рис. 4, *a* представлена схема 8-битного сумматора KSA в ПСС, сложение по модулю 2^k выполня-

Цифровая фильтрация в системе остаточных классов

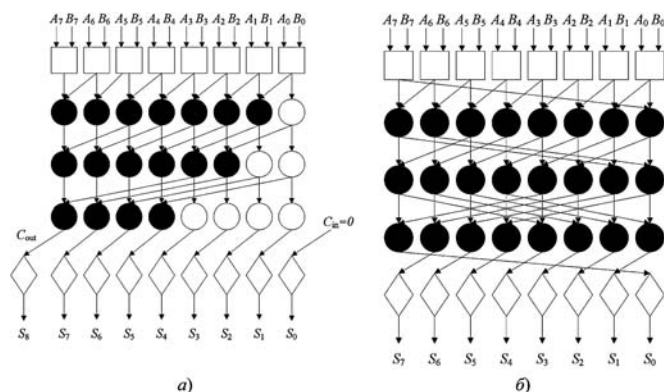


Рис. 4. Структура 8-битного параллельно-префиксного сумматора Когге—Стоуна:

а — для ПСС; б — по модулю $2^8 - 1$

ется аналогично, отличие заключается лишь в обрезке старшего бита переноса. Для сложения по модулю $2^k - 1$ с помощью КСА требуется использование техники ЕАС (рис. 4, б), обозначим такой сумматор как ЕАС-КСА [11].

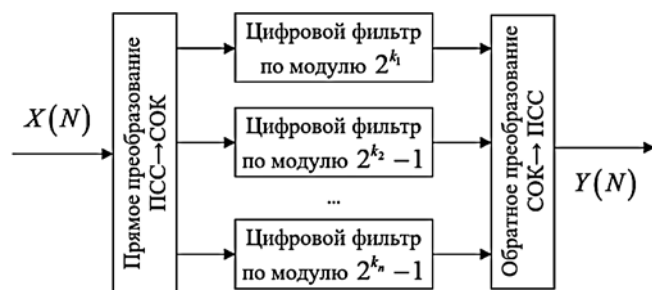


Рис. 5. Схема цифровой фильтрации в СОК $\{2^{k_1}, 2^{k_2} - 1, 2^{k_3} - 1, \dots, 2^{k_n} - 1\}$

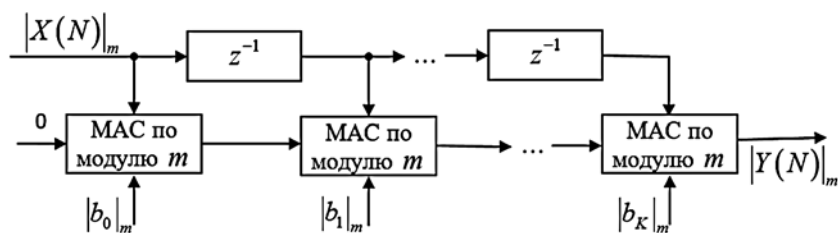


Рис. 6. Схема КИХ-ЦФ по модулю m порядка K на основе блоков МАС по модулю m

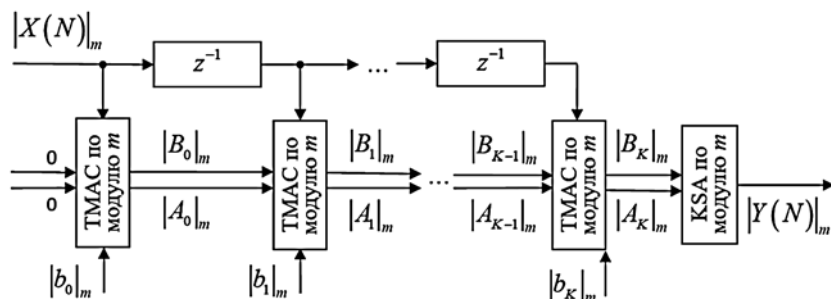


Рис. 7. Схема КИХ-ЦФ по модулю m порядка K на основе блоков ТМАС по модулю m

На рис. 5 представлена схема выполнения цифровой фильтрации в СОК с набором модулей $\{2^{k_1}, 2^{k_2} - 1, 2^{k_3} - 1, \dots, 2^{k_n} - 1\}$. Как видно из рис. 5, для осуществления цифровой фильтрации в указанной СОК необходимо использовать преобразователи ПСС $\rightarrow$ СОК и СОК $\rightarrow$ ПСС, а также представить ЦФ в модулярной форме. На вход КИХ-ЦФ подается последовательность отсчетов сигнала $X(N)$, формируемая аналого-цифровым преобразователем (АЦП) из аналогового сигнала, либо поступающая по вычислительной шине из цифрового источника. На выходе КИХ-ЦФ формируется сигнал $Y(N)$. Для каждого модуля СОК m_i , $i = 1, 2, \dots, n$, ЦФ по модулю выполняет преобразование сигнала по формуле

$$|Y(N)|_{m_i} = \left| \sum_{j=0}^K |b_j|_{m_i} |X(N-j)|_{m_i} \right|_{m_i}, \quad (11)$$

где b_j — коэффициенты фильтра; K — порядок фильтра.

Формула (11) предполагает перевод коэффициентов фильтра в модулярное представление и выполнение всех операций по модулю. Модулярное представление коэффициентов фильтра может быть получено путем преобразования его коэффициентов в формат с фиксированной точкой, с последующим масштабированием в диапазон целых чисел и нахождением остатков от деления на модуль. После того, как коэффициенты фильтра представлены в модулярной форме, фильтрация по формуле (11) может быть выполнена по схеме, изображенной на рис. 6.

Символами z^{-1} обозначены блоки задержки сигнала на один отсчет, которые на практике реализуются с помощью буферов. Сумматоры и умножители по модулю объединены в один блок — умножитель с накоплением по модулю (multiply and accumulate, MAC) [10].

В данной работе предлагается заменить МАС-блоки на модифицированные умножители с накоплением (truncated multiply and accumulate, ТМАС). На рис. 7 представлена схема КИХ-ЦФ по модулю m с использованием ТМАС-блоков по модулю m . Принцип работы блока ТМАС по модулю 2^k не отличается от принципа работы блока ТМАС в ПСС (рис. 8, а). Един-

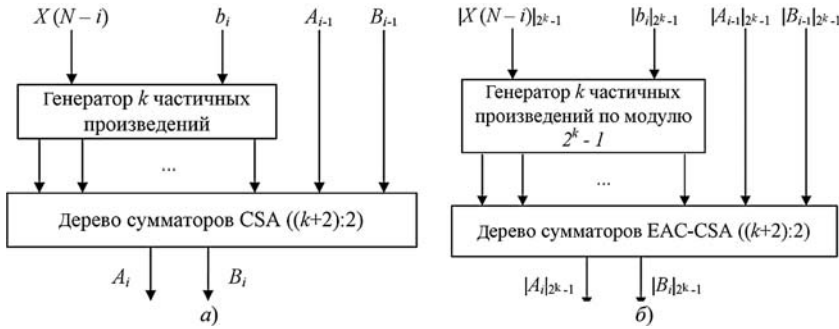


Рис. 8. Структура блока ТМАС:
а — в ПСС; б — по модулю $2^k - 1$

ственным отличием блока ТМАС по модулю 2^k является отсутствие устройств и шин для передачи старших значащих битов и переносов, начиная с k -го. Для генерации частичных произведений используется массив вентилях AND размером $k \times k$.

Принцип работы ТМАС-блока по модулю $2^k - 1$ схож с аналогом из ПСС, с той лишь разницей, что все старшие значащие переносы должны подаваться циклически на младший разряд, т. е. должна использоваться техника ЕАС. В дальнейшем будем обозначать такой блок ЕАС-ТМАС. Схема такого устройства, представлена на рис. 8, б. С помощью обозначения $((k+2):2)$ показано, что на вход дерева сумматоров ЕАС-CSA подаются $(k+2)$ слагаемых, а на выходе формируются два слагаемых [15].

Оценка параметров цифрового фильтра

Для оценки параметров цифровых устройств будем использовать абстрактную модель подсчета задержки и площади СБИС, известную как "unit-gate"-модель [16]. Если обозначить рассчитанную по указанной модели задержку логического устройства U_{delay} , а площадь логического устройства обозначить U_{area} , то будем иметь следующее описание для логических вентилях:

$$U_{\text{delay}}(NOT) = 0, U_{\text{area}}(NOT) = 0; \quad (12)$$

$$U_{\text{delay}}(AND) = 1, U_{\text{area}}(AND) = 1; \quad (13)$$

$$U_{\text{delay}}(OR) = 1, U_{\text{area}}(OR) = 1; \quad (14)$$

$$U_{\text{delay}}(XOR) = 2, U_{\text{area}}(XOR) = 2; \quad (15)$$

$$U_{\text{delay}}(XNOR) = 2, U_{\text{area}}(XNOR) = 2, \quad (16)$$

где вентиль *NOT* выполняет операцию отрицания, *AND* — операцию конъюнкции, *OR* — операцию дизъюнкции, *XOR* — операцию

сложения по модулю два, *XNOR* — операцию эквивалентности.

Тогда, с учетом формул (6) и (12)–(16), задержка и площадь *FA* может быть записана как

$$U_{\text{delay}}(FA) = 4, U_{\text{area}}(FA) = 7. \quad (17)$$

Параметры задержки и площади CSA и ЕАС-CSA совпадают и описываются следующим образом:

$$U_{\text{delay}}(CSA) = U_{\text{delay}}(FA) = 4; \quad (18)$$

$$U_{\text{area}}(CSA) = k U_{\text{area}}(FA) = 7k. \quad (19)$$

Для сумматоров *KSA* при выполнении условия $C_{\text{in}} = 0$, не требующего логической операции $\oplus$ вычисления S_0 по формуле (10), параметры задержки и площади *KSA* определяются по формулам

$$U_{\text{delay}}(KSA) = 2 + 2 \cdot \lceil \log_2 k \rceil + 2 \approx 2 \log_2 k + 4; \quad (20)$$

$$U_{\text{area}}(KSA) = 4k + 3(k \cdot \lceil \log_2 k \rceil - (2^{\lceil \log_2 k \rceil} - 1)) + 2(k - 1) \approx 3k \log_2 k + 3k + 1. \quad (21)$$

Знак приближенного равенства $\approx$ в формулах (20), (21) означает допущение $\lceil \log_2 k \rceil \approx \log_2 k$ и не вносит погрешности при рассмотрении наиболее распространенных на практике случаев суммирования 8-битных, 16-битных, 32-битных и т. д. чисел. Параметры задержки и площади ЕАС-*KSA* определяются по формулам

$$U_{\text{delay}}(EAC - KSA) = U_{\text{delay}}(KSA) \approx 2 \log_2 k + 4; \quad (22)$$

$$U_{\text{area}}(EAC - KSA) = 4k + 3(k \cdot \lceil \log_2 k \rceil) + 2k \approx 3k \log_2 k + 6k. \quad (23)$$

Таким образом, параметры U_{delay} и U_{area} блока ТМАС имеют вид

$$U_{\text{delay}}(TMAC) \approx 6,8 \log_2 k + 1; \quad (24)$$

$$U_{\text{area}}(TMAC) \approx k^2 + 7k^2 = 8k^2. \quad (25)$$

Параметры U_{delay} и U_{area} блока ЕАС-ТМАС имеют вид

$$U_{\text{delay}}(EAC - TMAC) \approx 6,8 \log_2 k + 1; \quad (26)$$

$$U_{\text{area}}(EAC - TMAC) \approx k^2 + 7k^2 = 8k^2. \quad (27)$$

Задержка и площадь вычислительной части КИХ-ЦФ, показанного на рис. 8, равна сумме задержек и площадей ТМАС-блоков и сумматора KSA, соответственно. Если обозначить $FIR_{TMAC}^{K,k}$ вычислительную часть КИХ-ЦФ K -го порядка с k -битными коэффициентами на основе ТМАС-блоков, то

$$\begin{aligned} U_{\text{delay}}(FIR_{TMAC}^{K,k}) &= \\ &= (K+1)U_{\text{delay}}(TMAC) + U_{\text{delay}}(KSA) \approx \quad (28) \\ &\approx 6,8K \log_2 k + 8,8 \log_2 k + K + 5, \end{aligned}$$

$$\begin{aligned} U_{\text{area}}(FIR_{TMAC}^{K,k}) &= \\ &= (K+1)U_{\text{area}}(TMAC) + U_{\text{area}}(KSA) \approx \quad (29) \\ &\approx 3k \log_2 k + 8k^2 K + 8k^2 + 3k + 1. \end{aligned}$$

Если обозначить $FIR_{EAC-TMAC}^{K,k}$ вычислительную часть КИХ-ЦФ K -го порядка с k -битными коэффициентами по модулю $2^k - 1$ на основе EAC-TMAC-блоков, то

$$\begin{aligned} U_{\text{delay}}(FIR_{EAC-TMAC}^{K,k}) &= \\ &= (K+1)U_{\text{delay}}(EAC - TMAC) + \quad (30) \\ &\quad + U_{\text{delay}}(EAC - KSA) \approx \\ &\approx 6,8K \log_2 k + 8,8 \log_2 k + K + 5, \end{aligned}$$

$$\begin{aligned} U_{\text{area}}(FIR_{EAC-TMAC}^{K,k}) &= \\ &= (K+1)U_{\text{area}}(EAC - TMAC) + \quad (31) \\ &\quad + U_{\text{area}}(EAC - KSA) \approx \\ &\approx 3k \log_2 k + 8k^2 K + 8k^2 + 6k. \end{aligned}$$

Сравнительный анализ цифровых фильтров

Для сравнения параметров задержки и скорости фильтров в различных системах счисления рассмотрим устройства одинакового порядка K в ПСС и СОК вида $\{2^{k_1}, 2^{k_2} - 1, 2^{k_3} - 1, \dots, 2^{k_n} - 1\}$, содержащие 3, 4 и 5 модулей, как наиболее распространенные случаи [9]. Несмотря на то, что при $n > 2$ невозможно существование идеально сбалансированного случая $k_1 = k_2 = \dots = k_n$, мы будем считать, что все эти степени примерно равны между собой. Это означает, что разрядность обрабатываемых данных в трехмодульной СОК будет приблизительно в три раза меньше, чем в ПСС соответствующего диапазона. Аналогичная картина будет наблюдаться в четырех- и пятимодульных СОК: уменьшение разрядности в 4 и 5 раз, соответственно. При расчетах не будут учитываться блоки прямого ПСС $\rightarrow$ СОК и обратного СОК $\rightarrow$ ПСС преобразований, так как их реализация является отдельной проблемой, не связанной непосредственно с ЦФ [17].

Для сравнительного анализа фильтров в ПСС и СОК зафиксируем поочередно порядок и разрядность фильтра в ПСС. Рассмотрим сначала случай фильтра 15-го порядка, т. е. $K = 15$. Для рассмотренного случая будем изменять разрядность k , перебирая наиболее популярные форматы данных: 8, 16, 32 и 64 бита. Для указанных форматов будем полагать в СОК $k_i > \frac{k}{n}$, т. е. $k_i \approx 3, 6, 11$ и 22 бит для трехмодульной СОК, $k_i \approx 3, 5, 9$ и 17 бит для четырехмодульной СОК, $k_i \approx 2, 4, 7$ и 13 бит для пятимодульной СОК. В табл. 1 приведены значения параметра U_{delay} для перечисленных случаев. В табл. 2 приведены значения параметра U_{area} для перечисленных случаев. Анализ данных, представленных в табл. 1 и 2, показывает, что применение СОК с тремя модулями позволяет сократить время работы КИХ-ЦФ 15-го порядка в 1,3...1,8 раз и аппаратные затраты на его реализацию в 2,3...2,8 раз. СОК с четырьмя модулями позволяет сократить время работы КИХ-ЦФ 15-го порядка в 1,4...1,8 раз и аппаратные затраты на его реализацию в 1,7...3,5 раза. СОК с пятью модулями позволяет сократить время работы КИХ-ЦФ 15-го порядка в 1,5...2,6 раз и аппаратные затраты на его реализацию в 3,1...4,8 раз. В целом, большее число модулей СОК обеспечивает большее преимущество по скорости работы и экономии аппаратных затрат. Кроме того, можно отметить, что с ростом разрядности обрабатываемых данных преиму-

Таблица 1

Значения параметра U_{delay} для КИХ-ЦФ 15-го порядка в ПСС и СОК с различным числом модулей

Разрядность данных, k	ПСС	СОК		
		3 модуля	4 модуля	5 модулей
8	352	196	196	131
16	463	306	277	242
32	574	403	371	331
64	685	514	473	430

Таблица 2

Значения параметра U_{area} для КИХ-ЦФ 15-го порядка в ПСС и СОК с различным числом модулей

Разрядность данных, k	ПСС	СОК		
		3 модуля	4 модуля	5 модулей
8	8289	3553	4737	2650
16	33009	14072	13059	10480
32	131649	47004	42030	31865
64	525633	187135	149210	109272

щества в скорости фильтрации в СОК несколько ослабевают, а преимущества в экономии аппаратных средств, наоборот, усиливаются.

В качестве второго подхода к анализу производительности фильтров в ПСС и СОК с различным числом модулей зафиксируем разрядность обрабатываемых данных в ПСС $k = 16$ и будем варьировать порядки КИХ-ЦФ: $K = 3, 7, 15$ и 31 . Снова будем полагать в СОК $k_i > \frac{k}{n}$, т. е. для трехмодульной СОК $k_i = 6$, для четырехмодульной СОК $k_i = 5$ и, наконец, для пятимодульной СОК $k_i = 4$. В табл. 3 приведены значения параметра U_{delay} для перечисленных случаев. В табл. 4 приведены значения параметра U_{area} для перечисленных случаев. Анализ данных, представленных в табл. 3 и 4, показывает, что применение СОК с тремя модулями позволяет сократить время работы КИХ-ЦФ с $k = 16$ в 1,4...1,5 раза и аппаратные затраты на его реализацию в 2,2...2,3 раза. СОК с четырьмя модулями позволяет сократить время работы КИХ-ЦФ с $k = 16$ в 1,6 раза и аппаратные затраты на его реализацию в 2,4...2,5 раза. СОК с пятью модулями позволяет сократить время работы КИХ-ЦФ с $k = 16$ в 1,8...1,9 раза и аппаратные затраты на его реализацию в 3,0...3,1 раза. Можно отметить, что рост порядка фильтра несколько увеличивает преимущества в скорости фильтрации и в экономии аппаратных средств при использовании СОК.

Аппаратное моделирование цифровых фильтров в системе остаточных классов

Аппаратное моделирование проведено на FPGA Artix xc7a200tffg1156-3 в Xilinx Vivado 18.3 с использованием языка описания аппаратуры VHDL.

Целью моделирования было сравнение технических характеристик КИХ-ЦФ, реализованных в ПСС и в СОК с различными наборами модулей. Для достижения данной цели было проведено аппаратное моделирование устройств, для которых был проведен теоретический анализ данных в табл. 1—4.

В табл. 5 представлены результаты аппаратного моделирования КИХ-ЦФ 15-го порядка с различными разрядностями. Применение СОК с тремя модулями позволяет увеличить частоту КИХ-ЦФ 15-го порядка в 2,0...2,5 раза и аппаратные затраты на его реализацию в 1,3...2,3 раза, при увеличении энергопотребления на 6...19 %. СОК с четырьмя модулями позволяет увеличить частоту КИХ-ЦФ 15-го порядка в 2,1...3,1 раза и аппаратные затраты на

его реализацию в 1,4...2,9 раза, при увеличении энергопотребления на 11...16 %. СОК с пятью модулями позволяет увеличить частоту КИХ-ЦФ 15-го порядка в 2,0...4,2 раза и аппаратные затраты на его реализацию в 1,1...2,6 раза, при увеличении энергопотребления на 7...33 %.

Результаты аппаратного моделирования, представленные в табл. 6, показывают, что применение СОК с тремя модулями позволяет увеличить частоту КИХ-ЦФ с $k = 16$ в 1,9...2,2 раза и сократить аппаратные затраты на его реализа-

Таблица 3
Значения параметра U_{delay} для КИХ-ЦФ в ПСС с $k = 16$ и СОК с различным числом модулей

Порядок фильтра, K	ПСС	СОК		
		3 модуля	4 модуля	5 модулей
3	125	83	76	66
7	238	158	143	125
15	463	306	277	242
31	914	604	546	475

Таблица 4
Значения параметра U_{area} для КИХ-ЦФ в ПСС с $k = 16$ и СОК с различным числом модулей

Порядок фильтра, K	ПСС	СОК		
		3 модуля	4 модуля	5 модулей
3	8433	3704	3459	2800
7	16625	7160	6659	5360
15	33009	14072	13059	10480
31	65777	27896	25859	20720

Таблица 5
Результаты моделирования для КИХ-ЦФ 15-го порядка в СОК с различным числом модулей

Характеристики	k	Система счисления			
		СОК			ПСС
		3 модуля	4 модуля	5 модулей	
Максимальная частота, МГц	16	278	285	283	139
	32	145	200	212	71
	64	71	90	123	29
Число LUT	16	638	588	740	801
	32	1644	1603	1388	2637
	64	4162	3348	3767	9645
Энергопотребление, Вт	16	0,335	0,353	0,413	0,315
	32	0,390	0,441	0,425	0,396
	64	0,464	0,445	0,560	0,376

Таблица 6

Результаты моделирования для КИХ-ЦФ в ПСС с $k = 16$
и СОК с различным числом модулей

Характеристики	K	Система счисления			
		СОК			ПСС
		3 модуля	4 модуля	5 модулей	
Максимальная частота, МГц	3	295	315	310	149
	7	292	315	305	132
	15	278	285	283	139
	31	258	279	266	135
Число LUT	3	241	193	261	433
	7	371	332	422	426
	15	638	588	740	801
	31	1157	1097	1382	1283
Энергопотребление, Вт	3	0,372	0,323	0,338	0,337
	7	0,386	0,340	0,377	0,331
	15	0,335	0,353	0,413	0,315
	31	0,433	0,456	0,469	0,372

цию на 10...44 %, при увеличении энергопотребления на 6...17 %. СОК с четырьмя модулями позволяет увеличить частоту КИХ-ЦФ с $k = 16$ в 2...2,4 раза и сократить аппаратные затраты на его реализацию на 15...55 % раз, при увеличении энергопотребления на 3...23 %. СОК с пятью модулями позволяет увеличить частоту КИХ-ЦФ с $k = 16$ примерно в два раза и сократить аппаратные затраты на его реализацию на 1...40 %, при увеличении энергопотребления на 0,3...31 %.

Разница в теоретических и практических результатах объясняется особенностями FPGA и недостатком "unit-gate" модели, который заключается в игнорировании эффектов нагрузочной способности выходов как отдельных логических элементов, так и микросхемы в целом.

Заключение

В работе рассмотрено выполнение цифровой фильтрации в СОК с использованием модифицированных умножителей с накоплением ТМАС. Результаты аппаратного моделирования на FPGA показали, что использование СОК позволяет увеличить частоту ЦФ примерно в четыре раза, а аппаратные затраты уменьшить в три раза, при увеличении энергопотребления на 33 %, по сравнению с реализацией ЦФ в ПСС.

Полученные результаты могут быть использованы для эффективной аппаратной реализации ЦФ на современных микроэлектронных устройствах, таких как FPGA и ASIC, и для решения практических задач цифровой обработки сигналов.

Список литературы

1. Bhaskar P. C., Bhaskar P. C., Uplane M. D. FPGA based digital FIR multilevel filtering for ECG Denoising // 2015 International Conference on Information Processing (ICIP). Pune. 2015. P. 733—738.
2. Хуако Р. А. Исследование возможности построения одноантенного ретранслятора с коэффициентом усиления больше единицы // Инфокоммуникационные технологии. 2012. Т. 10, № 2. С. 76—80.
3. Porshnev S. V., Kusaykin D. V., Klevakin M. A. On accuracy of periodic discrete finite-length signal reconstruction by means of a Whittaker-Kotelnikov-Shannon interpolation formula // 2018 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBERIT). Yekaterinburg, 2018. P. 165—168.
4. Tang F. et al. An Area-Efficient Column-Parallel Digital Decimation Filter With Pre-BWI Topology for CMOS Image Sensor // IEEE Transactions on Circuits and Systems I: Regular Papers. Aug. 2018. Vol. 65, N. 8. P. 2524—2533.
5. Kiran S. et al. Modeling of ADC-Based Serial Link Receivers With Embedded and Digital Equalization // IEEE Transactions on Components, Packaging and Manufacturing Technology. March 2019. Vol. 9, N. 3. P. 536—548.
6. Omondi A., Premkumar B. Residue Number Systems: Theory and Implementation. Imperial College Press, 2007. 296 p.
7. Червяков Н. И., Сахнюк П. А., Шапошников П. А., Ряднов С. А. Модулярные параллельные вычислительные структуры нейропроцессорных систем. М.: ФИЗМАТЛИТ, 2003. 288 с.
8. Акушский И. Я., Юдицкий Д. И. Машинная арифметика в остаточных классах. М.: Советское радио, 1968. 440 с.
9. Molahosseini A. S., Sorouri S., Zarandi A. A. Research challenges in next-generation residue number system architectures // 2012, 7th International Conference on Computer Science & Education (ICCSE). IEEE, 2012. P. 1658—1661.
10. Parhami B. Computer Arithmetic: Algorithms and Hardware Designs. Oxford University Press, Inc., 2000. 492 p.
11. Vergos H. T., Dimitrakopoulos G. On Modulo $2^n + 1$ Adder Design // IEEE Transactions on Computers. 2012. Vol. 61, N. 2. P. 173—186.
12. Rao K. R., Yip P. C. The Transform and Data Compression Handbook. CRC press, 2001. 399 p.
13. Živaljević D., Stamenković N., Stojanović V. Digital filter implementation based on the RNS with diminished-1 encoded channel // 2012 35th International Conference on Telecommunications and Signal Processing (TSP). Prague, 2012. P. 662—666.
14. Kogge P. M., Stone H. S. A Parallel Algorithm for the Efficient Solution of a General Class of Recurrence Equations // IEEE Trans. Comput. 1973. Vol. C-22, N. 8. P. 786—793.
15. Patel R. A. et al. Novel power-delay-area-efficient approach to generic modular addition // IEEE Transactions on Circuits and Systems I: Regular Papers. 2007. Vol. 54, N. 6. P. 1279—1292.
16. Zimmermann R. Binary adder architectures for cell-based VLSI and their synthesis. Hartung-Gorre, 1998.
17. Амербаев В. М., Соловьев Р. А., Тельпухов Д. В., Балака Е. С. Построение обратных преобразователей модулярной арифметики с коррекцией ошибок на базе полиадического кода // Нейрокомпьютеры: разработка, применение. 2014. № 9. С. 30—35.

P. A. Lyakhov, PhD of Physics and Mathematics Sciences, Assistant Professor,
Department of Applied Mathematics and Mathematical Modeling, e-mail: ljahov@mail.ru,
A. S. Ionisyan, PhD of Physics and Mathematics Sciences, Assistant Professor,
Department of Applied Mathematics and Mathematical Modeling, e-mail: asion@mail.ru,
M. V. Valueva, PhD student, e-mail: mriya.valueva@mail.ru,
A. S. Larikova, PhD student, e-mail: larikova@gmail.com,
North-Caucasus Federal University, Stavropol, 357736, Russian Federation

High Performance Digital Filtering on Modified Multiply and Accumulate Unit in Residue Number System with Special Type of Moduli

The paper proposes the implementation of digital filtering using residue number system and the modified truncated multiply and accumulate unit. The work was carried out a theoretical analysis of digital filters using residue number system arithmetic and implemented hardware simulation on FPGA. FPGA hardware simulation results show that the use of residue number system allows to increase the frequency of digital filters up to about 4 times and hardware costs reduce up to 3 times compared to using a common positional number system. The obtained results open up the possibility for efficient hardware implementation of digital filters on modern devices (FPGA, ASIC and etc.) to solve practical problems, such as noise reduction, amplification and suppression of the frequency spectrum, interpolation, decimation, equalization and many others.

Keywords: digital signal processing, digital filter, multiply and accumulate unit, residue number system

Acknowledgments: This work was supported by the Russian Federation for Basic Research (projects no. 19-07-00130A) and by the Presidential Grant of the Russian Federation (projects no. SP-126.2019.5 and no. MK-3918.2021.1.6).

DOI: 10.17587/it.27.171-179

References

1. Bhaskar P. C., Bhaskar P. C., Uplane M. D. FPGA based digital FIR multilevel filtering for ECG Denoising, *2015 International Conference on Information Processing (ICIP)*, Pune, 2015, pp. 733–738.
2. Huako R. A. Issledovanie vozmozhnosti postroeniya odno-antennogo retranslyatora s koefitsientom usileniya bol'she edinici, *Infokommunikacionnye tehnologii*, 2012, vol. 10, no. 2, pp. 76–80 (in Russian).
3. Porshnev S. V., Kusaykin D. V., Klevakin M. A. On accuracy of periodic discrete finite-length signal reconstruction by means of a Whittaker-Kotelnikov-Shannon interpolation formula, *2018 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT)*, Yekaterinburg, 2018, pp. 165–168.
4. Tang F., Wang Z., Xia Y., Liu F., Zhou X., Hu S., Lin Z., Bermak A. An Area-Efficient Column-Parallel Digital Decimation Filter With Pre-BWI Topology for CMOS Image Sensor, *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2018, vol. 65, no. 8, pp. 2524–2533.
5. Kiran S., Shafik A., Tabasy E. Z., Cai S., Lee K., Hoyos S., Palermo S. Modeling of ADC-Based Serial Link Receivers With Embedded and Digital Equalization, *IEEE Transactions on Components, Packaging and Manufacturing Technology*, March 2019, vol. 9, no. 3, pp. 536–548.
6. Omondi A., Premkumar B. Residue Number Systems: Theory and Implementation, Imperial College Press, 2007, p. 296.
7. Chervyakov N. I., Sahnyuk P. A., Shaposhnikov A. V., Ryadnov S. A. Modul'yarnie parallel'nye vychisleniya strukturi neyriprocessornih system, Moscow, FIZMATLIT, 2003, 288 p. (in Russian).
8. Akushskij I. Y., Yudichij D. I. Machine arithmetic in residual classes, Moscow, Sovetskoe radio, 1968, 440 p. (in Russian).
9. Molahosseini A. S., Sorouri S., Zarandi A. A. Research challenges in next-generation residue number system architectures, *7th International Conference on Computer Science & Education (ICCSE)*, IEEE, 2012, pp. 1658–1661.
10. Parhami B. Computer Arithmetic: Algorithms and Hardware Designs, Oxford University Press, Inc., 2000, 492 p. (in Russian).
11. Vergos H. T., Dimitrakopoulos G. On Modulo $2^n + 1$ Adder Design, *IEEE Transactions on Computers*, 2012, vol. 61, no. 2, pp. 173–186.
12. Rao K. R., Yip P. C. The Transform and Data Compression Handbook, CRC press, 2001, 399 p.
13. Živaljević D., Stamenković N., Stojanović V. Digital filter implementation based on the RNS with diminished-1 encoded channel, *35th International Conference on Telecommunications and Signal Processing (TSP)*, Prague, 2012, pp. 662–666.
14. Kogge P. M., Stone H. S. A Parallel Algorithm for the Efficient Solution of a General Class of Recurrence Equations, *IEEE Trans. Comput.*, 1973, vol. C-22, no. 8, pp. 786–793.
15. Patel R. A. et al. Novel power-delay-area-efficient approach to generic modular addition, *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2007, vol. 54, no. 6, pp. 1279–1292.
16. Zimmermann R. Binary adder architectures for cell-based VLSI and their synthesis, Hartung-Gorre, 1998.
17. Amerbaev V. M., Solov'ev R. A., Tel'pukhov D. V., Balaka E. S. Construction of inverse converters of modular arithmetic with error correction based on polyadic code, *NeuroKomp'yutery: razrabotka, primeneniye*, 2014, no. 9, pp. 30–35.

С. В. Шайтура, канд. техн. наук, доц., e-mail: swshaytura@gmail.com,
Российский государственный университет транспорта (МИИТ), Москва,
Институт гуманитарных наук, экономики и информационных технологий, г. Бургас, Болгария,
Д. А. Галкин, управляющий партнер, e-mail: julikas@inbox.ru,
Компания "Marketing Logic", Москва

Геомаркетинговый анализ больших данных

Накопление большого количества геопространственных данных требует новых подходов к их обработке и визуализации. Одним из таких подходов является создание геомаркетинговой системы с принципиально новым инструментарием, основанном на кластеризации данных. Возможности такой системы показаны на примерах оценки стоимости жилья, определения местоположения нового торгового центра, филиала банка и поликлиники.

Ключевые слова: геомаркетинговая система, пространственные данные, геомаркетинг, большие данные, кластеры, технологии геомаркетинга, информационные технологии

Введение

Цифровизация народного хозяйства и бурное развитие глобальных информационных сетей привело к появлению понятия больших данных [1–3]. Большинство данных, с которыми работают компании, имеют географические привязки. Например, продажи клиентам определяются географическим расположением целевых групп. Все клиенты имеют 2...3 адреса (проживания, регистрации, работы, учебы) и передвигаются достаточно логично между основными адресами. Выявление данных закономерностей позволяет повысить качество принимаемых решений и эффективность действий по текущим и потенциальным рынкам, клиентам, медиапланированию, способствующих росту продаж.

