

ВЕСТНИК ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА

УПРАВЛЕНИЕ,
ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА
И ИНФОРМАТИКА

TOMSK STATE UNIVERSITY
JOURNAL OF CONTROL AND COMPUTER SCIENCE

Научный журнал

2021

№ 55

Зарегистрирован в Федеральной службе по надзору в сфере
массовых коммуникаций, связи и охраны культурного наследия
(свидетельство о регистрации ПИ № ФС 77-29497 от 27 сентября 2007 г.)

Подписной индекс в объединённом каталоге «Пресса России» 44031

Журнал включен в «Перечень рецензируемых научных изданий,
в которых должны быть опубликованы основные научные результаты диссертаций
на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук»,
Высшей аттестационной комиссии

Founder – Tomsk State University

EDITORIAL BOARD

Alexander Gortsev – Editor-in-Chief, Doctor of Sc., Prof., Head of the Applied Mathematics Department Tomsk State University. Tel: +73822529599
Valery Smagin – Deputy Editor-in-Chief, Doctor of Sc., Prof. of the Applied Mathematics Department Tomsk State University. Tel: +73822529599
Lyudmila Nezhelskaya – Executive Editor, Doctor. of Sc., Prof. of the Applied Mathematics Department Tomsk State University. E-mail: vestnik_uvti@mail.tsu.ru

Sergey Vorobeychikov – Doctor of Sc., Prof. of the System Analysis and Mathematical Modeling Department Tomsk State University
Vladimir Vishnevsky – Doctor of Sc., Prof. Head of the laboratory Institute of Control Sciences of Russian Academy of Sciences (Moscow, Russia).
Gennady Koshkin – Doctor of Sc., Prof. of the System Analysis and Mathematical Modeling Department Tomsk State University
Yury Kostyuk – Doctor of Sc., Prof. of the Theoretical Informatics Department Tomsk State University
Anjela Matrosova – Doctor of Sc., Prof. of the Programming Department Tomsk State University
Anatoly Nazarov – Doctor of Sc., Prof., Head of the Probability Theory and Mathematical Statistics Department Tomsk State University
Konstantin Samouylov – Doctor of Sc., Prof., Head of the Applied Probability and Informatics Department RUDN University (Moscow, Russia)
Eugene Semenko – Doctor of Sc., Prof. System Analysis and Operations Research Department Reshetnev Siberian State University of Science and Technology (Krasnoyarsk, Russia)
Sergey Sushchenko – Doctor of Sc., Prof., Head of the Applied of Information Department Tomsk State University
Mais Farkhadov – Doctor of Sc., Head of the laboratory Institute of Control Sciences of Russian Academy of Sciences (Moscow, Russia).
Gurami Tsitsiashvili – Doctor of Sc., Prof., Chief researcher Institute for Applied Mathematics Far Eastern Branch of RAS, Prof. Far Eastern Federal University (Vladivostok, Russia)
Editorial address:
Institute of Applied Mathematics and Computer Science,
unit of Applied Mathematics
National Research Tomsk State University
36 Lenina Avenue, Tomsk, 634050
Telephone / fax: +73822529599
E-mail: vestnik_uvti@mail.tsu.ru

EDITORIAL COUNCIL

Ana Rosa Cavalli PhD, Prof. University VII Paris, France	Gilbert Saporta PhD, Prof. Pierre and Marie Curie University, Paris, France
Alexander Dudin Doctor of Sc., Prof. Belarusian State University Minsk, Republic Belorussia	Raimund Ubar Doctor of Sc., Prof. University of Technology Tallinn, Estonia
Reindert Nobel Doctor of Sc., Associate Prof. Vrije University, Amsterdam, Netherlands	Nina Yevtushenko Doctor of Sc., Prof. Ivannikov V.P. ISP RAS Moscow, Russia
Enco Orzinger PhD, Prof. University of Rome, Italy	Yervant Zorian PhD, Fellow & Chief Architect, Synopsys, Mountain View, CA, USA
Paolo Prinetto Prof. Politecnico Institute, Torino, Italy	

JOURNAL INFO

Tomsk State University Journal of Control and Computer Science is an independent peer-reviewed research journal that welcomes submissions from across the world.

Tomsk State University Journal of Control and Computer Science is issued four times per year, and can be subscribed to in the Russian Press Joint Catalogue (Subscription Index 44031)

The publication in the journal is free of charge and may be in Russian or in English.

The topics of the journal are the following:

- control of dynamical systems,
- mathematical modeling,
- data processing,
- informatics and programming,
- discrete function and automation,
- designing and diagnostics of computer systems.

Rules of registration articles are given in a site:

<http://journals.tsu.ru/informatics/>

ISSN 2311-2085 (Online), ISSN 1998-8605 (Print).

Учредитель – Томский государственный университет

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

Горцев Александр Михайлович – гл. редактор, проф., д-р техн. наук, зав. кафедрой прикладной математики ТГУ. Тел. +73822529599
Смагин Валерий Иванович – зам. гл. редактора, проф., д-р техн. наук, проф. кафедры прикладной математики ТГУ. Тел. +73822529599
Нежелская Людмила Алексеевна – ответственный секретарь, доц., д-р физ.-мат. наук, проф. кафедры прикладной математики ТГУ
E-mail: vestnik_uvti@mail.tsu.ru

Воробейчиков Сергей Эрикович – д-р физ.-мат. наук, проф. кафедры системного анализа и математического моделирования ТГУ
Вишневский Владимир Миронович – проф., д-р техн. наук, зав. лабораторией Института проблем управления РАН (г. Москва)
Кошкин Геннадий Михайлович – проф., д-р физ.-мат. наук, проф. кафедры системного анализа и математического моделирования ТГУ
Костюк Юрий Леонидович – проф., д-р техн. наук, проф. кафедры теоретической информатики ТГУ
Матросова Анжела Юрьевна – проф., д-р техн. наук, проф. кафедры программирования ТГУ
Назаров Анатолий Андреевич – проф., д-р техн. наук, зав. кафедрой теории вероятностей и математической статистики ТГУ
Самуйлов Константин Евгеньевич – проф., д-р техн. наук, зав. кафедрой прикладной информатики и теории вероятностей РУДН (г. Москва)
Семенкин Евгений Станиславович – проф., д-р техн. наук, проф. каф. системного анализа и исследования операций, СибГУ им. акад. М.Ф. Решетнева (г. Красноярск)
Сушенко Сергей Петрович – проф., д-р техн. наук, зав. кафедрой прикладной информатики ТГУ
Фархадов Маис Паша Оглы – д-р техн. наук, зав. лабораторией Института проблем управления РАН (г. Москва)
Цициашвили Гурами Шалвович – проф., д-р физ.-мат. наук, гл. науч. сотр. Института прикладной математики ДВО РАН, проф. ДВФУ (г. Владивосток)

Адрес редакции и издателя: 634050, Томск, пр. Ленина, 36
Национальный исследовательский Томский государственный университет,
Институт прикладной математики и компьютерных наук,
отделение прикладной математики
Телефон / факс: +73822529599
E-mail: vestnik_uvti@mail.tsu.ru

РЕДАКЦИОННЫЙ СОВЕТ

Ана Роза Ковали д-р. философии, проф. Университет VII Париж, Франция	Жильберт Сапорта д-р. философии, проф. Университет им. Пьера и Марии Кюри, Париж, Франция
Александр Дудин д.ф.-м.н., проф. БГУ, Минск, Республика Беларусь	Раймонд Убар д-р. проф. Технологический университет, Таллинн, Эстония
Рейндерт Нобель д-р., доцент Свободный университет, Амстердам, Нидерланды	Нина Евтушенко д-р. техн. наук, проф. ИСП РАН им. Иванникова В.П., Москва, Россия
Енцо Орзингер д-р. философии, проф. Римский университет, Италия	Ервант Зорян д-р. философии, гл. научный сотр. Фирмы «Синописис», США
Паоло Принетто проф. Политехнический институт Турин, Италия	

О ЖУРНАЛЕ

Журнал «Вестник Томского государственного университета. Управление, вычислительная техника и информатика» выходит ежеквартально и распространяется по подписке

Статьи публикуются на русском и английском языках.

Тематика публикаций журнала:

- управление динамическими системами,
- математическое моделирование,
- обработка информации,
- информатика и программирование,
- дискретные функции и автоматы,
- проектирование и диагностика вычислительных систем.

Журнал входит в систему Российского Индекса Научного Цитирования (РИНЦ).

Правила оформления статей приведены на сайте:

<http://journals.tsu.ru/informatics/>

ISSN 2311-2085 (Online), ISSN 1998-8605 (Print).

СОДЕРЖАНИЕ

УПРАВЛЕНИЕ ДИНАМИЧЕСКИМИ СИСТЕМАМИ

- Параев Ю.И., Колесникова С.И., Цветницкая С.А.**
Новый подход к решению задачи управления роботом-манипулятором 4

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

- Brokarev I.A., Farkhadov M.P., Vaskovskii S.V.**
Recurrent neural networks to analyze the quality of natural gas 11

ОБРАБОТКА ИНФОРМАЦИИ

- Вик К.В., Друки А.А., Григорьев Д.С., Спицын В.Г.**
Применение нейронных сетей глубокого обучения для решения задачи сегментации лесных пожаров на спутниковых снимках 18
- Воробьев А.В.**
Концепция многоуровневой системы цифровых двойников (на примере геомагнитных данных) 26
- Зенкова Ж.Н., Мусони У., Андриевская А.А.**
Статистическое оценивание показателей оборачиваемости оборотных средств с учетом дополнительной информации о квантиле 35
- Калашников С.Н., Мартусевич Е.А., Мартусевич Е.В., Рыбенко И.А.**
Кластеризация однородных производственных объектов на основе анализа технологических данных при формировании алюминиевых сплавов 45
- Нежелская Л.А., Першина А.А.**
Оценивание методом максимального правдоподобия параметра распределения случайного мертвого времени в рекуррентном обобщенном асинхронном потоке событий 53
- Rouban A.I.**
The sensitivity coefficients for multivariate dynamic systems described by ordinary difference equations with variable and constant parameters 65
- Рудько И.М.**
Оптимизация траектории движения объекта в конфликтной анизотропной среде по вероятностному критерию скрытности 71
- Фейзиев Ф.Г., Абаева Н.Б.**
Условия ортогональности входных последовательностей одного класса двоичных 4D-нелинейных модулярных динамических систем 80
- Шуленин В.П.**
Асимптотические свойства и робастность оценок урезанных вариантов стандартного отклонения и среднего абсолютных отклонений 91

ИНФОРМАТИКА И ПРОГРАММИРОВАНИЕ

- Шехалев Д.В.**
Структура матричного умножителя для декодеров QC-LDPC кода с минимальным ресурсом памяти ПЛИС 103

ПРОЕКТИРОВАНИЕ И ДИАГНОСТИКА ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

- Мухопад Ю.Ф., Мухопад А.Ю., Пунсык-Намжилов Д.Ц.**
Динамический контроль управляющих автоматов сложных систем дискретной автоматики 112
- Соловьев В.М., Сперанский Д.В.**
Алгоритм муравьиной колонии синтеза тестов для цифровых устройств 122
- Сведения об авторах** 127

CONTENTS

CONTROL OF DYNAMICAL SYSTEMS

- Paraev Ju.I., Kolesnikova S.I., Tsvetnitskaya S.A.**
A new approach to solving the problem of robot-manipulator control 4

MATHEMATICAL MODELING

- Brokarev I.A., Farkhadov M.P., Vaskovskii S.V.**
Recurrent neural networks to analyze the quality of natural gas 11

DATA PROCESSING

- Vik K.V., Druki A.A., Grigoriev D.S., Spitsyn V.G.**
Application of deep learning neural networks for solving the problem of forest fire segmentation on satellite images 18
- Vorobev A.V.**
The concept of a multi-level system of digital twins (on the example of geomagnetic data) 26
- Zenkova Z.N., Musoni W., Andrievskaya A.A.**
Analysis of assets turnover ratios using additional quantile information 35
- Kalashnikov S.N., Martusevich E.A., Martusevich E.V., Rybenko I.A.**
Clustering of homogeneous production facilities based on the analysis of technological data in the formation of aluminum alloys 45
- Nezhel'skaya L.A., Pershina A.A.**
Estimation of the uniform distribution parameter of unextendable dead time duration in a generalized recurrent asynchronous flow of events by the maximum likelihood method 53
- Rouban A.I.**
The sensitivity coefficients for multivariate dynamic systems described by ordinary difference equations with variable and constant parameters 65
- Rudko I.M.**
Optimization of the trajectory of the object in a conflict anisotropic medium according to the probabilistic criterion of stealth 71
- Feyziyev F.G., Abayeva N.B.**
The conditions of orthogonality of the input sequences of one class of binary 4D-nonlinear modular dynamic systems 80
- Shulenin V.P.**
Asymptotic properties and robustness of trimmed versions estimates of standard deviation and mean absolute deviations 91

INFORMATICS AND PROGRAMMING

- Shekhalev D.V.**
Matrix multiplier architecture for QC-LDPC code with minimum FPGA block RAM resource using 103

DESIGNING AND DIAGNOSTICS OF COMPUTER SYSTEMS

- Mukhopad Yu.F., Mukhopad Au.Y., Punsik-Namzhilov D.C.**
Dynamic control of control automata of complex discrete automation systems 112
- Solovye V.M., Speranskiy D.V.**
The ant colony algorithm for test syntesis for digital devices 122
- Brief information about the authors** 127

УПРАВЛЕНИЕ ДИНАМИЧЕСКИМИ СИСТЕМАМИ

УДК 62-50

DOI: 10.17223/19988605/55/1

Ю.И. Параев, С.И. Колесникова, С.А. Цветницкая

НОВЫЙ ПОДХОД К РЕШЕНИЮ ЗАДАЧИ УПРАВЛЕНИЯ РОБОТОМ-МАНИПУЛЯТОРОМ

Работа выполнена при финансовой поддержке РФФИ в рамках научного проекта № 20-08-00747.

Рассматривается задача управления роботом-манипулятором, состоящим из трех звеньев: стрела, рукоять и исполнительный механизм, соединенных между собой шарнирами. Конфигурация робота однозначно определяется углами между звеньями. Предполагается, что эти углы доступны измерению. Предполагается также, что на основании этих измерений можно вычислить моменты инерции и весовые моменты (произведения веса на плечо) звеньев. В результате получается простая математическая модель, описывающая движение робота, для которой получен простой алгоритм управления. Приведены результаты численного моделирования для задач перехода манипулятора из начального состояния в рабочее и движения манипулятора вдоль поверхности.

Ключевые слова: робот-манипулятор; управление движением центра масс; планировка.

В настоящее время существует большое количество различных типов подвижных роботов, или манипуляционных систем (экскаваторы, погрузчики, бурильные установки и пр.). Для каждого типа роботов имеют место проблемы с их управлением. В данной работе рассматриваются роботы-манипуляторы, состоящие из трех звеньев: стрела, рукоять и исполнительный механизм, соединенных между собой шарнирными механизмами. Теоретические вопросы управления подобными устройствами наиболее подробно изложены в [1–6]. В этих работах для управления роботами предложено использовать так называемые скользящие режимы [7. С. 12], что приводит к большим вычислительным трудностям. Следует также отметить работы [8–10] по синтезу управления на основе метода конструирования агрегированных регуляторов применительно к управлению движением центра масс подвижного объекта, описанным в виде системы дифференциальных и / или разностных уравнений. Этот метод привлекателен тем, что является аналитическим и приводит к созданию алгоритмов робастного нелинейного управления в условиях как систематических, так и случайных помех.

Несмотря на возрастающее число работ в направлении создания систем управления для существующих моделей манипуляторов, имеется ряд не вполне решенных вопросов: невысокая точность следящего управления; сложность технической реализации системы управления с учетом вопросов устойчивости управления.

В настоящей работе предложен новый и достаточно простой метод построения законов управления для объекта «трехзвенный робот-манипулятор» с целью его дальнейшего сравнения по эффективности с ранее полученными алгоритмами управления [1, 7, 9].

1. Схема робота-манипулятора

Схема робота-манипулятора приведена на рис. 1. Здесь OA – стрела, AB – рукоять, BC – исполнительный механизм. В точках O , A и B расположены шарнирные механизмы, которые выполняют

поворот звеньев. На рисунке l_1, l_2, l_3 – длины звеньев $0A, AB, BC$ соответственно. $G_1 = gm_1, G_2 = gm_2, G_3 = gm_3$ – вес этих звеньев, m_1, m_2, m_3 – массы звеньев, g – ускорение силы тяжести.

Конфигурация (состояние) робота однозначно определяется углами между звеньями α, β, γ или координатами точек $A(x_1, y_1), B(x_2, y_2), C(x_3, y_3)$. Между этими данными есть однозначное соответствие. Если заданы углы α, β, γ , то

$$\begin{aligned} x_1 &= l_1 \cos(\alpha), y_1 = l_1 \sin(\alpha), \\ x_2 &= x_1 + l_2 \sin(\rho), y_2 = y_1 - l_2 \cos(\rho), \\ x_3 &= x_2 + l_3 \cos(\tau), y_3 = y_2 + l_3 \sin(\tau), \end{aligned} \quad (1)$$

где

$$\rho = \alpha + \beta - \pi/2, \quad \tau = -\gamma + \rho + \pi/2 = \alpha + \beta - \gamma.$$

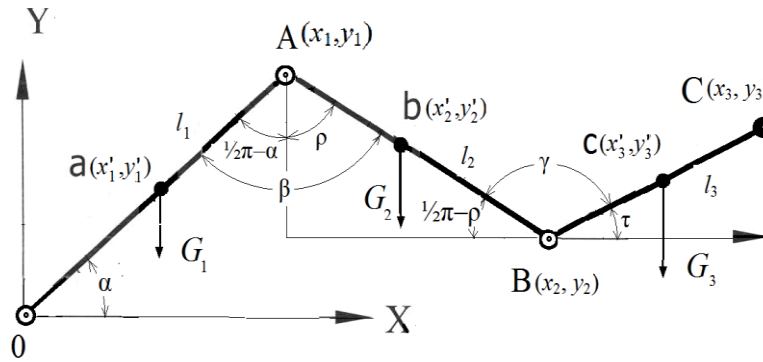


Рис. 1. Схема робота-манипулятора
Fig. 1. Robot manipulator diagram

Если заданы координаты точек $A(x_1, y_1), B(x_2, y_2), C(x_3, y_3)$, то

$$\begin{aligned} \alpha &= \arcsin(y_1 / l_1), \\ \beta &= \pi / 2 - \alpha + \arccos((y_1 - y_2) / l_2), \\ \gamma &= \alpha + \beta - \arcsin((y_3 - y_2) / l_3). \end{aligned} \quad (2)$$

Далее нам потребуется результат решения следующей задачи. Пусть заданы координаты точек $B(x_2, y_2)$ и $C(x_3, y_3)$. Эти точки должны удовлетворять условиям

$$x_2^2 + y_2^2 < (l_1 + l_2)^2, \quad (x_3 - x_2)^2 + (y_3 - y_2)^2 = l_3^2. \quad (3)$$

Для построения всей конфигурации робота нужно найти координаты точки $A(x_1, y_1)$. Она является точкой пересечения двух окружностей: одна с центром в точке O и радиусом l_1 , другая с центром в точке B и радиусом l_2 . Поэтому ее координаты находятся из условий

$$x_1^2 + y_1^2 = l_1^2, \quad (4)$$

$$(x_2 - x_1)^2 + (y_2 - y_1)^2 = l_2^2. \quad (5)$$

Решение этих уравнений сводится к следующему. Раскрывая скобки в (5) и вычитая (4), получаем

$$x_1 x_2 + y_1 y_2 = \frac{p}{2}, \quad (6)$$

где

$$p = l_1^2 - l_2^2 + l_4^2, \quad l_4^2 = x_2^2 + y_2^2.$$

Из (6) имеем

$$x_1 = \frac{1}{x_2} \left(\frac{p}{2} - y_1 y_2 \right). \quad (7)$$

Подставляя (7) в (4), получаем квадратное уравнение относительно y_1 :

$$y_1^2 - 2a y_1 + b = 0,$$

где

$$a = \frac{1}{l_4^2} \frac{p}{2} y_2, \quad b = \frac{1}{l_4^2} \left(\frac{p}{2}\right)^2 - \frac{1}{l_4^2} x_2^2 l_1^2.$$

Из двух решений этого уравнения выбираем большее:

$$y_1 = a + \sqrt{a^2 - b}. \quad (8)$$

Это следует из геометрических соображений. Подставляя (8) в (7), получаем

$$x_1 = \frac{1}{x_2} \left[\frac{p}{2} - y_2 (a + \sqrt{a^2 - b}) \right]. \quad (9)$$

Можно проверить, что корни уравнения вещественны, если выполняется (3)

2. Уравнения движения

Вращательное движение звеньев вокруг точек O , A и B описывается уравнениями

$$\begin{aligned} J_A \ddot{\alpha} &= M_A - R_A, & \alpha(0) &= \alpha_0, \dot{\alpha}(0) = 0, \\ J_B \ddot{\beta} &= M_B - R_B, & \beta(0) &= \beta_0, \dot{\beta}(0) = 0, \\ J_C \ddot{\gamma} &= M_C - R_C, & \gamma(0) &= \gamma_0, \dot{\gamma}(0) = 0, \end{aligned} \quad (10)$$

где M_A , M_B и M_C – управляющие моменты, J_A , J_B и J_C – моменты инерции, R_A , R_B и R_C – весовые моменты (произведение веса тела на плечо – проекцию центра тяжести звена на ось OX).

Постановка основной задачи. На интервале времени $(0, t_k)$ найти такие управляющие моменты M_A , M_B и M_C , при которых система из начального неподвижного состояния, определяемого углами

$$\alpha_0, \beta_0, \gamma_0 \quad (\dot{\alpha}_0 = \dot{\beta}_0 = \dot{\gamma}_0 = 0) \quad (11)$$

переходит в другое неподвижное состояние, определяемое углами

$$\alpha_1, \beta_1, \gamma_1 \quad (\dot{\alpha}_1 = \dot{\beta}_1 = \dot{\gamma}_1 = 0). \quad (12)$$

Заметим, что в неподвижном состоянии должно выполняться: $M_A = R_A$, $M_B = R_B$, $M_C = R_C$.

При решении задач управления нужно прежде всего определиться, какие параметры системы управления доступны измерению и могут быть использованы при формировании управляющих воздействий. В данной задаче естественно предположить, что измеряются углы между звеньями. Это, в свою очередь, позволяет вычислить координаты точек $A(x_1, y_1)$, $B(x_2, y_2)$, $C(x_3, y_3)$ согласно (1). Технически возможно создание датчиков, которые непосредственно измеряют эти координаты. Кроме того, можно предположить, что известны размеры и массы звеньев, а также расположение центров масс звеньев. Эти данные можно получить при проектировании робота и в результате натурных испытаний.

Пусть центры масс звеньев расположены в точках $a(x_1', y_1')$, $b(x_2', y_2')$, $c(x_3', y_3')$ (см. рис. 1), тогда

$$\begin{aligned} R_A &= G_1 x_1' + G_2 x_2' + G_3 x_3', \\ R_B &= G_2 (x_2' - x_1') + G_3 (x_3' - x_1'), \\ R_C &= G_3 (x_3' - x_2'). \end{aligned}$$

Если допустить, что плотность звеньев постоянна, то моменты инерции – произведение массы тела на квадрат расстояния между осью вращения и центром тяжести тела – согласно теореме о среднем равны

$$\begin{aligned} J_A &= m_1 (x_1'^2 + y_1'^2) + m_2 (x_2'^2 + y_2'^2) + m_3 (x_3'^2 + y_3'^2), \\ J_B &= m_2 ((x_2' - x_1')^2 + (y_2' - y_1')^2) + m_3 ((x_3' - x_1')^2 + (y_3' - y_1')^2), \\ J_C &= m_3 ((x_3' - x_2')^2 + (y_3' - y_2')^2). \end{aligned}$$

Таким образом, весовые моменты R_A , R_B и R_C и моменты инерции J_A , J_B и J_C могут быть вычислены. Вообще говоря, это утверждение может вызвать определенные сомнения. Поэтому авторы предполагают провести дальнейшие исследования для случая, когда последние величины известны с какими-то погрешностями.

3. Решение задачи

В (10) сделаем замену

$$M_A = J_A u_1 + R_A, \quad M_B = J_B u_2 + R_B, \quad M_C = J_C u_3 + R_C. \quad (13)$$

где u_1, u_2, u_3 – новые управления. Тогда из (10) получаем

$$\begin{aligned} \ddot{\alpha} &= u_1, & \alpha(0) &= \alpha_0, \dot{\alpha}(0) = 0, \\ \ddot{\beta} &= u_2, & \beta(0) &= \beta_0, \dot{\beta}(0) = 0, \\ \ddot{\gamma} &= u_3, & \gamma(0) &= \gamma_0, \dot{\gamma}(0) = 0. \end{aligned} \quad (14)$$

Теперь задача может быть сформулирована следующим образом. На интервале времени $(0, t_k)$ найти такие управления u_1, u_2, u_3 , при которых система (14) переходит из начального состояния (11) в конечное состояние (12).

Чтобы перевести систему из одного неподвижного состояния в другое неподвижное состояние, необходимо, чтобы скорость движения сначала увеличивалась, а затем убывала до нуля. Поэтому ускорение должно быть сначала положительным, а затем отрицательным. Или наоборот, в зависимости от того, в какую сторону происходит движение. Но в (14) ускорение равно управлению. Поэтому управление можно взять в виде

$$u(t) = \begin{cases} v & \text{для } 0 < t < t_1, \\ -v & \text{для } t_1 < t < t_k, \end{cases}$$

где v – какая-то искомая величина, $t_1 = t_k/2$.

Рассмотрим первое уравнение в (14). Положим

$$u_1(t) = \begin{cases} v_1 & \text{для } 0 < t < t_1, \\ -v_1 & \text{для } t_1 < t < t_k, \end{cases}$$

где v_1 – искомая величина. При таком управлении решение этого уравнения равно

$$\dot{\alpha}(t) = \begin{cases} v_1 t & \text{для } 0 < t < t_1, \\ v_1(2t_1 - t) & \text{для } t_1 < t < t_k, \end{cases} \quad (15)$$

и

$$\alpha(t) = \begin{cases} \alpha_0 + v_1 \frac{t^2}{2} & \text{для } 0 < t < t_1, \\ \alpha_0 + v_1 \left[2t_1 t - \frac{t^2}{2} - t_1^2 \right] & \text{для } t_1 < t < t_k. \end{cases}$$

Положим $\alpha(t_k) = \alpha_0 + v_1 t_1^2 = \alpha_1$. Тогда

$$v_1 = \frac{\alpha_1 - \alpha_0}{t_1^2}.$$

Видно, что если $\alpha_1 - \alpha_0 > 0$, то ускорение движения положительное и звено движется влево по часовой стрелке. Если $\alpha_1 - \alpha_0 < 0$, то ускорение движения отрицательное и звено движется вправо против часовой стрелки. Аналогичный результат можно получить и для других уравнений (14):

$$v_2 = \frac{\beta_1 - \beta_0}{t_1^2}, \quad v_3 = \frac{\gamma_1 - \gamma_0}{t_1^2}.$$

Остается определить момент времени $t_1 = t_k/2$. Естественнo потребовать, чтобы для всех звеньев этот момент был одинаковым (все звенья останавливаются одновременно) и каким-то минимальным. Такое значение для t_k можно найти из условия ограниченности угловой скорости движения. Как видно из (15), максимальная скорость достигается в точке $t = t_1$ и равна $v_1 t_1$.

В результате можем записать систему неравенств

$$\left| \frac{\alpha_1 - \alpha_0}{t_1} \right| \leq q_1, \quad \left| \frac{\beta_1 - \beta_0}{t_1} \right| \leq q_2, \quad \left| \frac{\gamma_1 - \gamma_0}{t_1} \right| \leq q_3,$$

где q_1, q_2, q_3 – какие-то ограничения на угловые скорости. Отсюда можно найти минимальное значение t_1 , при котором все эти неравенства выполняются.

4. Результаты численного моделирования

Приведем результаты моделирования сконструированного управления манипулятором со значениями длин звеньев: $l_1 = 7,5$ м, $l_2 = 5$ м, $l_3 = 2,5$ м.

4.1. Переход из начального состояния в рабочее

Конечное целевое состояние определяется углами: $\alpha_1 = \pi/3, \beta_1 = \pi/2, \gamma_1 = 2\pi/3$.

Рассмотрены два варианта начального состояния: а) в начале все звенья расположены на оси OY : $\alpha(0) = \pi/2, \beta(0) = 0, \gamma(0) = \pi$; б) в начале все звенья расположены на оси OX : $\alpha(0) = 0, \beta(0) = \pi, \gamma(0) = \pi$. Результаты моделирования приведены на рис. 2.

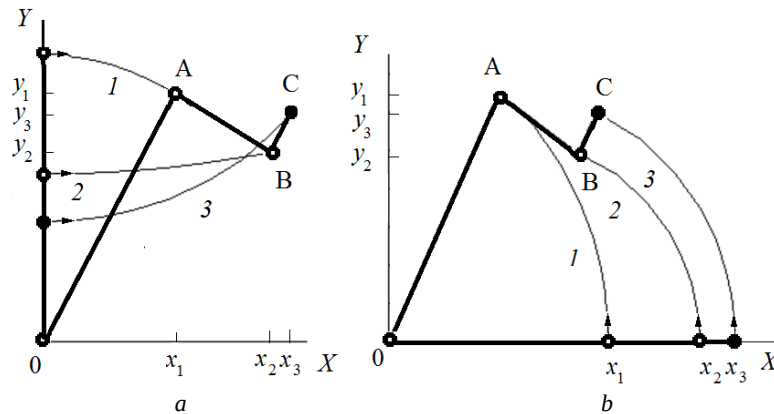


Рис. 2. Переход из начального состояния в рабочее:

a – начальное расположение – вертикальное, b – начальное расположение – горизонтальное

Fig. 2. Transient processes from the initial state to the active state for two cases:

a – initial position is vertical, b – initial position is horizontal

4.2. Планировка

Под планировкой понимается задача о движении исполнительного механизма манипулятора вдоль заданной горизонтальной поверхности. Это может быть движение электрода сварочного манипулятора вдоль шва, движение ковша экскаватора вдоль планируемой поверхности земли, движение режущего инструмента станка вдоль обрабатываемой поверхности и т.д. [1–6]. Необходимо, чтобы отклонения от поверхности лежали в заданных пределах. Наиболее подробно эта задача рассмотрена в [1, 3]. Здесь приводится новое решение.

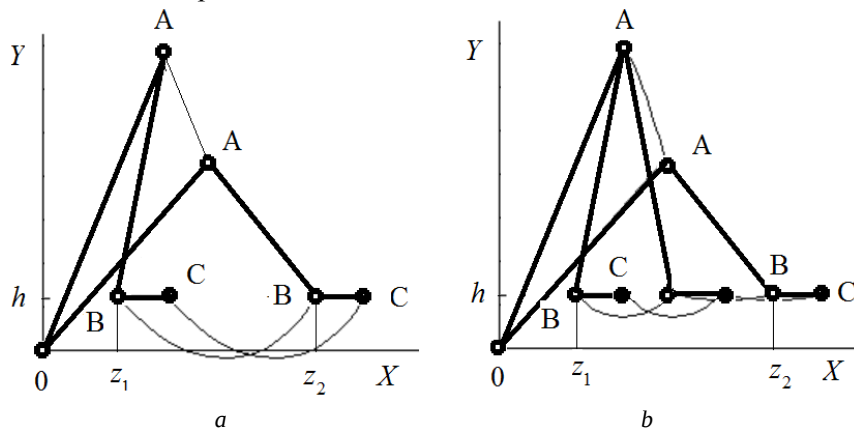


Рис. 3. Планировка: a – передвижение звена BC за один шаг, b – передвижение звена BC за два шага

Fig. 3. Movement of the manipulator actuator along a given horizontal surface:

a – movement of section BC in one step, b – movement of section BC in two step

В данном примере требуется перевести рабочий инструмент BC из положения $\{B(z_1, h), C(z_1+l_3, h)\}$ в конечное положение $\{B(z_2, h), C(z_2+l_3, h)\}$ и остановиться (рис. 3). На рис. 3, *a* передвижение звена BC осуществляется за один шаг. На рис. 3, *b* передвижение звена BC осуществляется за два шага. Видно, что при осуществлении одного шага звено BC движется значительно ниже уровня h . С другой стороны, при передвижении за два шага звено BC движется значительно выше. Очевидно, что при увеличении числа шагов можно добиться, чтобы звено BC двигалось практически на высоте h .

Заключение

В работе предложен новый и достаточно простой метод построения законов управления для робота-манипулятора. Этот метод связан с предположением, что при вычислении управляющих воздействий известны углы между звеньями, а также весовые моменты и моменты инерции звеньев, что значительно упрощает решение задачи.

Однако предположения, используемые для конструирования полученного управления, могут носить весьма ограничительный характер для определенного класса роботов-манипуляторов, функционирующих дистанционно и в условиях неизбежных в таких случаях измерительных помех и параметрических возмущений. В связи с этим авторами предполагаются дальнейшие исследования для случаев, когда в уравнения движения звеньев входят величины, измеряемые с какими-то погрешностями.

ЛИТЕРАТУРА

1. Матюхин В.И. Управление движением манипулятора. М. : Науч. изд. Ин-та проблем управления им. В.А. Трапезникова РАН, 2010. 95 с.
2. Матюхин В.И. Непрерывные универсальные законы управления манипуляционным роботом // Автоматика и телемеханика. 1997. № 4. С. 69–82.
3. Матюхин В.И. Стабилизация движений манипулятора вдоль заданной поверхности // Автоматика и телемеханика. 2011. № 4. С. 71–85.
4. Матюхин В.И., Пятницкий Е.С. Управление движением манипуляционных роботов на принципе декомпозиции при учете динамики приводов // Автоматика и телемеханика. 1989. № 9. С. 67–81.
5. Dwivedy S.K., Eberhard P. Dynamic analysis of flexible manipulators, a literature review // Mechanism and Machine Theory. 2006. № 41. P. 749–777.
6. Yu H., Liu Y., Hasan M.S. Review of modelling and remote control for excavators // International Journal of Advanced Mechatronic Systems. 2009. V. 2, № 1. P. 68–80.
7. Уткин В.И. Скользящие режимы в задачах оптимизации и управления. М. : Наука, 1981. 368 с.
8. Синергетика и проблемы теории управления : сб. науч. тр. / под ред. А.А. Колесникова. М. : Физматлит, 2004. 504 с.
9. Kolesnikov A.A., Kolesnikova S.I., Tsvetnitskaya S.A. Discrete Stochastic Regulator on a Manifold, Minimizing Dispersion of the Output Macrovariable // Mekhatronika, Avtomatizatsiya, Upravlenie. 2019. № 12 (20). P. 707–713. URL: <https://doi.org/10.17587/mau.20.707-713> (accessed: 17.12.2020).
10. Kolesnikova S., Tsvetnitskaya S., Pobegailo P. An Extension of the Method of Aggregated Regulators to a Discrete Stochastic Object // International Siberian Conference on Control and Communications, SIBCON. DOI: 10.1109/SIBCON.2019.8729595. URL: <https://ieeexplore.ieee.org/document/8729595>. (accessed: 17.12.2020).

Поступила в редакцию 20 декабря 2020 г.

Paraev Ju.I., Kolesnikova S.I., Tsvetnitskaya S.A. (2021) A NEW APPROACH TO SOLVING THE PROBLEM OF ROBOT-MANIPULATOR CONTROL. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika* [Tomsk State University Journal of Control and Computer Science]. 55. pp. 4–10

DOI: 10.17223/19988605/55/1

The problem of control of a robotic manipulator is considered, which consists of three links: an arrow, a handle and an actuator, connected by hinges. The configuration or state of the robot is uniquely determined by the angles α, β, γ between these links.

The rotational movement of the links around the centers of the hinges is described by the equations

$$\begin{aligned} J_A \ddot{\alpha} &= M_A - R_A, & \alpha(0) &= \alpha_0, \dot{\alpha}(0) = 0, \\ J_B \ddot{\beta} &= M_B - R_B, & \beta(0) &= \beta_0, \dot{\beta}(0) = 0, \\ J_C \ddot{\gamma} &= M_C - R_C, & \gamma(0) &= \gamma_0, \dot{\gamma}(0) = 0, \end{aligned}$$

where M_A , M_B and M_C are control moments, J_A , J_B and J_C are moments of inertia, R_A , R_B and R_C are weight moments (the projection of the center of gravity of the link onto the OX axis). The problem is to find, during a certain time interval $(0, t_k)$, such control moments M_A , M_B and M_C , at which the system passes from the initial stationary state to another stationary state.

It is assumed that the angles α , β , γ between the links are measured. It is also assumed that the moments of inertia and weight moments can be calculated from these measurements. So, we can do the following change of variables as a result

$$M_A = J_A u_1 + R_A, \quad M_B = J_B u_2 + R_B, \quad M_C = J_C u_3 + R_C.$$

where u_1 , u_2 and u_3 are new controls. Then the equations of rotational motion take the simple form

$$\begin{aligned} \ddot{\alpha} &= u_1, & \alpha(0) &= \alpha_0, & \dot{\alpha}(0) &= 0, \\ \ddot{\beta} &= u_2, & \beta(0) &= \beta_0, & \dot{\beta}(0) &= 0, \\ \ddot{\gamma} &= u_3, & \gamma(0) &= \gamma_0, & \dot{\gamma}(0) &= 0. \end{aligned}$$

To transfer the system from one stationary state to another stationary state, it is necessary that the speed of movement first increases and then decreases to zero. Therefore, the acceleration must first be positive and then negative, or vice versa, depending on which direction the movement is taking place. As can be seen from the last equations, acceleration equals control. Therefore, control can be represented in the form

$$u(t) = \begin{cases} v & \text{for } 0 < t < t_1, \\ -v & \text{for } t_1 < t < t_k, \end{cases}$$

where v is some unknown quantity, $t_1 = \frac{1}{2}t_k$. This value is chosen so that at the point t_k the speed of movement is equal to zero.

The results of numerical modeling for the problems of the transition of the robotic arm from the initial state to the working state and the movement of the manipulator along a given surface are presented.

Keywords: manipulation object; robotic arm; model of control over motion of center of mass of a moving object; layout.

PARAEV Jury Ivanovich (Doctor of Technical Sciences, Professor of the Applied Mathematics Department, Institute of Applied Mathematics and Computer Science, National Research Tomsk State University, Tomsk, Russian Federation).

E-mail: paraev@mail.ru

KOLESNIKOVA Svetlana Ivanovna (Doctor of Technical Sciences, Professor of the Department of Computer Technologies and Software Engineering of the Institute of Computing Systems and Programming, St. Petersburg State University of Aerospace Instrumentation, St. Petersburg, Russian Federation).

E-mail: skolesnikova@yandex.ru

TSVETNITSKAYA Svetlana Aleksandrovna (Candidate of Technical Sciences, Associate Professor of the Applied Mathematics Department, Institute of Applied Mathematics and Computer Science, National Research Tomsk State University, Tomsk, Russian Federation).

E – mail: svetasa@sibmail.com

REFERENCES

1. Matyukhin, V.I. (2010) *Upravlenie dvizheniem manipulyatora* [Manipulator motion control]. Moscow: RAS.
2. Matyukhin, V.I. (1997) Continuous universal laws of control of a manipulative robot. *Avtomatika i telemekhanika – Automation and Remote Control*. 4. pp. 69–82.
3. Matyukhin, V.I. (2011) Stabilization of manipulator movements along a given surface. *Avtomatika i telemekhanika – Automation and Remote Control*. 4. pp. 71–85.
4. Matyukhin, V.I. & Pyatnitskiy, E.S. (1989) Motion control of manipulation robots based on the principle of decomposition taking into account the dynamics of drives. *Avtomatika i telemekhanika – Automation and Remote Control*. 9. pp.67–81.
5. Dwivedy, S.K. & Eberhard, P. (2006) Dynamic analysis of flexible manipulators, a literature review. *Mechanism and Machine Theory*. 41. pp. 749–777. DOI: 10.1016/j.mechmachtheory.2006.01.014
6. Yu, H., Liu, Y. & Hasan, M.S. (2009) Review of modelling and remote control for excavators. *International Journal of Advanced Mechatronic Systems*. 2(1). pp. 68–80. DOI: 10.1504/IJAMECHS.2010.030850
7. Utkin, V.I. (2012) *Skol'zyschie rezhimy v zadachakh optimizatsii i upravleniya* [Sliding modes in control and optimization]. Berlin: Springer.
8. Kolesnikov, A.A. (ed.) (2004) *Sinergetika i problemy teorii upravleniya* [Synergetics and problems of control theory]. Moscow: Fismatlit.
9. Kolesnikov, A.A., Kolesnikova, S.I. & Tsvetnitskaya, S.A. (2019) Discrete Stochastic Regulator on a Manifold, Minimizing Dispersion of the Output Macrovariable. *Mekhatronika, Avtomatizatsiya, Upravlenie*. 12(20). pp. 707–713. DOI: 10.17587/mau.20.707-713
10. Kolesnikova, S., Tsvetnitskaya, S. & Pobegailo, P. (2019) An Extension of the Method of Aggregated Regulators to a Discrete Stochastic Object. *International Siberian Conference on Control and Communications, SIBCON*. DOI: 10.1109/SIBCON.2019.8729595

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

УДК 519.6

DOI: 10.17223/19988605/55/2

I.A. Brokarev, M.P. Farkhadov, S.V. Vaskovskii**RECURRENT NEURAL NETWORKS TO ANALYZE THE QUALITY OF NATURAL GAS**

Comparative analysis of various neural network models was carried out for natural gas quality analysis. Based on the results of such analysis, it was concluded that recurrent neural networks are main statistical models in this problem. This paper considers a recurrent neural network with a more complex architecture. The neural network with gated recurrent unit is used in the discussed task in particular. The comparison of the main recurrent neural network models (simple recurrent neural network, recurrent neural network with long short-term memory, recurrent neural network with gated recurrent unit) is shown. Models accuracy characteristics are shown for analyzing the models performance.

Keywords: recurrent neural networks; natural gas quality analysis; gated recurrent unit.

The natural gas quality analysis is an important task for the gas industry. Slight fluctuations of natural gas composition and energy characteristics can lead to unexpected difficulties in calculating its cost indicators. Currently, a wide variety of different natural gas analysis systems are developed. Moreover, many alternative systems that are based on the correlation methods are under development. The possibility to analyze gas quality in real time is the most significant benefit of this class of systems in comparison with systems based on the traditional gas chromatography methods. However, systems that are commonly used in gas industry have a number of drawbacks: expensive specialized equipment, significant amount of time of the analysis, the necessity of regular instrumentation calibration and checkout.

Various statistical models are used in correlation methods because of high complexity of solving the task with traditional computational methods. The choice of statistical model for the gas quality determination is made by heuristic methods in most cases due to the lack of a general algorithm. That is why comparative analysis of statistical models for the discussed task is an urgent problem that should be solved for reaching the required goal.

This paper provides a proposition of using recurrent neural network with gated recurrent unit for natural gas composition determination. The model differs from previously used neural network in the form of multi-layer perceptron and simple recurrent neural network. The main difference of the proposed system from previous papers [1, 2] is usage of more complex neural network architecture (network with gated recurrent unit).

1. Proposed recurrent neural network

Based on the results of correlation analysis and neural network models investigated in previous works [3], we choose the speed of sound, thermal conductivity coefficient and concentration of carbon dioxide as the model input parameters. As the output parameters, we choose the concentrations of the components included in the investigated gas mixture.

The Levenberg-Marquardt algorithm was chosen to train the model. This algorithm optimizes parameters of nonlinear regression models. Note that in this algorithm, the optimization criterion is the root-mean-

square error of the model on a training sample. The basic idea of the algorithm is as follows: to minimize the error locally, the initial values of the parameters are approximated. Let a regression sample be a set of pairs of an independent variable X and a dependent variable Y , and the regression model is a continuously differentiable function F . It is necessary to find the value of the parameter vector W , where the error function F_ε reaches its local minimum:

$$F_\varepsilon = \sum_{i=1}^N (Y_i - F(X_i, W)). \quad (1)$$

On the first iteration of the algorithm, the initial vector of parameters W_0 is specified. On each following iteration, the vector is replaced by the vector $W_0 + \Delta W$. To estimate the increment ΔW , the following approximation of F is used:

$$F(W_0 + W, X) - F(W_0, X) = J W, \quad (2)$$

where J is the Jacobian of F .

The increment ΔW at the minimum of F_ε is zero. So, to find the subsequent value of the increment ΔW , it is necessary to set the vector of partial derivatives of F_ε over W to zero. Then we differentiate this expression over W and set the partial derivative to zero. After all transformations, we get the expression for ΔW :

$$W = (J^T J)^{-1} J^T (Y - F(W)). \quad (3)$$

For this algorithm, the condition number of the matrix is important; the number shows how close the matrix is to the partial rank matrix (for square matrices it shows how close the matrix is to degeneracy). Since the condition number of the matrix $J^T J$ is equal to the squared condition number of the matrix J , the matrix $J^T J$ may turn out to be degenerate. For this reason, in this algorithm we introduced a regularization parameter λ ; the parameter is greater than or equal to zero. This parameter is selected on each iteration of the algorithm. Given the regularization parameter, expression for ΔW takes the following form:

$$W = (J^T J + \lambda E)^{-1} J^T (Y - F(W)), \quad (4)$$

where E is a unity matrix.

There is a modification of this method, where the regularization parameter is multiplied by the matrix D , a diagonal matrix with the elements that coincide with the diagonal elements of the matrix $J^T J$. This approach is used to reduce the effect of the regularization parameter on the ΔW value.

It is necessary, however, to note that this algorithm converges slower if the step is constant, which is a disadvantage. The problem is solved by introducing a coefficient K ; the coefficient determines the step length and makes the method converge faster. Given the two amendments described, the expression for ΔW takes the following form:

$$W = K (J^T J + \lambda E)^{-1} J^T (Y - F(W)). \quad (5)$$

The value of the vector W at the last iteration of the algorithm is the target. It is reached either if the calculated increment ΔW is less than the specified value, or if the error function F_ε is less than the specified value for the vector W .

To solve this problem, we use a recurrent neural network with a gated recurrent unit (GRU) [4] to model our system. Gated recurrent units are a specialized gate mechanism for recurrent neural networks. The structure of a gated recurrent unit is shown on Figure 1. The structure of a neural network with this block is similar to the previously presented simple recurrent network.

The main differences between neural networks with GRU and neural networks with LSTM are the number of gates and maintenance of cell states. Unlike neural networks with GRU, neural networks with LSTM have three gates (input, forget, output) and maintains an internal memory cell state, which makes it more flexible, but less efficient memory and time wise. However, since both of these networks are great at addressing the vanishing gradient problem required for efficiently tracking long term dependencies. When deciding between these two variants, it is recommended that firstly using neural networks with LSTM, since it has more parameters and is more flexible, followed by neural networks with GRU. If there are no sizable differences between the performance of this two types of RNN, then the simpler and more efficient neural networks with GRU are used.

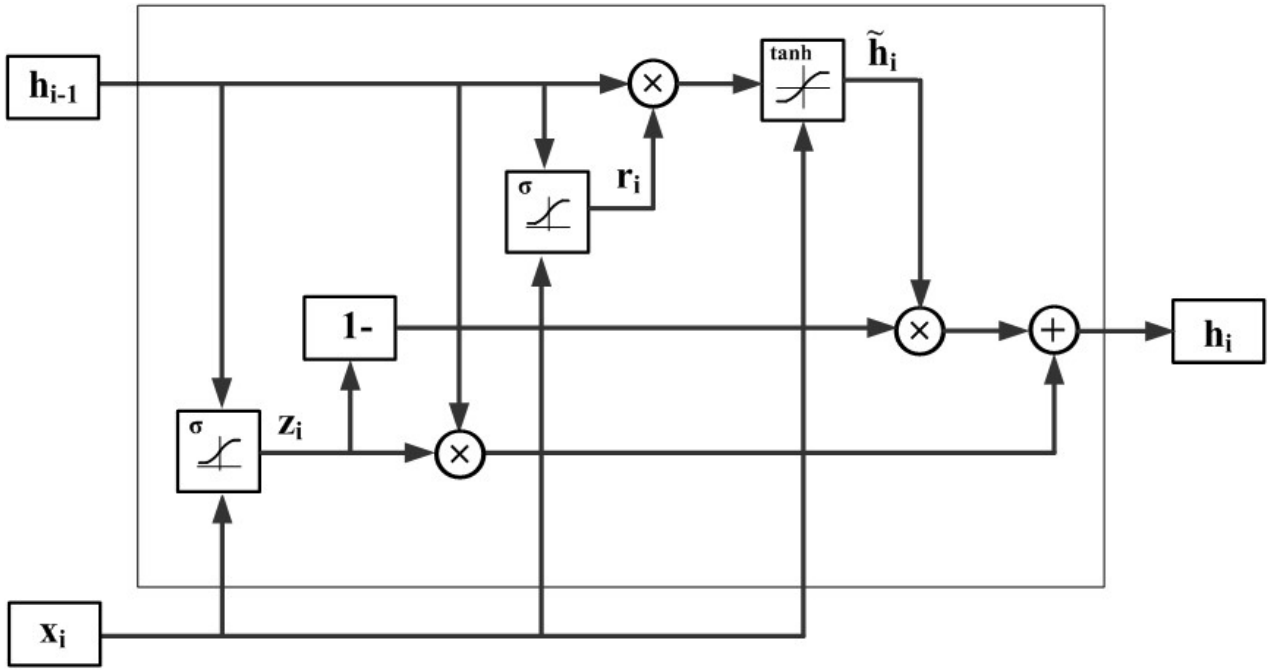


Fig. 1. The architecture of gated recurrent unit

Recurrent neural networks (RNN) is a class of neural networks that can use their internal memory when processing input data [5]. The functioning of this class of neural networks is based on the use of previous network state to calculate the current one. A recurrent network can be considered as several copies of the same network, each of which transfers information to a subsequent copy. Currently, there are a large number of architectures of recurrent neural networks. Taking into account the computational difficulties encountered in developing this class of neural networks, it was proposed to consider a simple recurrent neural network first. The hidden elements have links directed back to the input layer in such type of network. This allows to take into account the previous state of the network during training. Mathematically, the process of saving information about the previous training step is as follows: at each i -th training step, the output value of the RNN hidden layer h_i is calculated taking into account the output value of the hidden layer in the previous step.

The proposed model with gated recurrent unit is similar to recurrent neural network with long short-term memory [6]. Long short-term memory (LSTM) is a special type of architecture of recurrent neural networks, that is capable of learning long-term dependencies. A more complex method is used to calculate both the output value of the hidden layer and the output value of the network as a whole in neural networks with a similar architecture. This method involves use of so-called gates. A gate is a special unit in LSTM architecture, that is implemented as a logistic function and operation of elementwise multiplication (Hadamard's product). The logistic function layer shows how much of the information coming from a particular unit should be transmitted further along the network. This layer returns values in the range from zero (information does not go further along the network structure at all) to one (information completely goes further along the network structure). There are three such gates in traditional LTSM architecture: a forget gate, an input gate and an output gate. The sigmoid function is often used as a logistic function for gates.

This neural network architecture includes an update gate, a reset gate and a current memory container. Both gates are functions of the input vector and the state of the hidden layer on the previous step. The update gate determines the amount of information from the previous state of the hidden layer that should affect its current state. The reset gate determines the part of information from the previous state that will not be taken into account when we calculate the current state. The current memory container is used to store the relevant information from the previous state and depends on the reset gate.

Let us investigate how the gated recurrent unit operates. When the vector x_i is fed to the block, the vector is multiplied by the vector of weights W_z of the update gate. The vector of the previous state

of the hidden layer h_{i-1} , that stores information about the previous $i - 1$ steps, is multiplied by the vector of weights U_z of the update gate. Then the two products are added and fed to the sigmoidal activation function σ . As a result, the output value of the update gate z_i is calculated as:

$$z_i = \sigma(W_z x_i + U_z h_{i-1}). \quad (6)$$

The formula to calculate the output value of the reset gate r_i is similar to the formula for the update gate, except for the weight vectors W_r and U_r

$$r_i = \sigma(W_r x_i + U_r h_{i-1}). \quad (7)$$

We use the value of the reset gate to calculate the current value of the memory container. First, the corresponding vectors of the weights W and U . Multiply the input vector and the previous state vector of the hidden layer. Then the Hadamard product of the state reset filter value and the $U h_{i-1}$ value is calculated. This is necessary to determine the amount of information from the previous steps we do not consider. Then, we sum up the two previously obtained terms and apply the nonlinear hyperbolic tangent activation function ($\tanh$) to obtain the value of the current memory container H_i :

$$H_i = \tanh(W x_i + r_i U_r h_{i-1}). \quad (8)$$

If the value of the reset gate is close to one, then the memory container saves information from the previous hidden state. If the value of the reset gate is close to zero, then the memory container does not take into account the previous hidden state.

At the final stage of the block, the vector h_i is calculated; the vector contains the information about the current state of the hidden layer and is transmitted to the next layer of the neural network. For this procedure, we use an update gate; the filter determines how much information to take from the current memory container and from the previous state of the hidden layer. To do this, the Hadamard product of the update gate value and the vector of the previous state of the hidden layer, and the Hadamard product of the value of the current memory container and the value $(1 - z_i)$ are calculated. Finally, we sum up the results of the two products to get the block output value:

$$h_i = z_i h_{i-1} + (1 - z_i) H_i. \quad (9)$$

If the update gate value is close to one, then the current memory container is not taken into account, and the block output equals the previous value of the hidden layer. If the update gate value is close to zero, then the previous state of the hidden layer is not taken into account, and the block output depends only on the current memory container.

The main advantage of a managed recurrent block is the ability to store and filter information using update and reset gates. Such system has no vanishing gradients as the model stores the relevant information and passes the information on to the next time steps of the network.

2. Testing of neural network model

Based on the results of this study, we obtain a neural network model with a gated recurrent unit to analyze the quality of natural gas by measured physical parameters. To determine if the proposed neural network model is adequate, we studied a test sample of data. The results of composition determination and accuracy of the proposed model are shown in Table 1. Maximum absolute deviation (MaxAD), mean absolute deviation (MAD), mean square error (MSE) and determination coefficient (R^2) are determined.