Если данные привязаны к пространственно-временному континууму, то в этом случае говорят о пространственно-распределенных или геопространственных данных. Источником больших геопространственных данных являются данные дистанционного зондирования Земли и космического мониторинга, Интернет, разнородная не цифровая информация [4–8]. Понятие больших данных включает в себя не только большие объемы цифровых данных, но и средства работы с ними. Для обработки больших геопространственных дан-

ных используются, прежде всего, геоинформационные системы и методы геопространственного анализа: краудсорсинг, классификация, агрегация и интеграция разнородных данных, прогнозная аналитика, машинное обучение, распознавание образов, визуализация аналитических данных [9–14].

Имеется огромное число задач геопространственного анализа. Это поиск расположения различных мест, адресов, координат, и наоборот, объектов, которые привязаны к данным координатам на месте [15–20]. Этими задачами занимаются пространственные геопорталы типа Яндекс, Гугл, Бинг и другие. Также имеется большое число специализированных геопорталов: глобальных, региональных, муниципальных, которые решают задачи геопространственного анализа на данной территории. При этом для уточнения и обогащения цифрового содержания карт часто применяется метод краудсорсинга — привлечение широкого круга пользователей к созданию и обобщению контента.

Таким образом, создается своеобразный компьютерный мозг, который содержит огромное количество данных. Используя эти данные, можно решать многие частные задачи геомаркетинга, например, определение местоположения нового торгового центра, филиала банка, поликлиники и так далее [21–26].

1. Геомаркетинговые информационные системы

Геомаркетинговые информационные системы (ГМС) возникли на основе интеграции геоинформационных систем и маркетинговых информационных систем. Применение геомаркетинговых систем и технологий целесообразно там, где возникает потребность работы с пространственно-локализованными данными или необходимо применять для поддержки принятия решений тематические карты с деловой графикой. Пространственная локализация может быть грубой или точной.

ГМС позволяют применить визуальные методы представления и обработки статистической информации для поддержки принятия решений. Сюда следует отнести возможность обобщения (генерализации) однородных качественных явлений и характеристик. Визуальный анализ данных на 2...3 порядка быстрее анализа табличных данных, особенно при контроле критических или аномальных ситуаций.

Перечисленные дополнительные возможности обработки определяют эффективность применения геомаркетинга как рыночной информационной технологии особенно при анализе пространственно-распределенных характеристик рынка: политической ситуации, демографической ситуации, экономической ситуации, транспортных сетей, туристических маршрутов и т. д.

Геомаркетинг представляет собой технологию проведения маркетинговых исследований для принятия стратегических, концептуальных и управленческих решений, основанную на методах географического анализа различных пространственно-распределенных объектов и явлений. Такие исследования позволяют определить целевую аудиторию в нужной территориальной единице, провести конкурентный анализ, определить наилучшее местоположение нового объекта, дать прогноз оборота коммерческой недвижимости, разработать концепцию для существующего или планируемого объекта, оценить наилучшее использование земельного участка и многое другое. Геомаркетинговое исследование позволяет проанализировать внешние и внутренние геопространственные (имеющие географическую привязку) показатели компании, различные аспекты ее прошлой, текущей и будущей деятельности, включая инфраструктуру и конкурентную среду. Поскольку в течение геомаркетингового исследования заказчику необходимо предоставлять разнообразные данные компании (например, о выручке за несколько лет, о смене работников и т. д.), заказчик и ис-

полнитель максимально близко контактируют в процессе выполнения исследования. Это является несомненным плюсом геомаркетинга, ведь разного рода специалисты и эксперты, к которым обращаются за помощью бизнесмены, выполняют свою работу и просто передают работающий механизм заказчику.

2. Координатная привязка и кластеризация данных

Поскольку данные имеют координатную привязку, и это главное, что их объединяет, целесообразно координаты использовать как основные оси многомерного куба данных. При этом в качественном анализе точность привязки часто варьируется, поскольку усредненные данные принадлежат не к какой-то точке, а к территории. В задачах, где не требуются точное геокодирование данных, применяют модели кластерных растровых данных. Этот метод позволяет обрабатывать данные с заранее заданной точностью, избегая ненужного использования компьютерных мощностей. Так, например, спутниковый снимок Земли Landsat имеет 74 000 000 ячеек растра. Обработка такого количества данных является весьма трудоемкой задачей.

Кластеризация (или кластерный анализ) — это техника машинной обработки данных, задача которого состоит в разделении данных на несколько групп по определенным признакам. Цель кластерного анализа данных заключается в том, чтобы определить группы с одинаковыми характеристиками и распределить их по кластерам. В настоящее время актуальность такого процесса обуславливается большим потоком данных, которые необходимо грамотно структурировать для дальнейшей обработки.

Для того чтобы наглядно показать различия типов кластеризации, рассмотрим их на примере магазина "X". Предположим, директор магазина "X" заинтересован в том, чтобы проанализировать предпочтения клиентов для расширения бизнеса. Невозможно взглянуть на запросы каждого клиента и разработать уникальную бизнес-стратегию для каждого из них. С помощью статистического анализа возможно объединение всех своих клиентов, например, в 10 групп на основе их покупательских привычек. Далее можно разработать отдельные стратегии для клиентов 10 групп. Выделяют два типа кластеризации:

- *жесткая кластеризация*. В жесткой кластеризации каждая точка данных либо полностью

принадлежит кластеру, либо нет. Например, в приведенном выше примере каждый клиент помещается в одну группу из 10 групп;

- *мягкая кластеризация.* В мягкой кластеризации вместо того, чтобы помещать каждую точку данных в отдельный кластер, определяется вероятность того, что эта точка данных окажется в этих кластерах. Например, из приведенного выше сценария каждому клиенту назначается вероятность быть в любом из 10 кластеров магазина "Х".

Для реализации задачи кластеризации существует множество инструментов. Каждому методу следует различный набор правил для определения "сходства" между точками данных. В настоящее время известно более 100 алгоритмов кластеризации.

Иерархическая кластеризация представляет собой алгоритм, который строит иерархию кластеров. Этот алгоритм начинается со всех точек данных, назначенных их кластеру. Затем два ближайших кластера объединяются в один. Алгоритм подойдет к логическому завершению, когда остается только один кластер.

Результаты иерархической кластеризации могут быть показаны с помощью дендрограммы (графический метод, отображенный в виде дерева, построенного по матрице мер близости). Дендрограмму иллюстрирует рис. 1.

При квадротомическом представлении вся поверхность Земли представляется в виде древовидной структуры с разбиением каждого участка на 4 квадрата. Такое представление данных было использовано для создания стандартного протокола предварительно созданных и обработанных в реальном времени фрагментов карт с географической привязкой через сеть Интернет (рис. 2, см. вторую сторону обложки).

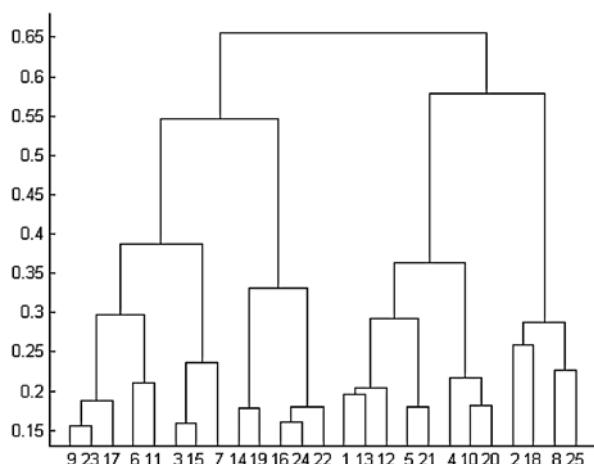


Рис. 1. Дендрограмма иерархической кластеризации

Кластерный подход позволяет использовать геопространственные данные с разной степенью точности пространственных координат.

3. Оценка привлекательности жилья на основе геомаркетинговых систем

При оценке привлекательности жилья следует учитывать ряд факторов [9, 10, 21]. Общее правило очевидно: с течением времени потребительские свойства недвижимости последовательно теряются, что приводит к снижению ее полезности, т. е. к уменьшению ее стоимости.

Разбиение региона на кластеры позволяет сформировать интеллектуальную оценочную модель на основе нейросетевых методов [9].

Данная модель реализует процесс оценки кластеров в виде многослойного персептрона. По исходным данным формируется нейросетевая модель, результатом которой являются карты Кохонена.

Этап визуализации ранее полученных результатов осуществлялся с использованием возможностей геоинформационной системы. Концепция оценки кластеров включает в себя следующие этапы: сбор исходных данных, создание математической модели оценки недвижимости, оценка параметров модели на основе непараметрических методов статистики, формулировка задачи в нейросетевом базисе, анализ параметров математической модели оценки недвижимости, основанный на результатах обучения нейронной сети.

Предлагаемый подход к оценке городских кластеров позволяет расширить круг решаемых оценочных задач. В этом случае в математической модели учитывается субъективность, и тем самым повышается точность расчетов и снижаются затраты на их корректировку с течением времени.

4. Прогнозирование финансовых показателей премиальных магазинов одежды

Одной из первых задач обработки геопространственных данных является их геокодирование. К целевой переменной прибавляются координаты. Далее мы должны привязать данные к каждому квадрату (кластеру).

С помощью машинного обучения мы можем предсказывать, какая будет прибыль магазина, если поставить его в тот или иной квадрат.

Для этого необходимо нарисовать портрет потенциального покупателя, у которого доста-

ток выше среднего. Нужно понять, где он живет, для этого используется геоинформационная система (богатые люди живут там, где самая дорогая недвижимость). Далее нужно понять, поедет ли он в потенциальный магазин или нет, — для этого используем модель Хаффа.

Возьмем один магазин и одну точку проживания и определим, с какой вероятностью человек из этой точки (квадрата) поедет в данный магазин? В качестве веса берем число магазинов дорогой одежды в данном квадрате и распределяем людей из данного квадрата по точкам вокруг него. Получаем, что с вероятностью 10 % человек поедет в магазин Д, с вероятностью 8 % он поедет в магазин С и аналогичным образом анализируем все магазины. Берем 20 % самой дорогой недвижимости и умножаем число людей, которое относится к данному магазину, на стоимость квадратного метра. Получаем своеобразный эквивалент людей и денег: богатый человек, который придет в магазин, принесет больше прибыли, нежели два небогатых человека.

В результате машинного обучения формируется модель, которая используется в данной задаче. Исходную информацию представляет заказчик в виде таблицы: адрес, площадь и продажи по годам. Далее проводится геокодирование данных, т.е. получение широты и долготы. Следующий этап — задание функции цели. По объему продаж за предыдущий период можно узнать, сколько магазин способен в идеале продавать. Например, магазин под номером 0 в год заработал 964,886 тысячи евро максимум. Однако у магазинов разная площадь, и некорректно сравнивать магазин площадью 30 м² с магазином площадью 300 м², поэтому необходимо учитывать прибыль на квадратный метр. Посчитаем середину нашей выборки. Считаем, что целевая функция — это количество денег на квадратный метр. Строим регрессионную модель — модель, которая будет получать на входе число тех или иных объектов для любой точки. Берем все квадраты, которые есть в Москве, и определяем для них прибыль, получаем таблицу прибыли и результаты заносим на карту (рис. 3, см. вторую сторону обложки).

5. Выбор оптимального решения для расположения отделения банка

Вначале создаются базовые слои, в которых для каждой ячейки указывается исходное значение для каждой переменной (например, для ячейки 100 × 100 м). Пример слоя: "Население" — в ячейке указано число, соответствующее числу людей, проживающих в данном квадрате (рис. 4).

Эти слои содержат информацию о внешних данных: благосостоянии, населении, трафике, остановках, маршрутах, конкурентах банка.

Далее создаются агрегирующие слои, в которых в одной ячейке содержится число, равное сумме ячеек базового слоя в соответствующем радиусе.

Пример агрегирующего слоя: "Население 1000" — в ячейке указано число, соответствующее числу людей, проживающему в радиусе 1000 м от данного квадрата (рис. 5, см. вторую сторону обложки).

Задаем для каждого слоя расстояние:

- население 1000;
- благосостояние 0;
- трафик 300;
- остановки 500;
- маршруты 500;
- конкуренты банка 500.

Далее создаются бинарные относительные слои, в которых балл ставится в соответствии



Рис. 4. Базовый слой — "Население"

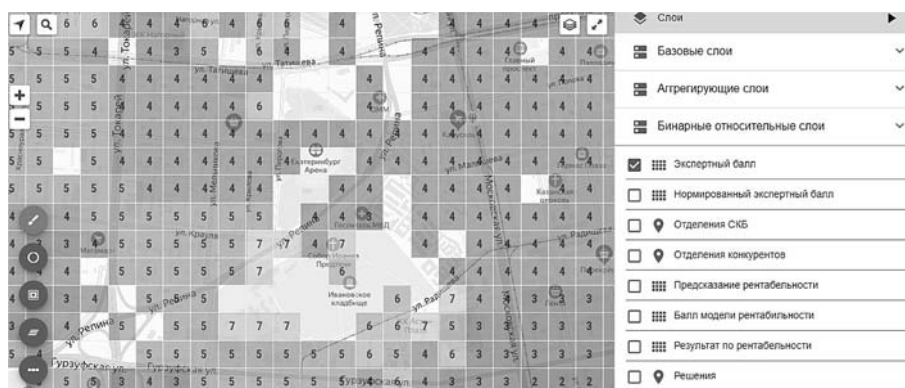


Рис. 6. Слой "Экспертный балл"

с принятым экспертным правилом, а затем слой "Экспертный балл", показывающий сум-му бинарных слоев (рис. 6). В приведенном на рис.6 примере используется шкала оценок от 0 до 7 (минимальная сумма равна -2, мак-симальная равна 5, далее они пересчитаны в шкалу 0—7 баллов).

Затем создается слой "Нормированный экс-пертный балл" (значения от 0 до 50), содержа-щий переменные, агрегирующие информацию

по шести экспертным правилам (выбраны под-рядчиком на основе опыта), заложенным в мо-дель (рис. 7).

Далее создается слой "Метки отделения банка", в котором данные отображаются в виде характеристик каждой метки (рис. 8). Каж-дая метка обладает обязательными атрибута-ми: адрес и геокоординаты, иконка (выбор из Google Material Design Icons), цвет (синий по умолчанию).

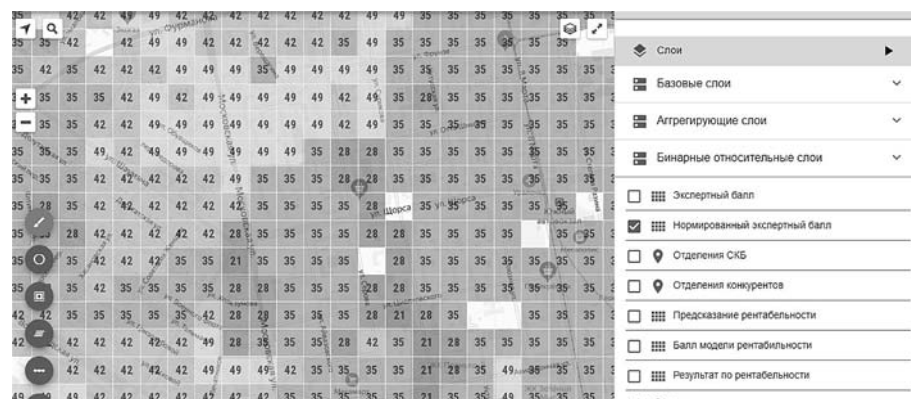


Рис. 7. Слой "Нормированный экспертный балл"

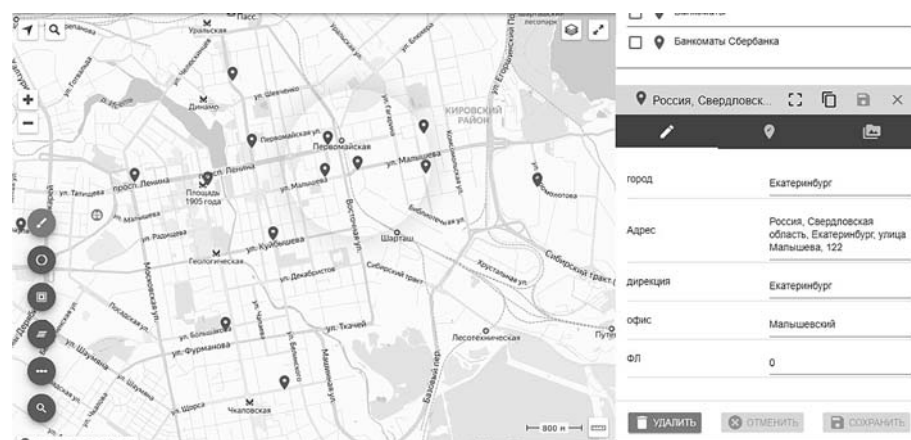


Рис. 8. Слой метки "Отделения банка"

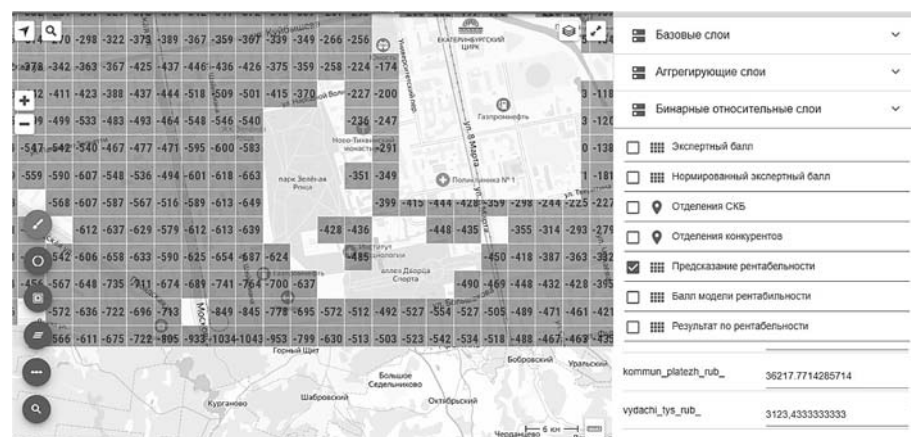


Рис. 9. Слой предсказание рентабельности

Аналогично создается слой "Отделения конкурентов". Да-лее реализуется слой "Пред-сказание рентабельности", в котором в каждой ячейке содержится предсказанное по модели на основе машинного обучения значение рентабель-ности офиса в случае присут-ствия в этой точке (рис. 9).

После этого создается слой "Балл модели рентабельно-сти", содержащий результат нормирования (преобразо-вание в сопоставимую шка-лу от 0 до 50) переменной "Предсказание рентабельно-сти", которая получена путем машинного моделирования показателя рентабельности для каждого офиса на основе внешних факторов (благосо-стояние, численность насе-ления, торговые центры в ра-диусе 2000 м, автостоянки/парковки в радиусе 2000 м) (рис. 10).

На основе анализа боль-шого массива данных (около 850 переменных — трафик, маршруты, точки притяже-ния потоков, население, кон-куренты и пр.) были выбраны 13 факторов, позволяющих предсказать рентабельность офиса в определенной лока-ции. Оценка рентабельности по модели — расчетный ана-лог показателя финансового результата до локации. Далее показатель преобразовыва-ется (нормируется) в шкалу 0—50 баллов.

Например, для отделения банка на основе моделирова-

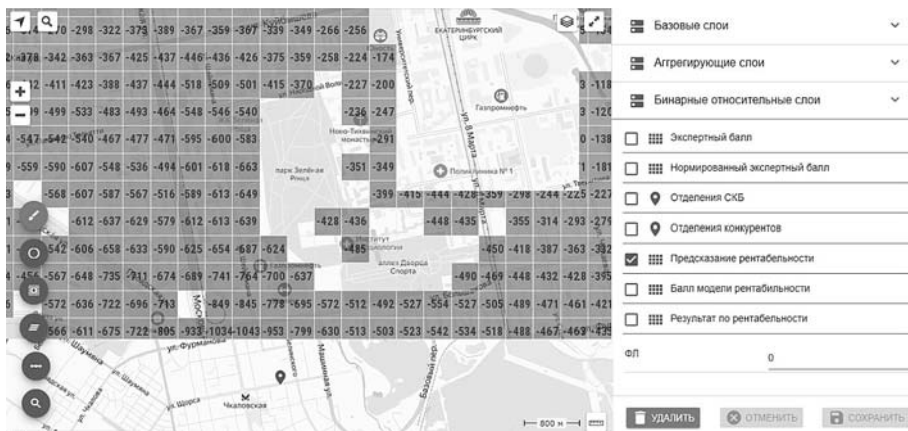


Рис. 10. Слой "Балл модели рентабельности"

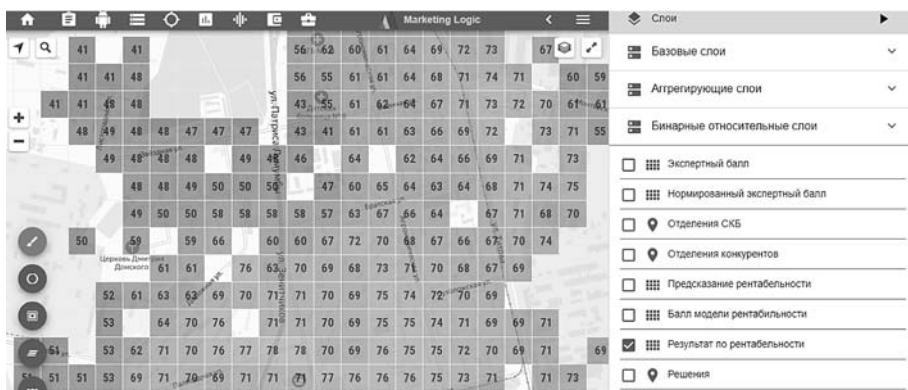


Рис. 11. Слой "Результат по рентабельности"

ния было получено предсказание рентабельности — 828, что после нормирования составило 24 балла.

Создается слой "Результат по рентабельности" (значения от 0 до 100), содержащий сумму баллов по переменным "Нормированный экспертный балл" и "Балл модели рентабельности", т. е. величину, учитывающую в равной пропорции результат экспертных оценок и машинного моделирования. Оценка удачности локации с точки зрения возможности достижения высоких показателей рентабельности показана на рис. 11.

Слой решения — результат анализа сети офисов банка — содержит для каждого офиса оценку его потенциала с учетом предпосылок по локации.

С помощью геоинформационной системы на карте визуально отображаем потенциальные локации для отделений банка (рис. 12, см. вторую сторону обложки).

Были апробированы комбинированные методы геомаркетинга, позволяющие оценить расположение отделения банка, а также решить задачи размещения новых отделений. Перспективными зонами для открытия отделений банка являются территории с высо-

кой плотностью населения, как проживающего на данной территории, так и работающего, с большим числом организаций и объектов торгово-развлекательного назначения. Применение данной методики способствует грамотной организации банковской сети.

6. Решение логистических задач медицинского обслуживания сотрудников

Система здравоохранения включает в себя взаимодействующие больницы, компании, занимающиеся медицинским обслуживанием, диагностикой, лечением, и отдельных практикующих специалистов и т. д. Планирование развития данной системы — непростая задача, так как одной из главных проблем является большой объем не структурированных данных.

Для оптимизации бизнес-процессов прохождения врачебной экспертной комиссии (ВЭК) была использована ГМС, разработанная специально для сотрудников горьковской железной дороги. С помощью ГМС решаются следующие задачи: доступное прохождение комиссии в установленные сроки, тщательное наблюдение за здоровьем сотрудников, работа которых связана с вредоносными факторами. Используя атлас, составленный с помощью ГМС, пользователи могут быстро увидеть, в какой срок необходимо пройти обследование, а также получить информацию о ближайшем медицинском учреждении.

Проект предоставляет разнообразные возможности исследования данных с помощью динамических таблиц, графиков и алгоритмов. Данная система содержит информацию о большом числе медицинских учреждений для прохождения ВЭК, а также информацию о сотрудниках горьковского направления железной дороги (место проживания; конкретная должность). Сайт логически организует около 40 тысяч сотрудников, предоставляя информацию о них. Для повышения эффективности прохождения ВЭК сотрудники распределены

по кластерам и "привязаны" к определенной больнице.

Цель иерархической кластеризации в проекте ГМС заключается в объединении сотрудников, которые расположены наиболее близко друг к другу, по принципу прохождения медицинского обследования в одно время. Все работники распределены по девяти группам, внутри каждой из них создаются новые кластеры по сроку прохождения осмотра для того, чтобы выбрать для конкретной группы оптимальную дату прохождения ВЭК.

Заключение

В статье рассмотрены ряд примеров использования ГМС для решения различных задач в области экономики. Кластеризация геопространственных данных позволяет эффективно проводить расчеты и с учетом точности пространственной локализации. Геопространственный анализ применяется во многих аспектах бизнеса. Он поможет проанализировать отток клиентов к конкурентам, оптимизировать штат менеджеров зала для эффективного обслуживания клиентов, определить загруженность транспортных узлов, потребность в строительстве небоскребов или парков и т. д.

Список литературы

1. Tsvetkov V. Ya., Shaytura S. V., Sultaeva N. L. Digital Enterprise Management in Cyberspace // Proceedings of the 2nd International Scientific and Practical Conference "Modern Management Trends and the Digital Economy: from Regional Development to Global Economic Growth" (MTDE 2020). Yekaterinburg, Russia. P. 361–365. doi: <https://doi.org/10.2991/aebmr.k.200502.059>.
2. Буравцев А. В., Цветков В. Я. Облачные вычисления для больших геопространственных данных // Информация и космос. 2019. № 3. С. 110–115.
3. Дышленко С. Г. Большие данные в науках о Земле // Славянский форум. 2015. № 3 (10). С. 88–96.
4. Shaitura S. V., Ordov K. V., Lesnichaya I. G., Romanova Yu. D., Khachaturova S. S. Services and mechanisms of competitive intelligence on the internet // Espacios. 2018. Т. 39, № 45. С. 24.
5. Маркелов В. М., Цветков В. Я. Геомониторинг // Славянский форум. 2015. № 2 (8). С. 177–184.
6. Чумаченко С. И., Князева М. Д., Митрофанов Е. М., Шайтура С. В. Космический мониторинг: Учеб. пособ. Бургас, 2017.
7. Шайтура С. В., Васкина М. Ю. Комплексная цифровая модель мониторинга района // Экология урбанизированных территорий. 2019. № 4. С. 71–76.
8. Шайтура С. В., Васкина М. Ю. Мониторинг земель районов Дальнего Востока // Землеустройство, кадастр и мониторинг земель. 2020. № 1. С. 28–33.
9. Баяндурова А. А., Розенберг И. Н., Шайтура С. В. Комплексный анализ крымских туристических дестинаций

// Ученые записки Крымского федерального университета имени В. И. Вернадского. Экономика и управление. 2016. Т. 2(68). № 1. С. 3–10.

10. Розенберг И. Н., Шайтура С. В. Кластерный анализ туристических дестинаций Крымского полуострова // Организационно-экономический механизм управления опережающим развитием регионов. 2016. С. 215–221.

11. Курейчик В. М., Кажаров А. А. Анализ и состояние задачи маршрутизации автотранспорта // Вестник Ростовского государственного университета путей сообщения. 2013. № 4 (52). С. 73–77.

12. Shaitura S. V., Stepanova M. G., Shaitura A. S., Ordov K. V., Galkin N. A. Application of information-analytical systems in management // Journal of Theoretical and Applied Information Technology. 2016. Т. 90, № 2. С. 10–22.

13. Петров Я. А., Степанов С. Ю., Сидоренко А. Ю., Мартын И. А., Петров А. Д. Геомаркетинговое исследование как инструмент анализа целевой аудитории при выборе расположения торговой точки // Информационные технологии и системы: управление, экономика, транспорт, право. 2019. № 4 (36). С. 44–48.

14. Кулагин В. П., Цветков В. Я. Геознание: представление и лингвистические аспекты // Информационные технологии. 2013. № 12. С. 2–9.

15. Shaitura S. V., Kozhaev Yu. P., Ordov K. V., Vintova T. A., Minitaeva A. M., Feoktistova V. M. Geoinformation services in a spatial economy // International Journal of Civil Engineering and Technology. 2018. Т. 9, № 2. С. 829–841.

16. Shaitura S. V., Minitaeva A. M., Ordov K. V., Shaparenko V. V. Virtual enterprises in a spatial economy // International Journal of Recent Technology and Engineering (IJRTE). 2019. Т. 7, № 6. С. 719–724.

17. Шайтура С. В., Цветков В. Я., Шайтура А. С., Кожаяев Ю. П., Харитонов С. В., Степаненко Н. В. Теория и практика геомаркетинга: Учебн. пособ. Бургас, 2016.

18. Тымченко Е. В. Организация данных в геомаркетинге // Перспективы науки и образования. 2014. № 6 (12). С. 160–165.

19. Майоров А. А. Геомаркетинговые исследования // Образовательные ресурсы и технологии. 2014. № 5 (8). С. 43–48.

20. Шарифьянов Т. Ф. Планирование социальной информационно-коммуникационной инфраструктуры региона на основе методов геомаркетинга // Практический маркетинг. 2017. № 12 (250). С. 29–34.

21. Гаврилова В. В., Шайтура С. В. Интеллектуальная обработка информации в области оценки недвижимости // Славянский форум. 2012. № 1 (1). С. 164–171.

22. Китова О. В., Шайтура С. В. Информационный маркетинг: Учеб. пособ. Бургас, 2016.

23. Маратканова О. Е. Геомаркетинговый подход к размещению объектов логистической инфраструктуры // Социально-экономическое управление: теория и практика. 2018. № 1 (32). С. 34–35.

24. Воробьева Д. Е., Дегтева Е. В. Геомаркетинг как инструмент изучения конкурентных преимуществ отелей // Российские регионы: взгляд в будущее. 2018. Т. 5, № 4. С. 47–56.

25. Петров Я. А., Степанов С. Ю., Сидоренко А. Ю., Мартын И. А., Петров А. Д. Геомаркетинговое исследование как инструмент анализа целевой аудитории при выборе расположения торговой точки // Информационные технологии и системы: управление, экономика, транспорт, право. 2019. № 4 (36). С. 44–48.

26. Герасименко О. А., Тхорилов Б. А., Наплекова Ю. А. Сущностное представление, роль, эволюционные этапы и подходы геомаркетинга // Вестник Белгородского университета кооперации, экономики и права. 2020. № 3 (82). С. 248–259.

S. V. Shaytura, Cand. Tech. Sciences, Associate Professor, e-mail: swshaytura@gmail.com,
Russian University of Transport" (MIIT), Moscow,
Institute of Humanities, Economics and Information Technologies, Burgas, Bulgaria,
D. A. Galkin, Managing Partner, e-mail: julikas@inbox.ru,
Marketing Logic Company, Moscow, Russian Federation

Geomarketing Big Data Analysis

The accumulation of a large amount of geospatial data requires new approaches to their processing and visualization. One of these approaches is the creation of a geomarketing system with a fundamentally new toolkit based on data clustering. The capabilities of such a system are shown using examples of housing cost assessment, determining the location of a new shopping center, a bank branch and a clinic.

Keywords: geomarketing system, spatial data, geomarketing, big data, clusters, geomarketing technologies, information technologies

DOI: 10.17587/it.27.180-187

References

1. Tsvetkov V. Ya., Shaytura S. V., Sultaeva N. L. Digital Enterprise Management in Cyberspace, *Proceedings of the 2nd International Scientific and Practical Conference "Modern Management Trends and the Digital Economy: from Regional Development to Global Economic Growth" (MTDE 2020)*, Yekaterinburg, Russia, pp. 361–365, doi: <https://doi.org/10.2991/aebmr.k.200502.059>.
2. Buravtsev A. V., Tsvetkov V. Ya. Cloud computing for large geospatial data, *Information and Space*, 2019, no. 3, pp. 110–115 (in Russian).
3. Dyshlisko S. G. Big Data in Earth Sciences, *Slavic Forum*, 2015, no. 3 (10), pp. 88–96 (in Russian).
4. Shaitura S. V., Ordov K. V., Lesnichaya I. G., Romanova Yu. D., Khachaturova S. S. Services and mechanisms of competitive intelligence on the internet, *Espacios*, 2018, vol. 39, no. 45, pp. 24.
5. Markelov V. M., Tsvetkov V. Ya. Geomonitoring, *Slavic Forum*, 2015, no. 2 (8), pp. 177–184 (in Russian).
6. Chumachenko S. I., Knyazeva M. D., Mitrofanov E. M., Shaytura S. V. Space Monitoring, Study Guide, Burgas, 2017 (in Russian).
7. Shaytura S. V., Vaskina M. Yu. Integrated digital model of monitoring the area, *Ecology Of Urbanized Territories*, 2019, no. 4, pp. 71–76 (in Russian).
8. Shaytura S. V., Vaskina M. Yu. Monitoring of lands in the regions of the Far East, *Land Management, Cadastre and Monitoring of Lands*, 2020, no. 28–33 (in Russian).
9. Bayandurova A. A., Rosenberg I. N., Shaytura S. V. Complex analysis of the Crimean tourist destinations, *Scientific notes of the Crimean Federal University named after V. I. Vernadsky. Economics and Management*, 2016, vol. 2 (68), no. 1, pp. 3–10 (in Russian).
10. Rosenberg I. N., Shaytura S. V. Cluster analysis of tourist destinations on the Crimean peninsula, *Organizational and economic mechanism for managing the advanced development of regions*, 2016, pp. 215–221 (in Russian).
11. Kureichik V. M., Kazharov A. A. Analysis and state of the problem of routing vehicles, *Bulletin of the Rostov State University of Communications*, 2013, no. 4 (52), pp. 73–77 (in Russian).
12. Shaitura S. V., Stepanova M. G., Shaitura A. S., Ordov K. V., Galkin N. A. Application of information-analytical systems in management, *Journal of Theoretical and Applied Information Technology*, 2016, vol. 90, no. 2, pp. 10–22.
13. Petrov Ya. A., Stepanov S. Yu., Sidorenko A. Yu., Martyn I. A., Petrov A. D. Geomarketing research as a tool for analyzing the target audience when choosing the location of a retail outlet, *Information Technologies and Systems: Management, Economics, Transport, Law*, 2019, no. 4 (36), pp. 44–48 (in Russian).
14. Kulagin V. P., Tsvetkov V. Ya. Geology: representation and linguistic aspects, *Information Technologies*, 2013, no. 12, pp. 2–9 (in Russian).
15. Shaitura S. V., Kozhaev Yu. P., Ordov K. V., Vintova T. A., Minitaeva A. M., Feoktistova V. M. Geoinformation services in a spatial economy, *International Journal of Civil Engineering and Technology*, 2018, vol. 9, no. 2, pp. 829–841.
16. Shaitura S. V., Minitaeva A. M., Ordov K. V., Shaparenko V. V. Virtual enterprises in a spatial economy, *International Journal of Recent Technology and Engineering (IJRTE)*, 2019, vol. 7, no. 6, pp. 719–724.
17. Shaytura S. V., Tsvetkov V. Ya., Shaytura A. S., Kozhaev Yu. P., Kharitonov S. V., Stepanenko N. V. Theory and Practice of Geomarketing, Study Guide, Burgas, 2016 (in Russian).
18. Tymchenko E. V. Organization of data in geomarketing, *Prospects for Science and Education*, 2014, no. 6 (12), pp. 160–165 (in Russian).
19. Mayorov A. A. Geomarketing research, *Educational Resources and Technologies*, 2014, no. 5 (8), pp. 43–48 (in Russian).
20. Sharifyanov T. F. Planning of social information and communication infrastructure of the region based on geomarketing methods, *Practical Marketing*, 2017, No. 12 (250), pp. 29–34 (in Russian).
21. Gavrilova V. V., Shaytura S. V. Intellectual processing of information in the field of real estate valuation, *Slavyanskiy forum*, 2012, no. 1 (1), pp. 164–171 (in Russian).
22. Kitova O. V., Shaytura S. V. Information Marketing, Study Guide, Burgas, 2016 (in Russian).
23. Maratkanova O. E. Geomarketing approach to the placement of logistic infrastructure facilities, *Socio-Economic Management: Theory and Practice*, 2018, no. 1 (32), pp. 34–35 (in Russian).
24. Vorobyova D. E., Degteva E. V. Geomarketing as a tool for studying the competitive advantages of hotels, *Russian regions: a look into the future*, 2018, vol. 5, no. 4, pp. 47–56 (in Russian).
25. Petrov Ya. A., Stepanov S. Yu., Sidorenko A. Yu., Martyn I. A., Petrov A. D. Geomarketing research as a tool for analyzing the target audience when choosing the location of a retail outlet, *Information technologies and systems: management, economics, transport, law*, 2019, no. 4 (36), pp. 44–48 (in Russian).
26. Gerasimenko O. A., Tkhorikov B. A., Naplekova Yu. A. Essential representation, role, evolutionary stages and approaches of geomarketing, *Bulletin of Belgorod University of Cooperation, Economics and Law*, 2020, no. 3 (82), pp. 248–259 (in Russian).

Д. В. Малаховецкий, студент, e-mail: dmitriymalax@yandex.ru,
Московский государственный технологический университет "СТАНКИН",
А. И. Разумовский, канд. техн. наук, ст. науч. сотр., e-mail: razumowsky@yandex.ru,
Институт проблем управления РАН, г. Москва

Синтаксический анализ символьных массивов методом рекурсивного охвата и структурирования на примере данных VRML-формата

Представлен новый подход к структуризации, сегментации и алгоритмическому оформлению синтаксического анализатора символьных массивов на примере данных VRML. Ключевой особенностью метода является возможность формирования иерархически-сложного объекта посредством рекурсивной структуризации данных, позволяющей охватывать в совокупности все содержимое объекта, включая его произвольную вложенность дочерних объектов. Это позволяет достичь высокой управляемости при разработке алгоритма синтаксического анализа и дает возможность сосредоточиться каждый раз на конкретном фрагменте данных, при этом не теряя из виду всей совокупной связности информации. Полученные результаты легко могут быть использованы при создании удобных структур хранения данных, связанных с обеспечением информационной безопасности, для решения проблемы сдерживания объема данных в файлах, в задачах управления большими данными в гетерогенных системах, а также при обработке иерархических данных в интернете вещей.

Ключевые слова: метод анализа данных, синтаксический анализ, рекурсия, VRML-формат, полиморфная иерархия

Введение

Всякий предметный или системный анализ необходим и востребован по причине содержательной сложности данных предмета или системы. Информационная сложность представляет собой нечто с трудом воображаемое либо не поддающееся мысленному охвату. Поэтому информацию перед непосредственным использованием необходимо соответствующим образом подготовить. Среди основных способов такой подготовки является анализ информации и ее практическая апробация. В современном мире обилие информации является нормой, поэтому востребованными остаются не только предоставляющие информацию ресурсы, но и методы, с помощью которых можно найти, соотносить, преобразовать нужную информацию, провести ее анализ и синтез.

В настоящем исследовании мы изучаем данные графического формата VRML (Virtual Reality Modeling Language), который был создан для моделирования виртуальной реальности. Формат, который набрал огромную популярность в 90-х годах и был вытеснен впоследствии своим преемником — языком моделирования виртуальной реальности X3D, по-прежнему вполне востребован. Область его применения простирается от интерактивной векторной графики и Web-технологий до приложений медицинской направленности, строительства, картографии и мультимедийных систем. Потребность в использовании этого формата

по-прежнему высока в системах автоматизированного проектирования, включая обмен данными между различными графическими системами [1—4]. Язык оперирует объектами, описывающими геометрические фигуры и их расположение в пространстве [5].

Используя файлы символьного формата, язык VRML позволяет определить не только геометрические свойства объектов в трехмерном пространстве, например, расположение и форму сложных поверхностей и многогранников, но и физические данные об их цвете, текстуре, блеске, прозрачности, источниках освещения. Кроме того, в качестве реакции на действия пользователя или другие внешние события существует возможность дополнить описание модели, задав параметры движения, звуков, освещения и иных аспектов виртуального мира [6].

Прежде чем искать характерные алгоритмические пути синтаксического анализа данных, важно понимать, что анализ соотносится с операциями последовательного считывания определенных порций данных из файла, поэтапной идентификации VRML-объектов по характерным признакам, сопоставления с уже имеющимися в хранилище данными и, наконец, формирования необходимой структуры для взаимодействия с данными объектами.

Как большинство языков, VRML имеет четкую конструкцию и определенный набор лексических, синтаксических и семантических правил, определяющих внешний вид програм-

мы и ее действия. Из этого можно сделать ложный вывод о том, что написание алгоритма синтаксического анализа не составит большого труда. Вся структура формата уже продумана разработчиками языка, и в таком случае необходимо только создать метод, который позволит ее правильно считать. Однако при написании программы возникают трудности, связанные с иерархическими особенностями языка.

Мы представляем новый подход к алгоритмической организации синтаксического анализа данных, который исключает проблемы, связанные с иерархической сложностью VRML информации. В основе метода лежат рекурсивное считывание объектов формата VRML, их последовательный анализ, идентификация и сопоставление с уже определенными данными. Для хранения содержательной структуры объектов используется специальный контейнер, который одновременно обеспечивает удобный доступ ко всему множеству объектов, а также возможность взаимодействия как с отдельными объектами, так и со всей структурой объектов в целом.

1. Методы исследования

Прежде чем выявить и сделать попытку разрешить проблемы синтаксического анализа сложноструктурированных данных, необходимо классифицировать виды их структурной организации.

- В формате VRML имеются три вида описаний:
- простые объекты (без спецификации), которые последовательно считываются из файла и сразу же помещаются в соответствующее место полиморфной иерархии;
 - "якорные" объекты (ЯО), выделяемые в формате ключевым словом "DEF". Они содержат определение и описание свойств объекта для последующего использования в иерархии. Порядок обработки таких объектов следующий: вначале считываются данные, имеющие полноту описания объекта, затем эти данные помещаются в отдельный контейнер DEF, которые хранит "якорные" объекты, идентифицируемые по ключу-имени. Далее указатель на такой объект помещается в соответствующее место иерархии;
 - пользовательские объекты, определяемые спецификацией "USE": содержат указание на использование определенного "якорного" объекта, имеющегося в иерархии. В момент считывания пользовательского объекта выполняется поиск по соответствующему ключу в контейнере DEF, после чего найденный указатель размещается в хранилище.

Первая и самая очевидная проблема, с которой можно столкнуться при анализе VRML-данных, заключается в многократной вложенности объектов, которая в некоторых случаях является избыточной (рис. 1, см. третью сторону обложки).

Средствами языка C++ имеется возможность описать и реализовать вложенность, однако для этого объекты должны быть разных типов. В данных VRML нередко можно встретить, как один и тот же объект (например, объект типа Group) несколько раз наследуется сам от себя, что делает невозможным последовательное формирование структуры вложенных данных. На иллюстрации красной рамкой обозначена ситуация многократной вложенности объекта Group (рис. 1, см. третью сторону обложки).

Следующая проблема происходит из синтаксиса формата VRML. В языке VRML, как и во многих других языках, допускаются ЯО, которые могут быть описаны в файле только единожды, а используются затем неопределенное число раз посредством именной ссылки. Иными словами, это определения комплексных типов данных с возможностью созидания сложно сгруппированных объектов. Затруднения здесь заключаются в вероятной потере контроля над динамической памятью, которая выделяется всякий раз при считывании данных. Так как ЯО может использоваться в алгоритме не один раз, необходимо определенно знать, в какой именно момент его следует из памяти удалять. Иначе говоря, за ЯО следует установить дополнительное наблюдение.

С последней проблемой можно столкнуться не всегда, а только при считывании многофайлового объекта — VRML-сборки. Синтаксический анализ и считывание сборок организуется аналогично однофайловым данным. Однако казус заключается в том, что один и тот же файл может использоваться в сборке многократно. Это означает, что однозначная идентификация объектов в такой сборке затруднена. Кроме того, возникает утечка памяти из-за

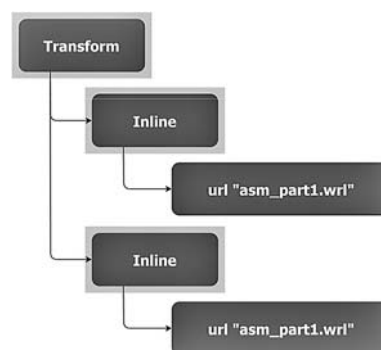


Рис. 2. Пример повторного использования файла в сборке

перезаписи указателей на ЯО, которые, имея одно и тоже имя, в разных файлах естественно ведут себя по-разному (рис. 2).

Память под такие объекты выделяется в общем порядке, однако она может быть вовремя не освобождена, так как указатели на них перезаписываются в ходе последующего считывания данных, ссылающихся на тот же самый объект.

1.1. Создание универсального контейнера VRML-объектов

Для того чтобы разрешить проблему вложенности объектов при создании полноценной структуры VRML, важно также понять семантику VRML-формата. Так, имеется определенный объект-родитель, в котором содержатся какие-либо свойства, и массив дочерних элементов (children). Такой массив включает, в свою очередь, множество подобных объектов, у которых имеются индивидуальные свойства и собственный массив наследуемых объектов. При этом такое представление по-прежнему не дает ответа на главный вопрос: как справиться с проблемой многократного наследования? Ведь массив дочерних объектов (children) может содержать практически любой произвольный объект формата VRML: Transform, Group, Shape и так далее. Мы наблюдаем здесь, как сложность охвата данных только возрастает. Необходим нетривиальный взгляд на многократно вложенные данные. Иначе говоря, требуется выразить вложенность абстрактно, сохраняя при этом индивидуальные черты формата данных.

Проведем классификацию типов. Вся структура VRML состоит из так называемых узлов (Node) [5,7], которые подразделяются на геометрические узлы (Box, Cone, Cylinder, IndexFaceSet) и группирующие узлы (Group, Transform, Collision). Кроме того, есть узлы, которые отвечают за графическое представление (Appearance), а также узлы, определяющие свойства геометрических объектов (geometry). Соответственно, ориентируясь на такую классификацию, несложно сформировать полиморфную иерархию на языке C++.

Определим абстрактный класс Node как обобщение всевозможных узлов формата VRML. Затем установим наследование прочих классов от абстрактного класса Node. Таким образом, мы получаем системную связность произвольных элементов данных посредством включения их типов в общую полиморфную иерархическую композицию.

Далее, поскольку любой тип, образованный на основании перечня ключевых слов формата VRML, может быть включен в иерархию, появляется возможность создать массив указателей

на объект базового типа данных (Node*) этой иерархии. Объект любого типа, входящего в иерархию, может быть помещен в этот массив в качестве указателя на динамически выделенную под объект память. Таким способом получится легко осуществить реализацию хранилища VRML-данных, сохраняя при этом их разнообразие.

1.2. Косвенный мониторинг "якорных" объектов

Для хранения ЯО и последующего их использования при построении полиморфной иерархии выбран ассоциативный контейнер стандартной библиотеки шаблонов std::map. Ключом в нем устанавливается имя считываемого объекта, а значением — указатель на динамический ЯО.

В процессе считывания данных выполняется проверка — является ли объект "якорным"? В положительном случае его указатель помещается в глобальный контейнер DEF по ключу — имени ЯО. Впоследствии при считывании пользовательского объекта выполняется поиск по ключу в контейнере DEF, после чего под него выделяется память, а сам объект помещается в иерархию данных. Если же объект не является "якорным", то под него сразу выделяется память, и он поступает в целевое хранилище.

Главным хранилищем данных является динамический массив nodes (узлы) стандартной библиотеки шаблонов, типа std::vector<Node*>. Он предназначен для сбора всей полиморфной иерархии считываемых объектов. Провести идентификацию — является ли текущий объект "якорным", можно посредством проверки именной строки def. Если она отсутствует у объекта, тогда объект просто помещается в главное хранилище. В противном случае объект квалифицируется как ЯО, он также помещается в хранилище std::map<string, Node*> DEF в целях косвенного мониторинга его последующего использования:

```
if (group->def! = "") {  
    DEF->insert(std::pair<string, Node*>(def, group));  
    def.clear();  
}
```

Если при считывании обнаруживается пользовательский объект ("USE"), тогда по его имени выполняется поиск в хранилище DEF по ключу (имя объекта + имя файла), и найденный там объект помещается в главный контейнер nodes:

```
if (str == "USE") {  
    file >> str;  
    shape->geometry = (Geometry*)DEF->at(str + "_" + cur_file);  
}
```


Имя	Значение	Тип
DEF	00015608 (size=22)	std::map<std::string, N...
↳ [comparator]	less	std::Compressed_pair...
↳ [allocator]	allocator	std::Compressed_pair...
↳ [FAIO_2_X_1217_MOTOR_ST]	000046058 (scale=000046078 (1.00000000, 1.00000000, 1.00000000) translat...	std::pair<std::string co...
↳ [FAIO_2_X_1218_MOTOR_ST]	000046750 (scale=000046780 (1.00000000, 1.00000000, 1.00000000) translat...	std::pair<std::string co...
↳ [FAIO_2_X_1219_MOTOR_ST]	000046748 (scale=000046618 (1.00000000, 1.00000000, 1.00000000) translat...	std::pair<std::string co...
↳ [FAIO_2_X_1220_MOTOR_ST]	000046678 (scale=000046688 (1.00000000, 1.00000000, 1.00000000) translat...	std::pair<std::string co...
↳ [FAIO_2_X_1221_MOTOR_ST]	000046650 (scale=000046670 (1.00000000, 1.00000000, 1.00000000) translat...	std::pair<std::string co...
↳ [FAIO_2_X_12_MOTOR_ST]	000142590 (scale=000142580 (1.00000000, 1.00000000, 1.00000000) translat...	std::pair<std::string co...
↳ [DRAWING_MOTOR_ST]	000157340 (scale=000157360 (1.00000000, 1.00000000, 1.00000000) translat...	std::pair<std::string co...
↳ [EA_OA_MOTOR_ST]	000157470 (scale=000157490 (1.00000000, 1.00000000, 1.00000000) translat...	std::pair<std::string co...
↳ [FEDAI_AQUE_AAE_MOTOR_ST]	000159640 (scale=000159660 (1.00000000, 1.00000000, 1.00000000) translat...	std::pair<std::string co...
↳ [FE_UOIA_IA_OA_A_NAI_A_MOTOR]	000142518 (scale=000142528 (1.00000000, 1.00000000, 1.00000000) translat...	std::pair<std::string co...
↳ [FE_UOIA_IA_OA_A_MOTOR_ST]	000142380 (scale=000142340 (1.00000000, 1.00000000, 1.00000000) translat...	std::pair<std::string co...
↳ [FAEO_A_NAI_A_MOTOR_ST]	000141180 (scale=000142010 (1.00000000, 1.00000000, 1.00000000) translat...	std::pair<std::string co...
↳ [FAEO_A_MOTOR_ST]	000142158 (scale=000142178 (1.00000000, 1.00000000, 1.00000000) translat...	std::pair<std::string co...
↳ [FAIAI_IA_OA_A_MOTOR_ST]	000156638 (scale=000156658 (1.00000000, 1.00000000, 1.00000000) translat...	std::pair<std::string co...
↳ [FAIAI_A_NAI_A_MOTOR_ST]	000142068 (scale=000142088 (1.00000000, 1.00000000, 1.00000000) translat...	std::pair<std::string co...
↳ [FI_OAIU_MOTOR_ST]	000142040 (scale=000142100 (1.00000000, 1.00000000, 1.00000000) translat...	std::pair<std::string co...
↳ [FI_OIAIE_AIAIAOAIU_MOTOR_ST]	000157388 (scale=000157248 (1.00000000, 1.00000000, 1.00000000) translat...	std::pair<std::string co...
↳ [I_BEAAIAI_MOTOR_ST]	000154298 (scale=000154268 (1.00000000, 1.00000000, 1.00000000) translat...	std::pair<std::string co...
↳ [I_BEAAIAI_MOTOR_ST]	000154208 (scale=000154228 (1.00000000, 1.00000000, 1.00000000) translat...	std::pair<std::string co...
↳ [I_moaia_MOTOR_ST]	000142608 (scale=000142628 (1.00000000, 1.00000000, 1.00000000) translat...	std::pair<std::string co...
↳ [NAIA_MOTOR_ST]	000142338 (scale=000142358 (1.00000000, 1.00000000, 1.00000000) translat...	std::pair<std::string co...
↳ [NIOAI_MOTOR_ST]	000141178 (scale=000141198 (1.00000000, 1.00000000, 1.00000000) translat...	std::pair<std::string co...
↳ [baseove прeдcтавлeнeе]	00015608 []	std::map<std::string, N...
↳ nodes	000157130 (size=0)	std::vector<Node*, std::...
↳ output	"Data/MOTOR_ST.x"	Q, < std::string

Рис. 3. Содержимое хранилища ЯО

Таким образом, все ЯО хранятся в контейнере DEF (рис. 3).

В главном хранилище, которое обеспечивает содержание иерархической структурной связности объектов, содержатся указатели на ЯО из DEF, а также объекты без специализации. Для того чтобы корректно выполнить удаление и очистку всей структуры данных, необходимо перед удалением объекта выполнять проверку — является ли он "якорным"? Если — нет, то удаление должно выполняться сразу, иначе удаление проводится по завершении всего синтаксического анализа.

1.3. Идентификация ЯО

Для осуществления однозначной идентификации по имени объекта необходимо образовывать комплексное именование, добавляя к имени некоторый суффикс. Такая дополнительная именная добавка позволит различать ЯО при считывании VRML-сборок, в которых названия объектов-переменных могут быть одинаковыми. Таким суффиксом удобно определить название файла, в котором находится считываемый объект.

Однако проблема с неоднозначной идентификацией переменных при использовании одного файла несколько раз при этом все же остается, поскольку имена файлов и имена объектов могут совпадать. Для решения данной проблемы необходимо создать дополнительное множество соотносимых через имя элементов. В стандартной библиотеке шаблонов существует весьма удобный для достижения этой цели ассоциативный контейнер `std::map`:

```
std::map<string, vector<Node*>*>* FILES
```

В объекте FILES ключом будет выступать название файла, считанного ранее, а значением — указатель на соответствующим об-

разом обработанный иерархический объект. При считывании имени файла выполняется поиск по контейнеру FILES. И если ранее это файл не был проанализирован, то проводится последовательное чтение объектов из этого файла с созданием соответствующей иерархии объектов, указатель на которую используется для добавления в хранилище FILES.

Данный способ идентификации переменных имеет ряд преимуществ:

- однозначная идентификация ЯО при считывании сборки с одинаковыми файлами;
- однозначная идентификация ЯО при считывании сборки с различными файлами, в которых названия переменных при этом являются одинаковыми;
- оптимизация времени выполнения программы за счет пропуска повторной процедуры считывания уже имеющихся данных.

1.4. Метод рекурсивного охвата данных

По мере выявления проблем и прояснения их особенностей важно правильно подобрать средства и способы достижения целевого результата. Итак, необходимо реализовать совокупность методов, которые позволят следующее:

- провести чтение данных из файла VRML-формата;
- выполнить разбор на лексемы считанных данных;
- идентифицировать лексемы в объекты и сопоставить их для установления связей между ними и прочих свойств;
- построить полиморфную иерархию на основе полученных объектов.

Проблема необходимости дополнительного мониторинга ЯО, а также неопределенности уровня и степени вложенности объектов на этапе чтения данных приводит к необходимости идентифицировать объекты так, словно уже известна их глубина вложенности и свойства. Для этого удобно использовать рекурсивный подход, который позволит охватить данные в их предположительной целостности.

Основной процедурой, которая осуществляет чтение формата VRML, установим функцию `parseNode` (рис. 4).

Настоящая функция принимает следующие параметры:

- объект файлового потока, откуда должно выполняться чтение данных;
- контейнер, в который должны быть помещены все считанные объекты;
- строка, определяющая текущее положение указателя на данные в файле;
- имя текущего файла для создания уникальных имен объектов.

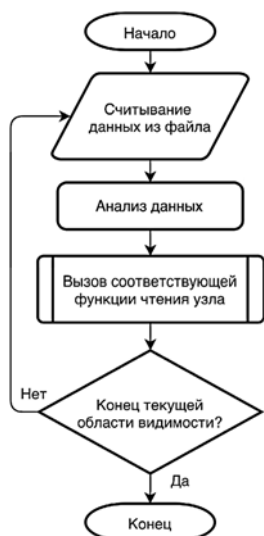


Рис. 4. Блок-схема алгоритма чтения VRML-файла *parseNode*

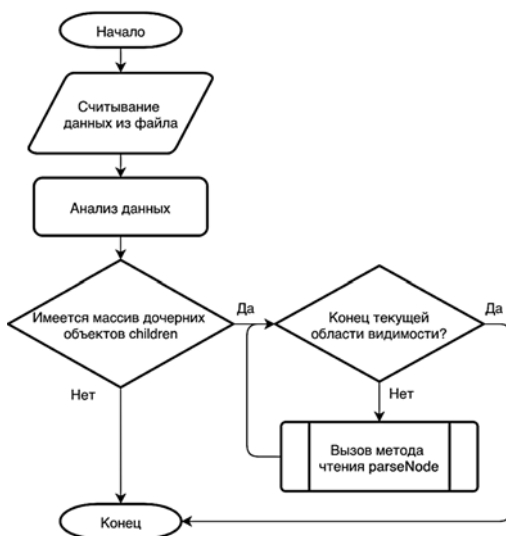


Рис. 5. Блок-схема алгоритма чтения группированных узлов *parseGroup*

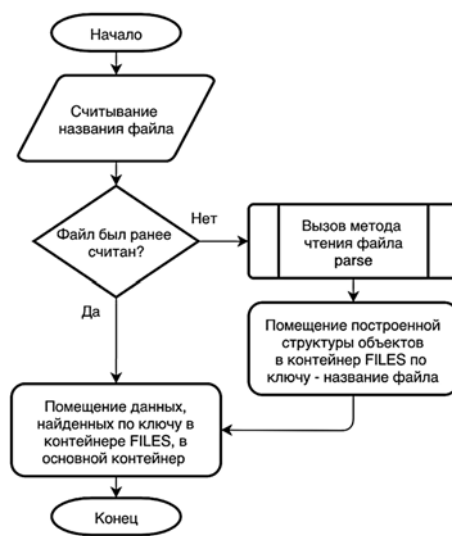


Рис. 6. Блок-схема алгоритма считывания VRML-сборок

В зависимости от квалификации объекта — Transform, Group, Shape, Inline — осуществляется вызов соответствующей функции для установления его свойств. Кроме того, выполняется запись наименования объекта для последующей ассоциации и идентификации ЯО. К имени объекта также добавляется имя текущего файла, из которого он был считан.

Реализация всех методов чтения группированных узлов — Group, Transform, Collision и прочих — наиболее наглядно видна на примере алгоритма рекурсивного чтения объекта Group — *parseGroup*, так как группирующие узлы отличаются только своими свойствами, и важно сосредоточиться на реализации чтения их дочерних объектов (рис. 5).

В случае, если объект имеет массив зависимостей *children*, организуется цикл, в котором итерационно вызывается метод *parseNode* и проводится рекурсивное чтение дочернего объекта с размещением затем в контейнере *children*.

Функция *parseNode* имеет следующие входные параметры:

- объект текущего файлового потока;
- строка, которая определяет позицию указателя в файле;
- название настоящего файла для создания уникальных имен объектов;
- контейнер, в который должны быть помещены считанные объекты. Здесь в качестве контейнера передается контейнер *children* текущего объекта.

Аналогично рекурсивно охватывается вся иерархическая цепочка дочерних объектов: если дочерний объект также является группировующим узлом, то при его чтении запускается аналогичный цикл.

Решение проблемы однозначной идентификации объектов при чтении сборочного VRML-файла отведено функции *Originality*. Эта операция отвечает за непосредственное наращивание полиморфной иерархии. Основной чертой функции является проведение однозначной идентификации объектов с применением контейнера FILES, описанного выше (рис. 6).

Реализация метода *Originality* делится на два этапа:

1. Считывание названия файла.
2. Проверка того факта, был файл считан ранее или нет.

2. Полученные результаты

По мере чтения VRML-файла (или сборки) программа рекурсивно выстраивает полиморфную иерархию всех объектов, их свойств и наследников (рис. 7).

Результирующая иерархия представлена в древовидном виде и хранится в объекте типа `std::vector`. Такой способ хранения позволяет наилучшим образом осуществлять удобный доступ ко всем объектам, а также простоту взаимодействия с ними. Кроме того, посредством метода рекурсивного охвата достаточно легко проводить дополнительный мониторинг ЯО, поскольку каждый ЯО наблюдается одновременно и как отдельный объект, и как содержащий множество, объем которого заранее неизвестен, прочих объектов.

Формат представления VRML-структуры в качестве полиморфной иерархии имеет ряд следующих преимуществ:

Node	{scale:0x012a9c38(1.00000000,1.00000000,1.00000000)}	Transform
Node	{def:"AcDbPolyFaceMesh_3D"}	Node
scale	0x012a9c38(1.00000000,1.00000000,1.00000000)	float[3]
translation	0x012a9c44(0.00000000,0.00000000,0.00000000)	float[3]
rotation	0x012a9c50(0.00000000,0.00000000,0.00000000)	float[4]
children	0x012a9c00 { size:1 }	std::vector<Node *,std::allocator<Node *>>
[capacity]	1	int
[allocator]	allocator	std::Compressed_pair<std::string,Node *>
[0]	0x012a9d20 {appearance:0x012a9d78 {materials:0x012a9d78 {materials:{ size:5 } } } }	Node * [Shape]
[Shape]	{appearance:0x012a9d78 {materials:0x012a9d78 {materials:{ size:5 } } } }	Shape
Node	{def:""}	Node
appearance	0x012a9d78 {materials:0x012a9d78 {materials:{ size:5 } } }	Appearance * (Material)
Material	{materials:{ size:5 } }	Material
Appearance	{def:""}	Appearance
materials	{ size:5 }	std::map<std::string,float *,std::Compressed_pair<std::string,Node *>>
[comparator]	less	std::pair<std::string,const float *,std::Compressed_pair<std::string,Node *>>
[allocator]	allocator	std::pair<std::string,const float *,std::Compressed_pair<std::string,Node *>>
[ambientIntensity]	0x012add00(0.117647000)	std::pair<std::string,const float *,std::Compressed_pair<std::string,Node *>>
[diffuseColor]	0x012b0f90(0.752941012)	std::pair<std::string,const float *,std::Compressed_pair<std::string,Node *>>
[emissiveColor]	0x012b0e78(0.000000000)	std::pair<std::string,const float *,std::Compressed_pair<std::string,Node *>>
[shininess]	0x012b0c90(0.0625000000)	std::pair<std::string,const float *,std::Compressed_pair<std::string,Node *>>
[specularColor]	0x012b0ac0(0.000000000)	std::pair<std::string,const float *,std::Compressed_pair<std::string,Node *>>
[базовое представление]	{-}	std::map<std::string,float *,std::Compressed_pair<std::string,Node *>>
Node	{def:""}	Node
def	""	std::string
geometry	0x012b1b0 {def:"" coordindex:{ size:10 } }	Geometry * (IndexedFaceSet)
[IndexedFaceSet]	{def:"" coordindex:{ size:10732 } normal:0x012b1b00 {def:"" points:{ size:3563 } } }	IndexedFaceSet
Geometry	{def:""}	Geometry
def	""	std::string
coordindex	{ size:10732 }	std::vector<std::vector<int,int>>
normalindex	{ size:10732 }	std::vector<std::vector<int,int>>
coord	0x012b1260 {def:"" points:{ size:3563 } }	Coordinate *
normal	0x012b1b00 {def:"" vector:{ size:3563 } }	Normal *
solid	false	bool
normalPerVertex	true	bool
Node	{def:""}	Node

Рис. 7. Фрагмент содержимого хранилища VRML-данных nodes

- возможность оперировать отдельными объектами в контексте всего их иерархического устройства;
- возможность обрабатывать весь массив накопленных данных единообразным способом, поскольку взаимодействие с каждым объектом использует функционал полиморфизма объектно-ориентированного подхода.

Последнее преимущество способствует более эффективному способу вывода всех накопленных данных в файл формата иной конфигурации.

Заключение

Всякая информация имеет индивидуальную сложность, которая начинает проявляться при анализе и разборе данных. Первоначальное предположение о простоте синтаксического анализа в связи с тем, что VRML видится четко структурированным языком, оказывается обманчивым. При более пристальном взгляде на задачу выявляются проблемы, связанные с поиском верного пути преодоления барьеров сложности.

В выборе метода синтаксического анализа важно обращать внимание на детали связности элементов информации, а также на последовательность их обработки и преобразований.

Метод рекурсивного охвата данных, представленный в статье, предлагает простое и элегантное решение проблем информационной связности и порядка обработки данных. Посредством рекурсивного способа чтения и структуризации объектов оказывается возможным создавать массивные иерархические структуры, с элементами которых впослед-

ствии можно легко взаимодействовать. Это открывает возможность оперировать объектами не только как отдельными единицами, но и с учетом всей иерархической структуры.

Представленный метод синтаксического анализа данных дает возможность, последовательно считывая данные, формировать "на лету" иерархию объектов непосредственно на основе структуры их формата. Иначе говоря, прямая выгода этого подхода заключается в отсутствии дополнительных преобразований и структуризации при считывании данных из файла. Это сокращает время считывания, а также понижает сложность программного кода.

Наиболее важный вывод состоит в том, что благодаря методу рекурсивного охвата данных оказывается возможным программно воспринимать и преобразовывать любые иерархические массивы символьных данных. Предложенный подход можно расширить в плане исследований параллельной обработки данных, а также для создания удобных структур хранения данных, связанных с обеспечением информационной безопасности [8]. Актуальными остаются проблемы сдерживания объема данных в файлах и повышения скорости считывания [9] — здесь также есть широкое поле приложений представленного метода. В плане управления большими данными [10, 11], в связи с появлением гетерогенных и гибридных систем хранения, таких как Hadoop и Spark, видится значительная перспектива применения данного метода. Наконец, широкий горизонт открывается в создании и обработке иерархических данных, используемых в интернете вещей (IoT) [12], что сулит значительные выгоды в восприятии распределенной информации.