The ranges of gas mixtures for the test sample were as follows: 85–100% for methane, 0–5% for propane, nitrogen and carbon dioxide. We took the concentration of carbon dioxide to be zero, since this parameter is an input for the model and is considered a known value.

Table 1

Accuracy of RNN model for determination of the natural gas composition

Component	MaxAD, %	MAD, %	MSE	R^2
Methane	0.73	0.33	0.32	0.9999
Propane	0.21	0.12	0.13	0.9999
Nitrogen	0.19	0.08	0.10	0.9999

Table 2

**Comparison of models accuracy characteristics for determination
of the natural gas composition on testing set**

Component	Characteristic	Model			
		ANN	RNN	LSTM	GRU
Methane	MaxAD, %	0.823	0.767	0.783	0.781
	MAD, %	0.456	0.381	0.356	0.343
	MSE	0.581	0.491	0.451	0.401
	R ²	0.998	0.9998	0.9998	0.9999
Propane	MaxAD, %	0.312	0.271	0.281	0.256
	MAD, %	0.161	0.183	0.171	0.128
	MSE	0.341	0.291	0.289	0.195
	R ²	0.998	0.9999	0.9999	0.9999
Nitrogen	MaxAD, %	0.512	0.331	0.305	0.245
	MAD, %	0.267	0.181	0.179	0.115
	MSE	0.356	0.201	0.197	0.191
	R ²	0.998	0.9998	0.9998	0.9999

The comparison of models accuracy characteristics for determination of the natural gas composition on training and testing set is shown in Table 2. The fact that the statistical model can show good results on a training set, but a high error on a testing set is taken into account.

3. Using the proposed model in natural gas quality determination system

The main structure of the proposed natural gas quality determination system is shown in Figure 2. The system consists of three main blocks. It is proposed to use commercially available and relatively inexpensive sensors for natural gas physical parameters measurements to obtain necessary measurement data. That data are input data of the proposed system. The aim of the system is to determine target natural gas quality parameters using input measurement data.

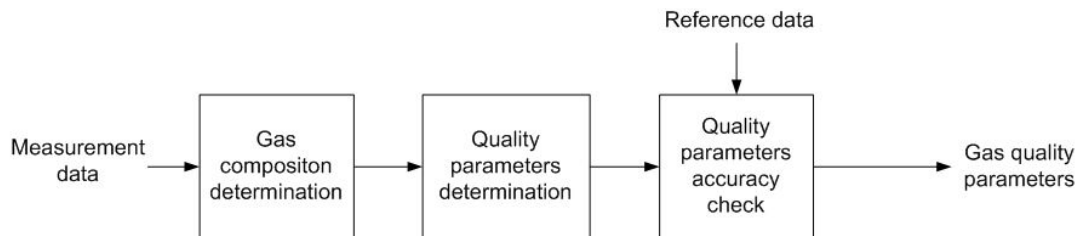


Fig. 2. Structure of the proposed natural gas quality determination system

The main advantages of the proposed system are possibility to operate in real time due to the usage of neural network model, the broad list of output quality parameters and the joint usage of the system with traditional gas chromatographs. The proposed recurrent neural network is located in the first block (gas composition determination). It is the main block of the designed system. It contains the majority of the proposed system features. The main goal of the proposed system is gas energy parameters determination. To calculate the energy parameters of the gas under study, NIST REFPROP software is used. The target energy parameters for the discussed task are volumetric superior calorific value and Wobbe index. These parameters along with partial gas composition and relative density are considered to be final gas quality parameters that system should determine. To calculate the quality parameters, the GERG-2008 gas state equation was used at standard temperature and pressure conditions. The amount of output parameters can be decreased to simplify the calculations or increased by adding volumetric inferior calorific value in special cases. The next step involves energy parameters accuracy check that occurs in the corresponding block. The calculated in previous block gas quality parameters are compared with reference data. Any data obtained from traditional natural gas analyzers, e.g. gas chromatographs, can be used as the reference data.

The volumetric calorific value determined using the proposed recurrent neural network is shown on Figure 3. The determination of target energy characteristics is conducted in the block of the designed gas quality determination system. That block is located after the gas composition determination block where the proposed neural network model determines the desired gas composition.

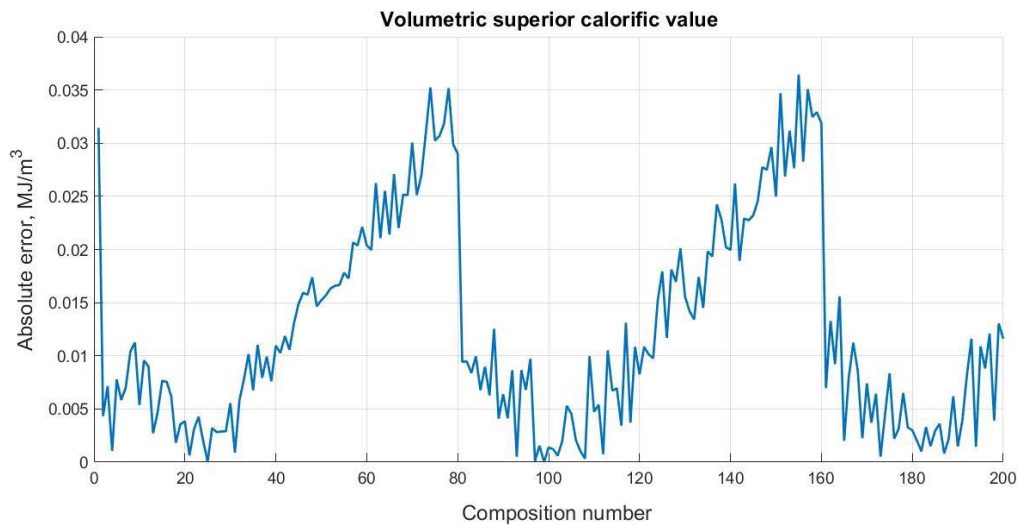


Fig. 3. The accuracy of determination of volumetric calorific value designed gas quality determination system

Theoretical values of natural gas energy parameters was used as reference data. The volumetric superior calorific value and Wobbe index were calculated using determined gas composition and compared with reference data. The maximum absolute error of gas quality parameters determination is less than the allowable error that is equal to 0.1 MJ/m³. The allowable error is permissible deviation of gas quality parameters determination for the first accuracy class according to current regulatory document [7].

Conclusion

Proposed neural network model with a gated recurrent unit determines the quality of natural gas by the values of the gas measured parameters. Our model is adequate to determine accurately the composition of the gas for the specified problem. The simulation results showed that the maximum absolute error in determining the gas quality parameters is less than the permissible error value determined by the current regulatory document.

Another advantage of the proposed gas quality analysis system is possibility to operate in real time due to the usage of recurrent neural network model.

Acknowledgments. The authors are grateful to the international cooperation projects BRISK II TA and Erasmus + 2017-1-SE01-KA107-034292 Staff Mobility between the universities for the opportunity to conduct this study.

REFERENCES

1. Vaskovskii, S.V. & Brokarev, I.A. (2020) Comparative analysis of statistical models for the task of natural gas composition analysis. *Journal of Information Technologies and Computing Systems*. 1. pp. 34–43. DOI: 10.14357/20718632200104
2. Brokarev, I.A. & Vaskovskii, S.V. (2020) Gas Quality Determination Using Neural Network Model-based System. *Proceedings of the 2nd International Workshop on Stochastic Modeling and Applied Research of Technology (SMARTY 2020)*. pp. 113–128.
3. Dörr, H., Koturbash, T. & Kutcherov, V. (2019) Review of impacts of gas qualities with regard to quality determination and energy metering of natural gas. *Meas. Sci. Technol.* 30(2). Article ID 022001.
4. Chung, J., Gulcehre, C., Cho, K. & Bengio, Y. (2014) Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*.
5. Callan, R. (1999) *The Essence of Neural Networks (The Essence of Computing Series)*. Prentice Hall Europe.
6. Hochreiter, S. & Schmidhuber, J. (1997) Long short-term memory. *Neural computation*. 9(8). pp. 1735–1780.

7. ISO 15971:2008. (2008) *Natural gas – Measurement of properties – Calorific value and Wobbe Index*. International Organization for Standardization, 2008. [Online] Available from: <https://www.iso.org/standard/44867.html> (Accessed: 2nd March 2021).

Received: March 2, 2021

Brokarev I.A., Farkhadov M.P., Vaskovskii S.V. (2021) RECURRENT NEURAL NETWORKS TO ANALYZE THE QUALITY OF NATURAL GAS. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika* [Tomsk State University Journal of Control and Computer Science]. 55. pp. 11–17

DOI: 10.17223/19988605/55/2

Брокарев И.А., Фархадов М.П., Васильковский С.В. РЕКУРРЕНТНЫЕ НЕЙРОННЫЕ СЕТИ ДЛЯ АНАЛИЗА КАЧЕСТВА ПРИРОДНОГО ГАЗА. *Вестник Томского государственного университета. Управление, вычислительная техника и информатика*. 2021. № 56. С. 11–17

В статье рассматривается рекуррентная нейронная сеть с управляемым рекуррентным блоком. Показано сравнение основных моделей рекуррентных нейронных сетей (простая рекуррентная нейронная сеть, рекуррентная нейронная сеть с долгой кратковременной памятью, рекуррентная нейронная сеть с управляемым рекуррентным блоком). Приведены точностные характеристики нейросетевых моделей для анализа качества природного газа.

Ключевые слова: рекуррентные нейронные сети; анализ качества природного газа; управляемый рекуррентный блок.

BROKAREV Ivan Andreevich (Post-graduate Student, Assistant, National University of Oil and Gas «Gubkin University», Moscow, Russian Federation).

E-mail: brokarev.i@gubkin.ru

FARKHADOV Mais Pashaevich (Doctor of Technical Sciences, V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russian Federation).

E-mail: mais@ipu.ru

VASKOVSKII Sergei Vladimirovich (Candidate of Technical Sciences, V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russian Federation).

E-mail: v63v@yandex.ru

ОБРАБОТКА ИНФОРМАЦИИ

УДК 004.855

DOI: 10.17223/19988605/55/3

К.В. Вик, А.А. Друки, Д.С. Григорьев, В.Г. Спицын

ПРИМЕНЕНИЕ НЕЙРОННЫХ СЕТЕЙ ГЛУБОКОГО ОБУЧЕНИЯ ДЛЯ РЕШЕНИЯ ЗАДАЧИ СЕГМЕНТАЦИИ ЛЕСНЫХ ПОЖАРОВ НА СПУТНИКОВЫХ СНИМКАХ

Работа выполнена при поддержке РФФИ в рамках научного проекта № 18-08-00977 и программы повышения конкурентоспособности Томского политехнического университета.

Целью работы является разработка алгоритмов семантической сегментации областей лесных пожаров на спутниковых снимках земной поверхности. При активном развитии алгоритмов компьютерного зрения сегодня существует ряд задач в данной области, которые не решены в полной мере и не обеспечивают требуемую точность работы. Поэтому существует потребность в разработке алгоритмов и программных средств, которые обеспечили бы высокое качество сегментации изображений. На основе анализа существующих методов и алгоритмов сегментации изображений было принято решение использовать нейросетевые алгоритмы. В процессе выполнения работы разработана сверточная нейронная сеть, а также сформирована обучающая выборка. Для разработки нейронной сети применялась библиотека машинного обучения Keras, также использовались оптимизации алгоритма обратного распространения ошибки. В результате была осуществлена программная реализация алгоритма, позволяющего выполнять сегментацию областей лесных пожаров на спутниковых снимках земной поверхности. Представлены результаты работы, а также сравнение их эффективности с существующими аналогами.

Ключевые слова: нейронные сети; семантическая сегментация; компьютерное зрение; обработка изображений; изображения поверхности Земли.

Семантическая сегментация изображений является процессом разделения изображения на отдельные составные части, которые соответствуют различным объектам, и последующего выполнения классификации данных объектов. Сегодня семантическая сегментация применяется в различных областях, например в системах видеонаблюдения для распознавания людей, в беспилотных автомобилях для анализа окружающей среды, в смартфонах для классификации объектов, в медицине для анализа рентгеновских снимков и т.д. Данная область является одной из самых активно развивающихся, однако качество работы большинства современных алгоритмов еще не достигло максимума. Поэтому зачастую часть работы, требующая высокой ответственности, выполняется людьми в ручном режиме, из-за чего могут возникать ошибки и увеличиваться время выполнения работы [1].

Сегодня существует ряд эффективных алгоритмов, которые применяются в задачах сегментации. Как правило, в основе большинства из них лежат поиск ключевых характеристик объектов на изображениях и последующая классификация данных объектов [2]. В качестве ключевых характеристик могут использоваться различные параметры: цвет, форма, размер, текстура и т.д. [3].

Алгоритмы, основанные на поиске особых точек для описания объектов на изображениях: SURF [2]; SIFT [3]; MSER [4]; FAST; FAST-9; FAST-ER [5]; HOG [6]. Особыми точками являются такие области изображения, которые сохраняют свои характеристики при различных изменениях условий съемки и смещениях камеры. К достоинствам методов поиска особых точек можно отнести их высокую устойчивость к масштабированию и незначительному повороту изображения объекта. Недостаток подобных методов – неустойчивость к специфичным изображениям объекта, на которых нельзя определить направления дескрипторов, к смене освещения [5, 6].

Алгоритмов классификаторов большое множество. Например, можно выделить следующие: SVM, Bag of words. Также к алгоритмам классификаторов можно отнести искусственные нейронные сети [7].

В основе алгоритма SVM лежит принцип нахождения гиперплоскости, которая разделяет объекты на несколько классов. Для выполнения качественной классификации необходимо нахождение данной гиперплоскости на достаточном расстоянии от объектов. Применение данного алгоритма не требует наличия большой обучающей выборки, так как используется не все множество, а лишь некоторая часть объектов на границах областей, что может быть причиной ошибок при классификации [8].

В основе работы алгоритма Bag of words лежат нахождение ключевых точек изображения и построение нормированной гистограммы встречаемости «слов», описывающей изображение. Недостаток данного алгоритма заключается в том, что он не учитывает пространственную информацию объектов, поэтому могут возникать ошибки классификации при наличии схожих ключевых точек объектов.

Классические полносвязные искусственные нейронные сети, также называемые многослойным перцептроном [9], не очень подходят для обработки изображений. Это связано с тем, что изображения, как правило, имеют большой размер и состоят из большого количества пикселей, поэтому необходимо, чтобы нейронная сеть состояла из большого количества нейронов и связей. Это приводит к тому, что нейронная сеть становится очень громоздкой и сложнообучаемой. Еще один недостаток заключается в том, что классические нейронные сети не учитывают двумерную структуру изображений и плохо запоминают пространственную взаимосвязь объектов и их составных частей [9].

В 1998 г. был разработан тип нейронных сетей, которые хорошо подходят для обработки изображений. Такие нейронные сети называются сверточными нейронными сетями, они обеспечивают высокую устойчивость к искажениям объектов, наличию шумов, сдвигам и смене ракурса съемки. Они имеют двумерную структуру, аналогичную структуре изображений, и анализируют изображение не целиком, а по отдельным частям. Этот процесс обеспечивается наличием специальных сверточных и подвыборочных слоев, которые от слоя к слою уменьшают размер входного сигнала. Таким образом, создается пирамида из слоев нейронной сети, где каждый отдельный слой содержит набор плоскостей. В рамках каждой плоскости применяется свой набор весов. Каждый нейрон выполняет операцию свертки двумерной области нейронов предыдущего слоя. Сверточные нейронные сети обеспечивают более быстрое и эффективное обучение, так как содержат меньшее количество настраиваемых весовых коэффициентов.

Целью работы является разработка алгоритмов семантической сегментации областей лесных пожаров на спутниковых снимках земной поверхности.

Основные задачи:

1. Выбор среды разработки и программных библиотек для реализации алгоритмов.
2. Создание и подготовка обучающей выборки снимков дистанционного зондирования Земли.
3. Разработка нейронной сети.
4. Сравнение реализованных алгоритмов с аналогами.

1. Выбор библиотеки машинного обучения

При выборе программных средств для разработки нейронных сетей было изучено несколько библиотек машинного обучения:

Keras – открытая библиотека машинного обучения, которая поддерживается компанией Google. Данная библиотека является надстройкой для более низкоуровневой библиотеки TensorFlow, содержит очень удобный функционал для реализации и обучения нейронных сетей, в том числе и в задачах компьютерного зрения [10].

PyTorch – библиотека машинного обучения, разработанная NEC Laboratories America и New York University. Библиотека включает широкий выбор алгоритмов машинного обучения и подходит для научных вычислений [11]. Использует динамический граф вычислений, что обеспечивает удоб-

ство и снижение временных затрат, так как позволяет изменять структуру нейронной сети перед каждым запуском без предварительной компиляции.

Эффективность применения библиотек исследовалась на наборе изображений CIFAR-10. Данный набор содержит изображения различных классов объектов: животные, автомобили, самолеты, техника и т.д. Количество изображений – 60 000. Для тестирования библиотек применялась классическая сверточная нейронная сеть VGG16.

В результате тестирования библиотеки Keras и TensorFlow показали результат скорости работы на уровне 236 с. Библиотека PyTorch показала результат 153 с. Все библиотеки показали точность работы на уровне 79%. Исходя из полученных результатов было принято решение использовать PyTorch для дальнейшей работы.

2. Структура нейронной сети

Как правило, архитектура и различные параметры нейронной сети выбираются экспериментальным путем. Это обусловлено тем, что сверточные нейронные сети содержат большое количество параметров, и на сегодняшний день не существует строго формализованных правил подбора этих параметров и структуры сети в целом. Рассматривались следующие варианты: U-Net [12], ENet [13], SegNet [14].

U-Net состоит из двух частей: сжимающейся части для вычисления характеристик и расширяющейся части для пространственной локализации объектов на изображении. В данной модели не используется полносвязный слой. Как следствие, число параметров модели уменьшается, и ее можно обучить с помощью небольшого набора данных.

SegNet, как и U-Net, состоит из двух частей. Сжимающая часть содержит несколько сверточных слоев с функцией активации ReLU, пакетную нормализацию и операции подвыборки. Это обеспечивает важные преимущества сохранения высокочастотных деталей в сегментированных изображениях, а также уменьшения общего количества настраиваемых параметров в декодерах.

Enet имеет более компактную структуру и содержит в несколько раз меньше параметров, что обеспечивает более высокую скорость работы и возможность применения на мобильных устройствах в режиме реального времени.

В процессе проведенных экспериментов архитектура SegNet показала лучшие результаты в сравнении с аналогами, поэтому было решено использовать ее в качестве основы с дальнейшей модификацией ее архитектуры. В итоге после проведения ряда вычислительных экспериментов была разработана архитектура нейронной сети (рис. 1).

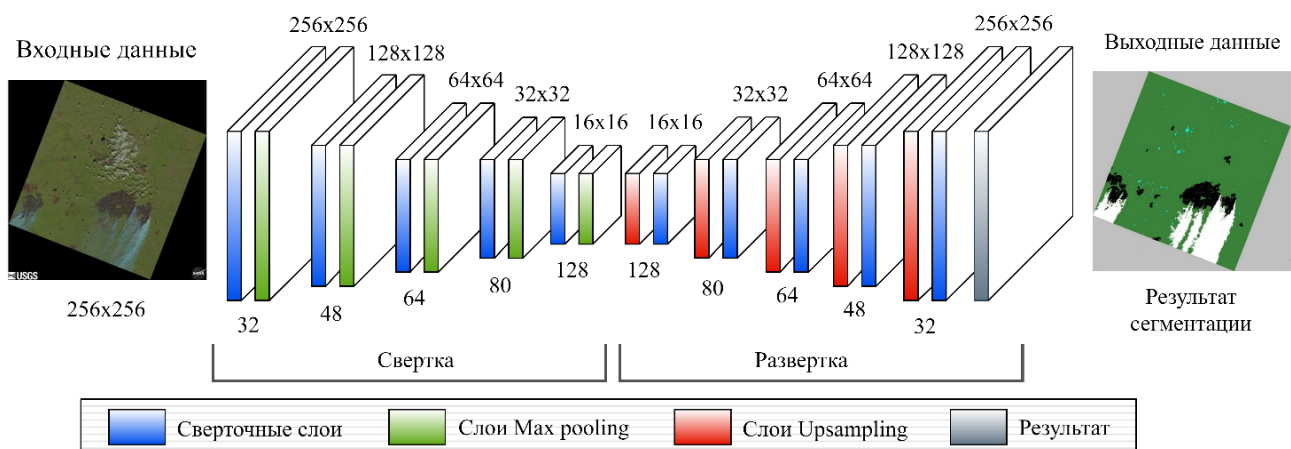


Рис. 1. Архитектура сверточной нейронной сети
Fig. 1. Architecture of the convolutional neural network

Нейронная сеть состоит из двадцати слоев: 10 сверточных слоев, 5 слоев Max pooling, 5 слоев Upsampling, и выходного слоя.

Работу нейронной сети можно разделить на два этапа: свертка входного изображения и развертка полученного вектора.

Как показано на рис. 1, входное изображение проходит через пять слоев (обозначены синим и зеленым цветом), и выполняется свертка изображения в вектор. Далее следуют пять слоев (обозначены синим и красным цветом), которые выполняют функцию Upsampling и предназначены для развертывания полученного вектора в исходный размер изображения. На последних этапах полученные данные переводятся в три цветовых канала, и на выходе получаем результат сегментации.

3. Обучение нейронной сети

Обучение нейронной сети выполнялось с помощью сформированного набора из 50 снимков дистанционного зондирования Земли с разрешением $8\,000 \times 8\,000$. Набор изображений был отобран из базы спутника Landsat 8 [15]. Основные критерии выбора касались размера сцены, а также количества снимков, которые делает спутник в течении суток.

Сформированный набор изображений содержит данные следующих классов:

- лесной пожар (красный цвет);
- выгоревшая площадь (черный цвет);
- дым от пожара (белый цвет);
- водоемы (синий цвет);
- лес (зеленый цвет).

Создание сегментированных масок изображений производилось с помощью преобразований и фильтрации в программе Adobe Photoshop (рис. 2).

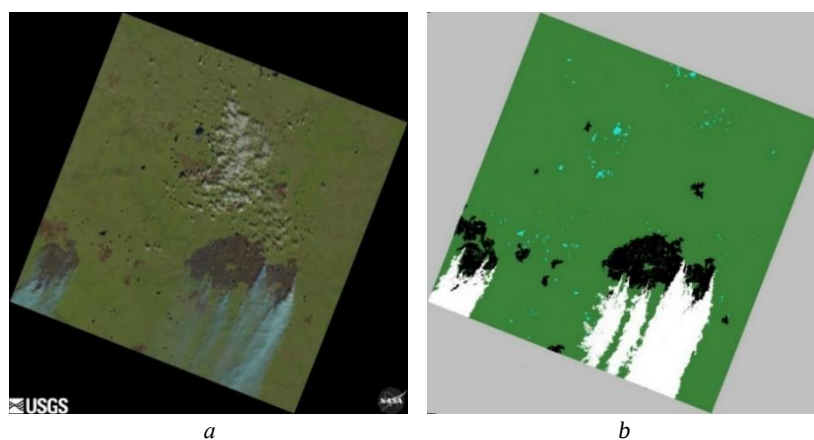


Рис. 2. Изображения для обучения нейронной сети: *a* – входное изображение; *b* – сегментированное изображение.

Fig. 2. Images for neural network training: *a* – input image; *b* – segmented image.

Так же для набора изображений была проведена аугментация, т.е. модификация данных для обучения. Использование данного метода улучшает обобщающую способность нейронной сети, добавляет новые примеры для обучения, которые нейросеть еще не видела, и не дает возможности переобучаться.

В качестве аугментации были выполнены следующие модификации: поворот изображения на произвольный градус; сжатие по осям; растяжение по осям; зеркальное отображение по осям; Гауссово размытие; изменение яркости и контраста. Обучение включало 50 эпох, каждая из которых содержит 2 000 итераций.

На сегодняшний день существует множество алгоритмов обучения нейронных сетей. Один из наиболее известных – алгоритм обратного распространения ошибки. Алгоритм применяется для минимизации отклонения текущего значения выходного сигнала сети от требуемого. Недостаток данно-

го алгоритма заключается в возможности переобучения при выборе малой скорости обучения или, наоборот, расходимости при выборе слишком большой скорости обучения.

Для устранения данных недостатков существуют различные модификации алгоритма обратного распространения ошибки. Проведен анализ трех модификаций данного алгоритма с целью выбора наилучшего: Adam – adaptive moment estimation; Adagrad – adaptive gradient; RMSProp – gradient descent with momentum. Основное отличие данных алгоритмов от классического алгоритма обратного распространения ошибки заключается в том, что в них используется адаптивная скорость обучения для каждого веса сети, которая адаптируется в процессе обучения [16, 17].

Для более стабильного обучения применялся метод регуляризации DropOut, который предназначен для уменьшения переобучения сети за счет случайного обновления весов в некоторых областях сети. Вероятность попадания нейронов в область обновления: 0,5. Для стабильной работы применялся метод нормализации Mini-batch. В данном методе на вход нейронной сети подаются данные, которые были предварительно обработаны и имеют нулевое математическое ожидание и единичную дисперсию. Для регуляризации весов сети применялся метод L2, который назначает размер штрафной коррекции для слишком больших и малых значений весов.

Параметры нейронной сети, использованные при обучении и тестировании:

- коэффициент обучения: 0,0005;
- изменение коэффициента обучения: 0,1;
- частота изменения коэффициента обучения: 10^4 .

Аппаратное обеспечение, используемое для экспериментов: Intel(R) Core(TM)6 CPU (3.7GHz), ОЗУ 32ГБ, GeForce RTX 2080 SUPER.

Результаты обучения представлены в табл. 1.

Таблица 1

Результаты обучения сверточной нейронной сети

Adam	Adagrad	RMSProp
Количество эпох: 50	Количество эпох: 50	Количество эпох: 50
Точность при обучении: 98,828	Точность при обучении: 95,761	Точность при обучении: 94,179
Точность при тестировании: 96,033	Точность при тестировании: 94,975	Точность при тестировании: 91,107

Очевидно, что алгоритм Adam показывает лучшие результаты по сравнению с двумя другими алгоритмами: точность классификации 96,033 %.

Преимущества алгоритма Adam в задачах оптимизации состоят в следующем: достаточно прост в реализации; вычислительно эффективен в применении; инвариантен к изменению масштаба градиента по диагонали; эффективен для задач с большим количеством параметров и данных; эффективен для задач с очень шумными или редкими градиентами.

Результаты обучения нейронной сети представлены на рис. 3.