Список литературы

1. Ziwar F., Elias R. VRML to WebGL Web-based converter application // 2014 International Conference on Engineering and Technology (ICET). IEEE. 2014, April. P. 1—6.
2. Whyte J., Bouchlaghem N., Thorpe A., McCaffer R. From CAD to virtual reality: modelling approaches, data exchange and interactive 3D building design tools // Automation in construction. 2000. Vol. 10, N. 1. P. 43—55.
3. Pazlar T., Turk Z. Interoperability in practice: geometric data exchange using the IFC standard // Journal of Information Technology in Construction (ITcon). 2008. Vol. 13, N. 24. P. 362—380.
4. Huber D. The ASTM E57 file format for 3D imaging data exchange // Three-Dimensional Imaging, Interaction, and Measurement. International Society for Optics and Photonics. 2011, January. Vol. 7864. P. 78640A.
5. Taubin G., Horn W. P., Lazarus F., Rossignac J. Geometry coding and VRML // Proceedings of the IEEE. 1998. Vol. 86, N. 6. P. 1228—1243.
6. Nadeau D. R. Building virtual worlds with VRML // IEEE Computer Graphics and Applications. 1999. Vol. 19, N. 2. P. 18—29.

7. **Pesce M.** VRML browsing and building cyberspace (No. 006 P473). New Riders Publishing, 1995.

8. **Tamassia R., Triandopoulos N.** Computational bounds on hierarchical data processing with applications to information security // *International Colloquium on Automata, Languages, and Programming*. Springer, Berlin, Heidelberg, 2005, July. P. 153–165.

9. **Schulze S. F., LaCour P., Buck P. D.** GDS-based mask data preparation flow: data volume containment by hierarchical data processing // *22nd Annual BACUS Symposium on Photomask Technology*. International Society for Optics and Photonics, 2002, December. Vol. 4889. P. 104–114.

10. **Krish K. R., Wadhwa B., Iqbal M. S., Rafique M. M., Butt A. R.** On efficient hierarchical storage for big data processing // *2016 16th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid)*. IEEE, 2016, May. P. 403–408.

11. **Wang X., Yang L. T., Kuang L., Liu, X., Zhang Q., Deen M. J.** A tensor-based big-data-driven routing recommendation approach for heterogeneous networks // *IEEE Network*, 2019, Vol. 33, N. 1. P. 64–69.

12. **Wan S., Zhao Y., Wang T., Gu Z., Abbasi Q. H., Choo K. K. R.** Multi-dimensional data indexing and range query processing via Voronoi diagram for internet of things // *Future Generation Computer Systems*, 2019, Vol. 91. P. 382–391.

D. V. Malakhovetsky, Student, e-mail: dmitriymalax@yandex.ru,

Moscow State University of Technology "STANKIN", Moscow, Russian Federation,

A. I. Razumowsky, PhD in Technical Sciences, Senior Researcher, e-mail: razumowsky@yandex.ru,
Trapeznikov Institute of Control Science RAS, Moscow, Russian Federation

Parsing Character Arrays by Recursive Scoping and Structuring Using the Example of VRML Data

Parsing character arrays by recursive scoping and structuring using the example of VRML data. The article presents a new method for structuring, segmentation and algorithmic design of the parser of character arrays using the example of VRML data. The key feature of the method is the ability to form a hierarchically complex object by means of recursive data structuring, which makes it possible to cover in aggregate the entire contents of the object, including its arbitrary nesting of child objects. This leads to high controllability of the development of the parsing algorithm, allowing you to focus each time on a specific piece of data, while not losing sight of the entire aggregate connectivity of information. The results obtained can easily be used in plans for creating convenient data storage structures related to information security, solving the problem of containing the amount of data in files, managing big data in heterogeneous systems, and processing hierarchical data in the Internet of Things.

Keywords: data analysis method, parsing, structuring, segmentation, recursion, VRML-format, polymorphic hierarchy

DOI: 10.17587/it.27.188-194

References

1. **Ziwar F., Elias R.** VRML to WebGL Web-based converter application, *2014 International Conference on Engineering and Technology (ICET)*, IEEE, 2014, April, pp. 1–6.

2. **Whyte J., Bouchlaghem N., Thorpe A., McCaffer R.** From CAD to virtual reality: modelling approaches, data exchange and interactive 3D building design tools, *Automation in Construction*, 2000, vol. 10, no. 1, pp. 43–55.

3. **Pazlar T., Turk Z.** Interoperability in practice: geometric data exchange using the IFC standard, *Journal of Information Technology in Construction (ITcon)*, 2008, vol. 13, no. 24, pp. 362–380.

4. **Huber D.** The ASTM E57 file format for 3D imaging data exchange, *Three-Dimensional Imaging, Interaction, and Measurement*, International Society for Optics and Photonics, 2011, January, vol. 7864, pp. 78640A.

5. **Taubin G., Horn W. P., Lazarus F., Rossignac J.** Geometry coding and VRML, *Proceedings of the IEEE*, 1998, vol. 86, no. 6, pp. 1228–1243.

6. **Nadeau D. R.** Building virtual worlds with VRML, *IEEE Computer Graphics and Applications*, 1999, vol. 19, no. 2, pp. 18–29.

7. **Pesce M.** VRML browsing and building cyberspace (No. 006 P473), New Riders Publishing, 1995.

8. **Tamassia R., Triandopoulos N.** Computational bounds on hierarchical data processing with applications to information security, *International Colloquium on Automata, Languages, and Programming*, Springer, Berlin, Heidelberg, 2005, July, pp. 153–165.

9. **Schulze S. F., LaCour P., Buck P. D.** GDS-based mask data preparation flow: data volume containment by hierarchical data processing, *22nd Annual BACUS Symposium on Photomask Technology*, International Society for Optics and Photonics, 2002, December, vol. 4889, pp. 104–114.

10. **Krish K. R., Wadhwa B., Iqbal M. S., Rafique M. M., Butt A. R.** On efficient hierarchical storage for big data processing, *2016 16th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGrid)*, IEEE, 2016, May, pp. 403–408.

11. **Wang X., Yang L. T., Kuang L., Liu, X., Zhang Q., Deen M. J.** A tensor-based big-data-driven routing recommendation approach for heterogeneous networks, *IEEE Network*, 2019, vol. 33, no. 1, pp. 64–69.

12. **Wan S., Zhao Y., Wang T., Gu Z., Abbasi Q. H., Choo K. K. R.** Multi-dimensional data indexing and range query processing via Voronoi diagram for internet of things, *Future Generation Computer Systems*, 2019, vol. 91, pp. 382–391.

Т. В. Нгуен, аспирант, e-mail: vietqn1987@gmail.com,
А. Г. Кравец, д-р техн. наук, проф., e-mail: agk@gde.ru,
Волгоградский государственный технический университет, г. Волгоград

Оценка и прогнозирование тенденций развития научных исследований на основе библиометрического анализа публикаций*

Предлагается подход к анализу и прогнозированию тематической эволюции исследований путем выявления восходящего тренда ключевых слов. Статистический анализ лексики публикаций позволяет проследить глубину проникновения новых идей и методов, которая может быть задана частотой появления слов, кодирующих целые концепции. В статье представлены разработанные метод анализа исследовательских тенденций и алгоритм ранжирования статей, основанный на структуре сети прямого цитирования. Данные для эксперимента были извлечены из Web of Science Core Collection, собраны 6696 публикаций в области искусственного интеллекта за период 2005–2016 гг. Для оценки предложенного метода было собрано 3211 публикаций с 2017 по 2019 гг. В результате метод оценивался путем проверки присутствия предсказанных ключевых слов в наборе самых частых терминов за период 2017–2019 гг. и обеспечил точность 73,33 %.

Ключевые слова: прогнозирование тенденций, тематическая эволюция, библиометрический анализ, ранжирование статей, искусственный интеллект, база данных Web of Science

Введение

В настоящее время число научных публикаций растет быстрыми темпами, и становится невозможным оставаться в курсе всего, что публикуется. Это затрудняет способность накапливать знания и анализировать данные предыдущих исследовательских работ. Один из самых важных исследовательских вопросов в вычислительном анализе научной литературы заключается в том, содержат ли обширные коллекции научного текста важные подсказки о динамике развития науки, которые помогут предсказать рост и падение научных идей, методов и даже областей знаний. Поэтому обзоры литературы все чаще принимают на себя решающую роль в обобщении результатов прошлых исследований, чтобы эффективно использовать существующую базу знаний, продвигать направление исследований и предоставлять основанное на фактических данных понимание практики применения и поддержания профессиональной экспертизы. При этом возможность заблаговременного прогнозирования научных

тенденций может потенциально революционизировать методы исследований, например, предоставляя финансирующим учреждениям возможность оптимизировать распределение ресурсов на перспективные направления.

Известно, что ученые используют различные качественные и количественные подходы к обзору литературы, чтобы понять и систематизировать более ранние результаты. Среди них библиометрия обладает потенциалом для введения систематического, прозрачного и воспроизводимого процесса обзора, основанного на статистических измерениях научной деятельности [1]. Во многих областях исследований используются библиометрические методы для изучения влияния научного направления, влияния научной школы или влияния конкретной статьи [2–4].

По этой причине основной целью данной работы является представление общего подхода к анализу и прогнозированию тематической эволюции конкретной области исследования путем выявления ключевых слов восходящего тренда. Этот подход представлен в виде метода анализа тенденций, алгоритма ранжирования публикаций и визуализации оценок терминов во временных рядах для обнаружения тематических направлений исследований.

*Исследование выполнено при финансовой поддержке РФФИ в рамках научных проектов № 19-07-01200 и № 20-37-90092.

В качестве примера исследования предложенный подход применяется для анализа тематической эволюции исследовательской области искусственного интеллекта (ИИ) с 2005 по 2016 гг. по наукометрическим данным платформы Web of Science (WoS).

Статья организована следующим образом: вначале представлен библиометрический анализ, собранный набор данных и некоторые полученные из библиометрического анализа результаты. Далее описаны алгоритмы ранжирования статей и извлечения значимых ключевых слов, а также результаты тематических тенденций исследования с указанием ключевых слов восходящего тренда. Затем метод оценивается путем обнаружения присутствий предсказанных ключевых слов в наборе истинных (наиболее часто встречающихся) терминов. В заключении приведены будущие направления исследования.

Анализ тенденций научных исследований на основе библиометрического анализа

В библиометрии есть две основные процедуры: анализ производительности и научное картирование. Анализ производительности нацелен на оценку групп научных субъектов (стран, университетов, департаментов, исследователей) и влияния их деятельности на основе библиографических данных. Научное картирование или библиометрическое картирование — это

пространственное представление того, как дисциплины, области, специальности и отдельные статьи или авторы связаны друг с другом [5, 6].

В данной работе рассмотрены данные из платформы WoS. Вначале по ключевому слову "Artificial intelligence" и категории "Computer science Artificial intelligence" отфильтровывались релевантные статьи, опубликованные за период 2005—2016 гг. и проиндексированные в ядре базы данных WoS (Web of Science Core Collection). На момент написания этой статьи были собраны для эксперимента 6696 публикаций. Кроме того, для оценки предложенного метода было собрано 3211 публикаций с 2017 по 2019 годы [7, 8].

Методы визуализации используются для представления научной карты и результатов различных аналитических обзоров. Например, исследовательские сети могут быть представлены для проведения анализа научных карт в целях выявления и визуализации тем исследований и их эволюции в течение периодов [9]. На рис. 1 (см. третью сторону обложки) показана эволюция в последовательные периоды времени тематических областей ИИ на базе информации об авторских ключевых словах, полученной из R-пакета Biblioshiny [10].

Альтернативно временной анализ предназначен для обнаружения концептуальной, интеллектуальной или социальной эволюции исследовательской области путем выявления закономерностей, тенденций и сезонности. Среди них обнаружение "взрыва слов" (word burst detection) представляет собой временной анализ, который нацелен на выявление терминов, имеющих высокую интенсивность использования в течение конечных периодов времени [11, 12].

Прогнозирование тенденций научных исследований на основе алгоритма ранжирования статей

Анализ вышеупомянутых методов показал, что их недостатком является отсутствие возможности агрегации данных из разделов публикации "Заголовки", "Аннотация", "Авторские ключевые слова" и "Ключевые слова плюс" (Keywords plus). В WoS "Ключевые слова плюс" являются индексными терминами, автоматически генерируемыми из названий статей, и включают фразы из одного или нескольких слов.

Кроме того, существует потребность в анализе и обобщении всех

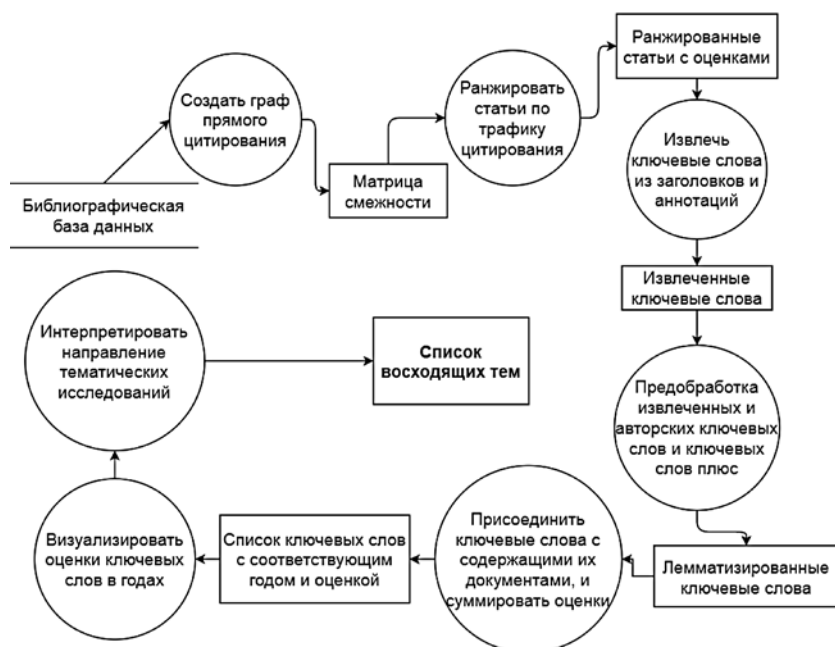


Рис. 2. Диаграмма потока данных метода анализа тенденций научных исследований

аспектов самых престижных и влиятельных публикаций (с учетом импакт-фактора). Темы тренда исследования и динамика слов просто идентифицируются по частоте терминов, не учитывая импакт-фактор статей, содержащих эти ключевые слова. При этом все упомянутые выше источники не показывают однозначно тематическую эволюцию через определенный промежуток времени.

Таким образом, предлагается новый метод анализа тенденций научных исследований путем отслеживания оценок ключевых слов в последовательные периоды времени. На рис. 2 продемонстрирована диаграмма потоков данных для предлагаемого метода.

Далее подробно описываются наиболее важные части предлагаемого метода.

Алгоритм ранжирования статей

Цитатный анализ или анализ библиографических ссылок в научных публикациях позволяет определять связи между публикациями, выявлять структуру областей знания и даже прогнозировать их развитие. Авторы работы [13] недавно предположили, что прямое цитирование более точно в представлении фронта исследования, чем библиографическое сочетание и совместное цитирование. Таким образом, предлагаемый алгоритм ранжирования основан на структуре сети прямого цитирования, сформированной из научных работ. Важность и влияние исследовательской работы хорошо отражается в ее цитировании в других публикациях. Исследовательские работы цитируют другие статьи, из которых принимают аргументы и доказательства. Влияние публикации прямо пропорционально важности, качеству и числу исследовательских работ, в которых она цитируется. Поэтому такое допущение используется в нашем алгоритме для ранжирования научных работ путем присвоения каждой из них авторитетной (импакт) оценки. Далее описываются процессы вычисления таких оценок и алгоритм ранжирования статей.

Шаг 1. Исходные данные. Пусть G — ориентированный граф n статей и m цитирований. Тогда $V(G)$ — множество статей, а $E(G)$ множество цитирований.

Далее, пусть для $p_i, p_j \in V(G)$ статья p_i цитирует p_j , если $e = \{p_i, p_j\} \in E(G)$.

Если использовать нотацию в виде графа для описания статьи p , то множество ссылок в списке литературы статьи p можно представить как набор $C_G^+(p) = \{r \in V(G) \mid (p, r) \in E(G)\}$, мощность которого определяется выходной

степенью вершины p ; $O(p)$ — число вершин, выходящих из p .

Число цитирований, полученных статьей p , равно $I(p)$ — это число входящих в p вершин, которое является мощностью множества $C_G^-(p) = \{r \in V(G) \mid (r, p) \in E(G)\}$. Отсюда следует, что $|C_G^-(p)| = I(p)$ и $|C_G^+(p)| = O(p)$.

Шаг 2. Инициализация. Пусть оценки статей обозначаются вектором $\mathbf{r} = [r_1, \dots, r_i, \dots, r_n]$, где r_i — импакт-оценка статьи i . Более того, нормализованным ограничением является уравнение $\sum_{i=1}^n r_i = 1$.

Инициализированные оценки на итерации $t = 0$: $r_i^{(t=0)} = \frac{1}{n}$, $i = \overline{1, n}$.

Шаг 3. Цикл вычисления. На каждой итерации (t) оценка статьи r_i рассчитывается последовательно по двум представленным ниже формулам:

$$r_i^{(t)} = \sum_{j \in C_G^-(i)} \frac{r_j^{(t-1)}}{O(j)};$$

$$r_i^{(t)} := r_i^{(t)} + \frac{1 - \sum_{k=1}^n r_k^{(t)}}{n}.$$

Шаг 4. Проверка остановки цикла. Цикл останавливается, когда вектор оценок $\mathbf{r}$ сходится с заданной точностью ε :

$$\mathbf{r}^{(t)} - \mathbf{r}^{(t-1)} < \varepsilon.$$

Если это условие не удовлетворено, то процесс вычисления возвращается к шагу 3.

Шаг 5. Вычисление итоговой оценки статей. Необходимо принять фактор времени при ранжировании исследовательских работ, чтобы уменьшить смещение по сравнению с недавними работами, которые получают меньше времени для изучения, следовательно, меньше цитируются исследователями по сравнению с более давними работами. Исходя из идеи m -индекса [14], который определяется как h/y , где h — индекс Хирша, y — число лет с момента публикации первой статьи ученого, для получения конечной оценки статьи делим каждую оценку на соответствующий возраст статей (множитель 100 был принят для того, чтобы последующая визуализация была более наглядна):

$$s_i = 100 \frac{r_i}{y_i}, \quad i = \overline{1, n},$$

где s_i — конечная оценка статьи; r_i — оценка статьи после выхода из цикла; y_i — возраст статьи.

Прогнозирование тенденции научных исследований в области искусственного интеллекта

При применении предложенного алгоритма для ранжирования статей к собранным данным области "Искусственный интеллект" за период 2005–2016 получим оценки каждой статьи и отсортируем их по оценкам в порядке убывания, которые представлены на рис. 3.

Далее создаем корпус текстов, агрегируя заголовки, аннотации, авторские ключевые слова, ключевые слова плюс от топ-20 лучших статей из приведенной выше таблицы рейтинга. После этого используется функция "Создание карты на основе текстовых данных" в библиометрическом программном обеспечении VOSviewer [15], чтобы создать карту терминов совместного использования (term co-occurrence map) на основе полученного корпуса текста, а затем извлечь все термины (ключевые слова) из этой карты. В программе VOSviewer этап идентификации терминов состоит из следующих пяти этапов:

- удаление заявлений об авторских правах в аннотациях;
- обнаружение и разбиение на предложения;
- маркировка части речи (Part-of-speech tagging): с использованием алгоритма маркировки части речи, предоставляемого библиотекой Apache OpenNLP, каждому слову присваивается часть речи, такая как глагол, существительное, прилагательное, предлог и т. д.;
- идентификация словосочетания: VOSviewer-программа определяет словосочетание как последовательность из одного или нескольких последовательных слов в предложении, последнее слово которой представляет собой существительное, а каждое из остальных слов является существительным или прилагательным. Чтобы определить фразы с существительными, VOSviewer рассматри-

вает только самые длинные фразы из существительных в предложении;

- объединение словосочетаний: объединение словосочетаний осуществляется путем удаления большинства не алфавитно-цифровых символов, удаления акцентов из символов, преобразования прописных символов в строчные и преобразования словосочетаний множественного числа в единственное число.

Описанная выше стадия идентификации термина дает набор терминов (ключевых слов), которые были идентифицированы в текстовых данных, доступных для VOSviewer. На следующем этапе, начиная с набора идентифицированных терминов, выполняется выбор терминов путем исключения терминов на основе порога частоты появления или с низким показателем релевантности, а также путем ручного исключения, а именно: оставляем только значимые термины, связанные с методом, алгоритмом, областью или разделом исследования, и исключаем в ручную все общие или неинформативные термины, такие как "direction", "first step", "action" и т. д.

Суммируя оценки статей, рассчитанные по разработанному алгоритму, сформируем итоговые оценки для ключевых слов, а затем сгруппируем полученные результаты по годам. Таким образом, можно продемонстрировать множество ключевых слов по годам во временных рядах и наблюдать, какие ключевые слова сохраняют восходящий тренд. Для ясности сохраняем только довольно частые и недавние ключевые слова (присутствуют в статьях за последние минимум четыре года) и разбиваем их на четыре подгруппы, одна из которых продемонстрирована на рис. 4 (см. четвертую сторону обложки).

Затем отбираем ключевые слова, число присутствия которых в исследуемом периоде более 9 (это "ant colony", "artificial bee colony", "honey bee", "particle swarm optimization", "swarm intelligence", "tree search"). Далее для проверки и прогнозирования тенденции изменения оценок ключевых слов необходимо приспособить модель полиномиальной регрессии [16]. Результаты такого приспособления модели показаны на рис. 5 (см. четвертую сторону обложки), из которого видно, что следующие ключевые слова сохраняют восходящие тенденции: "ant colony", "artificial bee colony", "particle swarm optimization", "swarm intelligence", "tree search".

Аналогичным способом в итоге собраны 30 явно восходящих по ключевым словам исследователь-

1	Title	Year	Score	Abstract	Author_keywords	Keyword_plus
2	On the performance of artificial bee colony (abc)	2008	0.195	Artificial bee colony (ABC) algorithm	swarm intelligence;	Differential Evo
3	A comprehensive survey: artificial bee colony	2014	0.054	Swarm intelligence (SI) is briefly d	Swarm intelligence;	Quantum Evolu
4	An artificial bee colony algorithm for the leaf-c	2009	0.043	Given an undirected, connected, w	Artificial bee colony	Abc Algorithm;
5	A survey of monte carlo tree search methods	2012	0.042	Monte Carlo tree search (MCTS) is	Artificial intelligence	Game
6	A modified artificial bee colony (abc) algorithm	2011	0.028	Artificial Bee Colony (ABC) algorit	Swarm intelligence;	Evolutionary Al
7	The 2014 general video game playing competit	2016	0.027	This paper presents the framework	Competitions; evolu	
8	The best-so-far selection in artificial bee colony	2011	0.026	The Artificial Bee Colony (ABC) al	Artificial Bee Colony	Image Registrat
9	A granular intrusion detection system using ro	2016	0.024	Security in computer networks is	Intrusion detection	Fuzzy; Decision
10	Affect control processes: intelligent affective in	2016	0.024	This paper describes a novel meth	Affect; Emotion; Soc	Emotion; Identi
11	Classification of dna microarrays using artific	2016	0.024	DNA microarray is an efficient nev	DNA microarrays; A	Particle Swarm
12	Computational interpretation of comic scenes	2016	0.024	Understanding intellectual produc	Computational mod	
13	Dynamic modeling based on a temporal-causal	2016	0.024	This paper presents a dynamic mo	Modeling; Dynamic;	Recurrent Neur
14	Future progress in artificial intelligence: a surv	2016	0.024	There is, in some quarters, concer	Artificial intelligence	AI
15	Optimization of electricity markets participati	2016	0.024	The electricity markets environme	Artificial intelligenc	
16	When thinking never comes to a halt: using for	2016	0.024	The recognition that human minds	Cognitive systems;	Approximabilit
17	Universal intelligence: a definition of machine	2007	0.022	A fundamental problem in artificia	AI; complexity the	Science
18	A universal measure of intelligence for artific	2005	0.022			
19	Artificial and natural intelligence integration	2015	0.021	The large amount of data generate	Data mining; visuali	
20	The coming of age of artificial intelligence in m	2009	0.019	This paper is based on a panel disc	Artificial intelligenc	Decision-Suppo
21	The primary language of ancient battles	2011	0.018	Linguistic Geometry (LG) is a type	Linguistic Geometry	

Рис. 3. Топ-20 лучших (влиятельных) статей, полученных по предлагаемому алгоритму

ских тенденций, и это также прогнозируемые тенденции после 2016 г.: "artificial neural network", "computational intelligence", "diagnosis", "disease", "emotion", "genetic algorithm", "machine learning", "ant colony", "cognitive science", "evolutionary algorithm", "evolutionary computation", "fuzzy set", "game theory", "granular computing", "human brain", "intrusion detection system", "machine intelligence", "markov decision process", "medicine", "multilayer perceptron", "particle swarm optimization", "pattern recognition", "radial basis function", "support vector machine", "swarm intelligence", "cognitive psychology", "health care", "rough set theory", "tree search", "fuzzy cognitive map", "pattern classification", "recurrent neural network", "artificial bee colony", "cognitive system", "differential evolution", "empirical evidence", "honey bee", "human reasoning", "network intrusion detection", "sociology", "tumor", "computer go", "electricity market", "gene-expression data".

Среди них такие ключевые слова, как "artificial neural network", "computational intelligence", "disease", "emotion", "genetic algorithm", "machine learning" показывают высокий темп эволюции, и можно логично предположить, что эти тематические исследования все еще сильно будут проявляться в последующие годы.

Экспериментальная оценка предложенного метода

Для проверки качества метода результаты оцениваются на основе статистического анализа лексики 3211 собранных статей из библиографической базы данных платформы WoS при поиске с запросом "Artificial Intelligence" и фильтром категории "Computer science Artificial Intelligence" за период 2017–2019 гг.

В первую очередь создаем корпус текста, комбинируя заголовки, аннотации, авторские ключевые слова, ключевые слова плюс всех статей из изложенного набора данных. Затем метод "Создание карты на основе текстовых данных" снова используется в библиометрической программе VOSviewer для создания карты

терминов совместного использования на основе полученного корпуса текста, тогда извлекаются все ключевые слова из этой карты совместного использования с порогом присутствия 50. Аналогично оставляем только значимые термины, касающиеся методов, алгоритмов, области или разделов исследования, и исключаем в ручную общие и неинформативные термины, такие как "argument", "distance", "thing", и т. д.

В результате получается набор топ-50 используемых ключевых слов в порядке убывания частоты с соответствующими показателями релевантности, которые представлены на рис. 6, а, б. Следует отметить, что термины с высоким показателем релевантности представляют конкретные темы в тексте, а термины с низкой релевантностью носят общий характер и не являются репрезентативными для какой-либо конкретной темы. После исключе-

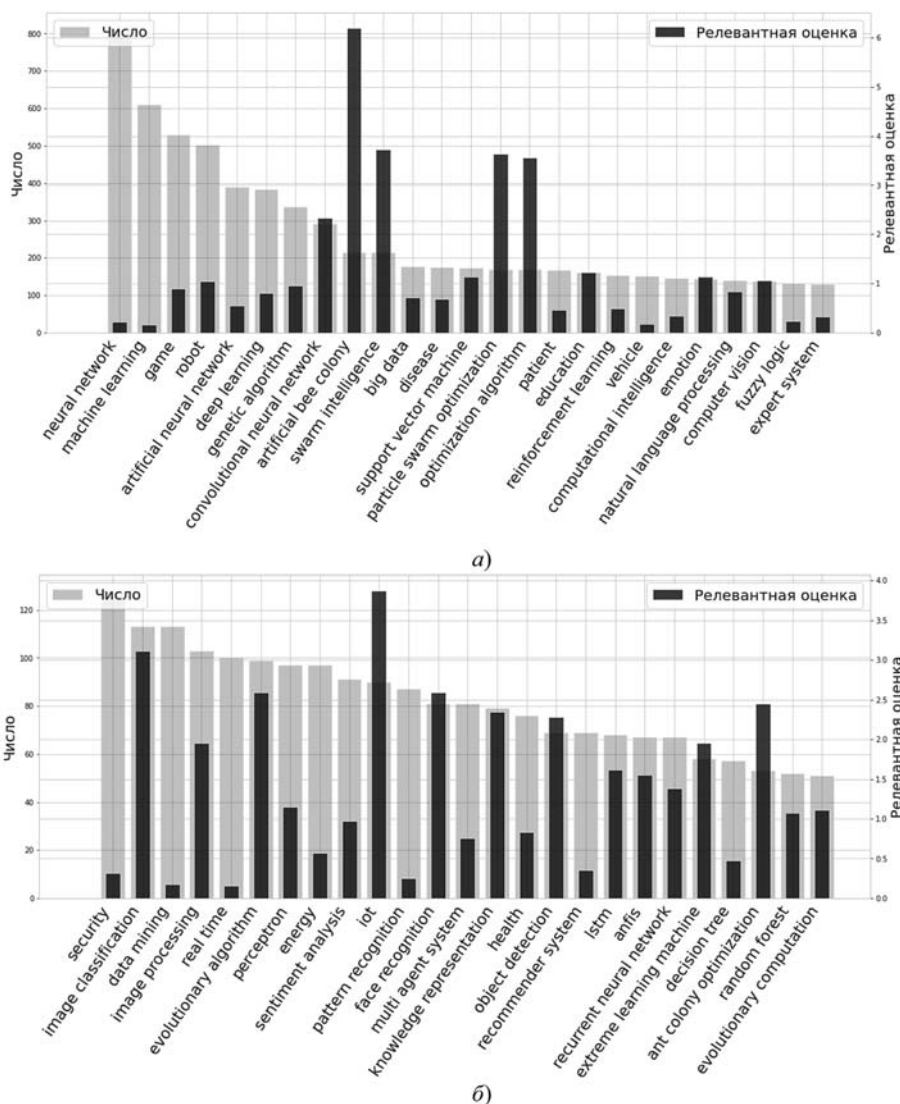


Рис. 6. Топ-50 (а — от 1 до 25, б — от 26 до 50) самых частых ключевых слов за период 2017–2019 гг.

ния терминов с низким показателем релевантности общие термины отфильтровываются, и фокус смещается на более конкретные и информативные термины [17].

Эти термины считаются принадлежащими фокусным исследованиям в период 2017–2019 гг. Необходимо определить точность нашего прогноза, сравнив число предсказанных ключевых слов с фактическим набором ключевых слов за выбранный период. Кроме того, сравнение ключевых слов выполняется не только по явному совпадению, но и по приблизительному совпадению (например, "game theory" и "game", "fuzzy cognitive map" и "fuzzy logic" считаются сходными ключевыми словами). В итоге обнаружено 22 ключевых слова (из всего 30 прогнозируемых ключевых слов восходящего тренда) в наборе 50 истинных ключевых слов фокусных исследований. Таким образом, точность прогноза составляет около 73,33 % (22/30). Однако полученную точность можно увеличить при расширении промежутка времени для оценки.

Заключение

Из-за неуклонного роста числа научных публикаций трудно вести обзор структуры, содержания и динамичного развития своей области науки и, тем более, связанных научных областей. Электронные ресурсы, в том числе базы данных публикаций, являясь реальным отображением процессов, происходящих в науке, содержат элементы, анализ которых позволяет оценивать и прогнозировать ее развитие.

В данной статье предложен общий подход к анализу и прогнозированию тематической эволюции области исследований ИИ путем определения ключевых слов восходящего тренда. Предложенный подход применен для анализа тематической эволюции исследований в области ИИ за период 2005–2016 гг. из базы данных платформы WoS. Оценка метода реализована путем обнаружения предсказанных ключевых слов в наборе истинных фокусных исследований ИИ за период 2017–2019 гг. Полученная точность прогноза составляет 73,33 %.

В будущем планируется повысить точность прогнозирования за счет улучшения алгоритма ранжирования статей и учета большего числа свойств, таких как факторы влияния авторов и журналов или ранжирование трудов конференций.