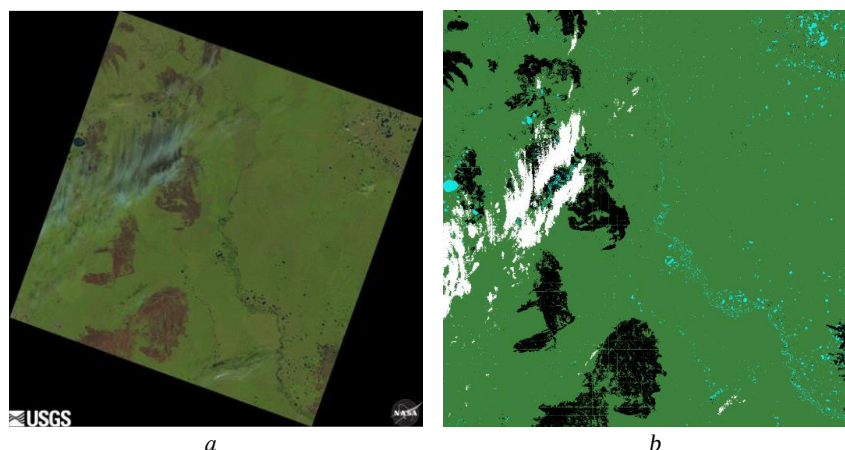


Рис. 3. Результат сегментации: *a* – исходное изображение; *b* – полученный результат
Fig. 3. Image segmentation result: *a* – original image; *b* – the result obtained

В табл. 2 представлено сравнение результатов работы предложенной нейронной сети с некоторыми аналогами.

Таблица 2

Сравнение результатов с аналогами

Название	Пожар (%)	Выгоревшая площадь (%)	Дым (%)	Водоемы (%)	Лес (%)	Итого (%)
Edge Flow [18]	90,6	93,2	82	88,8	86,6	88,24
PSO K-Means [Ibid.]	88,7	92,3	88,6	90,2	94,8	90,92
Multi scale [19]	93,7	96,4	88,2	89,5	96	92,76
Edison [Ibid.]	93,9	96	85,6	94,8	95,8	93,22
JSEG [Ibid.]	91,4	92,1	89,7	95,3	92,4	92,18
Разработанная нейронная сеть	93,6	95,7	87,6	96,9	97,3	94,22

Сравнительное исследование точности работы алгоритмов сегментации проводилось на наборе эталонных и тестовых изображений, подвергаемых шумовым искажениям. Для сравнения результатов сегментации использовались границы сегментированных объектов, которые представляет собой набор точек, не зависящих от закраски сегментов. Для измерения результатов сегментации использовались две метрики: среднее и Хаусдорфово расстояние. Исследование качества работы ряда алгоритмов показало, что они ведут себя неустойчиво при зашумлении и размытии изображения. Таким образом, можно сделать вывод, что целесообразно до процедуры сегментации выполнить очистку изображения от шума и повысить его четкость.

Точность работы разработанной нейронной сети составляет 94,22%. Данный результат является лучшим в сравнении с представленными аналогами. Как можно заметить, разработанная система несколько уступает некоторым аналогам по таким показателям, как пожар, выгоревшая площадь, дым. Однако по таким показателям, как леса, водоемы, она побеждает.

Заключение

В данной работе рассматривалась задача семантической сегментации областей лесных пожаров на спутниковых снимках земной поверхности. Изображение сегментировалось на пять различных объектов: лесной пожар; выгоревшая площадь; дым от пожара; водоемы; лес. Разработана сверточная нейронная сеть глубокого обучения, которая обеспечивает точность работы на уровне 94,22%. Также было проведено сравнение эффективности работы алгоритмов обучения, в результате чего лучшие показатели были получены с помощью алгоритма Adam.

ЛИТЕРАТУРА

1. Bundzel M., Hashimoto S. Object identification in dynamic images based on the memory-prediction theory of brain function // Journal of Intelligent Learning Systems and Applications. 2010. V. 2, № 4. P. 212–220.
2. Tawfiq A., Ahmed J. Object detection and recognition by using enhanced Speeded Up Robust Feature // International Journal of Computer Science and Network Security. 2016. V. 16, № 4. P. 66–71.
3. Park S., Yoo J.H. Realtime face recognition with SIFT based local feature points for mobile devices // The 1st International Conference of Artificial Intelligence, Modelling and Simulation (AIMS 13). Malaysia, 2013. P. 304–308.
4. Mammeri A., Boukerche A., Khiari E. MSER based text detection and communication algorithm for autonomous vehicles // IEEE Symposium of Computers and Communication. Messina, Italy. 2016. P. 456–460.
5. Tore V., Chawan P.M. FAST Clustering based feature subset selection algorithm for high dimensional data // International Journal of Computer Science and Mobile Computing. 2016. V. 5, № 7. P. 234–238.
6. Dalal N., Triggs B. Histograms of oriented gradients for human detection // IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR). San Diego, USA. 2005. V. 1. P. 886–893.
7. Mohey D.E. Enhancement bagofwords model for solving the challenges of sentiment analysis // International Journal of Advanced Computer Science and Applications. 2016. V. 7, № 1. P. 244–251.
8. Kecman V., Melki G. Fast online algorithms for Support Vector Machines // IEEE South East Conference. Virginia, USA. 2016. P. 26–31.
9. Le Cun Y., Bengio Y. Convolutional networks for images, speech and time series // The handbook of brain theory and neural networks. 1998. V. 7, № 1. P. 255–258.

10. TensorFlow: the Python Deep Learning library. URL: <https://www.tensorflow.org> (accessed: 02.12.2020).
11. Deep Learning Frameworks: a Survey of TensorFlow, Torch, Theano, Caffe, Neon, and the IBM Machine Learning Stack. URL: <https://www.microway.com/hpc-tech-tips/deep-learning-frameworks-survey-tensorflow-torch-theano-caffe-neon-ibm-machine-learning-stack> (accessed: 02.12.20).
12. Ronneberger O., Fischer P., Brox T. U-Net: Convolutional networks for biomedical image segmentation // International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI-2015). 2015. V. 93. P. 234–241.
13. Paszke A., Chaurasia A., Kim S., Culurciello E. ENet: a Deep Neural Network Architecture for Real-Time Semantic Segmentation // 5th International Conference on Learning Representations. 2017. Toulon, France. 2017. P. 1–10.
14. Badrinarayanan V., Kendall A., Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation // IEEE transactions on pattern analysis and machine intelligence. 2017. V. 39, № 12. P. 54–62.
15. Sharing Earth Observation Resources. URL: <https://directory.eoportal.org/web/eoportal/satellite-missions/l/landsat-8-ldcm> (accessed: 02.12.20).
16. Kingma D.P. Adam: a Method for Stochastic Optimization // International Conference of Learning Representations. San Diego, USA. 2015. P. 1–13.
17. Nguyen V., Kim H., Jun S. A Study on Real-Time Detection Method of Lane and Vehicle for Lane Change Assistant System Using Vision System on Highway // Engineering Science and Technology. 2018. V. 21. P. 822–833.
18. El-Khatib S.A. Image segmentation using a mixed and exponential particle-based algorithm // Computer Science and Cybernetics. 2015. № 1. P. 126–133.
19. Koltsov P.P. The use of metrics in a comparative study of the quality of work of image segmentation algorithms // Informatics and Its Applications. 2011. № 5. P. 53–63.

Поступила в редакцию 18 декабря 2020 г.

Vik K.V., Druki A.A., Grigoriev D.S., Spitsyn V.G. (2021) APPLICATION OF DEEP LEARNING NEURAL NETWORKS FOR SOLVING THE PROBLEM OF FOREST FIRE SEGMENTATION ON SATELLITE IMAGES. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika* [Tomsk State University Journal of Control and Computer Science]. 55. pp. 18–25

DOI: 10.17223/19988605/55/3

The aim of the work is to develop algorithms for the semantic segmentation of forest fire areas on satellite images. Despite the active development of computer vision algorithms, today there are a number of problems in this area that have not been fully solved and do not provide the required accuracy of work. Therefore, today there is a need for the development of algorithms and software that provide high quality image segmentation.

The analysis of existing algorithms for image segmentation was carried out and it was revealed that the most suitable algorithms for solving this problem are deep learning neural networks. The machine learning libraries Keras, TensorFlow and PyTorch were reviewed. The library performance was tested on a set of 60,000 images. In the process of research, the PyTorch library showed the best results, so it was decided to use it to develop algorithms. Convolutional neural network consisting of 20 layers has been developed.

The neural network was trained using a generated set of 50 images of Earth remote sensing with a resolution of 8000x8000. The set of images was selected from the Landsat 8 satellite database. The main selection criteria concerned the size of the scene, as well as the number of images taken by the satellite during the day.

The generated set of images contains data of the following classes: forest fire (red); burnt-out area (black); smoke from a fire (white); reservoirs (blue); forest (green).

For a set of images, augmentation was performed, that is, modification of the data for training. Using this method improves the generalizing ability of the neural network, adds new training examples that the neural network has not yet seen and does not provide an opportunity to retrain. As augmentation, the following modifications were performed: image rotation by an arbitrary degree; compression along the axes; stretching along the axes; mirroring along the axes; Gaussian Blur; change in brightness and contrast. The training included 50 epochs, each of which contains 2000 iterations.

When choosing an algorithm for learning a neural network, the following algorithms were considered: Adam - adaptive moment estimation; Adagrad - adaptive gradient; RMSProp - gradient descent with momentum. During the research, the best results were obtained using the Adam algorithm.

A comparison of the results of the proposed neural network with some analogues is presented. A comparative study of the accuracy of the segmentation algorithms was carried out on a set of reference and test images subjected to noise distortions. To compare the segmentation results, the boundaries of the segmented objects were used, which is a set of points that do not depend on the shading of the segments. To measure the segmentation results, two metrics were used: mean and Hausdorff distance. The study of the quality of work of a number of algorithms showed that they behave unstably when the image is noisy and blurred. Thus, we can conclude that it is advisable to clean the image from noise and increase its clarity before the segmentation procedure.

The accuracy of the developed neural network is 94.22%. For the classes of objects, the accuracy was the following: fire – 93.6%; burnt-out area - 95.7; smoke - 87.6; reservoirs - 96.9; forest - 97.3. This result is the best in comparison with the presented analogs. However, the developed system is somewhat inferior to some analogues in terms of such indicators as fire, burnt out area, smoke. However, in such classes as forests, reservoirs, it wins.

Keywords: neural networks; semantic segmentation; computer vision; image processing; images of the Earth's surface.

VIK Ksenia Vasilievna (Post-graduate Student of National Research Tomsk Polytechnic University, Tomsk, Russian Federation).
E-mail: kvv11@tpu.ru

DRUKI Alexey Alexeevich (Candidate of Technical Sciences, Associate Professor of Department of Information Technologies of National Research Tomsk Polytechnic University, Tomsk, Russian Federation).
Email: druki@tpu.ru

GRIGORIEV Dmitriy Sergeevich (Assistant of National Research Tomsk Polytechnic University, Tomsk, Russian Federation).
E-mail: trygx@tpu.ru

SPITSYN Vladimir Grigorievich (Doctor of Technical Sciences, Professor of the Department of Information Technologies of National Research Tomsk Polytechnic University, Tomsk, Russian Federation).
Email: spvg@tpu.ru

REFERENCES

1. Bundzel, M. & Hashimoto, S. (2010) Object identification in dynamic images based on the memory-prediction theory of brain function. *Journal of Intelligent Learning Systems and Applications*. 2(4). pp. 212–220. DOI: 10.4236/jilsa.2010.24024
2. Tawfiq, A. & Ahmed, J. (2016) Object detection and recognition by using enhanced Speeded Up Robust Feature. *International Journal of Computer Science and Network Security*. 16(4). pp. 66–71.
3. Park, S. & Yoo, J.H. (2013) Realtime face recognition with SIFT based local feature points for mobile devices. *The 1st International Conference of Artificial Intelligence, Modelling and Simulation (AIMS 13)*. pp. 304–308. DOI: 10.1109/AIMS.2013.56
4. Mammeri, A., Boukerche, A. & Khiari, E. (2016) MSER based text detection and communication algorithm for autonomous vehicles. *IEEE Symposium of Computers and Communication*. pp. 456–460. DOI: 10.1109/ISCC.2016.7543902
5. Tore, V. & Chawan, P.M. (2016) FAST Clustering based feature subset selection algorithm for high dimensional data. *International Journal of Computer Science and Mobile Computing*. 5(7). pp. 234–238. DOI: 10.1109/TKDE.2011.181
6. Dalal, N. & Triggs, B. (2005) Histograms of oriented gradients for human detection. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*. 1. pp. 886–893. DOI: 10.1109/CVPR.2005.177
7. Mohey, D.E. (2016) Enhancement bag-of-words model for solving the challenges of sentiment analysis. *International Journal of Advanced Computer Science and Applications*. 7(1). pp. 244–251. DOI: 10.14569/IJACSA.2016.070134
8. Kecman V. & Melki G. (2016) Fast online algorithms for Support Vector Machines. *IEEE South East Conference*. pp. 26–31. DOI: 10.1109/SECON.2016.7506733
9. Le Cun, Y. & Bengio, Y. (1998) Convolutional networks for images, speech and time series. *The Handbook of Brain Theory and Neural Networks*. 7(1). pp. 255–258.
10. TensorFlow: The Python Deep Learning library. [Online] Available from: <https://www.tensorflow.org> (Accessed: 2nd December 2020).
11. Deep Learning Frameworks. (n.d.) A Survey of TensorFlow, Torch, Theano, Caffe, Neon, and the IBM Machine Learning Stack. [Online] Available from: <https://www.microway.com/hpc-tech-tips/deep-learning-frameworks-survey-tensorflow-torch-theano-caffe-neon-ibm-machine-learning-stack> (Accessed: 2nd December 2020).
12. Ronneberger, O., Fischer, P. & Brox, T. (2015) U-Net: Convolutional networks for biomedical image segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI-2015)*. 93. pp. 234–241. DOI: 10.1007/978-3-319-24574-4_28
13. Paszke, A., Chaurasia, A., Kim, S. & Culurciello, E. (2017) ENet: A Deep Neural Network Architecture for Real-Time Semantic Segmentation. *5-th International Conference on Learning Representations*. pp. 1–10.
14. Badrinarayanan, V., Kendall, A. & Cipolla, R. (2017) SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*. 39(12). pp. 54–62. DOI: 10.1109/TPAMI.2016.2644615
15. Sharing Earth Observation Resources. (n.d.) [Online] Available from: <https://directory.eoportal.org/web/eoportal/satellite-missions/l/landsat-8-ldcm> (Accessed: 2nd December 2020).
16. Kingma, D.P. (2015) Adam: a Method for Stochastic Optimization. *International Conference of Learning Representations*. pp. 1–13.
17. Nguyen, V., Kim, H. & Jun, S. (2018) A Study on Real-Time Detection Method of Lane and Vehicle for Lane Change Assistant System Using Vision System on Highway. *Engineering Science and Technology*. 21. pp. 822–833. DOI: 10.1016/j.jestch.2018.06.006
18. El-Khatib, S.A. Image segmentation using a mixed and exponential particle-based algorithm. *Computer Science and Cybernetics*. 1. pp. 126–133.
19. Koltsov, P.P. (2011) The use of metrics in a comparative study of the quality of work of image segmentation algorithms. *Informatics and its Applications*. 5. pp. 53–63.

УДК 004.91

DOI: 10.17223/19988605/55/4

А.В. Воробьев

КОНЦЕПЦИЯ МНОГОУРОВНЕВОЙ СИСТЕМЫ ЦИФРОВЫХ ДВОЙНИКОВ (НА ПРИМЕРЕ ГЕОМАГНИТНЫХ ДАННЫХ)

Работа выполнена при финансовой поддержке РФФИ в рамках научного проекта № 20-07-00011.

Рассматриваются вопросы повышения эффективности пространственной интерполяции данных, описывающих процессы с выраженной пространственной анизотропией. Предложен подход к интерполяции, основанный на применении многоуровневой системы цифровых двойников, представленной совокупностью пространственных кластеров, которые включают физические прототипы соответствующих объектов и их виртуальные модели – цифровые двойники. На примере результатов мониторинга параметров геомагнитного поля показана эффективность применения предложенной концепции.

Ключевые слова: цифровые двойники; обработка данных; корреляционный анализ; фрактальные множества; пространственные кластеры.

В настоящее время, в условиях развития цифровой экономики, одним из магистральных направлений развития информационных технологий является разработка технологий обеспечения физических и технических систем различного уровня сложности их цифровыми двойниками.

Под цифровым двойником [1. С. 79] понимается программный аналог технического или физического объекта, процесса или системы, моделирующий их внутренние характеристики, свойства и поведение в динамически меняющихся условиях окружающей среды.

Основа концепции цифрового двойника независимо от прикладной области его применения представляет собой совокупность трех основных составляющих [2]:

- 1) физический объект-прототип в его реальном окружении;
- 2) виртуальный объект – цифровой двойник – в его виртуальном окружении [3. С. 2051];
- 3) информационные потоки и данные, обеспечивающие взаимодействие физического объекта-прототипа и его цифрового двойника [4. С. 86; 5. С. 1023].

При этом в зависимости от конкретных ситуаций использования в прикладных областях цифровые двойники могут выступать как в роли прототипа цифрового двойника, так и в качестве его цифрового экземпляра [6. С. 66; 7. С. 24]. В первом случае имеет место высокоточная модель, не связанная информационным потоком с физическим объектом или системой [8. С. 154; 9. С. 817; 10; 11]. Во втором случае предполагается описание реального объекта с непрерывной информационной связью с оригинальным объектом [12. С. 10; 13. С. 25].

Одним из перспективных направлений применения цифровых двойников является создание так называемых «виртуальных приборов (устройств)» [1. С. 79; 8. С. 154; 14. С. 1; 15. С. 109], целью которых являются непрерывная регистрация и мониторинг параметров объектов и систем различной природы происхождения в тех локациях, которые физически или по каким-либо иным причинам недоступны для размещения физических информационно-измерительных устройств и систем.

В этой связи указанное приложение цифровых двойников может являться эффективным решением проблемы пространственной интерполяции данных, имеющих пространственно-временную привязку и характеризующихся выраженной пространственной анизотропией описываемых ими процессов или явлений.

Примером тому являются неоднородные пространственно-временные геомагнитные данные, непрерывно регистрируемые сетью наземных магнитных обсерваторий и вариационных станций, а также искусственными спутниками в околоземном пространстве. Высокие требования к техническому и программному оснащению магнитных обсерваторий зачастую делают невозможным размещение магнитометрического оборудования в некоторых географических локациях, например в высокоширотных областях. Вместе с тем именно в таких регионах мониторинг геомагнитных параметров имеет принципиальное значение.

Известно, что задача пространственной интерполяции в настоящее время решается посредством применения широкого спектра геоинформационных технологий: крикинг, сплайн, обратновзвешенные расстояния, метод ближайшего соседа – основные методы, доступные сегодня как программно, так и с использованием инструментальных средств [16, 17]. Вместе с тем ни один из известных подходов не учитывает особенности пространственной анизотропии интерполируемых данных и принимает одни и те же модели и методы для различных географических областей. В этой связи имеет место погрешность, величина которой для многих задач неприемлема [2]. К примеру, для геомагнитных данных анализ показал, что применение известных методов пространственной интерполяции сопровождается среднеквадратической ошибкой в среднем от 7,3 до 11,2 нТл при допустимой стандартами погрешности в 1 нТл. В результате интерполированные данные могут повлечь за собой серьезные ошибки в моделировании и принятии решений, что, в свою очередь, сопряжено с финансовыми и временными затратами на устранение их последствий (например, при инклинометрических исследованиях с применением магниточувствительного оборудования).

В этой связи применение в обозначенных и им подобных локациях цифровых двойников технических объектов и систем позволит в значительной мере решить проблему их пространственной анизотропии, обеспечивая доступ к геомагнитной информации в том числе и в тех пространственно-временных координатах, для которых не предусмотрен или недоступен физический прототип.

1. Концепция многоуровневой системы цифровых двойников

Проблема пространственной анизотропии источников данных с точки зрения теории цифровых двойников имеет двойственную природу. С одной стороны, выраженный анизотропный характер пространственного распределения параметров геомагнитного поля и его вариаций в земном и околоземном пространстве объясняет тот факт, что для цифрового двойника не всегда гарантированно может быть найден и идентифицирован его физический прототип [4]. С другой стороны, физический прототип в совокупности со своим цифровым двойником не обеспечивают полного покрытия земного и околоземного пространства, что также усугубляется тем, что их работоспособное состояние в любой момент может быть нарушено в силу различных причин природного и техногенного характера, а также под воздействием человеческого фактора.

Для решения обозначенной проблемы представляется целесообразным представить цифровые двойники и их физические прототипы в виде многоуровневой системы, в которой каждый уровень образует пространственный кластер. Отличительной особенностью решения является то, что группа «цифровой двойник – физический прототип» представляет собой фрактал, многократно повторяемый как в рамках одного уровня, так и между несколькими уровнями системы. Тем самым на отличных от базового уровнях абстракции каждый цифровой двойник и его физический прототип также задаются парой «цифровой двойник – физический прототип». Взаимосвязь между парами осуществляется посредством организации дуплексных информационных потоков, имеющих место как локально (на одном уровне системы), так и глобально (между разными уровнями). В результате формируется многоуровневая система, где на каждом пространственном кластере задается собственная подсистема вида «цифровой двойник – физический прототип».

Представляется важным отметить, что агрегирование информационных потоков, поступающих из физических прототипов и их цифровых двойников, осуществляется на уровне одного родительского для указанных элементов пространственного кластера. К примеру, множество пар «цифровой

двойник – физический прототип» обмениваются друг с другом информацией о своем функционировании без дополнительного промежуточного сохранения, используя локальные средства и системы хранения данных. При этом совокупность всех одноуровневых пар передает информацию о значениях параметров в централизованное хранилище данных. В дальнейшем взаимодействие разноуровневых информационных потоков будет реализовано связями, организованными между хранилищами и базами данных. Интеграция данных, предполагающая сведение гетерогенных моделей хранимых данных в унифицированную форму (модель), может быть реализована практикуемыми в настоящее время при создании единого информационного пространства медиаторами и интерпретаторами.

Необходимо отметить, что в паре «физический прототип – цифровой двойник» любой из компонентов может отсутствовать. В этом случае создание виртуальной модели для цифрового двойника при отсутствии его физического прототипа осуществляется методами машинного обучения на основании доступных физических объектов, которые в этом случае могут выступать прототипами нескольких пар «физический прототип – цифровой двойник». Аналогичные рассуждения справедливы и для отсутствующих цифровых двойников: физический объект из пары может выступать в качестве прототипа для другой пары, но только того же уровня системы.

Структура многоуровневой системы цифровых двойников представляет собой объект некоторой размерности, обладающий свойством масштабной инвариантности. В общем виде при схематическом представлении система цифровых двойников, согласно предложенной концепции, представляет собой геометрический фрактал.

Применительно к техническим устройствам применение многоуровневой системы цифровых двойников обеспечивает качественно новый переход от аппаратных технических систем к программным системам, опирающимся в своем функционировании исключительно на возможности вычислительных машин.

2. Математическая модель многоуровневой системы цифровых двойников

Представляется целесообразным формализовать описание многоуровневой системы цифровых двойников в теоретико-множественном базисе. В качестве единицы моделирования в данном случае выступает уровень системы цифровых двойников, обозначенный как DW_n , что соответствует подсистеме цифровых двойников DW уровня n . Важно отметить, что для заданного уровня n допустимо соотношение

$$DW = \{DW_1, DW_2, \dots, DW_n\}, n \in N,$$

где DW_n – l -я подсистема множества подсистем уровня n .

Порядок обозначения смежных уровней подсистем может быть представлен последовательным присвоением соответствующих индексов. Каждая подсистема уровня n представляет собой совокупность элементов различного типа (как атомарных, так и составных). Базовым предлагается считать уровень с индексом $n = 1$, представленный совокупностью пар объектов вида «прототип – цифровой двойник», взаимодействие которых реализуется посредством информационных потоков:

$$dw_i^n = \{P_k, D_k\}_{k=1}^m : P_k \neq \emptyset, D_k \neq \emptyset,$$

где P, D – соответственно прототип и цифровой двойник объекта, k – индекс объекта, m – общее число объектов на уровне n , dw_i^n – i -я подсистема множества подсистем уровня n .

Соотношения между информационными потоками, представленными в качестве выходных для прототипа и его цифрового двойника, можно представить в виде соответствия между элементами соответствующих множеств. Пусть I_P – выходной информационный поток прототипа, представленный, к примеру, в виде набора временных рядов. Тогда I_D – соответствующий ему выходной информационный поток его цифрового двойника, представленного в той же подсистеме.

В этом случае имеет место соответствие между элементами множеств I_P и I_D , представленное в виде соотношения эквивалентности вида

$$\rho \subseteq I_P \times I_D,$$

где ρ – отношение эквивалентности, основанное на декартовом произведении соответствующих множеств (в данном случае I_P и I_D).

Указанное отношение по определению является рефлексивным, симметричным и транзитивным, т.е. имеет место система соотношения:

$$\begin{aligned} a \rho b, \forall a \in I_P, b \in I_D, \\ a \rho b \Rightarrow b \rho a, \forall a \in I_P, b \in I_D, \\ a \rho b \&\&b \rho c \Rightarrow a \rho c, \forall a \in I_P, b \in I_D. \end{aligned}$$

Очевидно, что I_P делит I_P . Если I_P делит I_D и I_D делит I_P , то они совпадают. И, наконец, если I_P делит I_D и I_D делит произвольно взятое I_X , то I_P делит I_X . Поэтому выполняются соответственно свойства рефлексивности, антисимметричности и транзитивности для данного отношения. Иными словами, имеет место соотношение следующего порядка:

$$I_P \leq I_D : a \leq b \parallel b \leq a, \forall a \in I_P, b \in I_D,$$

где $\leq$ – отношение порядка для множеств I_P и I_D , a, b – отдельные компоненты соответствующих кортежей данных для множеств I_P и I_D соответственно.

С учетом выделенных отношений справедливы следующие замечания. Пусть $(I_P, \leq)$ – отношение порядка на множестве I_P . Тогда для любого $B \subseteq I_P$ наименьшим в B будет такое $a_0 \in B$, что для любых $b \in B$ выполняется соотношение вида $a_0 \leq b$. И наоборот, для любого $B \subseteq I_P$ наибольшим в B будет такое $a_0 \in B$, что для любых $b \in B$ выполняется соотношение вида $a_0 \geq b$.

Тогда предполагается, что отношение линейного порядка на множестве I_P или I_D называется отношением полного порядка, если в любом его непустом подмножестве $B \subseteq I_P$ (или $B \subseteq I_D$) найдется минимальный элемент. Соответственно, согласно известным положениям теории множеств, в том случае, если множества I_P и I_D содержат минимальные элементы, то они, в свою очередь, являются вполне упорядоченными множествами с полным порядком на отношении полного порядка.

Каждая подсистема многоуровневой системы цифровых двойников демонстрирует поведение и обладает специфическими свойствами, характерными для фрактальных множеств. В общем виде пара «прототип – цифровой двойник» повторяется на каждом уровне системы. Так, пусть уровень с индексом n представлен подсистемой цифровых двойников вида, содержащей m пар «прототип – цифровой двойник» для объектов и процессов любой природы происхождения:

$$dw^n = \{P_k, D_k\}_{k=1}^m,$$

где P, D – соответственно прототип и цифровой двойник объекта, k – индекс объекта, m – общее число объектов на уровне n , dw^n – подсистема уровня n .