1. Iqbal W., Qadir J., Tyson G., Mian A. N., Saeed-ul H., Crowcroft J. A bibliometric analysis of publications in computer network research // *Scientometrics*. 2019. Vol. 119, N. 2. P. 1121–1155.
2. Merigo J. M., Pedrycz W., Weber R., de la Sotta C. Fifty years of Information Sciences: A bibliometric overview // *Information Sciences*. 2018. Vol. 432. P. 245–268.
3. Wang Q. A Bibliometric Model for Identifying Emerging Research Topics // *Journal of the Association for Information Science and Technology*. 2018. Vol. 69, N. 2. P. 290–304.
4. Robinson-Garcia N., Sugimoto C. R., Murray D., Yegros-Yegros A., Lariyiere V., Costas R. The many faces of mobility: Using bibliometric data to measure the movement of scientists // *Journal of Informetrics*. 2019. Vol. 13, N. 1. P. 50–63.
5. Cobo M. J., Jurgens B., Herrero-Solana V., Martinez M. A., Herrera-Viedma E. Industry 4.0: a perspective based on bibliometric analysis // *6th International Conference on Information Technology and Quantitative Management*. 2018. Vol. 139. P. 364–371.
6. Нгуен Т. В., Кравец А. Г. Метод прогноза исследовательских тенденций на основе алгоритма ранжирования статей // *Математические методы в технике и технологиях — ММТТ: сб. тр. XXXIII междунар. науч. конф. ММТТ—33 (г. Казань — г. Калининград — г. Минск (Беларусь) — г. Саратов). В 12 т. Т. 8 / Под общ. ред. А. А. Большакова; КНИТУ, КГТУ, СГТУ им. Гагарина Ю. А., БНТУ. Санкт-Петербург: Изд-во Политехнического ун-та, 2020. С. 123–127.*
7. Нгуен Т. В., Кравец А. Г. Автоматизация сбора информации из открытых интернет источников // *Математические методы в технике и технологиях (ММТТ—32): сб. тр. XXXII междунар. науч. конф. В 12 т. Т. 5 / Под общ. ред. А. А. Большакова; Санкт-Петербургский гос. технологический ин-т (техн. ун-т), Санкт-Петербургский ин-т информатики и автоматизации РАН, Санкт-Петербургский политехнический ун-т Петра Великого, Саратовский гос. техн. ун-т им. Гагарина Ю. А. [и др.]. Санкт-Петербург, 2019. С. 131–135.*
8. Нгуен Т. В., Кравец А. Г. Алгоритм работы веб-краулера для решения задачи сбора данных из открытых интернет источников // *Известия Санкт-Петербургского гос. технологического ин-та (технического ун-та)*. 2019. № 51 (77). С. 115–119.
9. Moral-Munoz J. A., Arroyo-Morales M., Piper B. F., Cuesta-Vargas A. I., Diaz-Rodriguez L., Cho W. C. S., Herrera-Viedma E., Cobo M. J. Thematic Trends in Complementary and Alternative Medicine Applied in Cancer-Related Symptoms // *Journal of Data and Information Science*. 2018. Vol. 3, N. 2. P. 1–19.
10. Aria M., Cuccurullo C. bibliometrix: An R-tool for comprehensive science mapping analysis // *Journal of Informetrics*. 2017. Vol. 11, N. 4. P. 959–975.
11. Sohrabi B., Vanani I. R., Jalali S. M. J., Abedin E. Evaluation of Research Trends in Knowledge Management: A Hybrid Analysis through Burst Detection and Text Clustering // *Journal of Information & Knowledge Management*. 2019. Vol. 18, N. 4. P. 27.
12. Нгуен Т. В., Кравец А. Г. Analyzing Recent Research Trends of Computer Science from Academic Open-access Digital Library // *8th International Conference on System Modeling and Advancement in Research Trends (SMART—2019, IEEE Conference ID: 46866) (22nd–23rd November, 2019): Proceedings / Eds.: A. Kr. Saxena, D. Parygin, D. Ather, V. Yadav; College of Computing Sciences & Information Technology, Teerthankar Mahaveer University (Moradabad, India), IEEE UP Section. New Delhi (India), 2019. P. 31–36.*
13. Klavans R., Boyack K. W. Which Type of Citation Analysis Generates the Most Accurate Taxonomy of Scientific and Technical Knowledge? // *Journal of the Association for Information Science and Technology*. 2017. Vol. 68, N. 4. P. 984–998.
14. Roldan-Valadez E., Salazar-Ruiz S. Y., Ibarra-Contreras R., Rios C. Current concepts on bibliometrics: a brief review about impact factor, Eigenfactor score, CiteScore, SCImago Journal

Rank, Source-Normalised Impact per Paper, H-index, and alternative metrics // *Irish Journal of Medical Science*. 2019. Vol. 188, N. 3. P. 939–951.

15. **Van Eck N. J., Waltman L.** Software survey: VOSviewer, a computer program for bibliometric mapping // *Scientometrics*. 2010. Vol. 84, N. 2. P. 523–538.

16. **Polynomial Regression**. [Электронный ресурс]. URL: <https://towardsdatascience.com/polynomial-regression-bbe8b9d97491/> (дата обращения: 20.11.2020).

17. **Van Eck N. J., Waltman L.** Citation-based clustering of publications using CitNetExplorer and VOSviewer // *Scientometrics*. 2017. Vol. 111, N. 2. P. 1053–1070.

T. V. Nguyen, Postgraduate Student, e-mail: vietqn1987@gmail.com,
A. G. Kravets, Dr. Tech. Sciences, Prof., e-mail: agk@gde.ru,
Volgograd State Technical University, Volgograd, Russian Federation

Evaluation and Prediction of Trends in the Development of Scientific Research Based on Bibliometric Analysis of Publications

The article proposes an approach to analyzing and predicting the thematic evolution of research by identifying an upward trend in keywords. Statistical analysis of the vocabulary of publications allows us to trace the depth of penetration of new ideas and methods, which can be set by the frequency of occurrence of words encoding whole concepts. The article presents a developed method for analyzing research trends and an article ranking algorithm based on the structure of a direct citation network. Data for the study was extracted from the Web of Science Core Collection, 6696 publications were collected for the experiment over the period 2005–2016 in the field of artificial intelligence. To evaluate the proposed method, 3211 publications were collected from 2017 to 2019. As a result, the method was evaluated by checking the presence of predicted keywords in the set of the most frequent terms for the period 2017–2019 and provided an accuracy of 73.33 %.

Keywords: trend forecasting, thematic evolution, bibliometric analysis, citation network, adjacency matrix, VOSviewer, Biblioshiny, article ranking, paper ranking, artificial intelligence, Web of Science database

Acknowledgments: This research was supported by the Russian Fund of Basic Research (grants No. 19-07-001200, No. 20-37-90092).

DOI: 10.17587/it.27.195-201

References

1. **Iqbal W., Qadir J., Tyson G., Mian A. N., Saeed-ul H., Crowcroft J.** A bibliometric analysis of publications in computer networking research, *Scientometrics*, 2019, vol. 119, no. 2, pp. 1121–1155.

2. **Merigo J. M., Pedrycz W., Weber R., de la Sotta C.** Fifty years of Information Sciences: A bibliometric overview, *Information Sciences*, 2018, vol. 432, pp. 245–268.

3. **Wang Q.** A Bibliometric Model for Identifying Emerging Research Topics, *Journal of the Association for Information Science and Technology*, 2018, vol. 69, no. 2, pp. 290–304.

4. **Robinson-Garcia N., Sugimoto C. R., Murray D., Yegros-Yegros A., Lariyiere V., Costas R.** The many faces of mobility: Using bibliometric data to measure the movement of scientists, *Journal of Informetrics*, 2019, vol. 13, no. 1, pp. 50–63.

5. **Cobo M. J., Jurgens B., Herrero-Solana V., Martinez M. A., Herrera-Viedma E.** Industry 4.0: a perspective based on bibliometric analysis, *6th International Conference on Information Technology and Quantitative Management*, 2018, vol. 139, pp. 364–371.

6. **Nguyen T. V., Kravets A. G.** Research trend forecasting method based on article ranking algorithm, *Matematicheskie metody v tekhnike i tekhnologiyah — MMTT: sb. tr. XXXIII mezhdunar. nauch. konf. MMTT–33* (g. Kazan — g. Kaliningrad — g. Minsk (Belarus) — g. Saratov), vol. 8, Sankt-Peterburg, Publishing house of Polytech. University, 2020, pp. 123–127 (in Russian).

7. **Nguyen T. V., Kravets A. G.** Automation of collection of information from open Internet sources, *Matematicheskie metody v tekhnike i tekhnologiyah (MMTT–32): sb. tr. XXXII mezhdunar. nauch. konf.*, vol. 5, Sankt-Peterburg, 2019, pp. 131–135 (in Russian).

8. **Nguyen T. V., Kravets A. G.** Algorithm of a web crawler to solve the problem of collecting data from open Internet sources, *Izvestiya Sankt-Peterburgskogo gos. tekhnologicheskogo in-ta (tehnicheskogo un-ta)*, 2019, no. 51 (77), pp. 115–119 (in Russian).

9. **Moral-Munoz J. A., Arroyo-Morales M., Piper B. F., Cuesta-Vargas A. I., Diaz-Rodriguez L., Cho W. C. S., Herrera-Viedma E., Cobo M. J.** Thematic Trends in Complementary and

Alternative Medicine Applied in Cancer-Related Symptoms, *Journal of Data and Information Science*, 2018, vol. 3, no. 2, pp. 1–19.

10. **Aria M., Cuccurullo C.** bibliometrix: An R-tool for comprehensive science mapping analysis, *Journal of Informetrics*, 2017, vol. 11, no. 4, pp. 959–975.

11. **Sohrabi B., Vanani I. R., Jalali S. M. J., Abedin E.** Evaluation of Research Trends in Knowledge Management: A Hybrid Analysis through Burst Detection and Text Clustering, *Journal of Information & Knowledge Management*, 2019, vol. 18, no. 4, pp. 27.

12. **Nguyen T. V., Kravets A. G.** Analyzing Recent Research Trends of Computer Science from Academic Open-access Digital Library, *8th International Conference on System Modeling and Advancement in Research Trends (SMART–2019, IEEE Conference ID: 46866) (22nd–23rd November, 2019): Proceedings*, College of Computing Sciences & Information Technology, Teerthanker Mahaveer University (Moradabad, India), IEEE UP Section, New Delhi (India), 2019, pp. 31–36.

13. **Klavans R., Boyack K. W.** Which Type of Citation Analysis Generates the Most Accurate Taxonomy of Scientific and Technical Knowledge?, *Journal of the Association for Information Science and Technology*, 2017, vol. 68, no. 4, pp. 984–998.

14. **Roldan-Valadez E., Salazar-Ruiz S. Y., Ibarra-Contreras R., Rios C.** Current concepts on bibliometrics: a brief review about impact factor, Eigenfactor score, CiteScore, SCImago Journal Rank, Source-Normalised Impact per Paper, H-index, and alternative metrics, *Irish Journal of Medical Science*, 2019, vol. 188, no. 3, pp. 939–951.

15. **Van Eck N. J., Waltman L.** Software survey: VOSviewer, a computer program for bibliometric mapping, *Scientometrics*, 2010, vol. 84, no. 2, pp. 523–538.

16. **Polynomial Regression**, available at: <https://towardsdatascience.com/polynomial-regression-bbe8b9d97491/> (date of access: 20.11.2020).

17. **Van Eck N. J., Waltman L.** Citation-based clustering of publications using CitNetExplorer and VOSviewer, *Scientometrics*, 2017, vol. 111, no. 2, pp. 1053–1070.

А. Н. Поляков, д-р техн. наук, проф., e-mail: anp_temos@mail.ru,

В. В. Позевалкин, аспирант, e-mail: pozevalkinvv@mail.ru,

Оренбургский государственный университет

Применение нейронной сети прямого распространения к прогнозированию тепловых характеристик станков*

Представлена методика исследования устойчивости результатов моделирования на основе искусственной нейронной сети применительно к тепловым характеристикам металлорежущих станков. Методика базируется на предложенной базовой математической модели, основанной на модальном подходе. Представлена архитектура типового программного средства, разработанного для изучения эффективности моделирования искусственной нейронной сети. Приведены результаты натурного и машинного экспериментов.

Ключевые слова: искусственная нейронная сеть, машинный эксперимент, тепловые характеристики станков, модальные параметры

Введение

Современные станки с ЧПУ используют методы искусственного интеллекта на различных этапах их жизненного цикла. Так, на этапе их проектирования методы искусственного интеллекта позволяют создавать оптимальную конфигурацию несущей системы станка, учитывающую относительные движения инструмента и заготовки [1]. На этапе эксплуатации и проведения опытно-конструкторских работ для достижения высокого уровня надежности станков создают различные системы диагностики, построенные на интеллектуальных моделях на основе нейронных сетей [2, 3]. Для компенсации флуктуации мощности резания, обусловленной традиционными факторами при механической обработке (колебания припуска и неоднородность механических свойств обрабатываемого материала), создают системы стабилизации на основе искусственных нейронных сетей [4, 5].

Сегодня современное высокотехнологичное производство активно использует высокоскоростные станки с ЧПУ. Применение современных систем смазывания, систем охлаждения, новых материалов, инновационных конструкторских решений позволяет создавать станки, работающие с большими скоростями резания. Однако даже сегодня проблема температурной погрешности станков остается актуальной и является предметом исследований и поиска новых технических и технологических разработок. Например, при испытании высоко-

скоростного обрабатывающего центра горизонтальной компоновки HDBS-63 (DALIAN MACHINE TOOLS GROUP Corp., КНР) даже на небольших частотах вращения (3000 мин^{-1}) была зафиксирована избыточная температура на переднем конце моторизованного шпинделя около 22°C и температурные смещения в осевом направлении около 35 мкм [6]. На максимальной частоте вращения шпинделя $24\,000 \text{ мин}^{-1}$ температурные смещения в осевом направлении составили более 110 мкм .

Особое место в достижении точности обработки на станках с ЧПУ занимают исследования в области систем компенсации температурной погрешности станков, создаваемых на основе искусственных нейронных сетей [7–15]. По разным оценкам применение таких систем позволяет сокращать до 80 % температурной погрешности станка, а в отдельных случаях их эффективность еще выше. Так, в работе [15] приведен результат компенсации температурной погрешности токарного обрабатывающего центра, что позволило уменьшить температурные смещения шпинделя от 19 до 1 мкм. Таким образом, снижение температурной погрешности станков с ЧПУ является одной из приоритетных задач современного высокотехнологичного производства, а применение методов искусственного интеллекта, в частности, искусственных нейронных сетей, в области обеспечения теплоустойчивости станков является одним из эффективных направлений решения этой задачи. Искусственные нейронные сети характеризуются тремя базовыми проблемами: необходимостью наличия экспериментальной базы данных, используемой для обучения сети; необходимостью применения

*Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 20-38-90045.

оптимизационной процедуры для отыскания весовых коэффициентов нейронов; необходимостью выбора среди многообразия архитектур искусственных нейронных сетей. Эти проблемы не позволяют безоговорочно говорить о превосходстве методов искусственного интеллекта над традиционными методами. Поэтому время от времени появляются исследования, посвященные анализу эффективности нейросетевых моделей, по сравнению, например, с регрессионным [16].

Необходимость использования нейросетевых технологий в режиме реального времени для управления станками накладывает строгие ограничения на однозначность генерируемых нейронной сетью результатов. Поэтому в данной работе было проведено исследование устойчивости получаемых решений по построению тепловых характеристик станка с ЧПУ на основе нейросетевой модели. Ранее построение таких моделей проводилось в работах [17, 18]. Были получены локальные результаты, которые не позволили выявить закономерности формирования устойчивого решения для построения тепловых характеристик.

1. Постановка задачи исследования устойчивости решения нейронной сети при построении тепловых характеристик станков

Несмотря на разнообразие архитектур искусственных нейронных сетей [19, 20] практика их применения к решению задач аппроксимации и прогнозирования показала, что сети прямого распространения с одним скрытым слоем нейронов достаточно эффективны. Поэтому в данном исследовании рассмотрена только одна архитектура искусственной нейронной сети.

Принятая архитектура нейронной сети. Многослойная сеть прямого распространения состоит из трех слоев: входного, одного скрытого и одного выходного. Входной сигнал распространяется только в прямом направлении. Для обучения сети используется алгоритм обратного распространения ошибки. Каждый нейрон скрытого слоя имеет сигмоидальную функцию активации. В этом случае выход нейрона может быть представлен в виде

$$y_j = \frac{1}{1 + \exp(-v_j)}, \quad (1)$$

где v_j — индуцированное локальное поле нейрона j .

При этом индуцированное локальное поле нейрона j для n -й итерации обучения сети имеет вид

$$v_j(n) = \sum_{i=1}^{m_1} w'_{ji}(n)x_i(n) + b_{1j}, \quad (2)$$

где m_1 — общее число входов; $x_i(n)$ — входной сигнал; $w'_{ji}(n)$ — вес синапса в скрытом слое; b_{1j} — значение смещения для нейрона скрытого слоя.

Выходной слой построен на одном линейном нейроне и описывается выражением

$$y_j(n) = \sum_{i=1}^{m_2} w''_{ji}(n)z_i(n) + b_{2j}, \quad (3)$$

где m_2 — число нейронов выходного слоя; $z_i(n)$ — входной сигнал для нейрона выходного слоя, который описывается выражением (1); $w''_{ji}(n)$ — вес синапса в выходном слое; b_{2j} — значение смещения для нейрона скрытого слоя.

В результате обучения сети генерируются веса синапсов и смещения для обоих слоев и соответствующих нейронов на основе выражений (2) и (3). Используемый при обучении сети алгоритм обратного распространения ошибки позволяет минимизировать функцию ошибки обучения вида

$$E = \frac{1}{2N} \sum_{n=1}^N \sum_{k=1}^M (d_k(n) - y_k(n))^2, \quad (4)$$

где N — число эпох или итераций обучения сети; M — число нейронов выходного слоя; $d_k(n)$ — желаемый отклик сети; $y_k(n)$ — фактический отклик сети.

2. Методы решения поставленной задачи

Особенности реализации тепловой модели станка. Для решения проблемы теплоустойчивости конкретного станка его тепловая модель может быть представлена своими тепловыми характеристиками [21].

Без учета запаздывания тепловая характеристика в конкретной точке несущей системы станка для температуры может быть представлена в виде

$$T_{ij}(t_j) = \sum_{k=1,4,\dots,3m-2} x_{k,i}(1 - e^{-t_j/x_{k+1,i}}) + \sum_{k=1,4,\dots,3m-2} x_{k+2,i} e^{-t_j/x_{k+1,i}}, \quad (5)$$

$$i = 1, \dots, N; j = 1, \dots, L,$$

где $x_{k,i}$, $x_{k+1,i}$, $x_{k+2,i}$ — модальные параметры; m — число температурных мод; i — номер датчика температуры; t_j — отдельный интервал времени, в который проводились измерения;

N — общее число термодатчиков; L — число интервалов времени, составляющих рассматриваемый диапазон времени $t \in [t_1, t_L]$.

Для температурных перемещений тепловая характеристика имеет вид

$$\delta_n(t_j) = \sum_{k=1,4,\dots,3m-2} x_{k,n} (1 - e^{-t_j/x_{k+1,n}}) + \sum_{k=1,4,\dots,3m-2} x_{k+2,n} e^{-t_j/x_{k+1,n}}, \quad (6)$$

где n может принимать значения X, Y, Z ; $\delta_n(t_j)$ — оцениваемое координатное перемещение вдоль оси n .

В выражениях (5), (6) модальные параметры для любой k -й моды с подстрочным первым индексом "1" соответствуют амплитуде соответствующей k -й экспоненциальной функции (максимальному значению этой функции или соответствующему значению ординаты ее асимптоты). Модальные параметры с подстрочным первым индексом "2" соответствуют тепловой постоянной времени. Начальный уровень соответствующей функции описывается модальным параметром с подстрочным первым индексом "3".

Вид выражений (5) и (6) объясняет выбор входного и выходного сигналов для нейронной сети. Представим множество наблюдаемых и прогнозируемых интервалов времени единым вектором t :

$$t = \{0, t_1, \dots, t_m, t_{m+1}, \dots, t_L\}, \quad (7)$$

где t_m — предельный момент времени, для которого сформирована обучающая выборка; t_L — предельный момент времени прогнозирования.

Аналогичные векторы можно сформировать для температур и температурных перемещений

$$T = \{0, T_1, \dots, T_m, T_{m+1}, \dots, T_L\} \\ \text{и } \delta = \{0, \delta_1, \dots, \delta_m, \delta_{m+1}, \dots, \delta_L\}, \quad (8)$$

где T_m и T_L — значения температуры в моменты времени t_m и t_L .

Если принять в качестве входного сигнала вектор t , а в качестве выходного сигнала принять или вектор температуры или вектор температурных перемещений, то при успешном решении с помощью нейронной сети задачи прогнозирования это позволит решить известную задачу сокращения длительности натурального эксперимента [22].

Скольльзящее окно. Очевидно, что в качестве входного сигнала можно принять вектор температуры или температурных перемещений вида

$$T = \{0, T_1, \dots, T_m\} \text{ или } \delta = \{0, \delta_1, \dots, \delta_m\}, \quad (9)$$

а в качестве выходного сигнала соответственно принять векторы вида

$$T = \{T_{m+1}, \dots, T_L\} \text{ и } \delta = \{\delta_{m+1}, \dots, \delta_L\}. \quad (10)$$

Для практической реализации (9) и (10) предложено использовать метод скользящего окна [23, 24]. Идея метода заключается в том, что задается некоторый интервал времени, в течение которого осуществляется наблюдение за процессом, в виде

$$t = \{0, t_1, \dots, t_m\}. \quad (11)$$

В течение этого интервала времени формируется вектор данных вида (9). Используя понятие скользящего окна, вводим ширину скользящего окна p , его смещение b и горизонт прогнозирования $Г$. Ширина скользящего окна p соответствует длине вектора входных данных — известных значений температуры или температурных перемещений. Размер смещения окна b определяет длину вектора целей или выходного сигнала в архитектуре нейронной сети. Обобщенный вектор X длиной m с учетом параметров p и b можно представить в виде матрицы:

$$\{X\} \rightarrow [X] = \begin{bmatrix} x_1 & x_{1+b} & \dots & x_{m-b-p+1} \\ x_2 & x_{2+b} & \dots & x_{m-b-p+2} \\ \dots & \dots & \dots & \dots \\ x_p & x_{p+b} & \dots & x_{m-b} \\ x_{p+1} & x_{p+b+1} & \dots & x_{m-b+1} \\ \dots & \dots & \dots & \dots \\ x_{p+b} & x_{p+2b} & \dots & x_m \end{bmatrix}. \quad (12)$$

Выражение (12) позволяет перейти от вектора длиной m к матрице размера $[(p+b) \times q]$. Очевидно, что число столбцов q зависит от соотношений трех чисел: m, p, b . Так как число столбцов — это натуральное число, то при соответствующем соотношении чисел m, p, b для параметра q можно записать соотношение вида

$$q = (m - p)/b. \quad (13)$$

Таким образом, имеющаяся выборка экспериментальных данных длиной m разбивается на q пар векторов: длиной p и b . На практике это означает, что прогнозирование искомых значений температуры или температурных перемещений в некоторые последующие b -е моменты времени осуществляется по соответствующим

p -м предшествующим по времени значениям температур или температурных перемещений. На первом шаге процедуры после определения первых b значений температур или температурных перемещений окно смещается именно на это число значений в выборке. Тем самым, спрогнозированные b значений искомой функции теперь попадают внутрь скользящего окна и принадлежат входному вектору, а последующие b значений искомой функции предстоит спрогнозировать, и они принадлежат выходному вектору данных. Используемая структура данных приводит к изменению архитектуры сети: входной слой сети будет представлен p нейронами; выходной слой — b нейронами. Проведя обучение нейронной сети, можно сформировать горизонт прогнозирования Γ за пределы обучающей выборки длиной m . При этом, следуя изложенной выше концепции метода скользящего окна, каждые b элементов нового ряда Γ прогнозируются последовательно, а весь ряд Γ формируется итерационно с последовательным смещением на b элементов.

Результаты экспериментальных исследований. Чтобы сформировать представление о зависимости тепловых характеристик от начальных условий, на рис. 1–3 приведено по пять реализаций экспериментальных тепловых характеристик температур и температурных перемещений. Испытания проводились в разные дни. Кривые температуры и температурного перемещения, сформированные в одном испытании, имеют совпадающие номера. Тепловые характеристики приведены для двух частот вращения шпинделя: 2500 и 4000 мин⁻¹.

Анализ экспериментальных кривых показал, что существует разброс экспериментальных данных как для температур, так и температурных смещений. На рис. 3 приведены тепловые характеристики, сформированные по данным термодатчика, фиксирующего изменение температуры кабинетной системы станка, и зависящие только от флуктуации температуры среды, окружающей станок, но не зависящие от режимов его работы. Их анализ показал

отсутствие жесткой зависимости разброса экспериментальных данных по температуре или температурным перемещениям от флуктуации начальных условий.

Сглаживание тепловых характеристик. Особенностью использования электронных измерительных приборов в натурных испытаниях является формирование ступенчатых тепловых характеристик. Это объясняется тем,

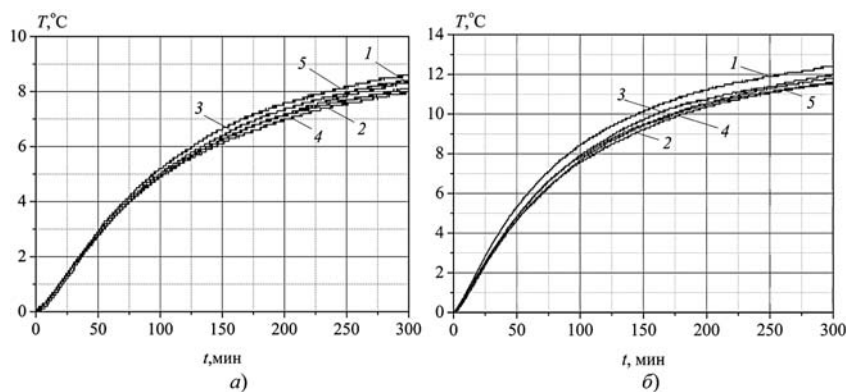


Рис. 1. Тепловые характеристики (температуры) для двух частот вращения шпинделя при различных начальных условиях: а — 2500 мин⁻¹; б — 4000 мин⁻¹

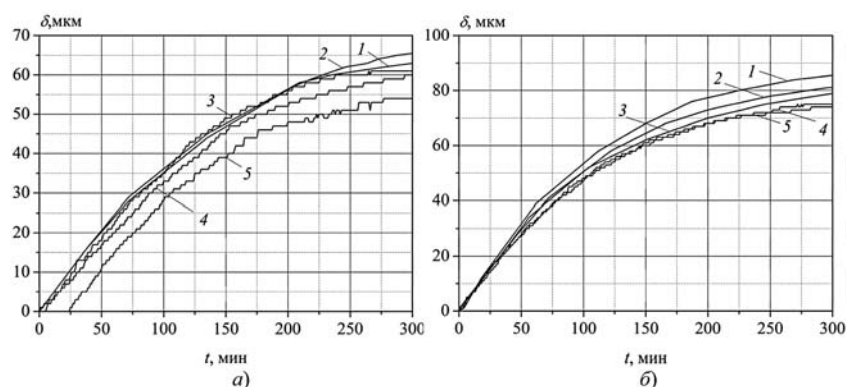


Рис. 2. Тепловые характеристики (температурные перемещения вдоль оси Z) для двух частот вращения шпинделя при различных начальных условиях: а — 2500 мин⁻¹; б — 4000 мин⁻¹

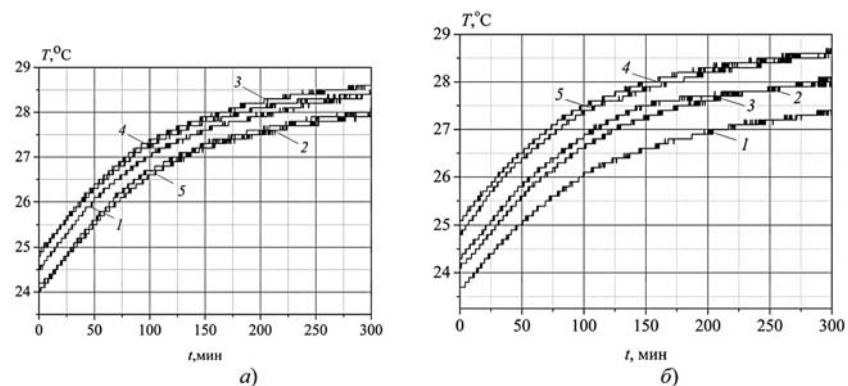


Рис. 3. Изменение начальных условий в процессе тепловых испытаний станка для двух частот вращения шпинделя при различных начальных условиях: а — 2500 мин⁻¹; б — 4000 мин⁻¹

что опрос и запись измеренных данных в файл выполняется с фиксированным шагом по времени. Поэтому непосредственное использование этих данных для обучения сети может создавать дополнительную ошибку, не связанную с реализацией тепловых процессов в станке. Для минимизации этой ошибки предлагается использовать процедуру сглаживания тепловых характеристик.

Для экспериментальных тепловых характеристик процедура сглаживания эквивалентна известной процедуре аппроксимации. Выше выражениями (5) и (6) уже представлен вид функции, который описывает эти кривые.

Однако все неудобство этих функций заключается в большой неопределенности искоемых коэффициентов, что порождает бесконечное множество решений [21, 22], в том числе и ошибочных. Поэтому для сглаживания экспериментальных кривых использована аппроксимация алгебраическими многочленами. В этом случае аппроксимирующая функция имеет вид

$$P_q(t) = \sum_{j=0}^q a_j t^j, \quad (14)$$

где $q \leq m - 1$ — степень полинома; m — число интервалов времени, в которые проводились измерения; t — аргумент аппроксимируемой функции (время).

Программное средство. Решение поставленной в работе задачи исследования устойчивости решения на основе нейросетевой модели для тепловых характеристик станка осуществлялось на основе машинного эксперимента. Для решения поставленной задачи было разработано программное средство (ПС) "Нейросетевое тепловое моделирование станков" в системе MATLAB. Главное окно ПС приведено на рис. 4. Архитектуру разработанного ПС можно отнести к слоистой. Каждый из слоев представляет собой комплекс программ, одинаковых по функциональности и на практике реализуемый в виде одного рабочего окна с элементами управления (рис. 4 и 5).

Основными модулями каждого слоя программ являются: модуль формирования исходных данных, расчетный модуль, модуль графического представления исходных данных, модуль графического и численного представления результатов обучения нейронной сети (промежуточный результат), модуль графического представления результатов имитационного моделирования сети для нового варианта данных в виде тепловых характеристик.

В качестве исходных данных для обучения нейронной сети могут использоваться экспериментальные данные, представляющие собой файл измеренных тепловых характеристик или температур в двенадцати точках несущей системы станка или температурных смещений

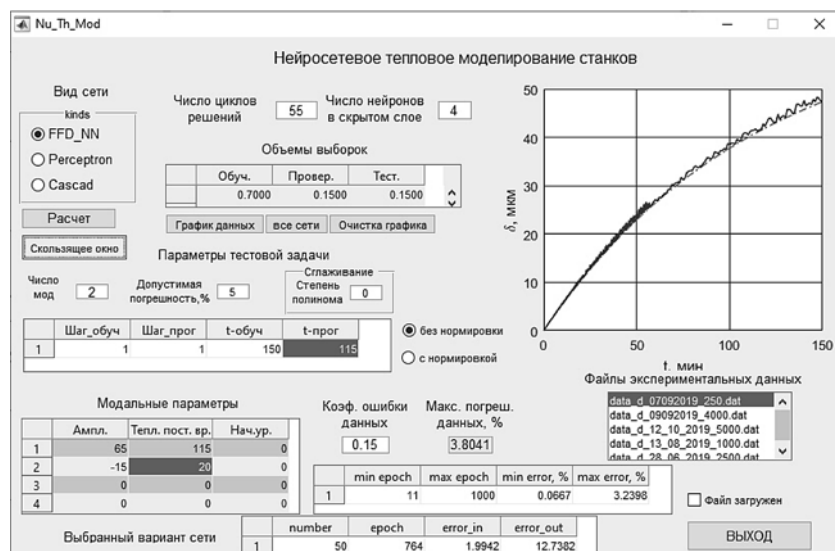


Рис.4. Главное окно программного средства "Нейросетевое тепловое моделирование станков"

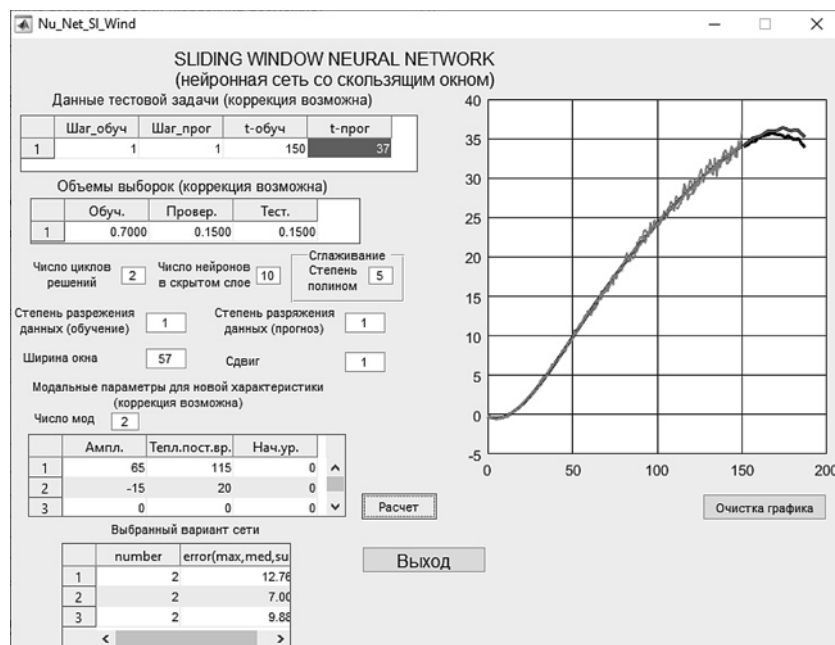


Рис. 5. Окно модуля "Скользящее окно"

шпиндельной головки, измеренных вдоль осей X , Y и Z . Погрешность исходных данных формируется автоматически из измерений. Также предусмотрено программное формирование исходных данных в соответствии с выражениями (5) и (6). В этом случае погрешность исходных данных задается программно.

С учетом этого коэффициента выражение (6) принимает вид

$$\delta_n(t_j) = \sum_{k=1,4,\dots,3m-2} x_{k,n}(1 - e^{-t_j/x_{k+1,n}}) + \sum_{k=1,4,\dots,3m-2} x_{k+2,n}e^{-t_j/x_{k+1,n}} + k\Delta_r t_j, \quad (15)$$

где Δ_r — случайная ошибка данных, генерируемая случайным образом, $\Delta_r \in [0, 1]$, k — масштабный коэффициент.