Соответственно, следующий – $(n + 1)$ -й – уровень системы цифровых двойников представляет собой очередную связку вида «прототип – цифровой двойник». При этом подсистема предыдущего – n -го – уровня представляет собой один из компонентов этой связки, выступая в качестве прототипа или его цифрового двойника. Иными словами, имеет место соотношение следующего вида:

$$dw^{n+1} = \{P_k, D_k\}_{k=1}^p : P_k = dw_i^n, D_k = dw_{i+1}^n,$$

где P, D – соответственно прототип и цифровой двойник объекта, k – индекс объекта, p – общее число объектов на уровне $(n + 1)$, dw^n – подсистема уровня n , dw^{n+1} – подсистема уровня $(n + 1)$.

Аналогичные рассуждения справедливы для всех уровней, предусмотренных в соответствующей системе цифровых двойников. В результате структура многоуровневой системы цифровых двойников представляет собой объект некоторой размерности n , обладающий свойством масштабной инвариантности. В общем виде при схематическом представлении система цифровых двойников, согласно предложенной концепции, представляет собой геометрический фрактал.

Тогда, в соответствии с теоремой о самоподобных множествах, для системы цифровых двойников DW существуют такие преобразования подобия $S_1, S_2, \dots, S_n$, что:

$$DW = S_1(DW) \cup S_2(DW) \cup \dots \cup S_n(DW).$$

Еще одно свойство, унаследованное системой цифровых двойников из теории фракталов, заключается в том, что для $DW S_i(DW)$ попарно не пересекаются. Размерность Минковского d для системы DW цифровых двойников определяется, согласно следующему выражению:

$$Nr^d = 1.$$

Иными словами, размерность подобия системы цифровых двойников DW как самоповторяющегося множества представляет собой решение следующего уравнения:

$$r_1^d + r_2^d + \dots + r_n^d = 1,$$

где r_i^d – i -й коэффициент подобия.

Далее, снова применяя к множеству цифровых двойников теорию самоповторяющихся множеств, представляется целесообразным отметить равенство нулю d -меры Хаусдорфа всех попарных пересечений составляющих его подмножеств.

3. Пример многоуровневой системы цифровых двойников

Рассмотрим в качестве примера систему цифровых двойников, представленную множествами прототипов технических объектов регистрации параметров магнитного поля Земли и их цифровых двойников (рис. 1) [19. С. 648; 20. С. 23]. На базовом уровне подсистема представлена совокупностью пар «прототип – цифровой двойник» для наземных магнитных обсерваторий. Каждый компонент вектора магнитного поля представляет собой связку прототипа объекта и его цифрового двойника.

Соответственно, имеет место следующее соотношение:

$$dw_i^d = \{MC_X_i, MC_Y_i, MC_Z_i, MC_F_i\}, i = 1, \dots, n,$$

$$MC_X_i = \{ЦД_X, Прототип_X\}, MC_Y_i = \{ЦД_Y, Прототип_Y\},$$

$$MC_Z_i = \{ЦД_Z, Прототип_Z\}, MC_F_i = \{ЦД_F, Прототип_F\},$$

где $MC_X_i, MC_Y_i, MC_Z_i, MC_F_i$ – объекты, соответствующие X, Y и Z -компонентам полного вектора геомагнитного поля в точке i , Прототип_ X , Прототип_ Y , Прототип_ Z , Прототип_ F – показатели на магнитометре значений X, Y, Z -компонент полного вектора поля в точке i , ЦД_ X , ЦД_ Y , ЦД_ Z , ЦД_ F – цифровые двойники показателя на магнитометре значений X, Y, Z -компонент полного вектора геомагнитного поля в точке i .

Следующий уровень сформирован магнитной обсерваторией, осуществляющей регистрацию параметров геомагнитного поля и его вариаций. Каждая магнитная обсерватория объединяет представленные выше связки «прототип – цифровой двойник» всех перечисленных компонент.

В свою очередь, обсерватории вместе со своими цифровыми двойниками объединяются в пары «прототип – цифровой двойник», и далее формируется новый уровень с индексом $n = 2$:

$$dw_i^2 = \{dw_i^1\}_{i=1}^n,$$

где dw_i^2 – магнитная обсерватория или станция (прототип или цифровой двойник соответствующего технического объекта).

Магнитные обсерватории и их цифровые двойники объединяются по пространственным характеристикам и образуют пространственные кластеры, в совокупности составляющие следующий уровень системы цифровых двойников с индексом $n = 3$:

$$dw_i^3 = \{dw_i^2\}_{i=1}^m,$$

где dw_i^3 – множество магнитных обсерваторий или станций (прототипов или цифровых двойников соответствующего технического объекта), составляющих один пространственный кластер.

Указанные пространственные кластеры подсистемы уровня $n = 3$ объединяются далее в подсистему с индексом $n = 4$ – наземный уровень:

$$dw_i^4 = \{dw_i^3\}_{i=1}^p,$$

где dw_i^4 – множество пространственных кластеров.

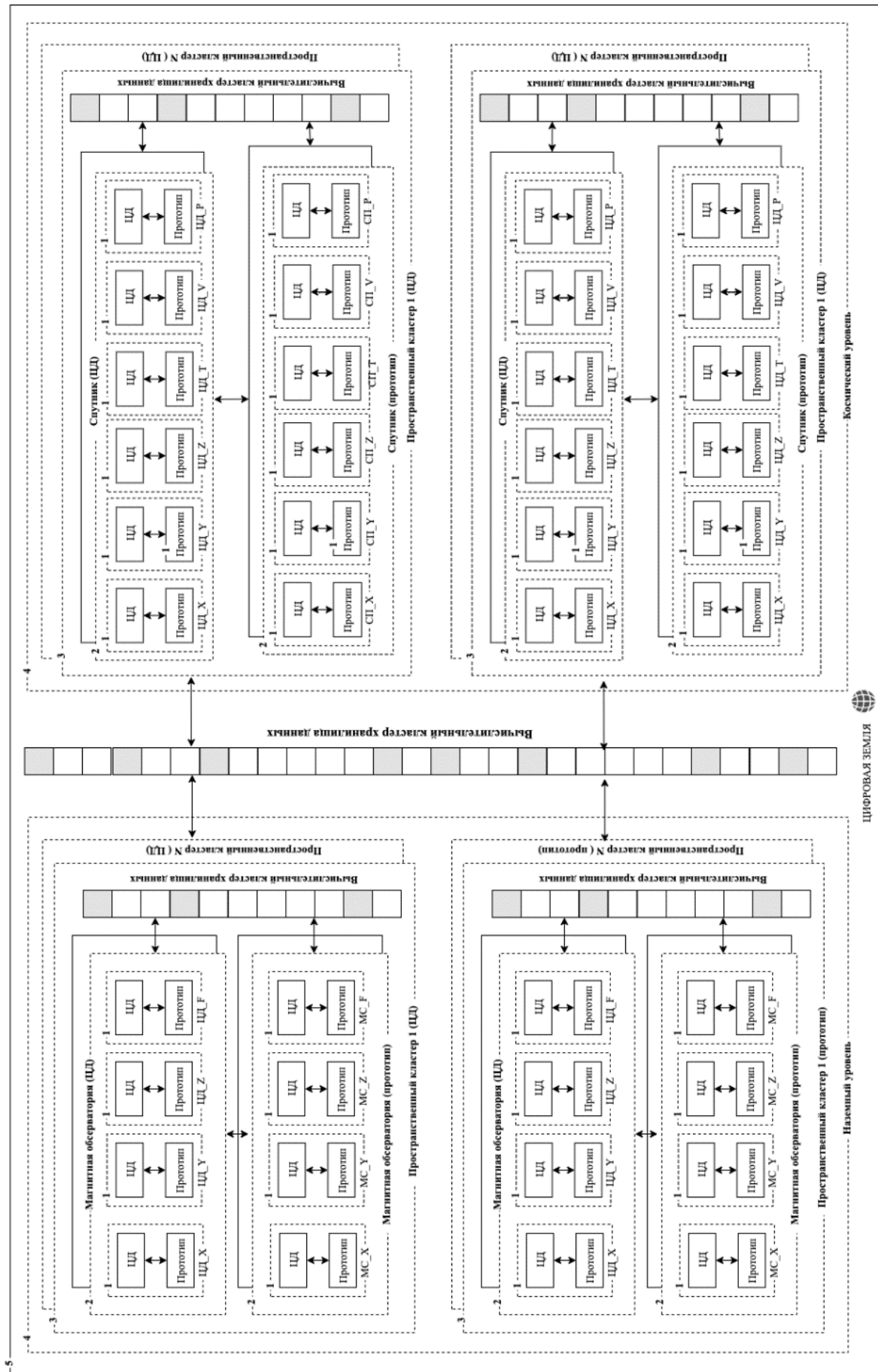


Рис. 1. Пример схемы многоуровневой системы цифровых двойников «Цифровая Земля»

Fig. 1. An example of multilevel digital twins system "Digital Earth"

Аналогичным образом формируется одноуровневый с ним космический уровень, интегрирующий информационные потоки, получаемые с искусственных спутников Земли, осуществляющих мониторинг геомагнитного поля. Если рассматривать декомпозицию подсистем в обратном направлении, то космический уровень декомпозируется на спутники (уровень $n = 3$).

Для оценки эффективности решения был проведен ряд вычислительных экспериментов по пространственной интерполяции геомагнитной информации по данным проекта SuperMAG (<https://supermag.jhuapl.edu/mag/>). Результаты показали, что применение системы цифровых двойников как виртуальных моделей магнитных обсерваторий обеспечивает среднеквадратическую ошибку от 0,53 до 1,02 нТл, что укладывается в регламентируемый стандартами диапазон.

Заключение

В результате проведенных исследований предложено решение проблемы пространственной интерполяции данных с выраженной пространственной анизотропией на основе концепции многоуровневой системы цифровых двойников. Последняя отличается тем, что для повышения эффективности и надежности процессов обработки и моделирования данных создаются пары вида «цифровой двойник – прототип», многократно повторяемые и на разных уровнях интегрируемые в отдельные пространственные кластеры, иерархически связанные друг с другом. Исследования показали значительное уменьшение величины среднеквадратической ошибки пространственной интерполяции при использовании указанного подхода на примере геомагнитных данных.

ЛИТЕРАТУРА

1. Аббасова Т.С. Развитие виртуальных инструментов для создания цифровых двойников // Информационно-технологический вестник. 2019. № 2 (20). С. 79–88.
2. Stark R., Freseman C., Lindow K. Development and operation of Digital Twins for technical systems and services // CIRP Annals. 2019. № 68. P. 129–132.
3. Alam K.M., El Saddik A. C2PS: a Digital Twin Architecture Reference Model for the Cloud-Based Cyber-Physical Systems // IEEE Access. 2017. P. 2050–2062.
4. Grieves M., Vickers J. Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems // Transdisciplinary Perspectives on Complex Systems: New Findings and Approaches / F.-J. Kahlen, S. Flumerfelt, A. Alves (eds.), Cham : Springer International Publishing, 2017. P. 85–113.
5. Bakhvalov Yu., Kopylov I. Training and assessment the generalization ability of interpolation methods // Computer Research and Modeling. 2015. № 7. P. 1023–1031. DOI: 10.20537/2076-7633-2015-7-5-1023-1031.
6. Ковалёв С.П. Проектирование информационного обеспечения цифровых двойников энергетических систем // Системы и средства информ. 2020. № 30:1. С. 66–81
7. Дмитриев В.М., Ганджа Т.В. Принцип формирования многоуровневых компьютерных моделей SCADA-систем для управления сложными технологическими объектами // Информатика и системы управления. 2013. № 2 (36). С. 24–36.
8. Дмитриев В.М., Ганджа Т.В., Панов С.А. Система виртуальных инструментов и приборов для автоматизации учебных и научных экспериментов // Программные продукты и системы. 2016. № 3 (29). С. 154–161.
9. Batty M. Digital twins // Environment and Planning B: Urban Analytics and City Science. 2018. № 45 (5). P. 817–820.
10. Boschert S., Rosen R. Digital Twin – The Simulation Aspect // Mechatronic Futures. Springer, 2016. P. 59–74. DOI: 10.1007/978-3-319-32156-1_5.
11. Cimino C., Negri E., Fumagalli L. Review of digital twin applications in manufacturing // Computers in Industry. 2019. № 113. P. 103–130. DOI: 10.1016/j.compind.2019.103130.
12. Habe H., Li B., Weissenberg N., Cirullies J., Otto B. Digital twin for real-time data processing in logistics // Hamburg International Conference of Logistics (HICL). 2019. № 27. P. 3–28. DOI: 10.15480/882.2462.
13. Josifovska K., Yigitbas E., Engels G. Reference Framework for Digital Twins within Cyber-Physical Systems // IEEE / ACM 5th International Workshop on Software Engineering for Smart Cyber-Physical Systems (SEsCPS). 2019. P. 25–31. DOI: 10.1109/SEsCPS.2019.00012.
14. Halenar I., Juhas M., Juhasova B., Borkin D. Virtualization of Production Using Digital Twin Technology // 20th International Carpathian Control Conference (ICCC), Krakow–Wieliczka, Poland, 2019. P. 1–5. DOI: 10.1109/CarpathianCC.2019.8765940.
15. Otto B., Hompel M., Wrobel S. International Data Spaces: Reference architecture for the digitization of industries // Digital Transformation. Springer, 2019. P. 109–128. DOI: 10.1007/978-3-662-58134-6_8.
16. Aliyev Z. Spatial data interpolation // International Journal of Medical and Biomedical Studies. 2018. № 2. P. 7–10. DOI: 10.32553/ijmbs.v2i4.36.

17. Albarghouthi A., Berdine J., Cook B., Kincaid Z. Spatial Interpolants // ESOP 2015: Programming Languages and Systems. 2015. P. 634–660. DOI: 10.1007/978-3-662-46669-8_26.
18. Negri E., Fumagalli L., Macchi M. A review of the roles of Digital Twin in CPS-based production // Procedia Manufacturing. 2017. V. 11. P. 939–948.
19. Воробьев А.В., Воробьева Г.Р. Подход к оценке относительной информационной эффективности магнитных обсерваторий сети INTERMAGNET // Геомагнетизм и аэрономия. 2018. № 58. С. 648–652. DOI: 10.1134/S0016794018050164.
20. Воробьева Г.Р., Воробьев А.В. Подход к повышению производительности программных процессов обработки и хранения больших объемов геомагнитных данных // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2020. № 50. С. 23–30.

Поступила в редакцию 22 декабря 2020 г.

Vorobev A.V. (2021) THE CONCEPT OF A MULTI-LEVEL SYSTEM OF DIGITAL TWINS (ON THE EXAMPLE OF GEOMAGNETIC DATA). *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika* [Tomsk State University Journal of Control and Computer Science]. 55. pp. 26–34

DOI: 10.17223/19988605/55/4

The problem of spatial interpolation of data with the required accuracy is especially relevant for processes and phenomena of natural origin, which are characterized by pronounced spatial anisotropy of information. An example of this is geomagnetic data as a result of observing the parameters of the Earth's magnetic field and its variations. Technical means of recording such data, as a rule, are limited by the possibilities of physical placement due to the high requirements for the hardware and software necessary for information and measurement monitoring, as well as the associated economic and time costs. The impossibility of placing monitoring technical means in certain geographic locations increases the information entropy of the processes that take place there and are relevant for study and observation.

The solution to this problem is partly possible by using modern geoinformation technologies, the algorithms of which allow interpolating spatial data by known values. However, the accuracy of spatial interpolation is not always satisfactory: in particular, the restoration of geomagnetic data in this way provides a root mean square error of 7.3 to 11.2 nT for various spatial regions, which significantly exceeds the error of 1 nT permissible by the standards.

In this regard, in order to solve the indicated problem, the concept of a multi-level system of digital twins of technical objects is proposed. The concept is based on the idea of multiple repetition of pairs "digital twin - physical prototype" to form a hierarchical multilevel structure of fractal type. Duplex inter-level information interaction is carried out through the use of specialized data stores that aggregate information coming from the components of the level (while the simultaneous presence of both elements of the pair is optional).

The physical prototypes are both directly technical objects and pairs of the form "digital twin - physical prototype", as a result of which the structure formed in this case functions according to the principles of information backup. The sequential creation of the levels of the system of digital twins allows to significantly increase the coverage area of the earth and near-earth surfaces by physical and virtual technical means, the functioning of which is modeled using machine learning methods that form interpolated values based on retrospective analysis of known data.

It is expected that for geomagnetic data, the use of a multilevel system of digital twins will allow collecting information on the state of the geomagnetic field and its variations in hard-to-reach geographic locations, for example, in the Arctic zone, where the placement of magnetometric equipment is extremely difficult. Studies have shown that the application of the proposed concept for solving the problem of spatial interpolation of geomagnetic data provides a root-mean-square error from 0.53 to 1.02 nT, which falls within the range regulated by the standards.

Keywords: digital twins; data processing; correlation analysis; fractal sets; spatial clusters.

VOROBIEV Andrei Vladimirovich (Candidate of Technical Sciences, Associate Professor, Ufa State Aviation Technical University, Ufa, Russian Federation).

E-mail: geomagnet@list.ru

REFERENCES

1. Abbasova, T.S. (2019) The development of virtual tools for creating digital doubles. *Informatsionno-tehnologicheskij vestnik – Information Technology Bulletin*. 2(20). pp. 79–88.
2. Stark, R., Freseman, C. & Lindow, K. (2019) Development and operation of Digital Twins for technical systems and services. *CIRP Annals*. 68. pp. 129–132. DOI: 10.1016/j.cirp.2019.04.024
3. Alam, K.M. & El Saddik, A. (2017) C2PS: A Digital Twin Architecture Reference Model for the Cloud-Based Cyber-Physical Systems. *IEEE Access*. pp. 2050–2062. DOI: 10.1109/ACCESS

4. Grieves, M. & Vickers, J. (2017) Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems. In: Kahlen, F.-J., Flumerfelt, S. & Alves, A. (eds) *Transdisciplinary Perspectives on Complex Systems: New Findings and Approaches*. Cham, Switzerland: Springer International Publishing. pp. 85–113.
5. Bakhvalov, Yu. & Kopylov, I. (2015) Training and assessment the generalization ability of interpolation methods. *Computer Research and Modeling*. 7. pp. 1023–1031. DOI: 10.20537/2076-7633-2015-7-5-1023-1031
6. Kovalev, S.P. (2020) Information architecture of the power system digital twin. *Sistemy i sredstva inform – Systems and Means of Informatics*. 30(1). pp. 66–81. DOI: 10.14357/08696527200106
7. Dmitriev, V.M. & Gandzha, T.V. (2013) Principle of formation of multilevel computer models of SCADA-systems for the control of complex technological objects. *Informatika i sistemy upravleniya*. 2(36). pp. 24–36.
8. Dmitriev, V.M., Gandzha, T.V. & Panov, S.A. (2016) VIP. The system of virtual instruments and devices for education and scientific experiment automation. *Programmnye produkty i sistemy – Software and Systems*. 3(29). pp. 154–161. DOI: 10.15827/0236-235X.115.154-162
9. Batty, M. (2018) Digital twins. *Environment and Planning B. Urban Analytics and City Science*. 45(5). pp. 817–820. DOI: 10.1177/2399808318796416
10. Boschert, S. & Rosen, R. (2016) Digital Twin–The Simulation Aspect. *Mechatronic Futures*. pp. 59–74. DOI: 10.1007/978-3-319-32156-1_5
11. Cimino, C., Negri, E. & Fumagalli, L. (2019) Review of digital twin applications in manufacturing. *Computers in Industry*. 113. pp. 103–130. DOI: 10.1016/j.compind.2019.103130
12. Habe, H., Li, B., Weissenberg, N., Cirullies, J. & Otto, B. (2019) Digital twin for real-time data processing in logistics. *Hamburg International Conference of Logistics (HICL)*. 27. pp. 3–28. DOI: 10.15480/882.2462
13. Josifovska, K., Yigitbas, E. & Engels, G. (2019) Reference Framework for Digital Twins within Cyber-Physical Systems. *IEEE/ACM 5th International Workshop on Software Engineering for Smart Cyber-Physical Systems (SEsCPS)*. pp. 25–31. DOI: 10.1109/SEsCPS.2019.00012
14. Halenar, I., Juhas, M., Juhasova, B. & Borkin D. (2019) Virtualization of Production Using Digital Twin Technology. *20th International Carpathian Control Conference (ICCC)*. Krakow-Wieliczka, Poland. pp. 1–5. DOI: 10.1109/CarpathianCC.2019.8765940
15. Otto, B., Hompel, M. & Wrobel, S. (2019) International Data Spaces: Reference architecture for the digitization of industries. *Digital Transformation*. pp. 109–128. DOI: 10.1007/978-3-662-58134-6_8
16. Aliyev, Z. (2018) Spatial data interpolation. *International Journal of Medical and Biomedical Studies*. 2. pp. 7–10. DOI: 10.32553/ijmbs.v2i4.36
17. Albarghouthi, A., Berdine, J., Cook, B. & Kincaid, Z. (2015) Spatial Interpolants. *ESOP 2015: Programming Languages and Systems*. pp. 634–660. DOI: 10.1007/978-3-662-46669-8_26
18. Negri, E., Fumagalli, L. & Macchi, M. (2017) A review of the roles of Digital Twin in CPS-based production. *Procedia Manufacturing*. 11. pp. 939–948.
19. Vorobev, A.V. & Vorobeve, G.R. (2018) Podkhod k otsenke otnositel'noy informatsionnoy effektivnosti magnitnykh observatoriy seti INTERMAGNET [An approach to assessing the relative information efficiency of magnetic observatories in the INTERMAGNET network]. *Geomagnetizm i aeronomiya – Geomagnetism and aeronomy*. 58. pp. 648–652.
20. Vorobeve, G.R. & Vorobev, A.V. (2020) Approach to improving the performance of software processes for processing and storing large volumes of geomagnetic data. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 50. pp. 23–30. DOI: 10.17223/19988605/50/3

УДК 519.2

DOI: 10.17223/19988605/55/5

Ж.Н. Зенкова, У. Мусони, А.А. Андриевская**СТАТИСТИЧЕСКОЕ ОЦЕНИВАНИЕ ПОКАЗАТЕЛЕЙ ОБОРАЧИВАЕМОСТИ
ОБОРОТНЫХ СРЕДСТВ С УЧЕТОМ ДОПОЛНИТЕЛЬНОЙ ИНФОРМАЦИИ
О КВАНТИЛЕ**

Предложен более точный подход к расчету показателей оборачиваемости оборотных средств предприятия с использованием дополнительной информации об известном квантиле заданного уровня функции распределения стоимости оборотных средств, основанный на модифицированной оценке математического ожидания, которая является несмещенной, асимптотически нормальной с дисперсией, в асимптотике не превосходящей дисперсию обычного выборочного среднего. С помощью имитационного моделирования исследовано поведение модифицированного среднего с учетом квантиля для малых объемов выборок, для ряда распределений показана достаточно большая скорость сходимости к асимптотическому результату; выявлено, что при малых объемах выборок и при уровне квантиля, близком к единице, а при нормальном распределении и близком к нулю, модифицированную оценку использовать не следует. Методика апробировалась на реальных данных. С помощью бутстреп-моделирования получены доверительные интервалы для средних значений показателей оборачиваемости.

Ключевые слова: дополнительная информация; квантиль; среднее с учетом квантиля; оборачиваемость.

Оборотные средства (ОС) любого производственного и торгового предприятия являются базовым источником его прибыли [1], именно благодаря успешному управлению ими можно существенно улучшить финансовые результаты предприятия [2], его рыночную стоимость [3], возврат инвестиций [4], сократить объемы складских запасов и помещений, а следовательно, себестоимость и конечные цены. Все это в значительной степени обуславливает важность максимально точного и адекватного расчета показателей оборачиваемости (ПО).

Существует ряд работ, в которых классическая система расчета ПО подвергается критике и модификации. В [5] авторы отмечают, что традиционный расчет коэффициента оборачиваемости (КО), не способен учесть колебаний их стоимости в течение рассматриваемого периода, и предлагают свой метод расчета на основе всех наблюдаемых значений ряда. В работах [6, 7] показано негативное влияние краткосрочных аномальных всплесков или падений спроса, которые рассматриваются как выбросы, для их ликвидации и получения более адекватной оценки использовалась урезанная оценка среднего. В [8] для анализа ПО привлекалась оценка среднего по интервальным данным с применением непараметрического алгоритма Тёрнбулла, что позволило максимально полно учесть всегда существующую на практике нестабильность текущего уровня ОС, например в течение месяца, когда фактически их уровень меняется в некоторых случайных пределах, а зафиксированные на конец месяца значения стоимости, используемые при подсчете классических ПО, никак не отражают внутримесячных колебаний. В [9] при расчете ПО учитывалась сезонность спроса на товары. В [10] рассмотрен способ расчета ПО по цензурированным справа данным, которые возникают в случае дефицита товара, когда все имеющиеся в наличии запасы одного или нескольких товаров были полностью распроданы и потребительский спрос не был удовлетворен в полной мере. Оценивание среднего уровня запаса осуществлялось с помощью непараметрической оценки Каплана–Мейера методом подстановки.

В работе предложен более точный способ расчета КО и оборота (О) с учетом дополнительной информации об известном квантиле заданного уровня функции распределения (ф.р.) стоимости оборотных средств с помощью модифицированной оценки среднего [11], которая является несмещенной,

асимптотически нормальной с дисперсией, которая в асимптотике не превосходит дисперсию обычного выборочного среднего. С помощью имитационного моделирования исследована точность оценки для малых объемов выборок, показано, что для ряда распределений скорость сходимости среднеквадратической ошибки MSE (Mean Squared Error) к асимптотическому значению нормированной на объем выборки дисперсии достаточно высокая, однако при очень малых объемах выборок и значениях уровня квантиля, близких к единице, а в нормальном случае также и близких к нулю, модификацию лучше не использовать. Апробация модифицированной методики расчета осуществлялась на реальных данных о стоимости остатков запасов производственного предприятия.

1. Показатели оборачиваемости запасов

Классически показатели оборачиваемости рассчитываются следующим образом:

1) коэффициент оборачиваемости:

$$KO = \frac{TR}{\bar{X}}, \quad (1)$$

где TR – суммарный объем выручки за год (руб./год), $\bar{X}$ – средняя стоимость вложений в оборотные фонды (руб.). Коэффициент KO показывает, сколько раз в течение года возвращались вложенные в оборотные фонды инвестиции,

2) оборот

$$O = \frac{365}{KO} \quad (2)$$

позволяет определить, на сколько дней были заморожены вложения в размере $\bar{X}$, при этом именно значение $\bar{X}$ является для менеджеров предприятия основным ориентиром при планировании размеров кредитования или факторинга на период не менее чем O . Данные показатели анализируются в динамике и в сравнении, например, со «средним» или эталонным показателем по отрасли. Задача каждого предприятия – повысить значение KO и снизить O .

Заметим, что от того, каким образом будет рассчитан $\bar{X}$, во многом зависит конечный результат анализа оборачиваемости. В большинстве источников средний уровень стоимости ОС за период определяется как полусумма их стоимостей на начало и на конец рассматриваемого периода. В [5, 10] данный подход подвергся критике за очевидную неадекватность, так как он дает приемлемый результат, только если уровень стоимости ОС существенно не меняется в течение года, что практически не наблюдается на реально работающем предприятии, кроме, возможно, случаев, когда спрос на все товары предприятия достаточно стабилен, поставки осуществляются регулярно и часто, успешно внедрена система «точно в срок» и / или VMS (Vendor Management System) – система управления запасов поставщиком, что требует многолетней работы и тесного контакта с поставщиками [12]. Для устранения недостатков классической системы Жу Ку и Бинг Жао [5] предложили для оценивания $\bar{X}$ новую методику расчета, которая фактически сводится к вычислению обычного арифметического среднего значений ряда динамики, наблюдаемых не только в конце и начале, но и внутри рассматриваемого периода (года).

2. Показатели оборачиваемости с учетом квантиля

Рассмотрим $X_1, X_2, \dots, X_N$ – стоимость ОС на начало i -го периода, N – количество рассматриваемых периодов в течение года (дни, недели, месяцы и пр.). Допустим, что это независимые, одинаково распределенные случайные величины с ф.р. $F(x)$. Тогда в качестве оценки средней стоимости ОС можно рассмотреть выборочное среднее

$$\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i. \quad (3)$$

Известно, что дополнительная информация позволяет сделать более точные расчеты статистических показателей, в том числе и среднего значения [11, 13–17]. В [11, 17] рассмотрена оценка

$$\bar{X}^q = \frac{1}{N(N-1)} \cdot \sum_{i=1}^N \sum_{j=1, i \neq j}^N X_i \cdot \left(1 - \frac{(I_{(X_i < x_q)} - q) \cdot (I_{(X_j < x_q)} - q)}{q(1-q)} \right), \quad (4)$$

позволяющая учесть информацию об известном квантиле ф.р. $F(x)$ (ниже x_q – известный квантиль заданного уровня q):

$$F(x_q) = q. \quad (5)$$

Заметим, что для задачи оценивания ПО дополнительная информация вида (5) может быть легко получена и иметь следующий вид: аналитикам предприятия известно, что, например, в течение года в 70% случаев стоимость вложений в оборотные средства на предприятии была меньше одного миллиона рублей (т.е. $x_q = 1$ млн руб., $q = 0,7$).

Оценка (4) является несмещенной, асимптотически нормальной с асимптотически нормированной на объем наблюдений дисперсией [11, 17]

$$\sigma_q^2 = \lim_{N \rightarrow \infty} N \cdot D\bar{X}^q = \sigma^2 - \frac{1}{q^2(1-q)^2} \left(aq - \frac{1}{q} \int_{-\infty}^{x_q} x dF(x) \right)^2, \quad (6)$$

где $\sigma^2 = N \cdot D\bar{X}$, a – истинное значение среднего. Из (6) очевидно, что для достаточно больших объемов наблюдений N $\sigma_q^2 \leq \sigma^2$, следовательно, в силу несмещенности получим снижение нормированной на N среднеквадратической ошибки $N \cdot \text{MSE}\{\bar{X}^q\} = N \cdot E(\bar{X}^q - a)^2$, что говорит о повышении точности оценивания среднего.

Заметим, что свойства оценки (4) в асимптотике совпадают со свойствами модифицированного с учетом квантиля среднего, рассмотренного в [16], которое вычисляется и является несмещенным, только если $\min\{X_1, \dots, X_N\} \leq x_q \leq \max\{X_1, \dots, X_N\}$. При этом преимуществом оценки (4) можно считать то, что она позволяет всегда учитывать дополнительную информацию вида (5), даже если $x_q < \min\{X_1, \dots, X_N\}$ или $x_q > \max\{X_1, \dots, X_N\}$.

Для случаев малых объемов выборок точность оценки (4) исследовалась с помощью имитационного моделирования, при этом достаточно высокая скорость сходимости при $N > 30$ наблюдалась для таких распределений, как равномерное, нормальное и экспоненциальное, при разных значениях параметров распределений и уровней квантиля. В качестве иллюстрации данного факта на рис. 1 приведены графики зависимости значений $N \cdot \text{MSE}_M\{\bar{X}^q\}$ от N , при этом $\sigma_q^2 \approx 0,0365$ и $\sigma^2 \approx 0,0833$ для равномерной в (0,1) ф.р. $F(x) = R_{(0,1)}(x)$ при $x_q = q = 0,75$, значение $N \cdot \text{MSE}\{\bar{X}^q\}$ получено путем имитационного моделирования с параметром $M = 10^5$. На рис. 2 приведены аналогичные графики $N \cdot \text{MSE}_M\{\bar{X}^q\}$, $\sigma_q^2 \approx 0,364$ и $\sigma^2 = 1$ для стандартной нормальной ф.р. $F(x) = N_{(0,1)}(x)$ при $x_q = 0$, $q = 0,5$. Очевидно, что значение $N \cdot \text{MSE}_M\{\bar{X}^q\}$ не превышает σ^2 и с ростом N убывает до σ_q^2 . Моделирование также позволило обнаружить влияние вида распределения и параметра q на среднеквадратическую ошибку $N \cdot \text{MSE}_M\{\bar{X}^q\}$ при малых объемах выборок ($N < 20-30$), при этом для стандартного нормального распределения $N \cdot \text{MSE}_M\{\bar{X}^q\}$ в зависимости от q изменяется симметрично относительно медианы, а для равномерного и экспоненциального – несимметрично; аномально высокие значения $N \cdot \text{MSE}_M\{\bar{X}^q\}$ выявлены при q , близких к 1, а для нормального распределения – и для q , близких к 0 (рис. 3–5). Данный факт позволяет сделать вывод, что модифицированную оценку вообще не следует применять при слишком малых объемах наблюдений $N < 10$; при отсутствии нормальности – при $10 \leq N \leq 20$ и значениях $q > 0,55$; при наличии нормальности – при $10 \leq N \leq 15$, $q < 0,15$ и $q > 0,85$. Выявленные особенности требуют дальнейшего исследования.

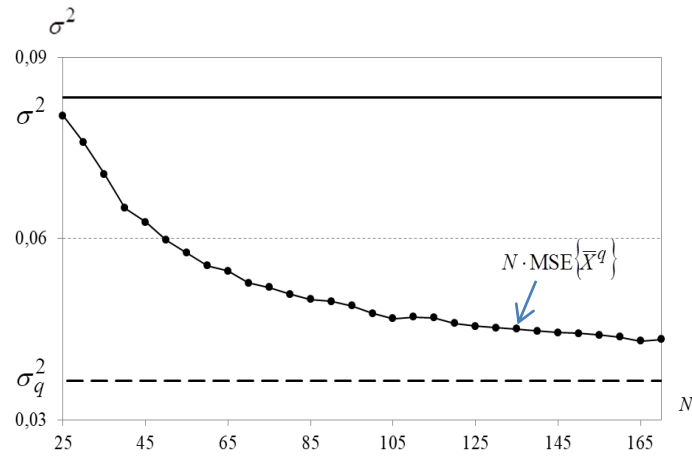


Рис. 1. График зависимости $N \cdot \text{MSE}_M\{\bar{X}^q\}$, σ_q^2 и σ^2 от N , $F(x) = R_{(0,1)}(x)$, $x_q = q = 0,75$

Fig. 1. Dependence of $N \cdot \text{MSE}_M\{\bar{X}^q\}$, σ_q^2 and σ^2 on N for $F(x) = R_{(0,1)}(x)$, $x_q = q = 0,75$

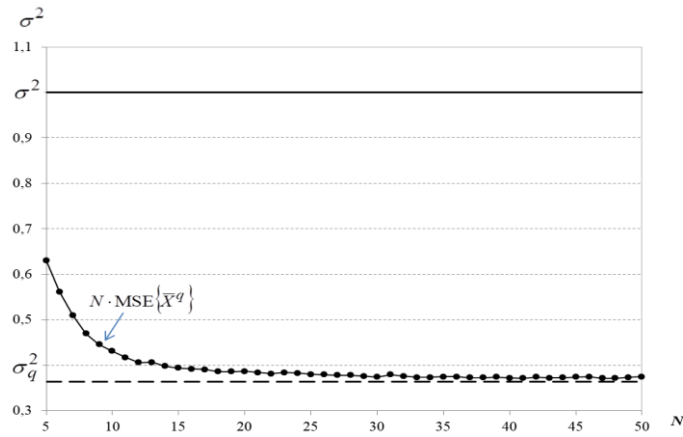


Рис. 2. График зависимости $N \cdot \text{MSE}_M\{\bar{X}^q\}$, σ_q^2 и σ^2 от N , $F(x) = N_{(0,1)}(x)$, $x_q = 0$, $q = 0,5$

Fig. 2. Dependence of $N \cdot \text{MSE}_M\{\bar{X}^q\}$, σ_q^2 and σ^2 on N for $F(x) = N_{(0,1)}(x)$, $x_q = 0$, $q = 0,5$

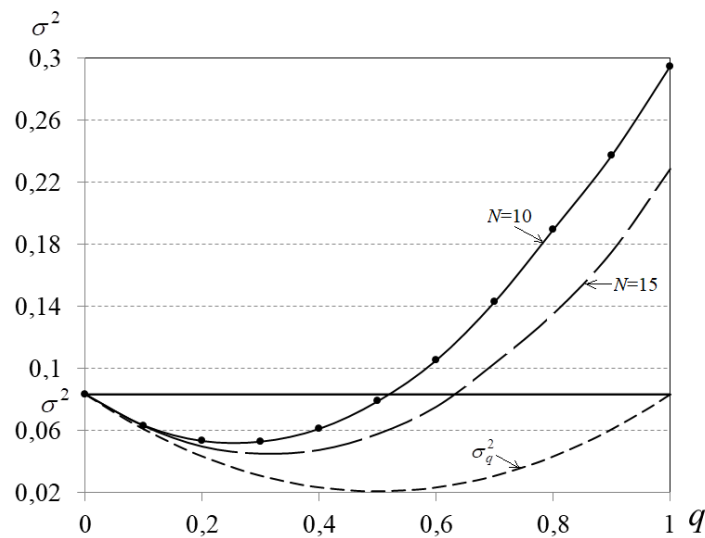


Рис. 3. График зависимости $N \cdot \text{MSE}_M\{\bar{X}^q\}$ для $N = 10$ и 15 , σ_q^2 и $\sigma^2 = 0,0833$ от q , $F(x) = R_{(0,1)}(x)$

Fig. 3. Dependence of $N \cdot \text{MSE}_M\{\bar{X}^q\}$ for $N = 10$ and 15 , σ_q^2 and $\sigma^2 = 0,0833$ on q for $F(x) = R_{(0,1)}(x)$

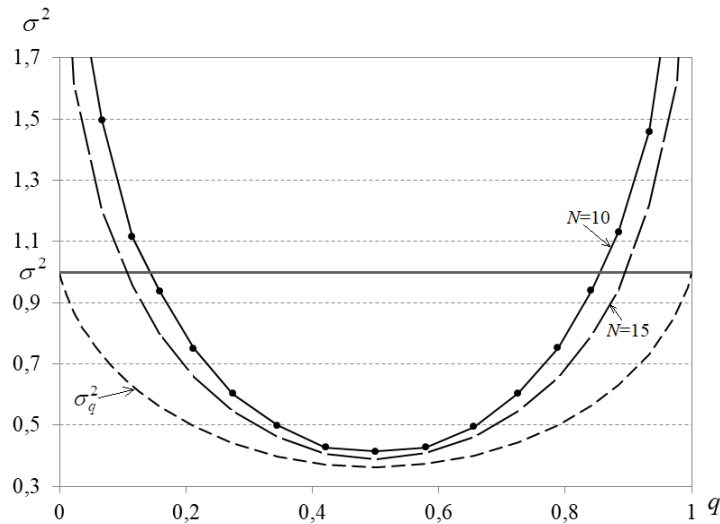


Рис. 4. График зависимости $N \cdot \text{MSE}_M \{ \bar{X}^q \}$ для $N = 10$ и 15 , σ_q^2 и $\sigma^2 = 1$ от q , $F(x) = N_{(0,1)}(x)$

Fig. 4. Dependence of $N \cdot \text{MSE}_M \{ \bar{X}^q \}$ for $N = 10$ and 15 , σ_q^2 and $\sigma^2 = 1$ on q for $F(x) = N_{(0,1)}(x)$

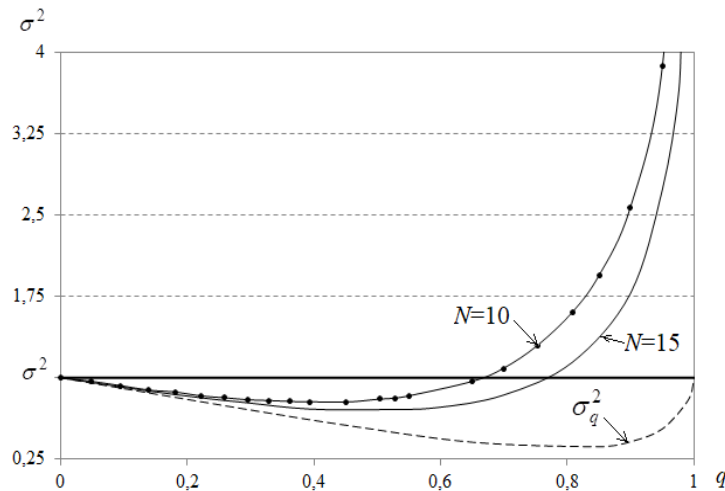


Рис. 5. График зависимости $N \cdot \text{MSE}_M \{ \bar{X}^q \}$ для $N = 10$ и 15 , σ_q^2 и $\sigma^2 = 1$ от q , $F(x) = 1 - e^{-x}, x \geq 0$

Fig. 5. Dependence of $N \cdot \text{MSE}_M \{ \bar{X}^q \}$ for $N = 10$ and 15 , σ_q^2 and $\sigma^2 = 1$ on q for $F(x) = 1 - e^{-x}, x \geq 0$

Таким образом, для достаточно больших объемов данных более точные показатели оборачиваемости с учетом квантиля могут быть рассчитаны с помощью следующих формул:

1) коэффициент оборачиваемости:

$$\text{КО}^q = \frac{TR}{\bar{X}^q}, \quad (7)$$

2) оборот:

$$\text{О}^q = \frac{365}{\text{КО}^q}. \quad (8)$$

Заметим, что на современных предприятиях системы учета позволяют использовать в анализе еженедельные и даже ежедневные данные о состоянии запасов, т.е. аналитики имеют доступ к выборкам очень больших объемов и могут применять данную методику, ориентируясь на асимптотические свойства новой оценки.

3. Апробация предложенной методики на реальных данных

Продemonстрируем работу классической и модифицированной систем расчетов ПО на реальных данных о стоимости остатков запасов предприятия за 12 месяцев, при этом количество наблюдений $N = 13$, так как стоимость учитывалась на первое число месяца, а последнее, 13-е значение, соответствует данным за последний рабочий день года. Это значение переносится затем на 1 января следующего года. Заметим, что выборка такого малого объема рассматривалась только для наглядности, но на практике детализация может быть куда больше, поскольку современные системы учета позволяют получать очень подробные отчеты о стоимости остатков запасов. В табл. 1 приведены данные о запасах и выручке производственного предприятия (см.: [15]).

Отметим, что, согласно критерию Шапиро–Уилка [18] с достигнутым уровнем значимости $p\text{-value} = 0,338$ данные являются нормально распределенными, выборочная оценка среднего $\bar{X} = 284\,122,861$ тыс. руб., выборочная дисперсия $S^2 = 84513,71$ тыс. (руб.)². Гипотеза случайности (независимости) данных проверялась критерием Вальда–Вольфовитца [19] (принята на уровне значимости 0,05), гипотеза зависимости отвергнута с помощью критерия Бартлетта [20] на уровне значимости более 0,01.

Таблица 1

Данные о стоимости остатков запасов и выручки

Месяц	Выручка, тыс. руб./мес.	Стоимость запасов на начало месяца, тыс. руб.
Январь	343 281,9	157 188,79
Февраль	500 587,2	211 566,90
Март	627 897,5	218 691,46
Апрель	653 847,7	345 808,36
Май	694 879,0	317 601,25
Июнь	644 220,6	331 117,79
Июль	774 122,4	490 150,71
Август	584 331,7	278 853,91
Сентябрь	556 282,6	277 191,10
Октябрь	521 414,9	275 095,37
Ноябрь	478 678,3	186 046,09
Декабрь	760 155,2	297 387,90
Январь		306 897,56

Таким образом, согласно традиционной методике расчета $KO = [25,13] = 25$ раз/год, так как на практике рекомендовано ориентироваться на пессимистический результат, то округление производится вниз. При этом $O = [14,53] = 15$ дней, так как период возврата вложений рассчитывается в полных днях. В итоге можно сделать вывод, что предприятие в течение года каждые 15 дней суммарно 25 раз в год возвращало инвестиции в запасы размером $\bar{X} = 284\,122,861$ тыс. руб. Далее аналитиками предприятия была предоставлена информация о том, что в 22% случаев стоимость запасов в течение года не превышала уровня $x_q = 216\,974,64$ тыс. руб. [24], т.е. известна дополнительная информация вида (5): $F(216\,974,64) = 0,22$. Это позволило переоценить среднюю стоимость запасов и показатели оборачиваемости: $O^q = [15,5] = 16$ дней, т.е. более точный способ расчета привел к более пессимистическому результату – для поддержания бесперебойной работы предприятие в среднем должно вкладывать в запасы на 6,91% больше денежных средств, вложения же возвращаются всего 23 раза в год при обороте уже в 16 дней.

Так как объем выборки $N = 13$ мал и не позволил использовать нормальность модифицированного среднего $\bar{X}^q$ для построения доверительных интервалов, то для более глубокого сравнительного анализа статистических характеристик показателей оборачиваемости с учетом и без учета дополнительной информации о квантиле применялся бутстреп-метод размножения выборок [20].

На рис. 5 приведены плотности бутстреп-распределений коэффициентов оборачиваемости КО и КО^q и соответствующие им графики подбора плотностей нормального распределения, параметр бутстреп-моделирования $R = 10^5$. Заметим, что бутстреп-распределение КО оказалось нормальным (гипотеза согласия проверялась с помощью χ^2 -критерия с достигнутым уровнем значимости $p\text{-value} \approx 1$), в то время как бутстреп-распределение КО^q ненормально, асимметрично с тяжелым правым хвостом (рис. 6). В табл. 2 представлены числовые характеристики бутстреп-распределений КО и КО^q (среднее и моды), их доверительные интервалы с уровнем доверия $\gamma = 0,9$ (здесь для адекватности сравнения рассмотрены 5 000-я и 95 000-я порядковые статистики бутстреп-распределений КО и КО^q), а также соответствующие им значения числовых характеристик и доверительные интервалы для О и О^q и средних вложений в запасы, полученные путем простых арифметических расчетов. Нетрудно видеть, что доверительный интервал для КО^q – уже со смещением целочисленной верхней границы влево, а для модифицированного среднего очевидно смещение всего интервала вправо.

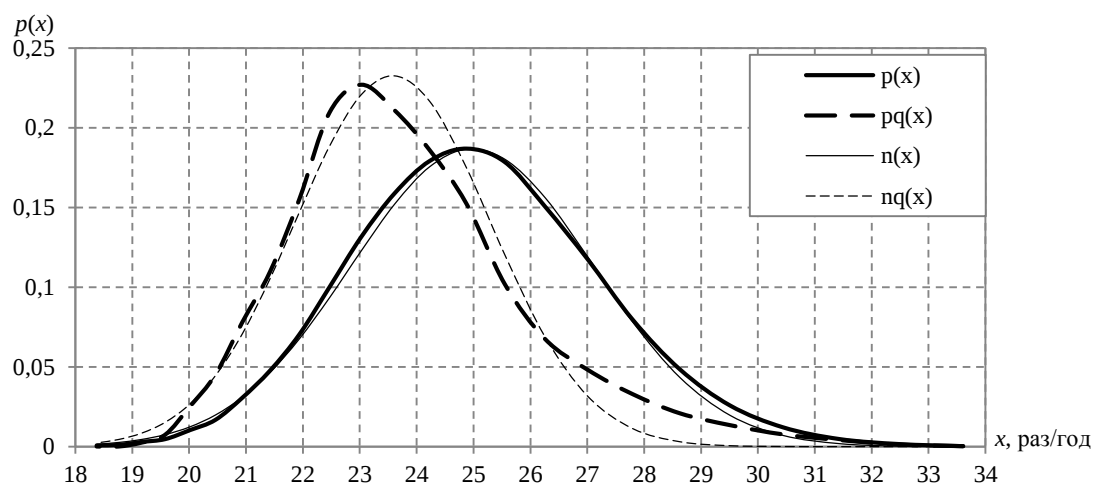


Рис. 6. Бутстреп-плотности $p(x)$ и $p^q(x)$ и соответствующие им плотности нормального распределения $n(x)$ и $n^q(x)$
Fig. 6. Bootstrap densities $p(x)$ and $p^q(x)$, and appropriate normal densities $n(x)$, $n^q(x)$

Таблица 2

**Характеристики показателей оборачиваемости и их доверительные интервалы
с уровнем доверия $\gamma = 0,9$, полученные бутстреп-методом с параметром $R = 10^5$**

Характеристика	Среднее	Мода	Среднеквадратическое отклонение	Нижняя граница интервала	Верхняя граница интервала
КО	[24,98] = 24	[24,82] = 24	2,135	[21,84] = 21	[29,27] = 29
КО ^q	[23,58] = 23	[22,98] = 22	1,715	[21,12] = 21	[28,33] = 28
О	[14,61] = 15	[14,71] = 15	–	[12,47] = 13	[16,71] = 17
О ^q	[15,48] = 16	[15,88] = 16	–	[12,88] = 13	[17,28] = 18
$\bar{X}$	285 816,61	287 659,10	–	243 911,18	326 933,66
$\bar{X}^q$	302 818,63	310 608,60	–	251 983,24	337 984,27

Таким образом, предложенная система расчета и бутстреп-анализ позволили сделать следующие выводы: классические показатели оборачиваемости давали завышенное значение коэффициента оборачиваемости 24 раза в год, притом если ориентироваться не на среднее (23 раза в год), а на самое вероятное согласно бутстреп-моделированию модальное значение КО^q, то можно сделать вывод, что предприятие возвращало вложенные в запас инвестиции только 22 раза в год. При наиболее вероятном исходе объем вложений можно оценить, как 310 608,60 против среднего 302 818,63 тыс. руб., в то время как классическое среднее давало оценку в 285 816,61 тыс. руб., т.е., при планировании бюджета возникал потенциальный риск дефицита финансовых средств в размере от 17 до 25 млн руб., которые нужно было инвестировать в запас не на 15, а на 16 дней. Такие неточности чреватые необхо-

димостью краткосрочного заимствования на очень невыгодных условиях, что очевидно негативно сказывается на прибыли предприятия.

Заключение

В работе предложен более точный подход к оцениванию показателей оборачиваемости с использованием дополнительной информации об известном квантиле заданного уровня, применимый для достаточно больших объемов наблюдений. При осуществлении расчетов использовалась модифицированная оценка математического ожидания, которая является несмещенной, асимптотически нормальной с дисперсией, в асимптотике не превосходящей дисперсию обычного выборочного среднего.

С помощью имитационного моделирования исследована среднеквадратическая ошибка модифицированного среднего для малых выборок, показано более высокое качество оценивания по сравнению с классическим средним для нормального и равномерного распределений для объемов наблюдений, больших 25–30. Для малых объемов выборок и близких к единице значений квантиля выявлены аномально большие значения среднеквадратической ошибки, которые требуют дополнительных исследований.

ЛИТЕРАТУРА

1. Fairfield P.M., Yohn T.L. Using asset turnover and profit margin to forecast changes in profitability // *Review of Accounting Studies*. 2001. V. 6. P. 371–385. DOI: 10.1023/A:1012430513430.
2. Supriati D., Kananto R., Kusriananda A. The effects of intellectual disclosures capital, debt to assets ratio, debt equity ratio, company size and assets turnover on company profitability // *Proceeding of the 5th Annual International Conference on Accounting Research (AICAR 2018): Advances in Economics, Business and Management Research*. 2019. P. 104–109. DOI: 10.2991/aicar-18.2019.23.
3. Patin J.-C., Rahman M., Mustafa M. Impact of total asset turnover ratios on equity returns: Dynamic panel data analyses // *Journal of Accounting, Business and Management*. 2020. V. 27, № 1. P. 19–29. DOI: 10.31966/jabminternational.v27i1.559.
4. Rahayu A., Supriadi Y. Analysis of net working capital turnover and total asset turnover against return on investment // *INA-Rxiv*. 2019. July. P. 1–18. DOI: 10.31227/osf.io/xtb8k.
5. Qu Z., Zhao B. The calculating model of inventory turnover based on time value // *Proceedings of the 2016 International Seminar on Education Innovation and Economic Management (SEIEM 2016)*. Series: Advances in Social Science, Education and Humanities Research (ASSEHR). 2016. V. 75. P. 51–54.
6. Гуров Н.В., Зенкова Ж.Н. Робастная оценка среднего в анализе оборачиваемости оборотных средств предприятия // *Математическое и компьютерное моделирование в экономике, страховании и управлении рисками : материалы V Междунар. молодежной науч.-практ. конф. (Саратов, 9–12 ноября 2016 г.)*. Саратов : Научная книга, 2016. С. 47–51.
7. Зенкова Ж.Н., Гуров Н.В. Доверительные интервалы для показателей оборачиваемости оборотных средств предприятия с использованием робастного оценивания // *XXI Век: итоги прошлого и проблемы настоящего плюс*. 2017. № 203 (36–37). С. 52–57.
8. Макеева О.Б., Зенкова Ж.Н. Доверительные интервалы для показателей оборачиваемости по интервальным данным // *Материалы VI Междунар. конф. «Логистические системы в глобальной экономике»*, Красноярск : СГАУ, 2016. С. 203–207.
9. Алтухова М.В. Устанавливаем норматив оборачиваемости запасов готовой продукции // *Справочник экономиста*. 2014. № 12: Управление материальными ресурсами.
10. Зенкова Ж.Н., Макеева О.Б. Применение методов обработки цензурированных данных при анализе оборачиваемости // *Вестник науки Казахского агротехнического университета*. 2014. № 3 (82). С. 21–30.
11. Зенкова Ж.Н., Мусони У. Экономичный размер заказа с учетом дополнительной информации об известном квантиле функции распределения объема продаж товара // *Бизнес-информатика*. 2020. Т. 14, № 3. С. 24–34.
12. HBR Harvard business review on supply chain management. Harvard Business School Press, 2006.
13. Dmitriev Yu.G., Tarasenko P.F. The use of a priori information in the statistical processing of experimental data // *Russian Physics Journal*. 1992. V. 35, № 9. P. 888–893. DOI: 10.1007/BF00560063.
14. Dmitriev Yu.G., Koshkin G.M. Nonparametric estimators of probability characteristics using unbiased prior conditions // *Statistical Papers*. 2018. V. 59, № 4. P. 1559–1575. DOI: 10.1007/s00362-018-1044-7.
15. Зенкова Ж.Н., Макеева О.Б. Использование информации о квантиле при анализе оборачиваемости оборотных средств // *Материалы III Всерос. молодежной конф. «Математическое и программное обеспечение информационных, технических и экономических систем»*. Томск : Изд. Дом Том. гос. ун-та, 2015. С. 82–87.
16. Zenkova Zh.N., Krainova E.A. Estimating the net premium using additional information about a quantile of the cumulative distribution function // *Business Informatics*. 2017. V. 4 (42). P. 55–63.