Исходные данные, задаваемые в виде (15), позволяют однозначно задать погрешность исходных данных и, соответственно, однозначно определить погрешность моделирования и установить соответствующие зависимости.

При реализации метода скользящего окна из главного окна ПС вызывается новое программное средство (рис. 5), включающее несколько программных модулей.

Машинный эксперимент. Машинный эксперимент проводили для двух вариантов входных данных. Для первого варианта на вход нейросети подавали вектор времени, а выходом являлись температуры или перемещения (ранее их вид описан в уравнениях (7) и (8)).

Во втором варианте исследовали метод скользящего окна, поэтому входной и выходной векторы имели один физический смысл: температура или температурные перемещения. В данном машинном эксперименте в качестве входного и выходного векторов сети были использованы температурные перемещения.

Для каждого варианта входных данных сети проводилась вариация следующих параметров: число нейронов в скрытом слое; размер входного и выходного вектора; погрешность входных векторов; размер обучающей, проверочной и тестовой выборки; функциональные особенности тепловых характеристик, подаваемые на вход сети или их мультимодальность; наличие и отсутствие нормировки входного вектора.

Особенностью построения искусственной нейронной сети (ИНС) являлась их циклическая генерация. Для этого в программном средстве предусмотрен параметр "Число циклов решений". Наличие различных вариантов ИНС обусловлено тем, что синаптические веса и смещение для каждого нейрона сети опре-

деляются из решения задачи оптимизации. Симуляция ИНС применительно к прогнозированию тепловых характеристик станков показала, что сложно найти стабильные, мало различающиеся решения. Поэтому для выбора одного варианта ИНС предложено использовать два похожих критерия вида

$$Kp_1 = \frac{1}{N_\lambda} \sum_{\zeta=1}^{N_\lambda} [(\delta_{\text{пр},\zeta} - \delta_{\lambda,\zeta})/\delta_{\lambda,\zeta}] \cdot 100; \quad (16)$$

$$Kp_2 = \max([(\delta_{\text{пр},\zeta} - \delta_{\lambda,\zeta})/\delta_{\lambda,\zeta}]) \cdot 100, \quad \zeta = \overline{1, N_\lambda}, \quad (17)$$

где $\delta_{\text{пр},\zeta}$, $\delta_{\lambda,\zeta}$ — прогнозные (по ИНС) и расчетные значения температурного перемещения; N_λ — число прогнозируемых элементов вектора $\delta_{\text{пр},\zeta}$.

Аналитические исследования и практика тепловых испытаний показали, что после 100 мин непрерывной работы станка все высокочастотные моды, характеризующиеся малым значением тепловой постоянной времени, уже не оказывают влияние на кинетику формирования тепловой характеристики [21, 22]. Поэтому в любой точке тепловой характеристики справедливо:

$$y = A_1(1 - e^{-\lambda_1 t}) + B_1 e^{-\lambda_1 t} + C + \varepsilon, \quad (18)$$

где A_1 , B_1 — амплитуда и начальный уровень первой температурной моды в физических координатах исследуемой функции y ; λ_1 — обратное значение тепловой постоянной времени первой моды; C — сумма амплитуд всех высокочастотных мод; ε — небольшая погрешность, обусловленная продолжающимися во времени процессами стабилизации высокочастотных мод.

Учитывая малость величины ε для станков с температурной погрешностью, существенно большей одного микрона, этой величиной можно пренебречь. Тогда можно аналитически определить тепловую постоянную времени тепловой характеристики как обратное значение от λ_1 по известным значениям ординаты тепловой характеристики в моменты времени t_1 и t_2 :

$$\lambda_1 = \ln(Y_1/Y_2)/(t_2 - t_1), \quad (19)$$

где $Y_1 = (A_1 + C - y_1)/(A_1 - B_1)$ и $Y_2 = (A_1 + C - y_2)/(A_1 - B_1)$.

Моменты времени t_1 и t_2 выбираются на участке тепловой характеристики, используемой для обучения сети, а момент времени t_2 разграничивает этот участок и участок прогнозирования. В этом случае любое прогно-

зируемое значение исследуемой функции y_3 в момент времени t_3 имеет вид

$$y_3 = A_1 + C - e^{\ln((A_1+C-y_2)/(A_1-B_1))-\lambda_1(t_3-t_2)}(A_1 - B_1). \quad (20)$$

Таким образом, при расчете критериев (16) и (17) в качестве прогнозных значений тепловой характеристики $\delta_{пр,\zeta}$ используются компоненты выходного вектора обученной ИНС, а расчетные значения $\delta_{\lambda,\zeta}$ определяются зависимостями (19) и (20).

Моделирование ИНС при различных комбинациях исходных данных осуществлялось при последовательном изменении одного параметра, но фиксации остальных и удовлетворении критериев (16) и (17).

Это позволило установить следующие результаты и закономерности.

Для первого варианта входных данных ИНС:

- для одномодальной функции наилучшие результаты моделирования ИНС достигаются при использовании одного нейрона в скрытом слое;

- при наличии двух мод предпочтительно использование для скрытого слоя нескольких нейронов, при этом не было выявлено строгих закономерностей о соответствии роста числа нейронов в скрытом слое и стабильности получаемых результатов моделирования ИНС;

- для одномодальной функции эффективность прогнозирования незначительно, но выше, так как размер входного вектора формировался за время, меньшее одной тепловой постоянной времени (устойчивые результаты фиксировались при построении прогнозной характеристики по входному вектору, полученному из наблюдений, по длительности составляющих 90 % тепловой постоянной времени тестовой тепловой характеристики); при моделировании двухмодальной функции устойчивые результаты фиксировались при построении прогнозной характеристики по входному вектору, полученному из наблюдений, по длительности составляющих 130 % тепловой постоянной времени первой моды тестовой тепловой характеристики;

- использование нормированных входных данных не способствовало повышению стабильности моделирования ИНС;

- при исследовании соотношений размеров обучающей, проверочной и тестовой выборок было установлено, что минимальный размер обучающей выборки составил 60 % длины вектора входных данных; минимальный размер проверочной и тестовой выборок не может быть менее 10 % длины вектора входных данных.

В качестве иллюстрации эффективности моделирования ИНС на рис. 6 приведены кривые для четырех вариантов вариации параметров моделирования ИНС: кривая 1 — зашумленный исходный вектор с известными модальными параметрами, сформированный в соответствии с формулой (15); кривая 2 — теоретическая тепловая характеристика, сформированная в соответствии с формулой (6); кривая 3 — прогнозная тепловая характеристика — результат моделирования ИНС, выделенный из множества решений критерием (17); кривая 4 — также прогнозная тепловая характеристика, представляющая результат моделирования ИНС и выделенная критерием (16). Рис. 6, а, б иллюстрируют эффективность моделирования ИНС для одномодальной тепловой характеристики, рис. 6, в, г — для двухмодальной. Из рис. 6, а видно, что в зависимости от выбранного критерия (16) или (17) можно получить различный результат. Однако на сегодняшний день обоснованного решения по приоритетности критерия (16) относительно (17) не существует. Поэтому при прочих равных условиях необходимо добиваться близких вариантов моделирования при использовании критериев (16) и (17), например,

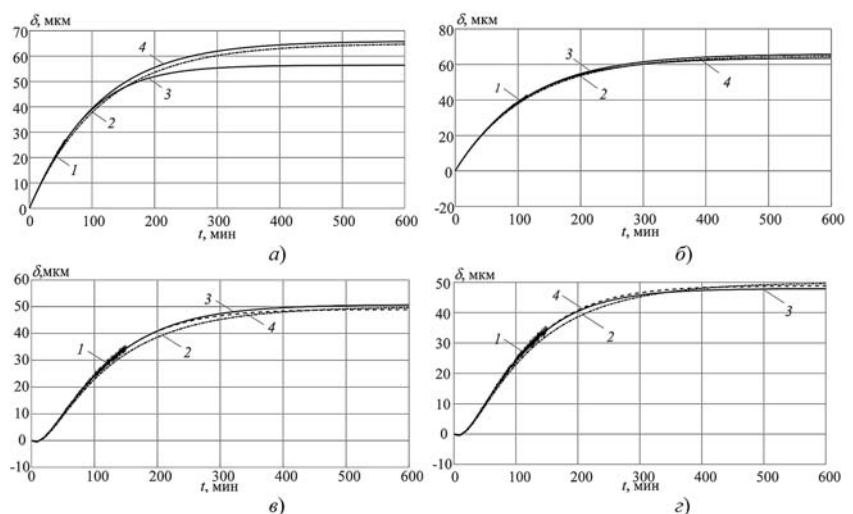


Рис. 6. Результаты моделирования ИНС:

а — 1 мода, длительность эксперимента $0,5\tau$ — 1 нейрон (в скрытом слое), $\Gamma > 3\tau$; б — 1 мода, 1τ , 1 нейрон, $\Gamma > 3\tau$; в — 2 моды, $1,3\tau$, $\Gamma > 3\tau$, 2 нейрона; г — 2 моды, $1,3\tau$, $\Gamma > 3\tau$, 3 нейрона (Γ — горизонт прогнозирования, т. е. интервал времени, на который осуществляется прогнозирование искомой функции; τ — тепловая постоянная времени первой моды тестовой тепловой характеристики)

за счет выбора длительности базового эксперимента, в ходе которого формируется обучающая выборка. Так, для тепловых характеристик, представленных на рис. 6, б, длительность базового эксперимента, в ходе которого формировалась обучающая выборка, была выбрана равной тепловой постоянной времени тестовой тепловой характеристики. Для тепловых характеристик, представленных на рис. 6, в, г, длительность базового эксперимента составляла 130 % тепловой постоянной времени первой моды тестовой тепловой характеристики.

Для второго варианта входных данных ИНС при использовании метода скользящего окна также были получены инвариантные результаты и закономерности:

- независимо от модальности исследуемой тепловой характеристики наилучшие результаты моделирования ИНС достигаются при использовании более двух нейронов в скрытом слое, но устойчивого снижения погрешности моделирования с ростом числа нейронов в скрытом слое установлено не было;

- для одномодальной тепловой характеристики при сопоставимой погрешности прогнозирования двухмодальной тепловой характеристики горизонт прогнозирования больше, но сопоставим с тепловой постоянной времени первой моды;

- для одномодальной тепловой характеристики более стабильные результаты моделирования ИНС достигаются при большем размере скользящего окна, для двухмодальной тепловой характеристики оптимальные соотношения для размера скользящего окна, горизонта прогнозирования и длительности базового эксперимента не однозначны и подбираются экспериментально.

На рис. 7 приведены пять кривых: кривые 1 — зашумленные тепловые характеристики на длине базового эксперимента. Длительность базового эксперимента на рис. 7, а составляла одну тепловую постоянную времени первой моды, а на рис. 7, б — 130 % тепловой постоянной времени первой моды тестовой тепловой характеристики. Чрезмерное увеличение длительности базового эксперимента автоматически приводит к снижению эффективности моделирования ИНС, так как из теории теплопроводности известно, что длительность базового эксперимента, равная 3τ , определяет более 95 % амплитуды тепловой характеристики. Поэтому в работе базовый эксперимент большой длительности не рассматривался. Кривая 2 — теоретическая тепловая ха-

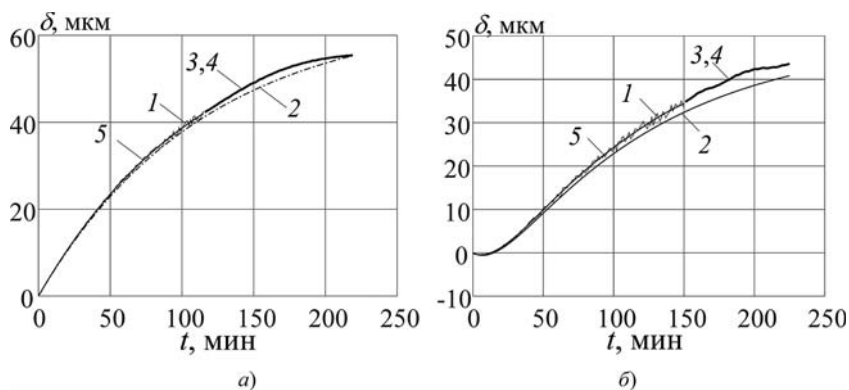


Рис. 7. Результаты моделирования ИНС:

а — 1 мода, $\Gamma = 0,9\tau$, 4 нейрона; б — 2 моды, $\Gamma = 0,65\tau$, 10 нейронов

рактеристика, сформированная в соответствии с (6); кривые 3 и 4 — прогнозные тепловые характеристики, полученные по результатам моделирования ИНС с использованием критериев (16) и (17), соответственно; кривые 5 — аппроксимированные функции с использованием соотношения (14).

Для обоих представленных на рис. 7 результатов моделирования ИНС ширина скользящего окна принята равной $0,5\tau$, горизонт прогнозирования Γ для одномодальной тепловой характеристики имеет чуть большие размеры. Представленные на рис. 7 результаты моделирования являются типовыми и показывают качество моделирования ИНС с использованием метода скользящего окна применительно к тепловым характеристикам станков.

Заключение

Представлена новая методика исследования устойчивости моделирования ИНС для тепловых характеристик станка. Новая методика позволила сформировать обобщенную концепцию изучения эффективности применения нейросетевых технологий в тепловом моделировании станков. Эта концепция определяет типовой набор варьируемых параметров моделирования, базовую математическую модель, основанную на модальном подходе, и архитектуру типового программного средства, которое может быть разработано для изучения эффективности моделирования ИНС. В работе исследован только один тип архитектуры ИНС, а с учетом их многообразия представленная методика показывает перспективность исследований ИНС в тепловом моделировании станков.

Практическим результатом использования нейронной сети прямого распространения в решении задачи прогнозирования тепловых харак-

теристик является возможность дальнейшего сокращения длительности натурного эксперимента станка; это сокращение оценивается приблизительно в 30 % от длительности натурного эксперимента, реализуемого в соответствии с подходом, основанном на экспериментальном модальном анализе, при котором устойчивые прогнозные решения получаются при длительности натурного эксперимента, равной удвоенному времени тепловой постоянной времени первой моды.

Представленная методика относится к решению задачи прогнозирования, но также может быть адаптирована к решению задачи идентификации тепловой модели и построению цифровых двойников. Однако при решении этих задач неизбежно появятся особенности, учет которых является предметом последующего изучения ИНС.

Список литературы

1. Heisel U., Pasternak S., Storchak M. et al. Optimal configurations of the machine tool structure by means of neural networks // *Prod. Eng. Res. Devel.* 2011. Vol. 5. P. 219–226.
2. Мунасыпов Р. А., Идрисова Ю. В., Масалимов К. А., Кудояров Р. Г., Фецак С. И. Способ оперативной диагностики модулей металлорежущих станков с использованием рекуррентных нейронных сетей с долгой краткосрочной памятью // *СТИН.* 2019. № 12. С. 5–11.
3. Масалимов К. А., Мунасыпов Р. А. Интеллектуальные модели на основе нейронных сетей с долгой кратковременной памятью для диагностики состояния станков в машиностроении // *Вестник Уфимского государственного авиационного технического университета.* 2018. Т. 22. № 2 (80). С. 138–145.
4. Белов М. П., Носиров И. С., Белов А. М. Исследование системы стабилизации мощности и усилия резания токарного станка с использованием нейронных сетей // *Известия СПбГЭТУ ЛЭТИ.* 2019. № 4. С. 77–84.
5. Белов М. П., Носиров И. С., Зувайдуллоев Ф. З., Бадалов Н. Ш. Система управления электроприводом подачи фрезерного станка с нейронными сетями // *Вестник Бохтарского государственного университета имени Носира Хусрава. Серия естественных наук.* 2017. № 2-2 (46). С. 41–45.
6. Du Z., Yao S., Yang J. Thermal behavior analysis and thermal error compensation for motorized spindle of machine tools // *Int. J. Precis. Eng. Manuf.* 2015. Vol. 16. P. 1571–1581.
7. Yang J., Shi H., Feng B. et al. Thermal error modeling and compensation for a high-speed motorized spindle // *Int. J. Adv. Manuf. Technol.* 2015. Vol. 77. P. 1005–1017.
8. Li Y., Zhao J., Ji S. Thermal positioning error modeling of machine tools using a bat algorithm-based back propagation neural network // *Int. J. Adv. Manuf. Technol.* 2018. Vol. 97. P. 2575–2586.
9. Ziegert J. C., Kalle P. Error compensation in machine tools: a neural network approach // *J. Intell. Manuf.* 1994. Vol. 5. P. 143–151.
10. Zhang Y., Yang J., Jiang H. Machine tool thermal error modeling and prediction by grey neural network // *Int. J. Adv. Manuf. Technol.* 2012. Vol. 59. P. 1065–1072.
11. Ma C., Zhao L., Mei X. et al. Thermal error compensation of high-speed spindle system based on a modified BP neural network // *Int. J. Adv. Manuf. Technol.* 2017. Vol. 89. P. 3071–3085.
12. Shi H., Jiang C., Yan Z. et al. Bayesian neural network-based thermal error modeling of feed drive system of CNC machine tool // *Int. J. Adv. Manuf. Technol.* 2020. Vol. 108. P. 3031–3044.
13. Guo Q., Yang J., Wu H. Application of ACO-BPN to thermal error modeling of NC machine tool // *Int. J. Adv. Manuf. Technol.* 2010. Vol. 50. P. 667–675.
14. Fu G., Gong H., Gao H. et al. Integrated thermal error modeling of machine tool spindle using a chicken swarm optimization algorithm-based radial basic function neural network // *Int. J. Adv. Manuf. Technol.* 2019. Vol. 105. P. 2039–2055.
15. El Ouafi A., Guillot M., Barka N. An Integrated Modeling Approach for ANN-Based Real-Time Thermal Error Compensation on a CNC Turning Center // *Advanced Materials Research.* 2013. Vol. 664. P. 907–915.
16. Majumder H., Maity K. P. Predictive Analysis on Responses in WEDM of Titanium Grade 6 Using General Regression Neural Network (GRNN) and Multiple Regression Analysis (MRA) // *Silicon.* 2018. Vol. 10. P. 1763–1776.
17. Поляков А. Н., Дьяконов П. И. Исследование теплового состояния станков с помощью нейронных сетей // *СТИН.* 2006. № 10. С. 10–11.
18. Поляков А. Н., Дьяконов П. И. Прогнозирование теплоустойчивости станков с помощью нейросетевого подхода // *Технология машиностроения.* 2003. № 6. С. 29–33.
19. Чичков Б. А., Раков П. И. Классификация искусственных нейронных сетей // *Научный вестник МГТУ ГА.* 2005. № 85. С. 82–87.
20. Хайкин С. Нейронные сети: полный курс / Пер. с англ. М.: Издательский дом "Вильямс", 2008. 1104 с.
21. Поляков А. Н., Гончаров А. Н., Марусич К. В. Исследование тепловых деформаций в металлорежущих станках // *Технология машиностроения.* 2011. № 2. С. 16–22.
22. Поляков А. Н. Методика сокращенных тепловых испытаний металлорежущих станков // *Техника машиностроения.* 2002. № 1 (35). С. 4–10.
23. Chou J.-S., Truong T. T. H. Sliding-window metaheuristic optimization-based forecast system for foreign exchange analysis // *Soft Comput.* 2019. P. 3545–3561.
24. Vafaeipour M., Rahbari O., Rosen M. A. et al. Application of sliding window technique for prediction of wind velocity time series // *Int. J. Energy Environ Eng.* 2014. Vol. 5. P. 105.

A. N. Polyakov, Professor, e-mail: anp_temos@mail.ru,
V. V. Pozevalkin, Postgraduate, e-mail: pozevalkinvv@mail.ru,
Orenburg State University, Orenburg, Russian Federation

Application of a Feedforward Neural Network to Predicting the Thermal Characteristics of Machine Tools

The paper presents a procedure for studying the stability of modeling an artificial neural network as applied to the thermal characteristics of machine tools. The topicality of this procedure is dictated by the ambiguity of the results generated by the neural network when constructing the predicted thermal characteristics of machine tools. Therefore, to select one of the possible solutions generated by the neural network, it was proposed to use two criteria. The effectiveness of their use is confirmed by the presented machine experiments. The methodology proposed in this work has made it possible to form a generalized concept for studying the effectiveness of the use of neural network technologies in thermal modeling of machine tools. This concept defines a typical set of variable modeling parameters, a basic mathematical model based on a modal approach, and an architecture of a typical software tool that can be developed to study the effectiveness of artificial neural network modeling. For each variant of the input data of the network, the following parameters were varied: the number of neurons in the hidden layer; the size of the input and output vectors; input vectors error; the size of the training, validation and test sample; functional features of thermal characteristics supplied to the network input or their multimodality; the presence and absence of normalization of the input vector. The paper presents experimental thermal characteristics for two spindle speeds of a vertical CNC machine. The results of the machine experiment are presented for six variants of the variable parameters of the mathematical model. The software tool used to carry out the machine experiment was developed in Matlab.

Keywords: artificial neural network, machine experiment, thermal characteristics of machine tools, modal parameters

Acknowledgments: The reported study was funded by RFBR according to the research project № 20-38-90045.

DOI: 10.17587/it.27.202-211

References

1. Heisel U., Pasternak S., Storchak M. et al. Optimal configurations of the machine tool structure by means of neural networks, *Prod. Eng. Res. Devel.*, 2011, vol. 5, pp. 219–226.
2. Munasypov R. A., Idrisova U. V., Masalimov K. A., Kudoyarov R. G., Fecak S. I. Technique for on-line diagnostics of metal-cutting machine tool modules using recurrent neural networks with long short-term memory, *STIN*, 2019, no. 12, pp. 5–11 (in Russian).
3. Masalimov K. A., Munasypov R. A. Intelligent models based on neural networks with long short-term memory for diagnosing the state of machine tools in mechanical engineering, *Bulletin of the Ufa State Aviation Technical University*, 2018, vol. 22, no. 2 (80), pp. 138–145 (in Russian).
4. Belov M. P., Nosirov I. S., Belov A. M. Investigation of the power and cutting force stabilization system of a lathe using neural networks, *Proceedings of Saint Petersburg Electrotechnical University*, 2019, no. 4, pp. 77–84 (in Russian).
5. Belov M. P., Nosirov I. S., Zuvaydullov F. Z., Badalov N. S. Control system for electric drive of milling machine with neural networks, *Bulletin of Bokhtar State University named after Nosir Khusrav. Natural Science Series*, 2017, no. 2(46), pp. 41–45.
6. Du Z., Yao S., Yang J. Thermal behavior analysis and thermal error compensation for motorized spindle of machine tools, *Int. J. Precis. Eng. Manuf.*, 2015, vol. 16, pp. 1571–1581.
7. Yang J., Shi H., Feng B. et al. Thermal error modeling and compensation for a high-speed motorized spindle, *Int. J. Adv. Manuf. Technol.*, 2015, vol. 77, pp. 1005–1017.
8. Li Y., Zhao J., Ji S. Thermal positioning error modeling of machine tools using a bat algorithm-based back propagation neural network, *Int. J. Adv. Manuf. Technol.*, 2018, vol. 97, pp. 2575–2586.
9. Ziegert J. C., Kalle P. Error compensation in machine tools: a neural network approach, *J. Intell. Manuf.*, 1994, vol. 5, pp. 143–151.
10. Zhang Y., Yang J. & Jiang H. Machine tool thermal error modeling and prediction by grey neural network, *Int. J. Adv. Manuf. Technol.*, 2012, vol. 59, pp. 1065–1072.
11. Ma C., Zhao L., Mei X. et al. Thermal error compensation of high-speed spindle system based on a modified BP neural network, *Int. J. Adv. Manuf. Technol.*, 2017, vol. 89, pp. 3071–3085.
12. Shi H., Jiang C., Yan Z. et al. Bayesian neural network-based thermal error modeling of feed drive system of CNC machine tool, *Int. J. Adv. Manuf. Technol.*, 2020, vol. 108, pp. 3031–3044.
13. Guo Q., Yang J., Wu H. Application of ACO-BPN to thermal error modeling of NC machine tool, *Int. J. Adv. Manuf. Technol.*, 2010, vol. 50, pp. 667–675.
14. Fu G., Gong H., Gao H. et al. Integrated thermal error modeling of machine tool spindle using a chicken swarm optimization algorithm-based radial basic function neural network, *Int. J. Adv. Manuf. Technol.*, 2019, vol. 105, pp. 2039–2055.
15. El Ouafi A., Guillot M., Barka N. An Integrated Modeling Approach for ANN-Based Real-Time Thermal Error Compensation on a CNC Turning Center, *Advanced Materials Research*, 2013, vol. 664, pp. 907–915.
16. Majumder H., Maity K. P. Predictive Analysis on Responses in WEDM of Titanium Grade 6 Using General Regression Neural Network (GRNN) and Multiple Regression Analysis (MRA), *Silicon*, 2018, vol. 10, pp. 1763–1776.
17. Polyakov A. N., Dyakonov P. I. Investigation of the thermal state of machine tools using neural networks, *STIN*, 2006, no. 10, pp. 10–11 (in Russian).
18. Polyakov A. N., Dyakonov P. I. Predicting the thermal stability of machine tools using a neural network approach, *Tekhnologiya Mashinostroeniya*, 2003, no. 6, pp. 29–33 (in Russian).
19. Chochkov B. A., Rakoc P. I. Classification of artificial neural networks, *Scientific Nauchnyi Vestnik MSTU GA*, 2005, no. 85, pp. 82–87 (in Russian).
20. Haykin S. *Neural Networks: A Comprehensive Foundation*, Moscow, Williams Publishing House. 2008, 1104 p. (in Russian).
21. Polyakov A. N., Goncharov A. N., Marusich K. V. Investigation of thermal deformations in metal-cutting machine tools, *Tekhnologiya Mashinostroeniya*, 2011, no. 2, pp. 16–22 (in Russian).
22. Polyakov A. N. Methodology for reduced thermal testing of metal-cutting machine tools, *Tekhnika Mashinostroeniya*, 2002, no. 1(35), pp. 4–10 (in Russian).
23. Chou J-S, Truong T. T. H. Sliding-window metaheuristic optimization-based forecast system for foreign exchange analysis, *Soft Comput.*, 2019, pp. 3545–3561.
24. Vafaeipour M., Rahbari O., Rosen M. A. et al. Application of sliding window technique for prediction of wind velocity time series, *Int. J. Energy Environ. Eng.*, 2014, vol. 5, p. 105.

В. А. Коквин, канд. техн. наук, доц., e-mail: kokovin@uni-protvino.ru,

А. А. Евсиков, канд. техн. наук, доц., e-mail: eaa@uni-protvino.ru,

А. П. Леонов, канд. техн. наук, доц., e-mail: anatoli.leonov@googlemail.com,

Государственный университет "Дубна", филиал "Протвино", г. Протвино, Московская область

Особенности организации и взаимодействия функциональных сетевых компонентов в распределенных управляющих системах

Рассматриваются особенности организации и взаимодействия устройств, названных функциональными сетевыми компонентами (ФСК). Анализируются тенденции в развитии промышленных сетевых экосистем, направленные на повышение интеллектуальных возможностей производственного оборудования. Обсуждаются возможности использования стандартизованных программных платформ для разработки управляющих приложений сетевых контроллеров распределенных технологических систем. Обоснована необходимость использования гибридных управляющих вычислителей на базе ARM контроллеров и программируемых логических интегральных схем. Предложена многоуровневая структура коммуникационной сети для взаимодействия ФСК в реальном времени, разработана и испытана модель в пакете AnyLogic 8.6.

Ключевые слова: распределенные управляющие системы, функциональные сетевые компоненты, мехатронные модули, МЭК-61499, ПЛИС, AnyLogic 8.6

Введение

Разработка распределенных систем, управляющих технологическими процессами промышленных или научных комплексов, связана с построением модели на основе формализации входных параметров и параметров текущего состояния этих процессов. Под распределенной управляющей системой (РУС) будем понимать не только географическое распределение отдельных подсистем, но и распределение алгоритма управления, направленного на решение общей технологической задачи.

При разработке многих РУС необходимо учитывать управление быстрыми технологическими процессами, которые требуют реакции исполнительных устройств в реальном времени. Эти требования задают жесткие ограничения на время отработки входных воздействий в программно-аппаратных приложениях элементов РУС.

Другая особенность распределенных систем заключается в том, что в РУС отдельные подсистемы могут быть взаимозависимыми по ал-

горитмам. В работе [1] представлена методика определения зависимости управляющих алгоритмов через построение графов зависимостей. Алгоритмы управления подсистем могут быть зависимы по данным или по управлению. Пример взаимосвязанных технологических процессов приведен в работе [2] для задачи выравнивания поверхности (в заданной плоскости) географически распределенных технологических платформ, находящихся в неустойчивой водной среде. При разработке РУС при решении этой задачи было необходимо определить зависимость алгоритмов подсистем, что напрямую влияло на особенности организации телекоммуникационной системы как части РУС.

Еще одним примером взаимосвязанных алгоритмов управления взаимосвязанных технологических процессов могут служить алгоритмы системы управления технологических подсистем циклического ускорителя заряженных частиц. Каждый ускоритель оснащен определенными средствами таймирования, которые необходимы для обеспечения режима реально-

го времени систем управления и синхронизации технологических процессов на ускорителе. Технологические процессы отдельных подсистем являются взаимосвязанными с точки зрения решения главной задачи — получения устойчивого пучка заряженных частиц с заданными параметрами [3].

Успешно внедряемая в производство концепция *Internet of Things (IoT)* [4], названная *Industrial IoT (IIoT)*, широко применяется ведущими фирмами, такими как *Intel*, *Bosch Si*, *ABB* и другими, обеспечивая успешную интеграцию компонентов *IIoT* в промышленную автоматизацию [5]. Интеллектуальные производственные решения этих компаний используют подключенные датчики и устройства для повышения производительности машины и человека и передачи данных в облако для более глубокого анализа в режиме реального времени. Развитие *IoT* и *IIoT* дало толчок появлению нового направления, которое с помощью сетевых технологий кооперирует работу роботов или роботизированных устройств. Это направление, названное *Internet of Robotic Things (IoRT)*, нацелено на реализацию робототехнических технологий путем расширения функциональности устройств *IoT* и *IIoT*. В работе [6] представлена концепция *IoRT*, подчеркивается огромная гибкость в процессе разработки и внедрения новых приложений для сетевой робототехники при достижении цели предоставления распределенных вычислительных ресурсов в качестве основной утилиты. Промышленные сетевые экосистемы, такие как *IoRT*, представляют собой распределенные технологические системы со сложными и многоуровневыми коммуникационными сетями, участники которых могут объединяться в кластеры (подсистемы).

При анализе состава подсистем и устройств РУС (например, [2, 3, 6]) среди участников могут быть не только мехатронные компоненты с сетевыми возможностями, но и, например, самодостаточные электротехнические устройства со встроенным интеллектом, т. е. устройства, в составе которых нет узлов точной механики (что является признаком мехатронных устройств). Под самодостаточностью будем понимать способность устройств решать самостоятельно часть распределенной технологической задачи, возложенной на них. При этом необходим обмен информацией с другими участниками через формирование событий или сообщений для решения всей задачи. Условно такие устройства можно назвать функциональными сетевыми компонентами (ФСК). К данному классу устройств будем относить устройства, которые имеют возможность по-

лучать и обрабатывать информацию (имеют вычислитель с сетевыми портами), имеющие физическую природу (а не только программный ресурс) и способные воздействовать на окружающую физическую среду [7]. Последнее условие по отношению к ФСК может быть определяющим, поскольку выделяет свойства данных устройств по отношению к внешней среде их использования. Кроме того, ФСК разной "природы" (мехатронные, электротехнические и т. д.) могут по-разному воздействовать на физическую среду, что отражается в особенностях их организации как подсистем РУС.

Целью данной работы является формулирование требований по организации аппаратно-программного обеспечения ФСК как участников РУС, предложение и обоснование решений по конфигурации коммуникационной сети для взаимодействия отдельных ФСК, моделирование коммуникационной сети с заданными критериями эффективности для определения предельных возможностей ее как масштабируемой системы.

1. Анализ работ по данной тематике

Анализ работ по аппаратно-программному обеспечению участников РУС показал большое разнообразие аппаратных и программных платформ в рамках одной системы. Это разнообразие обусловлено задачами (функциями) участников РУС и, как было сказано выше, разной природой ФСК, что создает дополнительные трудности при разработке РУС. В работах по данной тематике авторы предлагают унификацию требований, выделяя сетевые компоненты как основу для взаимодействия в рамках РУС.

Например, в работе [8] рассмотрены основные требования к мехатронным компонентам (МК) как к объектам распределенных технологических систем (РТС). В настоящее время мехатронные устройства и мехатронные системы все чаще используются как объекты автоматизации РТС. Наиболее широко мехатронные системы используются в тех отраслях, где требуется точное позиционирование исполнительных механизмов, быстрая реакция вычислителей на внешние и внутренние события, повышенная надежность и ограниченные массогабаритные параметры этих устройств. По мнению авторов работы [8], мехатронные компоненты должны иметь:

- коммуникационные интерфейсы сигналов (событий) и данных для взаимодействия с другими МК;

- интерфейсы ввода-вывода параметров управляемого процесса;
- набор данных и память для хранения информации о состоянии объекта;
- вычислительные ресурсы для реализации алгоритма управления этим объектом.