17. Dmitriev Y., Zenkova Z., Musoni W. Statistical Estimation with a Known Quantile and Its Application in a Modified ABC-XYZ Analysis // Book of Abstracts 8th International Conference on Risk Analysis and Design of Experiments. Vienna, Austria, 23–27 April 2019. P. 143–144.
18. Shapiro S.S., Wilk M.B. An analysis of variance test for normality (complete samples) // *Biometrika*. 1965. V. 52, № 3–4. P. 591–611.
19. Кобзарь А.И. Прикладная математическая статистика для инженеров и научных работников. М. : Физматлит, 2006. 816 с.
20. Efron B. Bootstrap methods: Another look at the jackknife // *The Annals of Statistics*. 1979. V. 7 (1). P. 1–26. DOI: 10.1214/aos/1176344552.

Поступила в редакцию 20 января 2021 г.

Zenkova Z.N., Musoni W., Andrievskaya A.A. (2021) ANALYSIS OF ASSETS TURNOVER RATIOS USING ADDITIONAL QUANTILE INFORMATION. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika* [Tomsk State University Journal of Control and Computer Science]. 55. pp. 35–44

DOI: 10.17223/19988605/55/5

The paper proposes more accurate approach to calculate the turnover ratios of an enterprise using additional information about the known quantile x_q with the given level q of the cumulative distribution function (CDF) $F(x) = P(X \leq x)$ of assets value X :

$$F(x_q) = q$$

for big sample size $N (X_1, X_2, \dots, X_N)$ from $F(x)$.

The method is based on a modified estimation of the mean value taking into account the quantile

$$\bar{X}^q = \frac{1}{N(N-1)} \cdot \sum_{i=1}^N \sum_{j=1, i \neq j}^N X_i \cdot \left(1 - \frac{(I_{(X_i < x_q)} - q) \cdot (I_{(X_j < x_q)} - q)}{q(1-q)} \right),$$

which is unbiased, asymptotically normal with a variance

$$\sigma_q^2 = \lim_{N \rightarrow \infty} N \cdot \text{var}\{\bar{X}^q\} = \sigma^2 - \frac{1}{q^2(1-q)^2} \left(aq - \frac{1}{q} \int_{-\infty}^{x_q} x dF(x) \right)^2,$$

where σ^2 is the variance of the sample mean $\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i$, $a = \int_{-\infty}^{+\infty} x dF(x)$.

Thus, for big sample size N the modified turnover ratios are calculated as:

1) turnover ratio:

$$TR^q = \frac{\text{Total Revenue for a Year}}{\bar{X}^q},$$

2) turnover:

$$T^q = \frac{365}{TR^q}.$$

The technique was tested on real data set, which contains the monthly inventories' value of an industrial enterprise. Using bootstrap modelling, the densities of turnover ratios with and without quantile, as well as average values of turnover ratios and their confidence intervals were obtained. Finally, a more accurate calculation led to pessimistic results: classical indicators of turnover gave more rosy values, and the average level of investment in inventories was underestimated, which entailed making inadequate management decisions.

Keywords: additional information; quantile; mean value estimation taking into account the quantile; turnover ratios.

ZENKOVA Zhanna Nikolaevna (Candidate of Physical and Mathematical Sciences, Associate Professor, Department of Systems Analysis and Mathematical Modeling, Institute of Applied Mathematics and Computer Science, National Research Tomsk State University, Senior Researcher, Russian Presidential Academy of National Economy and Public Administration (RANEPA), Moscow, Russian Federation).

E-mail: zhanna.zenkova@mail.tsu.ru

MUSONI Wilson (Post-graduate Student, Department of Systems Analysis and Mathematical Modeling, Institute of Applied Mathematics and Computer Science, National Research Tomsk State University, Russian Federation).

E-mail: wmusoni@uok.ac.rw

ANDRIEVSKAYA Anna Andreevna (Post-graduate Student, Department of Systems Analysis and Mathematical Modeling, Institute of Applied Mathematics and Computer Science, National Research Tomsk State University, Russian Federation).
E-mail: anna.andriev@mail.tsu.ru

REFERENCES

1. Fairfield, P.M. & Yohn, T.L. (2001). Using asset turnover and profit margin to forecast changes in profitability. *Review of Accounting Studies*. 6. pp. 371–385. DOI: 10.1023/A:1012430513430
2. Supriati, D., Kananto, R. & Kusriandana, A. (2019). The effects of intellectual disclosures capital, debt to assets ratio, debt equity ratio, company size and assets turnover on company profitability. *Proceeding of the 5th Annual International Conference on Accounting Research (AICAR 2018): Advances in Economics, Business and Management Research*. pp. 104–109. DOI: 10.2991/aicar-18.2019.23
3. Patin, J.-C., Rahman, M. & Mustafa, M. (2020). Impact of total asset turnover ratios on equity returns: Dynamic panel data analyses. *Journal of Accounting, Business and Management*. 27(1). pp. 19–29. DOI: 10.31966/jabminternational.v27i1.559
4. Rahayu, A. & Supriadi, Y. (2019) Analysis of net working capital turnover and total asset turnover against return on investment. *INA-Rxiv*. pp. 1–18. DOI: 10.31227/osf.io/xtb8k
5. Qu, Z. & Zhao, B. (2016) The calculating model of inventory turnover based on time value. *Proceedings of the 2016 International Seminar on Education Innovation and Economic Management (SEIEM 2016). Series: Advances in Social Science, Education and Humanities Research (ASSEHR)*. 75. pp. 51–54. DOI: 10.2991/seiem-16.2016.12
6. Gurov, N.V. & Zenkova, Zh.N. (2016) Robastnaya otsenka srednego v analize oborachivaemosti oborotnykh sredstv predpriyatiya [Robust estimation of mean value in turnover analysis of current assets]. In: Balash, V.A. (ed.) *Matematicheskoe i komp'yuternoe modelirovaniye v ekonomike, strakhovanii i upravlenii riskami* [Mathematical and Computer Modelling in Economics, Insurance and Risk Management]. Saratov: Nauchnaya kniga. pp. 47–51.
7. Zenkova, Zh.N. & Gurov, N.V. (2017) Confidential intervals for turnover ratios of current assets using robust estimating. *XXI Vek: itogi proshlogo i problemy nastoyashchego plyus* [XXI Century: Resumes of the Past and Challenges of the Present plus]. 203(36–37). pp. 52–57.
8. Makeeva, O.B. & Zenkova, Z.N. (2016) Doveritel'nye intervaly dlya pokazateley oborachivaemosti po interval'nym dannym [Confidence intervals for the inventory turnover ratios based on interval data]. *Logisticheskie sistemy v global'noy ekonomike* [Logistic Systems in Global Economy]. Proc. of the 6th International Conference. Krasnoyarsk: SSAU. pp. 203–207. (In Russian.)
9. Altukhova, M. (2014) Ustanavlivayem normativ oborachivaemosti zapasov gotovoy produktsii [Establishing a standard for the turnover ratio of finished goods]. *Spravochnik ekonomista*. 12.
10. Zenkova, Z.N. & Makeeva, O.B. (2014) Primeneniye metodov obrabotki tsenzurirovannykh dannykh pri analize oborachivaemosti [Application of methods for processing censored data in the analysis of turnover]. *Vestnik nauki Kazakhskogo agrotekhnicheskogo universiteta*. 3(82). pp. 21–30.
11. Zenkova, Z. & Musoni, W. (2020) The economic order quantity taking into account additional information about the known quantile of the cumulative distribution function of the product's sales volume. *Biznes-informatika – Business Informatics*. 3(14). pp. 24–34. DOI: 10.17323/2587-814X.2020.3.24.34
12. HBR. (2006) *Harvard Business Review on Supply Chain Management*. Harvard Business School Press.
13. Dmitriev, Yu.G. & Tarasenko, P.F. (1992) The use of a priori information in the statistical processing of experimental data. *Russian Physics Journal*. 35(9). pp. 888–893. DOI: 10.1007/BF00560063
14. Dmitriev, Yu.G. & Koshkin, G.M. (2018) Nonparametric estimators of probability characteristics using unbiased prior conditions. *Statistical Papers*. 59(4). pp. 1559–1575. DOI: 10.1007/s00362-018-1044-7.
15. Zenkova, Z.N. & Makeeva, O.B. (2015) Ispol'zovaniye informatsii o kvantile pri analize oborachivaemosti oborotnykh sredstv [Using quantile information in the analysis of working capital turnover]. *Matematicheskoe i programnoye obespecheniye informatsionnykh, tekhnicheskikh i ekonomicheskikh sistem* [Mathematical and software support for information, technical and economic systems]. Proc. of the Third All-Russian Youth Conference. Tomsk, Russia, May 22–23. pp. 82–87.
16. Zenkova, Z.N. & Krainova, E.A. (2017) Estimating the net premium using additional information about a quantile of the cumulative distribution function. *Biznes-informatika – Business Informatics*. 4. pp. 55–63. DOI: 10.17323/1998-0663.2017.4.55.63
17. Dmitriev, Y., Zenkova, Z. & Musoni, W. (2019) Statistical estimation with a known quantile and its application in a modified ABC-XYZ analysis. *Book of Abstracts 8th International Conference on Risk Analysis and Design of Experiments*. Vienna, Austria, April 23–27. pp. 143–144.
18. Shapiro, S.S. & Wilk, M.B. (1965) An analysis of variance test for normality (complete samples). *Biometrika*. 52(3–4). pp. 591–611.
19. Kobzar, A.I. (2006) *Prikladnaya matematicheskaya statistika dlya inzhenerov i nauchnykh rabotnikov* [Applied mathematical statistics for engineers and scientists]. Moscow: Fizmatlit.
20. Efron, B. (1979) Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*. 7(1). pp. 1–26. DOI: 10.1214/aos/1176344552

УДК 004.672

DOI: 10.17223/19988605/55/6

С.Н. Калашников, Е.А. Мартусевич, Е.В. Мартусевич, И.А. Рыбенко

КЛАСТЕРИЗАЦИЯ ОДНОРОДНЫХ ПРОИЗВОДСТВЕННЫХ ОБЪЕКТОВ НА ОСНОВЕ АНАЛИЗА ТЕХНОЛОГИЧЕСКИХ ДАННЫХ ПРИ ФОРМИРОВАНИИ АЛЮМИНИЕВЫХ СПЛАВОВ

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-37-90087\19.

Рассматривается применение метода классификационного кластерного анализа для осуществления разбиения некоторого множества однородных производственных объектов (литейных ковшей) по их формальным признакам на соответствующие кластерные группы (кластеры). При решении данной технологической задачи целесообразно использование метода кластеризации на основе оценки степени «отдаленности» химического состава исходного алюминия-сырца от установленных значений выбранной марки алюминиевого сплава в результате расчета метрических расстояний в пространстве состояний, что позволит снизить количество излишних эмпирических вариантов неправильной компоновки литейных ковшей.

Ключевые слова: технологические данные; химический состав; кластерный анализ; метрика; функция Хевисайда.

Ключевым вариантом модернизации действующего производственного процесса является совершенствование технологического этапа формирования алюминиевых сплавов в литейном отделе [1] с помощью определения необходимых вариантов сочетания доступных литейных ковшей на основе детального анализа технологических данных, отражающих действительные параметры химического состава разлитого алюминия-сырца в имеющемся наборе литейных ковшей.

Так, совершенствование технологии производства алюминия является актуальной задачей и обусловлено отсутствием единого подхода к рационализации установившихся технологических процессов, а также связано с наличием избыточного количества циклов модификации алюминиевого расплава [2]. Многошаговая модификация расплава происходит в реальных условиях и зависит от уровня квалификации технологического персонала, т.е. является причиной неточных расчетов необходимых значений параметров смешивания содержимого из литейных ковшей, определяя технико-экономические последствия неверно выбранных управляющих воздействий сотрудников литейного отделения [3, 4]. Все это приводит к повышению себестоимости производства, повышенному износу оборудования, сдвигу плановых сроков сдачи готовой продукции заказчикам и снижению конкурентоспособности промышленного предприятия на рынке цветных металлов [5, 6].

Одним из подходов для решения представленной проблемы является применение метода классификационного кластерного анализа для осуществления разбиения множества однородных производственных объектов (литейных ковшей) на кластерные группы (кластеры), соответствующие установленным требованиям реализации той или иной технологии выплавки алюминиевых сплавов с целью получения наиболее точных вариантов смешивания содержимого литейных ковшей и минимизации дополнительных корректировок формируемого химического состава алюминиевого сплава. Этот метод имеет все преимущества комбинационной группировки и открывает широкие перспективы его применения в многомерном статистическом анализе, классификации объектов, исследовании связей, типизации выборок, а также является эффективным и удобным инструментом в научно-практических исследованиях [7, 8]. При этом исходные технологические данные могут иметь значительный объем отличительных формальных признаков, однако необходимо учитывать только наиболее важные осо-

бенности исследуемых объектов. Также к преимуществам кластерного анализа можно отнести возможность разделения объектов по ряду выбранных признаков или критериев с потенциалом последующей реализации программно-алгоритмического обеспечения на языке программирования высокого уровня. Кроме этого, в отличие от большинства математических методов, кластерный анализ не накладывает никаких ограничений на вид и происхождение изучаемых объектов, что позволяет исследовать объекты практически произвольной природы [9].

В частности, реализация метода кластерного анализа возможна с помощью различных алгоритмов, которые имеют собственные отличительные особенности и применяются авторами в зависимости от условий выбранной задачи, что обеспечивает нахождение наиболее рационального варианта разбиения доступного множества имеющихся объектов одного типа на несколько конкретных кластеров, образованных с использованием соответствующего алгоритма. Например, к таким алгоритмам относятся: метод К-средних, ЕМ-алгоритм, алгоритмы семейства FOREL, метод нечеткой кластеризации С-средних, генетический алгоритм, иерархическая кластеризация [10].

Однако необходимо отметить, что не существует уникального алгоритма кластерного анализа объектов, который бы мог использоваться повсеместно и отвечал заданным условиям задачи, в том числе соответствовал требованиям быстродействия вычислительного оборудования, а полученные при этом результаты полноценно соотносились с практическими ожиданиями. Поэтому для достижения наилучшего результата кластеризации необходимо экспериментировать с выбором метрических расстояний и периодически менять используемый алгоритм. Так, в результате исследования существующих алгоритмов и перебора соответствующих подходов к расчету метрических расстояний возникла необходимость в формировании набора оценок значений метрического отклонения исходного химического состава алюминия-сырца в литейных ковшах относительно заданного химического состава установленной марки алюминиевого сплава.

На основе полученных значений этих оценок разработан и протестирован алгоритм кластеризации технологических данных, позволяющий получить новые практические результаты, удобные для дальнейшей интерпретации и осуществления визуально-событийного анализа изменяющихся значений параметров химического состава алюминия-сырца при смешивании в сравнении с действующим эмпирическим методом компоновки литейных ковшей с целью принятия взвешенных решений технологическим персоналом.

1. Кластерный анализ литейных ковшей с алюминием-сырцом

Для реализации анализа технологических данных с дальнейшим выделением классификационных групп литейных ковшей необходимо выбрать наиболее точные отличительные особенности, чтобы объекты, относящиеся к одному кластеру, отличались от объектов, принадлежащих другим кластерам. Сходство между объектами определяется через расстояние между их соответствующими точками в n -мерном пространстве, т.е. чем больше расстояние между точками, тем дальше друг от друга находятся исследуемые объекты, и, следовательно, чем меньше расстояние между точками, тем ближе располагаются рассматриваемые объекты.

С целью определения расстояний между объектами в конкретной метрике необходимо сформировать уникальный вектор характеристик для каждого исследуемого объекта, например в виде некоторого набора числовых значений. Вектор характеристик приводится к определенному диапазону значений, описывающих основные признаки представленных объектов. При таком описании каждый признак дает одинаковый вклад при проведении расчетов метрических расстояний, определяя взаимосвязь между исследуемыми объектами, но при этом для каждой выбранной метрики необходимо учитывать исходные условия и действующие ограничения решаемой задачи [11].

Как правило, перед решением задачи кластеризации исходных объектов, количество кластеров заранее не известно и выбирается по субъективным критериям автора, занимающегося вопросами математического анализа. В общем случае применение классификационного кластерного анализа сводится к формализации и последовательному выполнению следующих алгоритмических этапов:

- 1) формирование выборки необходимых объектов для кластеризации;
- 2) определение множества субъективных признаков для оценки объектов в выборке;
- 3) выбор метрики и вычисление значений расстояний (сходства) между объектами;
- 4) реализация алгоритма работы метода кластерного анализа для создания групп кластеров;
- 5) представление результатов проведенного анализа и их интерпретация.

При этом после получения и анализа имеющихся результатов возможна корректировка выбранной метрики и сопутствующих исполняемых параметров метода кластеризации для получения наиболее точного решения поставленной задачи.

Возвращаясь к вопросу решения технологических задач, связанных с формированием алюминиевого расплава, использование кластерного анализа является целесообразным, так как позволяет осуществлять разбиение доступных литейных ковшей на группы кластеров при помощи оценки химического состава расплава в некотором диапазоне значений с учетом установленных требований и ограничений.

Таким образом, при анализе исходных технологических данных и имеющихся ограничений определены следующие кластеры литейных ковшей:

1) *к первому кластеру* относятся ковши с алюминием-сырцом, химический состав которых полностью соответствует заданным требованиям исходной марки алюминиевого сплава, т.е. значения химических элементов находятся в допустимых границах и их не нужно модифицировать;

2) *ко второму кластеру* относятся ковши с алюминием-сырцом, химический состав которых не совпадает с заданными требованиями исходной марки алюминиевого сплава, но значения концентраций химических элементов можно откорректировать либо путем смешивания, либо с помощью применения модифицирующих компонентов в виде лигатур и флюсов;

3) *к третьему кластеру* относятся оставшиеся литейные ковши с алюминием-сырцом, химический состав которых не может быть оптимально скорректирован доступными способами.

Данные о химическом составе алюминия-сырца в литейных ковшах представлены в виде упорядоченных наборов значений X , отражающих концентрации химических элементов в разных алюминиевых расплавах, где $X = (x_1, x_2, \dots, x_n)^T = (\text{Fe}, \text{Si}, \text{Ti}, \text{Al}, \text{Cu}, \text{Zn}, \text{Mn}, \text{Mg}, \text{Pb}, \text{Sn})^T$.

Кластеризация литейных ковшей с алюминием-сырцом осуществляется на основе полученных с помощью соотношения $S_k = \|V_k\|_m$ оценок значений метрического отклонения S_k химического состава X_k алюминия-сырца в k -м литейном ковше с учетом заданного химического состава установленной марки алюминиевого сплава.

Оценка метрического отклонения S_k химического состава X_k алюминия-сырца в k -м литейном ковше от заданного химического состава установленной марки алюминиевого сплава осуществляется с помощью вычисления значения нормы вектора V_k , т.е. $V_k = (V_{1k}, V_{2k}, \dots, V_{nk})^T$, что соответствует соотношению

$$V_{ik} = \begin{cases} 0, & \text{при } (x_{ik} - x_{i\min})(x_{ik} - x_{i\max}) \leq 0, \\ \min(|x_{ik} - x_{i\min}|, |x_{ik} - x_{i\max}|), & \text{при } (|x_{ik} - x_{i\min}|, |x_{ik} - x_{i\max}|) > 0. \end{cases} \quad (1)$$

Компонентами V_{ik} вектора V_k являются значения отклонений концентраций i -го химического элемента в k -м литейном ковше от допустимых значений согласно заданным требованиям марки формируемого расплава в миксере. Компоненты V_{ik} формируются на основе функции Хевисайда [12]

$$\theta(x) = \begin{cases} 0, & x < 0, \\ 1, & x \geq 0, \end{cases} \quad (2)$$

с помощью соотношения

$$V_{ik} = (x_{i\min} - x_i)\theta(x_{i\min} - x_i) + (x_i - x_{i\max})\theta(x_i - x_{i\max}). \quad (3)$$

В конечном итоге в сформированных кластерах S_k принимает следующие значения:

1) $S_k = 0$, при соответствии содержимого алюминия-сырца заданной марке алюминиевого сплава (первый кластер);

2) $0 < S_k \leq \varepsilon$, при незначительном несоответствии содержимого алюминия-сырца заданной марке алюминиевого сплава и возможности изменения текущего расплава до заданных значений концентраций химических элементов с помощью модифицирующих материалов и технологических операций (второй кластер);

3) $S_k > 0$, при несоответствии содержимого алюминия-сырца заданной марке алюминиевого сплава и отсутствии возможности модификации формируемого расплава до заданных значений концентраций химических элементов (третий кластер).

В результате анализа технологических данных и интерпретации полученных результатов компоновки производственных объектов был разработан математический метод кластеризации доступного набора литейных ковшей, характеризующий степень «отдаленности» химического состава исходного алюминия-сырца от установленных значений выбранной марки алюминиевого сплава на основе расчета метрических расстояний в некотором пространстве состояний.

2. Интерпретация результатов кластеризации литейных ковшей по их химическому составу

На рис. 1–3 представлен исходный вариант значений концентраций магния Mg (горизонтальная ось) и кремния Si (вертикальная ось) одной из марок алюминиевого деформируемого сплава для некоторого набора литейных ковшей. При этом установлено, что $\varepsilon = 0,2$.

Так, внутренний прямоугольник на рис. 1, а, 2, а, 3, а соответствует установленным требованиям заданной марки алюминиевого сплава, а область между ним и последующим прямоугольником соответствует первоначальным значениям концентраций Mg и Si в литейных ковшах, однако значения этих концентраций можно откорректировать путем смешивания или с помощью использования модифицирующих компонентов (лигатур и флюсов).

Далее, на рис. 1, б, 2, б, 3, б, представлены значения метрического отклонения S_k концентраций химического состава X_k алюминия-сырца в k -м литейном ковше от установленного химического состава заданной марки алюминиевого сплава.

На рис. 1 представлены значения концентраций Mg и Si во всех литейных ковшах, что полностью соответствует требованиям заданной марки алюминиевого расплава и выражается нулевыми значениями метрического отклонения S_k (первый кластер).

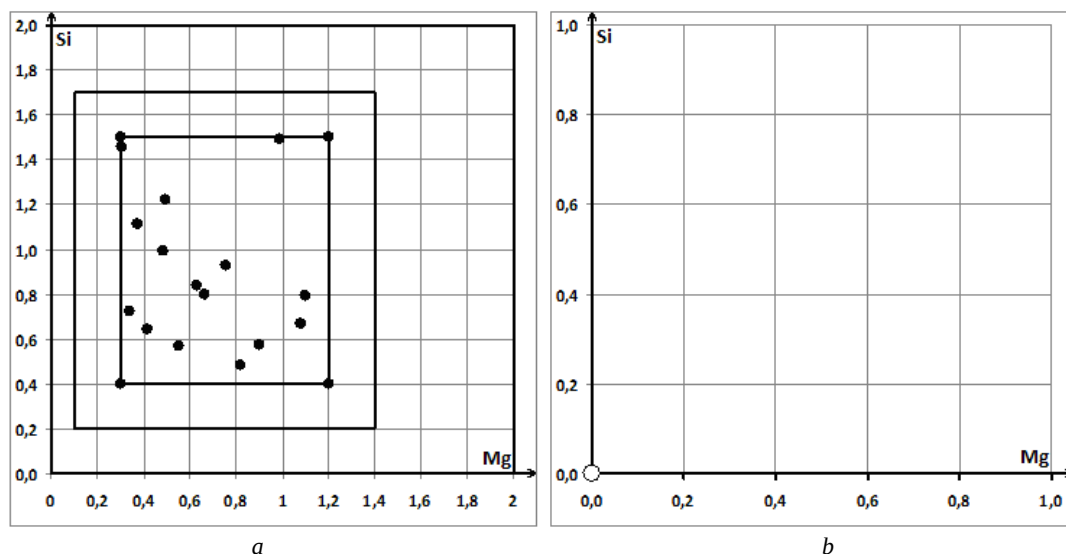


Рис. 1. Значения концентраций Mg, Si и метрических отклонений $S_k(\text{Mg})$ и $S_k(\text{Si})$ при соответствии содержимого алюминия-сырца заданным условиям марки деформируемого сплава

Fig. 1. Values of concentrations of Mg, Si and metric deviations $S_k(\text{Mg})$ and $S_k(\text{Si})$ when the content of raw aluminum corresponds to the specified conditions of the grade of the wrought alloy

На рис. 2 представлены значения концентраций Mg и Si во всех литейных ковшах, что уже не соответствует требованиям заданной марки алюминиевого сплава и выражается ненулевыми значениями метрического отклонения S_k в границах квадратной области, но при этом значения концентраций этих химических элементов можно оптимально скорректировать (*второй кластер*).

Аналогично на рис. 3 представлены значения концентраций Mg и Si во всех литейных ковшах, что также не соответствует требованиям заданной марки алюминиевого сплава и выражается ненулевыми значениями метрического отклонения S_k в границах несимметричной области, но при этом значения концентраций этих химических элементов нельзя скорректировать оптимальным образом (*третий кластер*).