В недавнем прошлом такие характеристики могли себе "позволить" средние по производительности электронно-вычислительные машины, а не вычислители отдельных мехатронных компонентов. Но сегодня для создания производительных решений и этого уже недостаточно. Необходимы аппаратные возможности для формирования видеопотоков с помощью видеокамер и анализа этих видеопотоков на платформе нейросетевых ускорителей.

В работе [9] дано понятие интеллектуального мехатронного компонента (*Intelligent Mechatronic Component, IMC*) и сформулированы условия использования таких компонентов в РТС, разработанной на основании стандарта IEC 61499 [10]. Любой *IMC* может

- являться мехатронным устройством, т. е. представлять физическое функциональное устройство с датчиками, исполнительными механизмами и электронными схемами;
- содержать встроенное устройство управления, представляющее собой вычислительное устройство, имеющее в своем составе интерфейсы к датчикам, исполнительным приводам и коммуникационным сетям, для взаимодействия к другими *IMC*;
- иметь программное обеспечение с поддержкой данных и логики управления для реализации функций автоматизации стандарта IEC 61499.

Особенностью организации взаимодействия отдельных *IMC* является использование событийного интерфейса, через который осуществляются вызовы на исполнение определенного кода. Интерфейс события хорошо подходит для моделирования межпроцессного обмена сообщениями и хорошо формализуем. Вместе с тем данная форма организации может вызвать недетерминированность управляющей системы.

В работе [11] обсуждается концепция *IoRT*, возможности использования стандартных сетевых протоколов и разнообразных технологий, включая облачные технологии и искусственный интеллект. Обсуждается архитектура участников промышленной сетевой экосистемы, которая играет важную роль в разработке многоцелевых роботизированных систем для *IoRT*.

Организация распределенных систем управления коллективами роботов обсуждается в работе [12]. Подчеркивается, что распределенное управление уменьшает коммуникационный

трафик между отдельными роботами и центральным вычислителем (центральным командным пунктом [12]), но делает необходимой большую интеллектуализацию каждого робота. Это требует повышения вычислительной мощности управляющей системы робота для выполнения поставленных задач.

Любой ФСК может быть не только законченной подсистемой, как, например, система синхронизации быстрого вывода (ССБВ) пучка распределенного комплекса ускорителей Института физики высоких энергий, но и отдельными элементами этих подсистем [3]. В этом случае в установке ССБВ сетевое взаимодействие с другими подсистемами ускорительного комплекса осуществляется через общую таймерную систему [13], через которую транслируются кодированные таймерные сообщения и события. При разработке ССБВ как элемента технологической подсистемы не было ограничений по потребляемой энергии и конструктивным решениям, но были жесткие требования по точности "привязки" момента вывода к заданному уровню поля (энергии пучка) на участке ускорения. Для повышения точности синхронизации быстрого вывода, а главное детерминированности управляющих сигналов аппаратные решения ССБВ были в основном реализованы на базе цифровых матриц — программируемых логических интегральных схемах (ПЛИС). В приведенной работе ключевым требованием к ФСК является точность временной привязки и детерминированность управляющих сигналов.

Другие варианты реализации ФСК — это ФСК как отдельные элементы компактных технологических подсистем. В работе [14] проанализирована возможность повышения интеллектуализации отдельных подсистем промышленного робота, в частности руки робота-манипулятора, которая может взаимодействовать как устройство *IoRT* с другими участниками через облачные технологии. Создание таких ФСК ставит перед разработчиками задачи выбора архитектуры и алгоритмов встроенных приложений в условиях ограниченных ресурсов.

Разработчик при ограниченности ресурсов данного типа ФСК сталкивается с проблемами выбора между многими критериями и оптимизации решений. Поскольку сетевые взаимодействия ФСК определяют режим реального времени РУС, то можно выделить следующие критерии для оптимизации обработки данных, сжатия и передачи результата по сети во встраиваемых системах, реализованных на контроллерах и ПЛИС:

- быстродействие как функция сложности реализации алгоритма обработки данных;

- минимизация потребления энергии при обработке данных как функция быстродействия;
- минимизация потребления энергии как функция использованного объема ПЛИС;
- эффективная загрузка сетевого контроллера (критерий работы без простоев);
- степень сжатия данных как функция времени.

В задачах многокритериального выбора не всегда необходимо найти максимальное или минимальное значение функции. В некоторых случаях необходимо усреднить значение некоторого параметра, т.е. определить компромиссное решение. Интеллектуализация отдельных ФСК и всей системы управления позволяет использовать не формализованную модель управления, а модель, основанную на получении информации о реальном внешнем мире через датчики и систему, которая обладает способностью адаптироваться к внешним изменениям.

2. Программно-аппаратная организация ФСК как участника сетевых промышленных экосистем

ФСК как часть промышленной сетевой экосистемы должны отвечать определенным программно-аппаратным и коммуникационным требованиям. Анализ работ по выбранной тематике показывает, что основные тенденции по организации ФСК сводятся к унификации управляющих и коммуникационных подсистем, повышению вычислительной мощности каждого компонента и наличию на борту как проводных, так и беспроводных коммуникационных средств. Важным фактором вычислителей ФСК как участника распределенного процесса является обеспечение детерминированности передаваемых сообщений и событий [7].

Организация программного обеспечения ФСК

Важной задачей при создании и эксплуатации ФСК как части РТС является разработка программных управляющих приложений. Наиболее оптимальным решением по созданию программного обеспечения (ПО) промышленных систем становится использование общепринятых промышленных стандартов. До недавнего времени большинство разработок ПО базировалось на языках и спецификациях международного стандарта *International Electrotechnical Commission (IEC) IEC 61131-3* [15]. Принятие данного стандарта позволило унифицировать языки разработки управляющих приложений для программируемых логических контроллеров (ПЛК), что дало возмож-

ность портировать разработанные проекты на ПЛК различных производителей. Стандарт *IEC 61131-3* определяет централизованную модель управления и имеет существенные недостатки. Например, при добавлении в управляющую систему нового контроллера или при изменении алгоритма взаимодействия используемых контроллеров необходимо заново изменять программу центрального вычислителя. Кроме того, масштабирование и реконфигурирование такой системы требует сложных процедур и времени, а ФСК РТС не могут взаимодействовать между собой.

В 2005 г. *IEC* был принят новый стандарт *IEC 61499* [10], который определяет способ построения систем управления распределенными технологическими процессами. Архитектура *IEC 61499* строится на основе определений языка *Function Block Diagram (FBD) IEC 61131-3*, что позволяет выстраивать преемственность разрабатываемых программных приложений. В основе архитектуры *FBD* лежит *Function Block (FB)* с расширенными интерфейсными возможностями. Одно из основных расширений *FB* — событийный интерфейс, который позволяет явно определить последовательности выполнения *FB*. Каждый *FB* может содержать несколько инкапсулированных алгоритмов, к которым нет прямого доступа со стороны других *FB*.

Система управления, разработанная на основании стандарта *IEC 61499*, представляет собой набор устройств, взаимодействующих между собой посредством коммуникационной сети. Система управления реализует функции, описываемые с помощью приложений (*applications*). Эти приложения могут распределяться среди нескольких устройств (*devices*), в качестве которых могут использоваться ПЛК, программируемые контроллеры автоматизации и цифровые вычислители на платформе ПЛИС (*FPGA — Field-Programmable Gate Array*) [16]. Каждое устройство состоит из одного или нескольких ресурсов. Ресурс — это функциональная единица, которая имеет независимое управление своими операциями, включая выполнение алгоритмов. Приложение является программной функциональной единицей, предназначенной для решения определенной задачи в системе управления. Приложение представляется в виде сети связанных между собой *FB*, которые могут выполняться на различных ресурсах и устройствах системы управления (рис. 1). Для поддержки системы *IEC 61499* разработаны различные программные платформы. Наиболее известный коммерческий проект *ISaGRAF* [17] в виде средства поддержки проектирования (*workbench*)

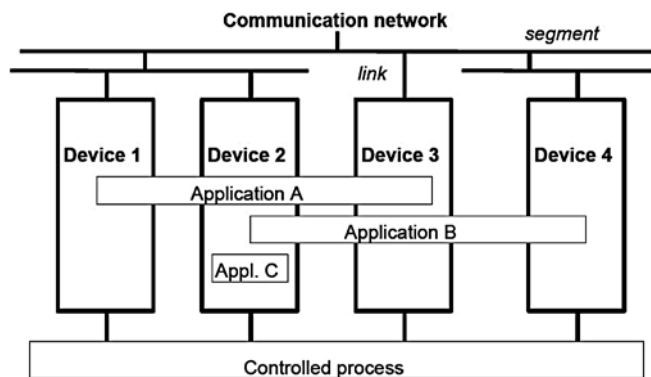


Рис. 1. Модель организации приложений и взаимодействие устройств системы IEC 61499 [10]

работает как со стандартом IEC 61131-3, так и с IEC 61499.

Проект системы *nxtSTUDIO* [18] ориентирован исключительно на IEC 61499 и позволяет совместить распределенную систему управления с HMI/SCADA. Кроме того, *nxtSTUDIO* дает возможность автоматизировать процесс построения каналов связи между контроллерами мехатронных устройств распределенной системы.

Программные платформы на базе стандарта IEC 61499, используемые для разработки проектов распределенных приложений, отвечают большинству требований, сформулированных выше для ФСК. Для управления технологическими процессами, где требуется быстрая реакция и строгая детерминированность управляющих сигналов, требуются дополнительные решения. Одним из таких решений может быть использование гибридных вычислителей на разных аппаратных платформах.

Аппаратные решения по организации ФСК

Аппаратные вычислительные решения ФСК зависят в первую очередь от функциональной задачи, решаемой данными устройствами. Но независимо от функциональной задачи в аппаратной архитектуре каждого ФСК есть вычислительное ядро с достаточно общими для всех ФСК вычислительными и коммуникационными требованиями. Анализ статей по данной тематике показал, что большой популярностью пользуются 32- или 64-разрядные контроллеры на основе ARM-ядер [19]. Например, на интернет-ресурсе [20] приведен список ПЛК отечественной разработки (начало 2020 года), которые поддерживаются ISaGRAF платформой (5 и 6 версия ISaGRAF). Из 30 ПЛК, представленных в списке, более половины реализованы на ARM-контроллерах. Популярность ARM-контроллеров объясняется разнообрази-

ем архитектуры, малым потреблением энергии (что во многих решениях достаточно критично), большим быстродействием и поддержкой программных платформ.

Рассмотрим в качестве примера организацию архитектуры ФСК, приведенную в работе [21]. В данной статье рассмотрено создание прототипа интеллектуального силового преобразователя (ИСП) с возможностью удаленного управления и сетевого взаимодействия с другими устройствами распределенной технологической системы. В различных отраслях производства ответственное место занимают силовые преобразователи, которые используются в составе управляющих приводов систем позиционирования, в формирователях мощных гармонических сигналов для ультразвуковых систем очистки поверхности изделий различного назначения [22, 23] и во многих других приложениях. Рассмотрим основные требования к аппаратным возможностям ИСП при использовании его в приложениях, отвечающих требованиям функционального сетевого устройства. Силовой преобразователь должен иметь следующие возможности:

- 1) с помощью встроенных датчиков воспринимать изменения внешней среды и компенсировать управляющими воздействиями эти изменения для поддержания заданных параметров;
- 2) самостоятельно принимать решения с учетом условий безопасности и заложенных в интеллект вычислителей ИСП правил;
- 3) взаимодействовать с другими устройствами системы РУС;
- 4) иметь возможность удаленной системной (например, отвечающей требованиям IEC 61499) конфигурации (желательно "на лету") или выполнять реконфигурацию на локальном уровне.

Силовая часть ФСК с функцией силового преобразователя (ИСП) построена на базе запатентованной технологии с двумя взаимосвязанными резонансными контурами [24], которая позволяет формировать инфранизкий гармонический сигнал большой мощности. В работах [25, 26] выполнено компьютерное моделирование процессов в силовых цепях ИСП и разработана методика настройки. В результате моделирования ИСП было получено семейство параметрических кривых выходного напряжения на нагрузке генератора с выходным колебательным контуром. Эти кривые позволили оценить быстродействие переключения параметров работы ИСП при удаленном управлении. В динамической системе моделирования *VisSim* [27] проведены исследования электропривода, работающего в связке с ИСП в различных режимах.

При разработке прототипа ИСП была поставлена задача создания конфигурируемой

платформы с универсальным вычислительным ядром, где в качестве силового преобразователя мог бы выступать, например, мощный генератор амплитудно-модулированных гармонических сигналов с регулируемыми выходными параметрами или стабилизированный блок питания. Кроме того, возможен вариант исполнения ИСП с несколькими силовыми преобразователями различного функционального назначения и одним блоком управления. Управление процессами несколькими силовыми преобразователями в реальном времени требует быстрой реакции вычислителя. В этом случае целесообразно контуры управления ИСП реализовать на ПЛИС.

Выбор управляющего контроллера осуществлялся исходя из требований пониженного энергопотребления (хотя в случае с ИСП это не является определяющим требованием), расширенных функциональных возможностей, наличия сетевых портов и достаточного быстродействия. Как было сказано выше, таким требованиям отвечают контроллеры с ARM-ядрами. Важным аргументом при выборе контроллера для ИСП являлась поддержка его средствами *open source (hardware и software)*. ARM-контроллер выбирался в составе платы прототипирования отвечающей следующим требованиям:

- 1) наличие на борту LAN с быстродействием ≥ 100 Мбит/с;
- 2) наличие беспроводной сети WiFi с двумя диапазонами: 2,4/5 ГГц, 802.11n;
- 3) наличие дополнительной шины ввода/вывода данных для подключения периферийных устройств (например, дисплея с сенсорным экраном);
- 4) возможность работать с различными операционными системами.

Указанным требованиям отвечают многие платы прототипирования [28]. В качестве контроллера для прототипа ИСП был выбран *Raspberry Pi 3* и *Raspberry Pi 4* [29] (для усовершенствованной версии ИСП). Дополнительным аргументом для выбора этих устройств послужила возможность работать не только под различными версиями операционной системы *Linux*, но и *Windows 10 IoT Core* [30].

Взаимодействие устройств распределенной промышленной сети необходимо для координации их работы, включая синхронизацию при реализации распределенного алгоритма. Применение LAN с протоколами стека *TCP/IP* и со случайным доступом к сети для целей взаимодействия не всегда оправдано, поскольку LAN, как правило, используется для системных целей: загрузка конфигураций, реализация облачных технологий и т. д. В условиях сильных электромагнитных помех от технологического

оборудования и внешних возмущений использование беспроводных технологий для решения этих задач затруднительно. Поэтому необходима дополнительная быстрая локальная связь на уровне платы управления, которая не занимала бы ресурсы микроконтроллера. Для разделения системной коммуникации ФСК (конфигурация, задание параметров в реальном времени) и формирования обмена данными для решения задачи управления в работе [21] предложена структура гибридного вычислителя: для решения первой задачи использован контроллер с ARM-ядрами, а для второй — ПЛИС. Структура ПЛИС позволяет сформировать большое число параллельно работающих коммуникационных линков как для формирования детерминированной локальной сети, так и для связи с датчиками по стандарту *IEC 61131-9:2013 (IO-Link)* [31].

Хорошей альтернативой сетевого интерфейса *Ethernet* на небольших расстояниях может быть интерфейс стандарта *IEEE-1355* [32]. Стандарт *IEEE 1355* является стандартом передачи данных для гетерогенного межсоединения. Эта сеть надежно отправляет пакеты данных переменной длины с большим быстродействием без предварительного согласования скорости. На рис. 2 представлена структура гибридного вычислителя ИСП с двумя портами коммуникационных дуплексных DS-линков.

Микроконтроллер реализован на основе 64-битного четырехъядерного процессора ARM Cortex-A53 с частотой 1,4 ГГц. Микроконтроллер оснащен 1 Гбайт оперативной памяти, что позволяет успешно использовать в качестве операционной системы различные клоны *Linux*. Программа взаимодействия микроконтроллера и *Field-Programmable Gate Array (FPGA)* через шину ввода/вывода написана на языке *Python 3.0*.

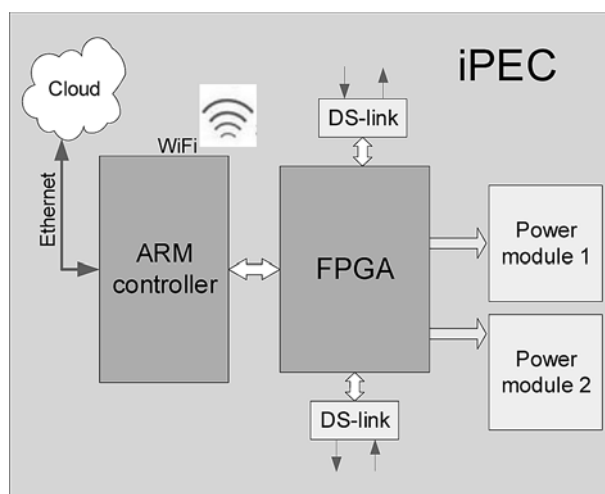


Рис. 2. Структура гибридного вычислителя ИСП с быстрой локальной магистралью [21]

Плата управления реализована на *FPGA Cyclone III EP3C16E144C8N* фирмы *ALTERA (INTEL)*. В улучшенной версии ИСП используется плата *EK-10CL025U256* на базе *FPGA* семейства *Cyclone 10* фирмы *INTEL* с *Gigabit Ethernet*-портом.

Разработка проекта для *FPGA* и моделирование протокола взаимодействия двух ИСП выполнено в пакете *Quartus Prime Lite Edition* [33]. Формировалось 64-разрядное сообщение и через *DS*-линк отправлялось другому ИСП. Транспортная задержка передачи на расстояние 30 м по витой паре категории 5е (с номинальной скоростью распространения сигнала 66 % от скорости света в вакууме) принималась минимальной и составляла порядка 10 мкс. Время работы автомата (с тактовой частотой 100 МГц) по кодированию и отправке 64-разрядного сообщения в первом ИСП составило 820 нс, а время приема и декодирования во втором ИСП — 910 нс. Общее время передачи сообщения составило порядка 12 мкс.

3. Выбор конфигурации сети и организация обмена функциональных сетевых компонентов

Выбор типа и числа коммуникационных сетей в РУС определяется многими факторами: решаемой технологической задачей, выбранной конфигурацией сети, степенью географической распределенности взаимодействующих подсистем и так далее. Выше было сказано, что обеспечить детерминированность передачи событий и сообщений (которые составляют основу распределенного управления) по сетям со случайным доступом невозможно с приемлемым джиттером. Например, в цитируемых выше работах [3, 13] описана общая таймерная система, предназначенная для передачи таймерных сообщений и событий по распределенным подсистемам ускорительного комплекса У-70 с ограничением джиттера, не превышающим 5 мкс. Данные требования исключают использование сети *LAN*.

В работе [34] обосновывается круговая конфигурация сети, состоящая из последовательно соединенных сетевых контроллеров (СК), реализованных на ПЛИС. Последовательная сеть формируется с помощью одного дуплексного *DS*-линка (сигналы данных и синхронизации). Сообщения передаются последовательно от одного СК к другому с помощью *DS*-интерфейса и могут иметь широ-

ковещательный статус, либо содержать адрес конкретного СК. Для контроля прохождения сообщений последний СК соединен с первым. Преимущества кольцевой последовательной сети заключаются в следующем:

- *DS*-линк каждого СК является сетевым повторителем-усилителем, что позволяет поддерживать высокую скорость с большим числом сетевых контроллеров ФСК;
- возможен контроль переданного сообщения по достоверности и времени передачи при возврате после прохождения всей сети;
- время передачи событий строго детерминировано.

На рис. 3 представлена структура распределенной технологической системы с двумя коммуникационными сетями — *LAN (TCP/IP)* и дополнительной локальной (*DS*-линки). Каждое ФСК-устройство может представлять собой определенную функциональную технологическую единицу. Это может быть мехатронный модуль, отдельные узлы роботизированной системы или силовой преобразователь. Устройства ФСК2, ФСК2.1, ФСК2.2, ФСК2.3 и ФСК4, ФСК4.1, ФСК4.2 объединены в кластеры с однонаправленной передачей данных внутри кластера. Устройства ФСК1, ФСК1.1, ФСК1.2 имеют дуплексную организацию передачи событий и сообщений с помощью *DS*-линков. Хотя бы одно ФСК-устройство из каждого кластера имеет доступ к сети *LAN* для выполнения задач, требующих доступа к общим базам данных: для реконфигурации, передачи измеренных параметров технологического процесса или видеоданных и другого информационного обмена. На рис. 3 это "головные" устройства ФСК1, ФСК2, ФСК3 и

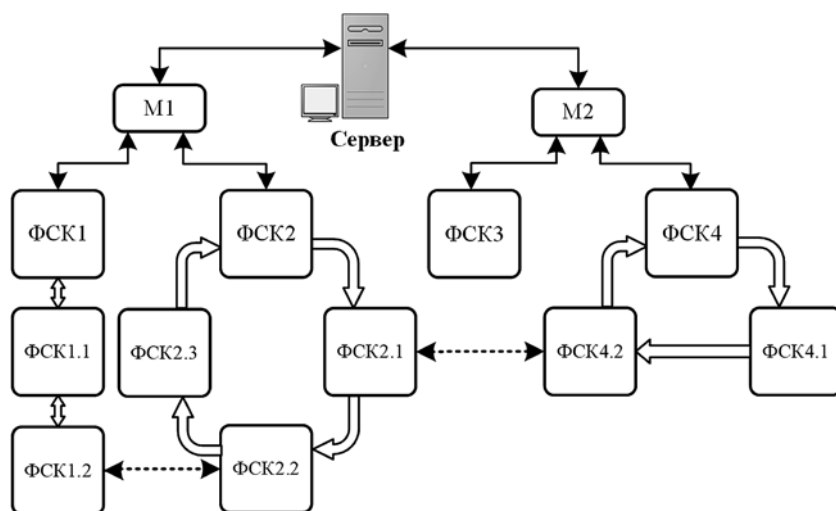


Рис. 3. Структурная схема организации коммуникационной сети для взаимодействия ФСК

ФСК4, подключенные к маршрутизаторам М1 и М2. Информационный обмен любого ФСК с сервером идет через эти "головные" ФСК. На рис. 3 штриховыми линиями показаны дополнительные коммуникационные связи между ФСК разных кластеров. Эти связи могут образовывать третий уровень сетевого взаимодействия. На рис. 3 не показан четвертый сетевой уровень, который обеспечивает ФСК информацией от датчиков (*IO-Link*) технологического процесса.

Для проверки эффективности предложенных решений была построена модель коммуникационной сети, включающей три ФСК: ФСК4, ФСК4.1 и ФСК4.2

4. Разработка и испытания модели конфигурационной сети в пакете *Anylogic 8.6*

Коммуникационную сеть как разновидность сетевой системы можно отнести, с точки зрения построения модели, к системе массового обслуживания. Это позволяет в качестве модели выбрать имитационный способ, который имеет ряд преимуществ над аналитическим способом моделирования [35]. В нашем случае важно то, что имитационная модель будет полностью отображать структуру коммуникационной сети и давать возможность исследовать поведение во времени. Для построения модели использован программный пакет *AnyLogic 8.6* [36].

Для большей достоверности создаваемой модели за основу возьмем исходные данные и требования, предъявляемые к общей таймерной системе (ОТС) ускорительного комплекса, описанной в работе [13]. Авторы принимали активное участие в создании ОТС, и многие решения предложенной конфигурационной сети были отработаны при разработке ОТС [37].

В процессе опытной эксплуатации одного из сегментов ОТС на ускорительном комплексе были получены экспериментальные данные распределения входных таймерных сообщений, поступающих на узлы ОТС. Анализ интервалов времени между сообщениями показал, что поток сообщений обладает тремя признаками простейшего или стационарного потока с пуассоновским распределением (т. е. поток стационарен, ординарен и без последствия) [35]. Таким образом, для построения модели в качестве входного потока возьмем поток с экспоненциальным распределением. Для упрощения вычисления суммарной интенсивности потоков сделаем допущение, что все входные потоки представляют собой стационарные потоки с пуассоновским распределением. Это

позволяет найти суперпозицию всех входных потоков методом простого сложения.

Определим структуру модели. Структура модели должна максимально соответствовать реальной структуре коммутационной сети. В предложенной конфигурации сети на рис. 3 ФСК объединены в три кластера. Наибольший интерес для построения модели представляют кластеры с круговой организацией сети, поскольку кластер с дуплексным обменом имеет удвоенное сетевое аппаратное обеспечение и, соответственно, большую производительность сетевого трафика.

Возьмем за основу структуры модели кластер, состоящий из трех ФСК: ФСК4, ФСК4.1 и ФСК4.2. ФСК в этом кластере взаимодействуют через однонаправленную *DS*-сеть, передавая сообщения друг другу по кольцу. Источниками сообщений в каждом ФСК являются потоки от датчиков и от соседнего ФСК. Сообщения, передаваемые по круговой *DS*-сети, имеют наивысший приоритет прохождения, а сообщения, поступающие в ФСК от "своих" датчиков, обрабатываются в первую очередь.

Будем считать, что во всех ФСК моделируемого кластера набор функциональных блоков одинаков. В модели ФСК4.2 добавлен блок конвейера, через который передаются сообщения в ФСК2.1 с определенным тактом, что может быть использовано для синхронизации работы различных кластеров ФСК. Сообщения по *LAN* принимаются все время, а передаются в отсутствие выполняемого алгоритма (для реконфигурации, передачи измеренных параметров технологического процесса или видеоданных). Линия связи с другим кластером может иметь дуплексную связь, но основное ее назначение — синхронизация, поэтому будем считать, что моделируемый кластер является мастером для остальных. В модели не учитываются сообщения, поступающие из *LAN* и других кластеров. Ниже представлены блоки программы *AnyLogic 8.6*, использованные для построения модели (рис. 4).

- *SourceSenN* — источник сообщений от датчиков (где $N = 1$ — для ФСК4, $N = 2$ — для ФСК4.1 и $N = 3$ — для ФСК4.2) с экспоненциальным распределением интервалов поступающих событий;
- *SinkDSN* — уничтожает "свои" сообщения, вернувшиеся по *DS*-сети;
- *SinkFNCN* — уничтожает все сообщения после обработки, поступившие в данный ФСК;
- *SinkFNC2I* — имитирует прием сообщений ФСК2.1 от текущего кластера;
- *SplitDSN* — создает копии сообщений, поступающие в ФСК из *DS*-сети (вся опера-

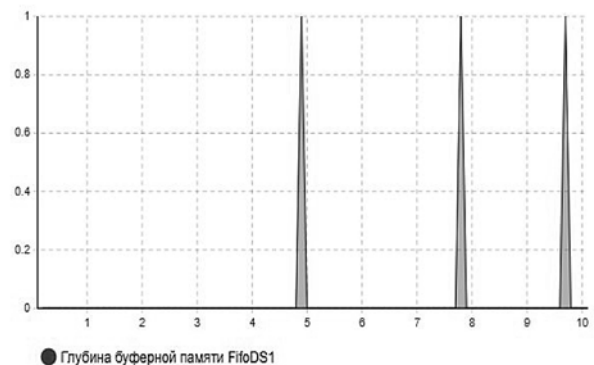
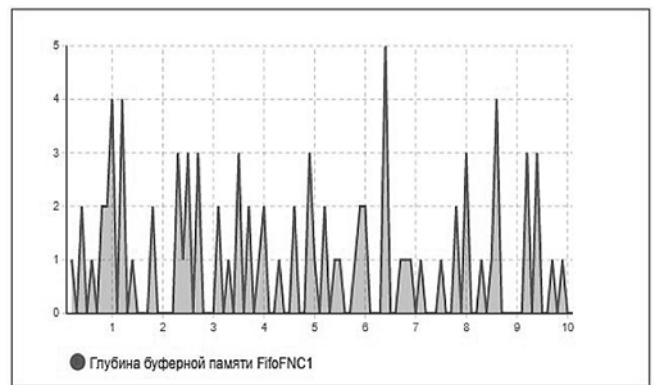
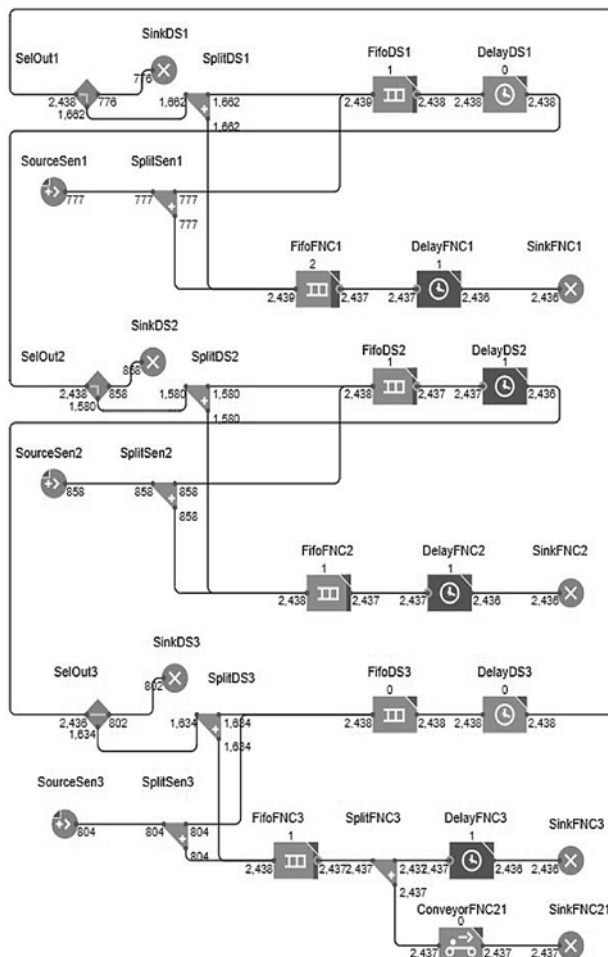


Рис. 4. Структурная схема модели кластера с круговой организацией сети (слева), график $K(t)$ (в блоках $FifoDS1$ и $FifoFNC1$) при предельной пропускной способности

ция копирования выполняется в течение нулевого времени);

- $SplitSenN$ — создает копии сообщений, поступающие в ФСК от датчиков;
- $SplitFNC3$ — создает копии сообщений, поступающие в ФСК4.2, для передачи в ФСК2.1;
- $SelOutDSN$ — принимает сообщения из DS-сети, а затем сортирует в зависимости от заданного условия ("свои" — уничтожить, "чужие" — пропустить);
- $FifoDSN$ — моделирует накопитель типа $FIFO$ для сообщений, поступающих из DS-сети и от датчиков данного ФСК (с приоритетом от DS-сети);
- $FifoFNCN$ — моделирует накопитель типа $FIFO$ для сообщений, поступающих от датчиков данного ФСК и из DS-сети (с приоритетом от датчиков);
- $DelayDSN$ — транспортная задержка (фиксированное время) для передачи по DS-сети, одинаковая для всех ФСК кластера;

- $DelayFNCN$ — задержка для обработки сообщений данного ФСК (из DS-сети и от датчиков ФСК);
 - $ConveyorFNC21$ — блок, передающий сообщения в ФСК2.1 с определенным тактом.
- Параметры моделирования при выполнении экспериментов:
- случайное начальное число сообщений на информационных входах каждого ФСК при каждом прогоне (уникальные эксперименты);
 - время моделирования в каждом эксперименте 10 модельных единиц;
 - интенсивность поступления (по умолчанию) сообщений от $SourceSenN$ при экспоненциальном распределении интервалов соответствует интенсивности сообщений ОТС и равна 1,6 С/е.м.в. (сообщений за единицу модельного времени). Суммарное число сообщений — 48 (за 10 е.м.в.);
 - число прогонов в каждом эксперименте — 10;
 - число ФСК по умолчанию — 3 (правый кластер ФСК на рис. 3).

Время доставки любого сообщения можно выразить как

$$T_{NET} = (P-1) \cdot DelayDSN + DelayFNCN, \quad (1)$$

где P — число ФСК, а $DelayFNCN = KT_{PROC}$ (T_{PROC} — процессное, фиксированное время обработки сообщений, определяемое только аппаратными возможностями вычислителя ФСК). Исходя из требований к ОТС со стороны технологического оборудования ускорительного комплекса и физической пропускной способности магистралей общее время доставки T_{NET} не должно превышать времени, соответствующего обработке 8 сообщений, а $T_{PROC}/DelayDSN = 3$. Последнее соотношение определяется скоростью коммуникационной сети и производительностью обработки сообщений. Это соотношение может меняться при изменении сетевых возможностей, разных аппаратных платформах вычислителей, моделях и алгоритмах обработки сообщений. С учетом перечисленных требований глубина емкости буферной памяти K (*FifoFNCN*) не должна превышать 8 ($K \leq 8$). При увеличении K уменьшается вероятность потерь, но увеличивается время от возникновения сообщения до обработки в N -м ФСК, что снижает достоверность сообщений.

Сделаем подстановку в соотношение (1) с учетом требований и соотношений, представленных выше:

$$T_{NET} = (P-1)(T_{PROC}/3) + KT_{PROC} \leq 8T_{PROC}. \quad (2)$$

Выполнение соотношения (2) при прогонах модели определяет эффективность работы коммуникационной сети с заданными критериями. Упростив выражение (2), получим

$$P + 3K \leq 25 \quad (P, K > 0). \quad (3)$$

Из выражения (3) видно, что при $P = 3$ глубина памяти *FifoFNCN* не должна превышать 7 ($K \leq 7$), а при $P = 10$, глубина памяти *FifoFNCN* не должна превышать 5 ($K \leq 5$)

Результаты экспериментов на модели представлены в таблице, где указаны максимальные значения глубины заполнения буферной памяти при экспериментах со следующими условиями ($a/b/c$): a — интенсивность по умолчанию, $P = 3$; b — суммарное число сообщений, обработанное в каждом ФСК за 10 е.м.в. ≤ 2400 , $P = 3$; c — интенсивность по умолчанию, $P = 10$.

На основании анализа результатов моделирования можно сделать следующие выводы: прогоны модели конфигурационной сети соответствуют показателям эффективной работы как при существенном превышении значений интенсивности сообщений, заданной по умолчанию, так и при изменении конфигурации сети, т. е. увеличении числа ФСК до 10 ($P = 10$). Преимущества предложенной конфигурации сети (например, по сравнению с ОТС [13]) заключаются еще и в том, что предложенные линейные связи между отдельными кластерами ФСК позволяют синхронизировать работу подсистем и с меньшими задержками доставлять сообщения. Разработанную модель можно использовать и при моделировании полнодуплексного линейного кластера, преобразовав его в кольцевой.

Заключение

Решение задачи взаимодействия ФСК в рамках распределенной системы управления требует анализа характера технологических процессов, анализа зависимости алгоритмов ФСК по данным или управлению и т. д. Этот анализ позволяет выявить особенности организации ФСК с точки зрения управляющих и коммуникационных требований для решения поставленной задачи. В статье проанализированы тенденции

Результаты прогона модели

Буферная память	K_m — глубина буферной памяти, m — номер прогона модели ($m = 1, 2, \dots, 10$)									
	K_1	K_2	K_3	K_4	K_5	K_6	K_7	K_8	K_9	K_{10}
<i>FifoDS1</i>	1/1/1	1/1/2	1/1/1	1/1/1	1/1/1	1/1/1	1/1/1	1/1/1	1/1/1	1/1/1
<i>FifoFNC1</i>	1/4/3	1/4/3	1/5/4	1/4/3	1/4/2	1/4/3	1/4/3	1/5/4	1/4/3	1/4/3
<i>FifoDS2</i>	1/1/1	1/1/1	1/1/1	1/2/1	1/1/1	1/1/1	1/1/1	1/1/1	1/1/1	1/1/1
<i>FifoFNC2</i>	1/4/4	1/5/4	1/4/3	1/5/4	1/5/4	1/4/3	1/5/4	1/4/4	1/4/3	1/5/3
<i>FifoDS3</i>	1/1/1	1/1/1	1/1/1	1/1/1	1/1/1	1/1/2	1/1/1	1/1/1	1/1/1	1/1/1
<i>FifoFNC3</i>	1/4/3	1/5/3	1/5/4	1/4/5	1/4/3	1/4/4	1/5/4	1/4/3	1/4/3	1/4/4

развития промышленных сетевых экосистем, программно-аппаратный инструментарий реализации участников сетевых объединений. В частности, рассмотренные возможности программной платформы на базе стандарта IEC 61499 для разработки управляющих программ ФСК отвечают требованиям, предъявляемым к сетевым компонентам. Возможную недетерминированность взаимодействия (как отмечено в работе [9]) предложено компенсировать путем использования ПЛИС в составе контроллеров ФСК при создании дополнительных быстрых коммуникационных линков для реализации детерминированных сетевых приложений. Предложенная дополнительная вычислительная платформа на базе ПЛИС вместе с ARM-контроллером образует гибридный вычислитель в составе ФСК, который позволяет управлять быстрыми технологическими процессами.

Предложена и обоснована многоуровневая конфигурация промышленной сети с объединением отдельных ФСК в кластеры. Данная конфигурация позволяет гибко совмещать загрузочные и коммуникационные задачи ФСК, при необходимости выполняя обмен информацией между кластерами. Разработанная модель в программном пакете *AnyLogic* 8.6 показала эффективность предложенной конфигурации. Результаты экспериментов позволили сделать вывод, что увеличение интенсивности сообщений в большей степени влияет на эффективность работы, чем увеличение числа ФСК.

Развитие данной тематики предполагается по следующим направлениям: улучшение характеристик ФСК по вычислительным и коммуникационным возможностям. Разработка алгоритма оптимизации пути передачи сообщений между кластерами, с анализом времени ожидания занятых DS-линков.

Список литературы

1. Дубинин В. Н., Вяткин В. В., Дроздов Д. Н. Опережающие вычисления в системах функциональных блоков IEC 61499 // Тр. Междунар. науч.-техн. конф. "Современные информационные технологии". Пенза, 2013. Вып. 17. С. 9—24.
2. Коквин В. А. Анализ зависимостей алгоритмов управления взаимосвязанными технологическими процессами // Системный анализ в науке и образовании. 2015. № 4. С. 23—27.
3. Воеводин В. П. Автоматизация У-70 и ОС Linux. Препринт ИФВЭ 2008-2. Протвино, 2008. 23 с.
4. Javed B., Iqbal M. W., Abbas H. Internet of Things (IoT) Design Considerations for Developers and Manufacturers // ICC Workshops. 2017. P. 834—842.
5. IIoT Companies [Электронный ресурс]. URL: <https://www.postscapes.com/iiot-companies/> (дата обращения: 16.03.2020).
6. Ray P. P. Internet of robotic things: Concept technologies and challenges // IEEE Access. Jan. 2017. Vol. 4. P. 9489—9500.
7. Kokovin V. A., Evsikov A. A., Uvaysov S. U., Uvaysova S. S. Event-based Cooperation of Functional Networking Components in Distributed Technological Systems // Proceedings of MWENT-2020. Moscow, Russia. March 11—13, 2020, in print.
8. Panjaitan S., Frey G. Functional control objects in distributed automation systems // IFAC Proceedings Volumes. 2007. Vol. 40, Iss. 3. P. 259—264.
9. Pang C., Vyatkin V. IEC 61499 Function Block Implementation of Intelligent Mechatronic Component // 8th IEEE Conference on Industrial Informatics (INDIN 2010). 2010. P. 1124—1129.
10. International Standard IEC 61499. Function blocks for industrial-process measurement and control systems (edition 2.0). Part 2: Software tool requirements: ed2.0. International Electrotechnical Commission. Geneva, 2012. 97 p.
11. Batth R. S., Nayyar A., Nagpal A. Internet of Robotic Things: Driving Intelligent Robotics of Future — Concept, Architecture, Applications and Technologies // 2018 4th International Conference on Computing Sciences (ICCS). Jalandhar, 2018. P. 151—160. doi: 10.1109/ICCS.2018.00033.
12. Каляев И. А., Капустян С. Г., Усачев Л. Ж. Основы построения распределенных систем управления коллективами роботов // Информационные технологии. 1998. № 5. С. 13—18.
13. Комаров В. В. Базовые принципы построения общей таймерной системы ускорительного комплекса У-70 // Приборы и системы. Управление, контроль, диагностика. 2005. № 5. С. 23—26.
14. Arefin S. E., Ashrafi Heya T., Uddin J. Real-life Implementation of Internet of Robotic Things Using 5 DoF Heterogeneous Robotic Arm // 2018 Joint 7th International Conference on Informatics, Electronics & Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision & Pattern Recognition (icIVPR). Kitakyushu, Japan. 2018. P. 486—491. doi: 10.1109/ICIEV.2018.8641019.
15. International Standard IEC 61131-3 (edition 2.0): Programmable Controllers / International Electrotechnical Commission. Geneva, 2003. 230 p.
16. Pearce H. A., Roop P. S. Synthesizing IEC 61499 Function Blocks to hardware // Proceedings of International Conference on Electronics, Information, and Communication (ICEIC-2019). Auckland, New Zealand, 22—25 Jan. 2019. P. 1—6. doi: 10.23919/ELINFocom.2019.8706345
17. ISaGRAF Technology. [Электронный ресурс]. URL: <https://isagraf.ru/> (дата обращения: 16.03.2020).
18. nxtSTUDIO. [Электронный ресурс]. URL: <https://www.nxtcontrol.com/en/engineering/> (дата обращения: 16.03.2020).
19. ARM cores. [Электронный ресурс]. URL: <https://www.arm.com/> (дата обращения: 16.03.2020).
20. Отечественные ISaGRAF-контроллеры (2020) [Электронный ресурс]. URL: https://isagraf.ru/images/industry_avt/soft/isagraf/PLC%20-%20Russian.pdf/ (дата обращения: 16.03.2020).
21. Kokovin V., Diagilev V., Uvaysov S., Uvaysova S. Intelligent power electronic converter for wired and wireless distributed applications // Proceedings of the IEEE International Conference SED-2019. 23—24 April 2019. Prague: IEEE, 2019. P. 1—5.
22. Strzelecki R., Zinoviev Z., Strzelecki R., Benysek G. Overview of power electronic converters and controls // Power Electronics in Smart Electric Energy Networks. Power Systems. London: Springer, 2008. P. 55—105.
23. Достанко А. П., Аваков С. М., Агеев О. А. и др. Технологические комплексы интегрированных процессов производства изделий электроники. Минск: Белорусская наука. 2016. 251 с.
24. Дягилев В. И., Евсиков А. А., Коквин В. А. Силовой преобразователь. Патент на изобретение RU 2692687, 26.06.2019.
25. Евсиков А. А., Леонов А. П. Об использовании системы моделирования VisSim для исследования электропривода, работающего при скалярном управлении в повторно-кратковременном номинальном режиме с частыми пусками и электрическим торможением // Известия Института инженерной физики. 2019. № 1 (51). С. 42—48.

26. Дягилев В. И., Коковин В. А., Увайсов С. У., Увайсова С. С. Компьютерное моделирование работы силового преобразователя с выходным синусоидальным напряжением // Информационные технологии. 2016. Т.22, № 4. С. 261–266.

27. Система динамического моделирования *VisSim* [Электронный ресурс]. URL: <http://pro-spo.ru/winmat/480-vissim/> (дата обращения: 16.03.2020).

28. IoT Hardware Guide [Электронный ресурс]. URL: <https://www.postscapes.com/internet-of-things-hardware/> (дата обращения: 16.03.2020).

29. Raspberry PI 4 [Электронный ресурс]. URL: <https://www.raspberrypi.org/products/raspberry-pi-4-model-b/> (дата обращения: 16.03.2020).

30. Windows IoT platform [Электронный ресурс]. URL: <https://www.microsoft.com/en-us/software-download/win-dows10IoTCore/> (дата обращения: 16.03.2020).

31. International Standard IEC 61131-9:2013: Programmable controllers — Part 9: Single-drop digital communication interface for small sensors and actuators (SDCI) [Электронный ресурс]. URL: <https://webstore.iec.ch/publication/4558/> (дата обращения: 16.03.2020).

32. IEEE Standard for Heterogeneous Interconnect (HIC) (Low-Cost, Low-Latency Scalable Serial Interconnect for Parallel System Construction)", IEEE Standard 1355 — 1995, IEEE, June 1996 SpaceWire Standard. ECSS — Space Engineering. "SpaceWire — Links, Nodes, Routers and Networks". ECSS-E -ST-50-12C, Rev1 Draft D: November 2014.

33. Quartus Prime Lite Edition [Электронный ресурс]. URL: <https://fpgasoftware.intel.com/?edition=lite> (дата обращения: 16.03.2020).

34. Kokovin V. A., Evsikov A. A. Event-Related Interaction of Mechatronic Modules in Distributed Technological Installations // Special Issue on "Mechanical Engineering Research and Education. 2018. Vol. 12, N. 5.

35. Хинчин А. Я. Работы по математической теории массового обслуживания. М.: "Едиториал УРСС", 2019. 240 с.

36. Anylogic 8 University Researcher 8.6 [Электронный ресурс]. URL: <https://www.anylogic.ru/>

37. Коковин В. А., Комаров В. В. Контроллер таймерной сети общей таймерной системы ускорительного комплекса ИФВЭ // Приборы и системы. Управление, контроль, диагностика. НАУЧТЕХЛИТИЗДАТ. 2005. № 6. С. 15–20.

V. A. Kokovin, Assistant Professor, e-mail: kokovin@uni-protvino.ru,

A. A. Evsikov, Assistant Professor, e-mail: eaa@uni-protvino.ru,

A. P. Leonov, Assistant Professor, e-mail: anatoli.leonov@googlemail.com,
State University "Dubna", Branch "Protvino", Moscow Region, Protvino

Features of the Organization and Interaction of Functional Network Components in Distributed Control Systems

Solving the problem of FNC interaction within a distributed control system requires an analysis of the nature of technological processes, analysis of the dependence of FNC algorithms on data or control, etc. This analysis makes it possible to identify the features of the FNC organization from the point of view of control and communication requirements for solving the task. The article analyzes the development trends of industrial network ecosystems, software and hardware tools for the implementation of network associations participants. In particular, the considered capabilities of the software platform based on the IEC 61499 standard for the development of FNC control programs meet the requirements for network components. Possible non-determinism of interaction (as noted in [9]) is proposed to be compensated by using FPGAs as part of FNC controllers when creating additional fast communication links for the implementation of deterministic network applications. The proposed additional FPGA-based computing platform together with the ARM controller forms a hybrid computer as part of the FNC, which allows you to control fast technological processes. A multi-level configuration of an industrial network with the unification of individual FNCs into clusters is proposed and substantiated. This configuration allows you to flexibly combine the boot and communication tasks of FNC, if necessary, performing the exchange of information between clusters. The developed model in the AnyLogic 8.6 software package showed the effectiveness of the proposed configuration. The results of the experiments led to the conclusion that an increase in the intensity of messages has a greater effect on work efficiency than an increase in the number of FNCs. The development of this topic is expected in the following areas: improving the characteristics of FNC in terms of computing and communication capabilities. Development of an algorithm for optimizing the message transmission path between clusters, with an analysis of the waiting time of busy DS-links.

Keywords: distributed control systems, Functional Networking Components, mechatronic modules, IEC-61499, Field-Programmable Gate Array, AnyLogic 8.6

DOI: 10.17587/it.27.212-224

References

1. Dubinin V. N., Vyatkin V. V., Drozdov D. N. Advance Computing in IEC 61499 Function Block Systems, *Trudy Mezhdunarodnoj nauchno-tehnicheskoy konferencii "Sovremennye informacionnye tekhnologii"*, Penza, 2013, no. 17, pp. 9–24 (in Russian).

2. Kokovin V. A. Analysis of the dependence of algorithms interrelated management processes, *Sistemnyj analiz v nauke i obrazovanii*, 2015, no. 4, pp. 23–27 (in Russian).

3. Voevodin V. P. The Automation of U-70 and OS Linux, IHEP Preprint 2008-2, Protvino, 2008, p. 23 (in Russian).

4. Javed B., Iqbal M. W., Abbas H. Internet of Things (IoT) Design Considerations for Developers and Manufacturers, *ICC Workshops*, 2017, pp. 834–842.

5. IIoT Companies, available at: <https://www.postscapes.com/iiot-companies/> (accessed March 16, 2020).

6. Ray P. P. Internet of robotic things: Concept technologies and challenges, *IEEE Access*, Jan. 2017, vol. 4, pp. 9489–9500.

7. Kokovin V. A., Evsikov A. A., Uvaysov S. U., Uvaysova S. S. Event-based Cooperation of Functional Networking Components in Distributed Technological Systems, *Proceedings of MWENT-2020*, Moscow, Russia, March 11–13, 2020 (in print).
8. Panjaitan S., Frey G. Functional control objects in distributed automation systems, *IFAC Proceedings Volumes*, 2007, vol. 40, iss. 3, pp. 259–264.
9. Pang C., Vyatkin V. IEC 61499 Function Block Implementation of Intelligent Mechatronic Component, *8th IEEE Conference on Industrial Informatics (INDIN 2010)*, 2010, pp. 1124–1129.
10. International Standard IEC 61499. Function blocks for industrial-process measurement and control systems (edition 2.0). Part 2: Software tool requirements: ed2.0 / International Electrotechnical Commission, Geneva, 2012, 97 p.
11. Batth R. S., Nayyar A., Nagpal A. Internet of Robotic Things: Driving Intelligent Robotics of Future — Concept, Architecture, Applications and Technologies, *2018. 4th International Conference on Computing Sciences (ICCS)*, Jalandhar, 2018, pp. 151–160, doi: 10.1109/ICCS.2018.00033.
12. Kaljaev I. A., Kapustjan S. G., Usachev L. Zh. Creation bases of distributed control systems of robot collective, *Journal Information Technologies*, 1998, no. 5, pp. 13–18 (in Russian).
13. Komarov V. V. Basic Principle Timing System Creation for U-70 Accelerators Complex, *Instruments and Systems: Monitoring, Control, and Diagnostics*, 2005, no. 5, pp. 23–26 (in Russian).
14. Arefin S. E., Ashrafi Heya T., Uddin J. Real-life Implementation of Internet of Robotic Things Using 5 DoF Heterogeneous Robotic Arm, *2018 Joint 7th International Conference on Informatics, Electronics & Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision & Pattern Recognition (icIVPR)*, Kitakyushu, Japan, 2018, pp. 486–491, doi: 10.1109/ICIEV.2018.8641019.
15. International Standard IEC 61131-3 (edition 2.0): Programmable Controllers, International Electrotechnical Commission, Geneva, 2003, 230 p.
16. Pearce H. A., Roop P. S. Synthesizing IEC 61499 Function Blocks to hardware, *Proceedings of International Conference on Electronics, Information, and Communication (ICEIC-2019)*, Auckland, New Zealand, 22–25 Jan. 2019, pp. 1–6. doi: 10.23919/ELINFOCOM.2019.8706345
17. ISaGRAF Technology, available at: <https://isagraf.ru/> (accessed March 16, 2020).
18. nxtSTUDIO, available at: <https://www.nxtcontrol.com/en/engineering/> (accessed March 16, 2020).
19. ARM cores, available at: <https://www.arm.com/> (accessed March 16, 2020).
20. Russian ISaGRAF-PLC (2020), available at: https://isagraf.ru/images/industry_avt/soft/isagraf/PLC%20-%20Russian.pdf/ (accessed March 16, 2020).
21. Kokovin V., Diagilev V., Uvaysov S., Uvaysova S. Intelligent power electronic converter for wired and wireless distributed applications. 2019, In: *Proceedings of the IEEE International Conference SED-2019*, 23–24 April 2019; Prague: IEEE; 2019. pp. 1–5.
22. Strzelecki R., Zinoviev Z., Strzelecki R., Benysek G. Overview of power electronic converters and controls, *Power Electronics in Smart Electric Energy Networks. Power Systems*, London, Springer, 2008, pp. 55–105.
23. Dostanko A. P., Avakov S. M., Ageev O. A. et al. Technological systems integrated production processes of electronics products, Minsk, Belorusskaya nauka, 2016, 251 p. (in Russian).
24. Diagilev V. I., Evsikov A. A., Kokovin V. A. Power converter. Russian Federation Patent RU 2692687, 26.06.2019 (in Russian).
25. Evsikov A. A., Leonov A. P. About using the simulation system VisSim for the investigation of the electric drive, working under scalar control in repeatedly-short-term nominal mode with frequent starts and electric braking, *Izvestiya Instituta inzhenernoj fiziki*, 2019, no. 1 (51), pp. 42–48 (in Russian).
26. Diagilev V. I., Kokovin V. A., Uvaysov S. U., Uvaysova S. S. Computer simulation of the power converter with harmonic wave output, *Information Technologies*, 2016, no. 4, vol. 22, pp. 261–266 (in Russian).
27. The system of dynamic modeling VisSim, available at: <http://pro-spo.ru/winmat/480-vissim/> (accessed March 16, 2020).
28. IoT Hardware Guide, available at: <https://www.postscapes.com/internet-of-things-hardware/> (accessed March 16, 2020).
29. Raspberry PI 4, available at: <https://www.raspberrypi.org/products/raspberry-pi-4-model-b/> (accessed March 16, 2020).
30. Windows IoT platform, available at: <https://www.microsoft.com/en-us/software-download/windows10IoTCore/> (accessed March 16, 2020).
31. International Standard IEC 61131-9:2013: Programmable controllers — Part 9: Single-drop digital communication interface for small sensors and actuators (SDCI), available at: <https://webstore.iec.ch/publication/4558/> (accessed March 16, 2020).
32. IEEE Standard for Heterogeneous Interconnect (HIC) (Low-Cost, Low-Latency Scalable Serial Interconnect for Parallel System Construction), IEEE Standard 1355 — 1995, IEEE, June 1996.
33. Quartus Prime Lite Edition, available at: <https://fpgasoftware.intel.com/?edition=+lite> (accessed March 16, 2020).
34. Kokovin V. A., Evsikov A. A. Event-Related Interaction of Mechatronic Modules in Distributed Technological Installations, *Special Issue on "Mechanical Engineering Research and Education"*, 2018, vol. 12, no. 5.
35. Khinchin A. Ya. Work on the mathematical theory of queuing, Moscow, "Editorial URSS", 2019, 240 p. (in Russian).
36. Anylogic 8 University Researcher 8.6, available at: <https://www.anylogic.ru/> (accessed March 16, 2020).
37. Kokovin V. A., Komarov V. V. Timing Network Controller of IHEP Accelerators Complex General Timing System, *Instruments and Systems: Monitoring, Control and Diagnostics*, 2005, no. 6, pp. 15–20 (in Russian).

Адрес редакции:

107076, Москва, Стромьинский пер., 4

Телефон редакции журнала (499) 269-5510

E-mail: it@novtex.ru

Технический редактор Е. В. Конова.

Корректор М. Ю. Безменова.

Сдано в набор 04.02.2021. Подписано в печать 23.03.2021. Формат 60×88 1/8. Бумага офсетная.

Усл. печ. л. 8,86. Заказ ИТ421. Цена договорная.

Журнал зарегистрирован в Министерстве Российской Федерации по делам печати, телерадиовещания и средств массовых коммуникаций.

Свидетельство о регистрации ПИ № 77-15565 от 02 июня 2003 г.

Оригинал-макет ООО "Авансд солишнз". Отпечатано в ООО "Авансд солишнз".

119071, г. Москва, Ленинский пр-т, д. 19, стр. 1. Сайт: www.aov.ru

Рисунок к статье Д. В. Малаховецкого, А. И. Разумовского

«СИНТАКСИЧЕСКИЙ АНАЛИЗ СИМВОЛЬНЫХ МАССИВОВ МЕТОДОМ РЕКУРСИВНОГО ОХВАТА И СТРУКТУРИРОВАНИЯ НА ПРИМЕРЕ ДАННЫХ VRML-ФОРМАТА»

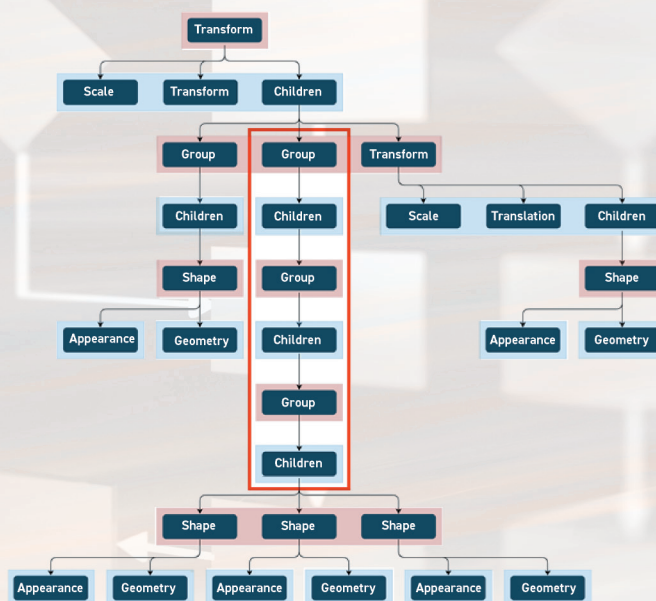


Рис.1. Пример структуры VRML

Рисунок к статье Т. В. Нгуен, А. Г. Кравец

«ОЦЕНКА И ПРОГНОЗИРОВАНИЕ ТЕНДЕНЦИЙ РАЗВИТИЯ НАУЧНЫХ ИССЛЕДОВАНИЙ НА ОСНОВЕ БИБЛИОМЕТРИЧЕСКОГО АНАЛИЗА ПУБЛИКАЦИЙ»

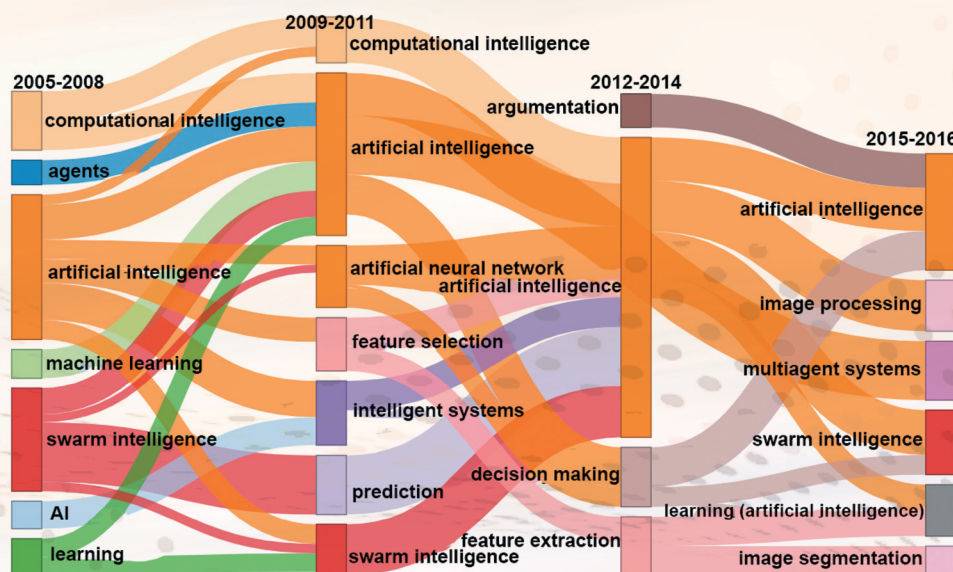


Рис. 1. Карта тематической эволюции авторских ключевых слов:

- computational intelligence (2005 – 2008, 2009 – 2011); — agents;
- artificial intelligence (2005 – 2008, 2009 – 2011, 2012 – 2014, 2015 – 2016), artificial neural network (2009 – 2011);
- machine learning; — swarm intelligence (2005 – 2008, 2009 – 2011, 2015 – 2016); — AI; — learning;
- feature selection (2009-2011), feature extraction (2012-2014), image segmentation (2015-2016);
- intelligent systems; — prediction; — decision making; — argumentation;
- image processing; — multi-agent systems; — learning (artificial intelligence)

«ОЦЕНКА И ПРОГНОЗИРОВАНИЕ ТЕНДЕНЦИЙ РАЗВИТИЯ НАУЧНЫХ ИССЛЕДОВАНИЙ НА ОСНОВЕ БИБЛИОМЕТРИЧЕСКОГО АНАЛИЗА ПУБЛИКАЦИЙ»

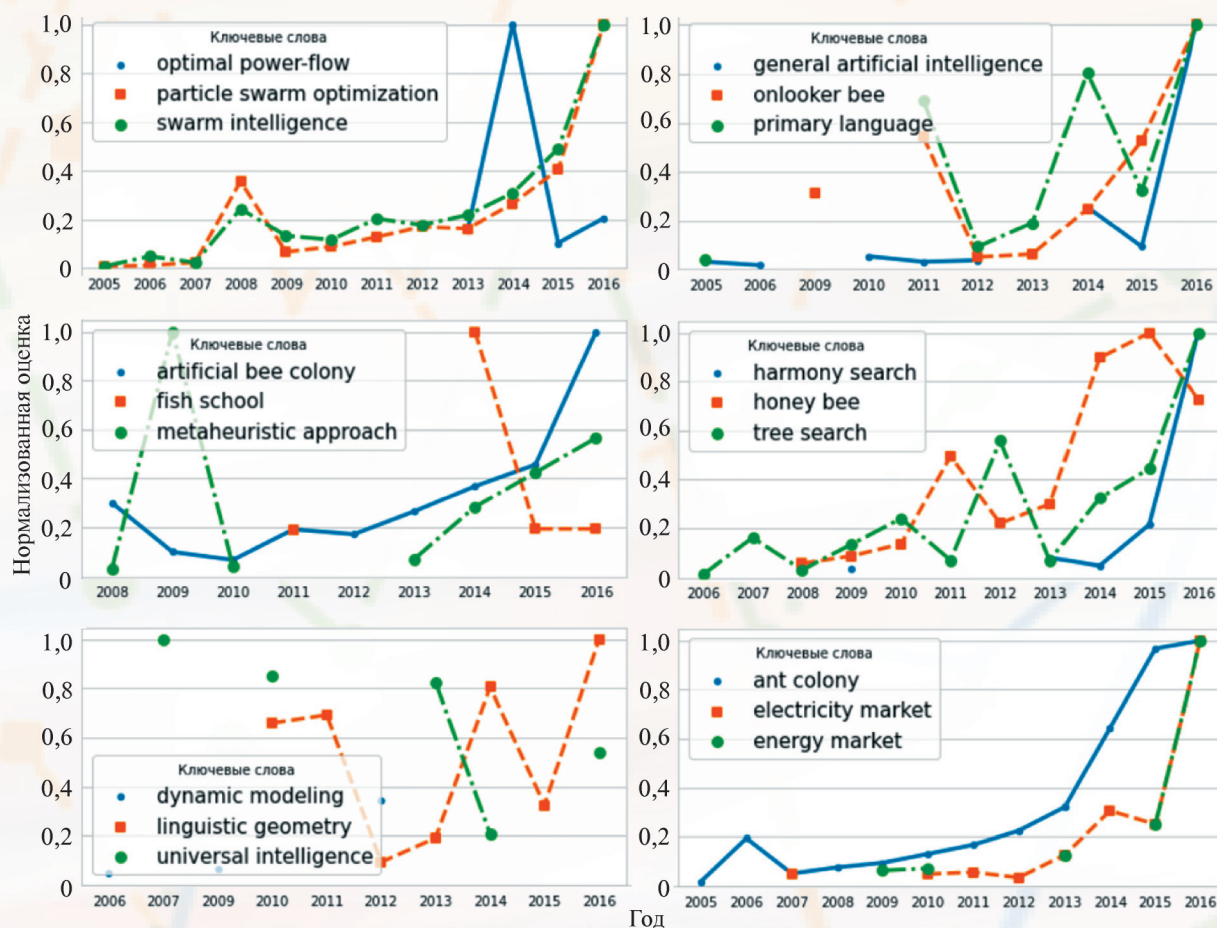


Рис. 4. Динамика оценок ключевых слов во временных рядах

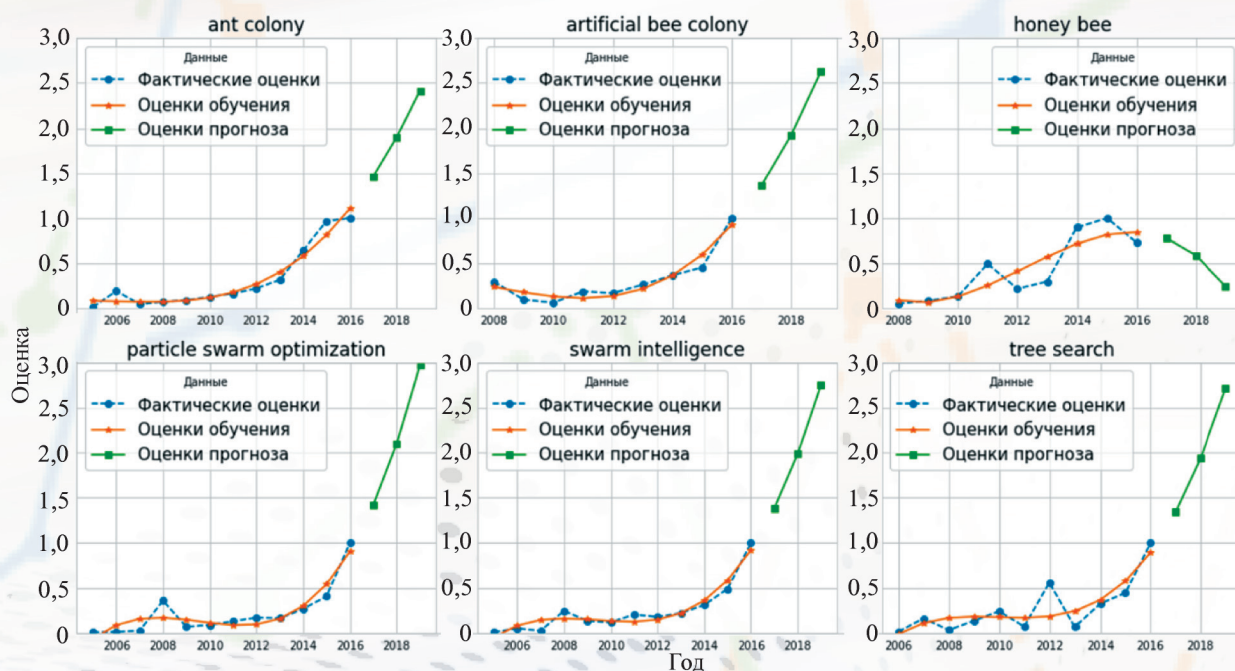


Рис. 5. Результаты приспособления оценок с помощью модели полиномиальной регрессии