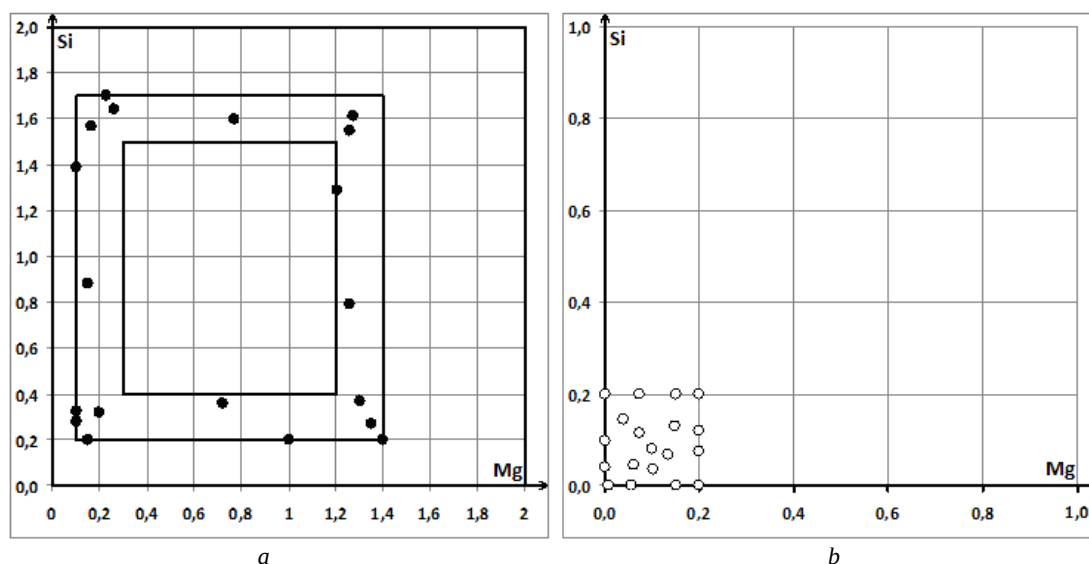


Рис. 2. Значения концентраций Mg, Si и метрических отклонений $S_k(\text{Mg})$ и $S_k(\text{Si})$ при несоответствии содержимого алюминия-сырца заданным требованиям марки алюминиевого деформируемого сплава, но с возможностью оптимального модифицирования расплава

Fig. 2. Values of concentrations of Mg, Si and metric deviations $S_k(\text{Mg})$ and $S_k(\text{Si})$ when the content of raw aluminum does not correspond to the specified requirements for the grade of aluminum wrought alloy, but with the possibility of optimal modification of the melt

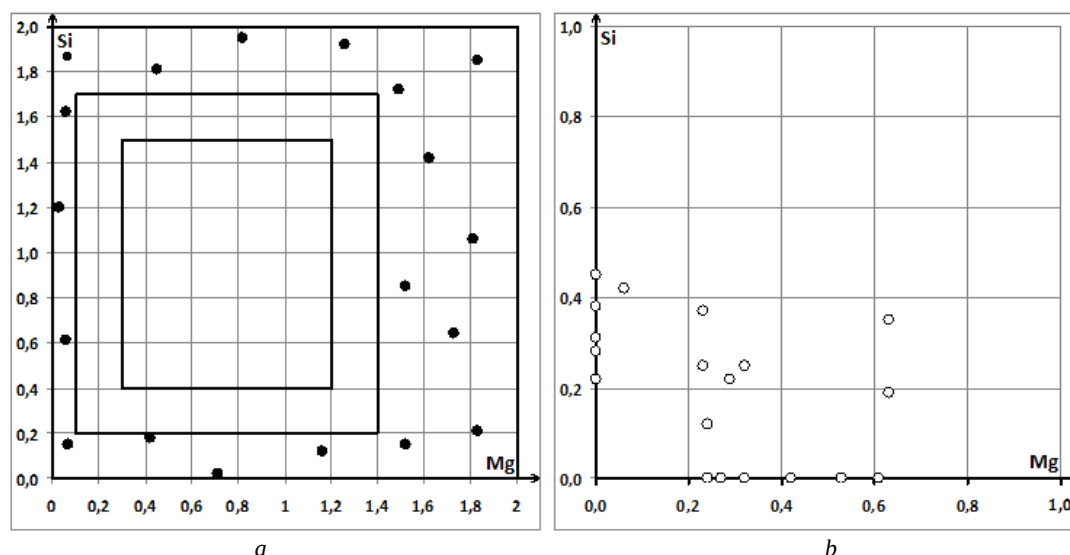


Рис. 3. Значения концентраций Mg, Si и метрических отклонений $S_k(\text{Mg})$ и $S_k(\text{Si})$ при несоответствии содержимого алюминия-сырца заданным требованиям марки алюминиевого деформируемого сплава и без возможности оптимального модифицирования содержимого расплава

Fig. 3. Values of concentrations of Mg, Si and metric deviations $S_k(\text{Mg})$ and $S_k(\text{Si})$ when the content of raw aluminum does not correspond to the specified requirements of the grade of aluminum wrought alloy and without the possibility of optimal modification of the content of the melt

В результате применение кластерного анализа позволило осуществить рациональное разбиение доступного набора литейных ковшей на три кластера с помощью детальной оценки химического состава исходного алюминия-сырца в рамках реализации выбранной технологии изготовления алюминиевых сплавов, исключив использование эмпирического подхода к компоновке литейных ковшей и уменьшив количество дальнейших корректировок формируемого химического состава алюминиевого расплава в миксере литейного отделения.

Заключение

В статье разработан и протестирован способ кластеризации литейных ковшей с алюминием-сырцом на основе применения математического аппарата, предназначенного для расчетов соответствующих метрических отклонений начальных значений концентраций химических элементов от установленных значений, определяемых выбранной маркой алюминиевого сплава согласно спецификациям портфеля заказов.

Представленный способ кластеризации однородных объектов, рассмотренный в рамках алюминиевого производства, целесообразно использовать при реализации информационных программных комплексов, предназначенных для обработки большого массива данных в условиях непрерывно изменяющихся технологических процессов для мониторинга основных показателей производства и интеллектуального контроля управляющих воздействий технологического персонала.

ЛИТЕРАТУРА

1. Скуратов А.П., Шахрай С.Г., Фомичев И.В., Белянин А.В. Повышение экологической и энергетической эффективности производства алюминия // Исследование энергетической эффективности и экологических показателей оборудования алюминиевой промышленности. Красноярск : Сиб. фед. ун-т, 2018. 179 с.
2. Володькина А.А. Роль аналитических методов в металлургии алюминия // Молодежный вестник ИРГТУ. 2020. № 10 (1). С. 77–80.
3. Дмитрюк А.И., Григорьев А.А. Совершенствование технологии прессования алюминиевых заготовок // Заготовительные производства в машиностроении. 2020. Т. 18, № 8. С. 353–358.
4. Шестаков А.К., Садыков Р.М., Петров П.А. Состояние производства алюминия на начало 2020 года // Международный научно-исследовательский журнал. 2020. № 7-1 (97). С. 114–119.
5. Васечко Д.Ю. Конъюнктура мировых рынков меди и алюминия // Производство цветных металлов и сплавов. М. : Кабели и Провода, 2011. С. 10–12.
6. Александров А.В., Немчинова Н.В. Расчет ожидаемой экономической эффективности производства алюминия за счет увеличения применения глинозема отечественного производства // Вестник Иркутского государственного технического университета. 2020. Т. 24, № 2 (151). С. 408–420.
7. Хальцинова Ж., Требуна П. Методы кластерного анализа // Формирование экономического порядка. 2012. № 2 (30). С. 31–34.
8. Барышникова Н.А. Методика кластеризации на основе иерархических и неиерархических методов кластерного анализа // Вестник Таганрогского института управления и экономики. 2012. № 1 (15). С. 81–83.
9. Кельманов А.В. О сложности некоторых задач кластерного анализа // Журнал вычислительной математики и математической физики. 2011. Т. 51, № 11. С. 2106–2112.
10. Тюрин А.Г., Зуев И.О. Кластерный анализ, методы и алгоритмы кластеризации // Вестник МГТУ МИРЭА. 2014. № 2 (3). С. 86–97.
11. Buhaienko Y., Globa L.S., Liashenko A., Grebinechenko M. Analysis of clustering algorithms for use in the universal data processing system // Открытые семантические технологии проектирования интеллектуальных систем. 2020. № 4. С. 101–104.
12. Chikayama E. Decomposition of multivariate function using the Heaviside step function // SpringerPlus. 2014. № 3 (1). P. 1–11.

Поступила в редакцию 3 февраля 2021 г.

Kalashnikov S.N., Martusevich E.A., Martusevich E.V., Rybenko I.A. (2021) CLUSTERING OF HOMOGENEOUS PRODUCTION FACILITIES BASED ON THE ANALYSIS OF TECHNOLOGICAL DATA IN THE FORMATION OF ALUMINUM ALLOYS. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika* [Tomsk State University Journal of Control and Computer Science]. 55. pp. 45–52

DOI: 10.17223/19988605/55/6

The continuous development of the metallurgical industry, including the current trend of improving the technological process of forming aluminum alloys, determines the need for additional research work aimed at obtaining safe and effective methods for modernizing existing technological measures for the production of aluminum alloys, considering specific requirements for their chemical composition. Modernization of the established modes of operation of the technological stages of the production process, as well as continuous monitoring of the main features of the technological operations in the manufacture of aluminum alloys, determines the achievement of new practical results in the production of modern high-quality materials with higher environmental and operational properties. All these factors directly increase the competitiveness of the industry in the world market of non-ferrous metals, increasing consumer demand for products of this kind.

One of the available options for upgrading the current production cycle of aluminum alloy manufacturing is to improve the technological stage of forming aluminum melt in the melting mixer of the foundry department. In particular, it is proposed to optimize the established operating modes of production units and control actions. At the same time, there is a need to use mathematical tools designed to formalize the main features of the studied production facilities in order to continuously monitor dynamic changes in technological parameters. Proper processing of process data contributes to the application of classification cluster analysis of the existing set of casting buckets with raw aluminum based on the initial values of the concentrations of elements of the chemical composition.

The cluster analysis provides an effective division of the available set of casting buckets with raw aluminum into appropriate clusters, including considering the established requirements for the selected brand of aluminum alloy according to the specifications of the order portfolio. This principle of cluster analysis, followed by software and algorithmic implementation of appropriate quality, especially with the use of a high-speed high-level programming language, will ensure the creation of a workable way to obtain the most correct options for allocating cluster sets of casting buckets for their subsequent layout and mixing.

Thus, the use of cluster analysis is not only a convenient analytical tool for the study of a large array of technological data, reducing the complexity of the perception of a large amount of information about the production process by technological personnel, but also helps to reduce the number of unnecessary empirical options for incorrect layout of available casting buckets at optimal technical and economic costs.

The reported study was funded by RFBR, project number № 19-37-90087/19.

Keywords: process data; chemical composition; cluster analysis; metric; Heaviside function.

KALASHNIKOV Sergey Nikolaevich (Doctor of Technical Sciences, Associate Professor, Professor of the Department of Applied Information Technologies and Programming, Siberian State Industrial University, Novokuznetsk, Russian Federation).

E-mail: s.n.kalashnikov@yandex.ru

MARTUSEVICH Efim Aleksandrovich (Junior Researcher, Post-graduate Student of the Department of Applied Information Technologies and Programming, Siberian State Industrial University, Novokuznetsk, Russian Federation).

E-mail: program.pro666@yandex.ru

MARTUSEVICH Elena Vladimirovna (Candidate of Technical Sciences, Associate Professor of the Department of Natural Science Disciplines named after Professor V.M. Finkel, Siberian State Industrial University, Novokuznetsk, Russian Federation).

E-mail: science_nvkh@yandex.ru

RYBENKO Inna Anatolyevna (Doctor of Technical Sciences, Associate Professor, Head of the Department of Applied Information Technologies and Programming, Siberian State Industrial University, Novokuznetsk, Russian Federation).

E-mail: rybenkoi@mail.ru.

REFERENCES

1. Skuratov, A.P., Shakhrai, S.G., Fomichev, I.V. & Belyanin, A.V. (2018) Povyshenie ekologicheskoy i energeticheskoy effektivnosti proizvodstva alyuminiya [Improving the environmental and energy efficiency of aluminum production]. In: *Issledovanie energeticheskoy effektivnosti i ekologicheskikh pokazateley oborudovaniya alyuminiyevoy promyshlennosti* [Research of energy efficiency and environmental performance of equipment for the aluminum industry]. Krasnoyarsk: Siberian Federal University.
2. Volodkina, A.A. (2020) The role of analytical methods in aluminum metallurgy. *Molodezhny Vestnik IrGTU – Young Researchers' Journal of ISTU*. 10(1). pp. 77–80.
3. Dmitryuk, A.I. & Grigoriev, A.A. (2020) Enhancement of aluminum billets extrusion technology. *Zagotovitel'nye proizvodstva v mashinostroyeni*. 18(8). pp. 353–358.
4. Shestakov, A.K., Sadykov, R.M. & Petrov, P.A. (2020) Sostoyaniye proizvodstva alyuminiya na nachalo 2020 goda [The state of aluminum production at the beginning of 2020]. *Mezhdunarodnyy nauchno-issledovatel'skiy zhurnal – International Research Journal*. 7-1(97). pp. 114–119.
5. Vasechko, D.Yu. (2011) Kon'yunktura mirovykh rynkov medi i alyuminiya [The conjuncture of the world copper and aluminum markets]. In: *Proizvodstvo tsvetnykh metallov i splavov* [Production of non-ferrous metals and alloys]. Moscow: Kabeli i Provoda. pp. 10–12.

6. Aleksandrov, A.V. & Nemchinova, N.V. (2020) Calculation of the expected economic efficiency of aluminum production by increasing the use of domestic alumina. *Vestnik Irkutskogo gosudarstvennogo tekhnicheskogo universiteta – Proceedings of Irkutsk State Technical University*. 24 - 2(151). pp. 408–420. DOI: 10.21285/1814-3520-2020-2-408-420
7. Khaltsinova, J. & Trebuna, P. (2012) Cluster Analysis Method. *Sever i Rynok. Formirovanie ekonomicheskogo poryadka – The North and the Market: Forming the Economic Order*. 2(30). pp. 31–34.
8. Baryshnikova, N.A. (2012) Metodika klasterizatsii na osnove ierarkhicheskikh i neierarkhicheskikh metodov klasterного analiza [Clustering technique based on hierarchical and non-hierarchical methods of cluster analysis]. *Vestnik Taganrogskego instituta upravleniya i ekonomiki – Bulletin of the Taganrog Institute of Management and Economics*. 1(15). pp. 81–83.
9. Kelmanov, A.V. (2011) On the complexity of some cluster analysis problems. *Zhurnal vychislitel'noy matematiki i matematicheskoy fiziki – Computational Mathematics and Mathematical Physics*. 51(11). pp. 2106–2112.
10. Tyurin, A.G. & Zuev, I.O. (2014) Klasternyy analiz, metody i algoritmy klasterizatsii [Cluster analysis, methods and algorithms for clustering]. *Vestnik MGTU MIREA*. 2(3). pp. 86–97.
11. Buhaienko, Y., Globa, L.S., Liashenko, A. & Grebinechenko, M. (2020) Analysis of clustering algorithms for use in the universal data processing system. *Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh sistem*. 4. pp. 101–104.
12. Chikayama, E. (2014) Decomposition of multivariate function using the Heaviside step function. *SpringerPlus*. 3(1). pp. 1–11. DOI: 10.1186/2193-1801-3-704

УДК 519.21

DOI: 10.17223/19988605/55/7

Л.А. Нежелская, А.А. Першина

ОЦЕНИВАНИЕ МЕТОДОМ МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ ПАРАМЕТРА РАСПРЕДЕЛЕНИЯ СЛУЧАЙНОГО МЕРТВОГО ВРЕМЕНИ В РЕКУРРЕНТНОМ ОБОБЩЕННОМ АСИНХРОННОМ ПОТОКЕ СОБЫТИЙ

Изучается рекуррентный обобщенный асинхронный поток событий, являющийся распространенной математической моделью информационных потоков сообщений, функционирующих в телекоммуникационных сетях, и относящийся к классу дважды стохастических потоков событий. Функционирование потока рассматривается в стационарном режиме и в условиях непродлевающегося случайного мертвого времени, распределенного по равномерному закону на отрезке $[0, T^*]$. Рассматриваются общий и особый случаи задания параметров рекуррентного обобщенного асинхронного потока событий. Приводятся численные результаты (в общем и особом случаях) по оцениванию параметра T^* методом максимального правдоподобия.

Ключевые слова: рекуррентный обобщенный асинхронный поток; непродлевающееся случайное мертвое время; оценка параметра; метод максимального правдоподобия.

Подавляющее число статей прошлого века посвящено изучению систем и сетей массового обслуживания (СМО, СеМО) с входящими простейшими потоками событий (сообщений, запросов, заявок). Однако в связи с интенсивным развитием (с конца прошлого века) телекоммуникационных сетей, беспроводных и мобильных сетей связи математическая модель простейшего потока перестала быть адекватной реальным информационным потокам запросов. Требования практики послужили стимулом к использованию дважды стохастических потоков в качестве математической модели реальных потоков запросов в телекоммуникационных системах и сетях. Подчеркнем, что основным свойством дважды стохастических потоков является их коррелированность. Термин «дважды стохастические потоки» связан с тем, что события потока наступают в случайные моменты времени, а интенсивность потока представляет собой случайный процесс. В связи с этим дважды стохастические потоки можно разделить на два класса: первый класс составляют потоки, интенсивность которых есть непрерывный случайный процесс [1, 2]; второй – потоки, интенсивность которых есть кусочно-постоянный случайный процесс с конечным числом состояний. Впервые результаты исследований потоков второго класса опубликованы практически одновременно, в 1979 г., в работах [3, 4] и работе [5]. В [3, 4] отмеченные потоки получили название МС(Markov chain)-потоки. В [5] – MVP(Markov versatile processes)-потоки. В работе [6] эти потоки называются также МАР(Markovian Arrival Process)-потоками событий.

Подчеркнем, что МС-потоки событий являются наиболее характерной и подходящей математической моделью потоков в реальных телекоммуникационных сетях, в частности в широкополосных сетях беспроводной связи вдоль протяженных транспортных магистралей [7–11].

Большинством авторов исследования СМО и СеМО осуществляются в условиях, когда все события входящего дважды стохастического потока доступны наблюдению. В реальности же зарегистрированное событие может создать период мертвого времени для регистрирующего прибора [12], в течение которого другие события потока становятся ненаблюдаемыми для регистрирующего прибора (теряются). В этой связи можно считать, что мертвое время выступает искажающим фактором при решении различного рода задач оценивания по измерениям моментов наступления наблюдаемых сообщений исходного дважды стохастического потока (эффект мертвого времени влечет за собой потери событий исходного потока, что отрицательно сказывается на решении задач оценивания). Все

устройства регистрации делятся на две группы [13]. Первую группу составляют устройства с непродлевающимся мертвым временем, вторую – устройства с продлевающимся мертвым временем. Период ненаблюдаемости событий потока (период мертвого времени) может продолжаться некоторое фиксированное время, а также может быть случайным. В настоящей работе в качестве искажающего фактора рассматривается непродлевающееся случайное мертвое время.

В настоящее время в мировой литературе имеется, по-видимому, единственная монография [14], где приведено систематизированное изложение теории очередей с коррелированными (дважды стохастическими) потоками применительно к телекоммуникационным сетям. Подчеркнем, что изложенная в [14] теория и ее применение в телекоммуникационных сетях рассмотрены без искажающих факторов (непродлевающегося либо продлевающегося мертвого времени), воздействующих на входящий дважды стохастический поток событий.

Математические модели дважды стохастических потоков событий с непродлевающимся детерминированным мертвым временем широко использовались и используются при решении задач оценивания состояний и параметров дважды стохастических потоков событий по измерениям моментов наступления событий наблюдаемых потоков [15–21].

Однако достаточно открытым остается вопрос изучения дважды стохастических потоков событий, когда мертвое время является случайной величиной. Здесь отметим работу [22], в которой решается задача оценки параметров асинхронного потока событий в условиях случайного мертвого времени, работу [23], в которой решается задача оценки параметра распределения непродлевающегося случайного мертвого времени в пуассоновском потоке, и работу [24], в которой находятся формулы для начальных моментов общего периода ненаблюдаемости в пуассоновском потоке событий при продлеваемомся случайном мертвом времени.

В настоящей статье рассматривается рекуррентный обобщенный асинхронный дважды стохастический поток событий (рекуррентный обобщенный ММРР-поток), являющийся обобщением асинхронного потока событий [25], функционирующий в условиях непродлевающегося случайного мертвого времени. Случайное мертвое время распределено по равномерному закону. На параметры обобщенного асинхронного потока событий накладываются ограничения, приводящие его к рекуррентному потоку (изучаются общий и особый случаи). Данная статья непосредственно примыкает к исследованиям, проведенным в [26, 27].

1. Математическая модель наблюдаемого потока

Рассматривается рекуррентный обобщенный асинхронный дважды стохастический поток событий, сопровождающий процесс которого есть кусочно-постоянный стационарный случайный процесс $\lambda(t)$ с двумя состояниями λ_1 и λ_2 ($\lambda_1 > \lambda_2 \geq 0$). В течение временного интервала, когда $\lambda(t) = \lambda_i$, имеет место пуассоновский поток событий с интенсивностью λ_i , $i = 1, 2$. Переход из первого состояния процесса $\lambda(t)$ во второе (из второго в первое) может осуществляться в произвольный момент времени, не связанный с моментами наступления событий пуассоновского потока интенсивности λ_i , $i = 1, 2$ (свойство асинхронности потока). При этом длительность пребывания процесса $\lambda(t)$ в i -м состоянии распределена по экспоненциальному закону с параметром α_i , $i = 1, 2$. При переходе процесса $\lambda(t)$ из первого состояния во второе инициируется с вероятностью p ($0 \leq p \leq 1$) дополнительное событие во втором состоянии. Наоборот, при переходе процесса $\lambda(t)$ из второго состояния в первое инициируется с вероятностью q ($0 \leq q \leq 1$) дополнительное событие в первом состоянии. В сделанных предположениях $\lambda(t)$ – скрытый (принципиально ненаблюдаемый) марковский процесс. После каждого зарегистрированного события в момент времени t_k наступает период мертвого времени случайной длительности, так что другие события исходного потока, наступившие в течение этого периода мертвого времени (периода ненаблюдаемости), недоступны наблюдению (теряются) и не вызывают его продления (непродлевающееся мертвое время). Принимается, что случайная длительность мертвого времени распределена по равномерному закону с плотностью $p(T)=1/T^*$, $0 \leq T \leq T^*$. В результате формируется наблюдаемый поток событий, отличный от исходного (часть событий исходного потока теряется).

Рассматривается стационарный режим функционирования наблюдаемого потока событий (переходными процессами на полуинтервале наблюдения $(t_0, t]$, где t_0 – начало наблюдений, t – окончание наблюдений, пренебрегаем). Необходимо в момент времени t на основании выборки $t_1, t_2, \dots, t_n$ ($t_n < t$) наблюдаемых моментов наступления событий оценить методом максимального правдоподобия параметр T^* (МП-оценка).

2. Приближенная МП-оценка параметра T^* в общем случае

Предварительно отметим, что в [28] получены результаты по оценке параметра T^* для коррелированного обобщенного асинхронного потока событий при непродлеваемом случайном мертвом времени. Для того чтобы перейти от коррелированного потока к рекуррентному, необходимо установить условия рекуррентности, при которых коррелированный поток становится рекуррентным. В [29] получены условия рекуррентности для наблюдаемого потока: 1) $\lambda_1\lambda_2 - p q \alpha_1\alpha_2 = 0$, 2) $\lambda_1 - \lambda_2 + p\alpha_1 - q\alpha_2 = 0$, 3) $\lambda_1 - \lambda_2 + q\alpha_1 - p\alpha_2 = 0$, когда T – детерминированная величина, т.е. при выполнении одного из этих условий коррелированный наблюдаемый поток становится рекуррентным.

Обозначим $\tau_k = t_{k+1} - t_k$, $k = 1, 2, \dots$, значение длительности k -го интервала между соседними событиями наблюдаемого потока ($\tau_k > 0$). Так как рассматривается стационарный режим, то плотность вероятности значений длительности k -го интервала есть $p(\tau_k) = p(\tau)$, $\tau \geq 0$, для любого k в наблюдаемом потоке событий. В силу этого момент времени t_k без потери общности можно положить равным нулю, т.е. момент наступления события есть $\tau = 0$.

В [29] с учетом первого условия рекуррентности получено выражение для условной плотности вероятности $p(\tau | T)$, когда длительность мертвого времени является детерминированной величиной:

$$\begin{aligned} p(\tau | T) &= 0, \quad 0 \leq \tau < T; \quad p(\tau | T) = \gamma(T)z_1 e^{-z_1(\tau-T)} + [1 - \gamma(T)]z_2 e^{-z_2(\tau-T)}, \quad \tau \geq T; \\ \gamma(T) &= [z_2 - (\lambda_1 + p\alpha_1)\pi_1(T) - (\lambda_2 + q\alpha_2)\pi_2(T)] / (z_2 - z_1); \quad z_2 - z_1 \neq 0; \\ \pi_1(T) &= \pi_1 \left[1 + \alpha_1 \frac{\lambda_1 - \lambda_2 + q\alpha_1 - p\alpha_2}{\lambda_1\alpha_2 + \lambda_2\alpha_2 + (p+q)\alpha_1\alpha_2} e^{-(\alpha_1+\alpha_2)T} \right], \quad \pi_2(T) = 1 - \pi_1(T); \\ \pi_1 &= \alpha_2 / (\alpha_1 + \alpha_2), \quad \pi_2 = 1 - \pi_1; \end{aligned} \quad (1)$$

$$z_{1,2} = \frac{1}{2} \left[(\lambda_1 + \alpha_1 + \lambda_2 + \alpha_2) \mp \sqrt{(\lambda_1 - \lambda_2 + \alpha_1 - \alpha_2)^2 + 4\alpha_1\alpha_2(1-p)(1-q)} \right], \quad 0 < z_1 < z_2.$$

Тогда плотность $p(\tau)$ примет вид:

$$p(\tau) = \int_{(T)} p(T) p(\tau | T) dT = \begin{cases} p_1(\tau) = \int_0^\tau p(T) p(\tau | T) dT, & 0 \leq \tau < T^*; \\ p_2(\tau) = \int_{T^*}^\tau p(T) p(\tau | T) dT, & \tau \geq T^*. \end{cases} \quad (2)$$

Подставляя в (2) выражение (1), учитывая, что $p(T) = 1/T^*$, находим

$$p_1(\tau) = \frac{1}{T^*} \left[1 - e^{-z_1\tau} - e^{-z_2\tau} + e^{-(\alpha_1+\alpha_2)\tau} \right], \quad 0 \leq \tau < T^*; \quad (3)$$

$$\begin{aligned} p_2(\tau) &= \frac{1}{T^* (z_2 - z_1)(\alpha_1 + \alpha_2)} \left\{ [-(\alpha_1 + \alpha_2)(z_2 - z_1) + z_2(\alpha_1 + \alpha_2 - z_1)e^{z_1T^*} - z_1(\alpha_1 + \alpha_2 - z_2) \times \right. \\ &\quad \times e^{-(\alpha_1+\alpha_2-z_1)T^*}] e^{-z_1\tau} - [(\alpha_1 + \alpha_2)(z_2 - z_1) + z_1(\alpha_1 + \alpha_2 - z_2)e^{z_2T^*} - \\ &\quad \left. - z_2(\alpha_1 + \alpha_2 - z_1)e^{-(\alpha_1+\alpha_2-z_2)T^*}] e^{-z_2\tau} \right\}, \quad \tau \geq T^*. \end{aligned} \quad (4)$$

Отметим, во-первых, что $\int_0^{\infty} p(\tau) d\tau = \int_0^{T^*} p_1(\tau) d\tau + \int_{T^*}^{\infty} p_2(\tau) d\tau = 1$, во-вторых, $\lim_{\tau \rightarrow \infty} p_1(\tau) = p_2(T^*)$

при $\tau \rightarrow T^* - 0$, т.е. в точке $\tau = T^*$ плотность $p(\tau)$ является непрерывной функцией переменной τ , в-третьих, $\lim p'_1(\tau) \neq p'_2(T^*)$ при $\tau \rightarrow T^* - 0$, т.е. в точке $\tau = T^*$ плотность $p(\tau)$ претерпевает излом.

В упомянутых выше статьях [26–28] для нахождения оценки параметра T^* был использован метод моментов [30–32]. В настоящей статье для случая рекуррентного обобщенного асинхронного потока событий для оценки параметра T^* воспользуемся методом максимального правдоподобия [31, 32]. Применение последнего оправдано тем, что если параметры потока обеспечивают реализацию одного из условий рекуррентности, например первого, то случайные величины – длительности интервалов между соседними событиями потока – становятся взаимонезависимыми.

Пусть в процессе наблюдения за потоком на полуинтервале времени $(t_0, t]$ измерены n значений: $\tau_1, \tau_2, \dots, \tau_n$. Тогда функция правдоподобия запишется в виде:

$$L(T^* | \tau_1, \tau_2, \dots, \tau_n) = \prod_{k=1}^n p(\tau_k | T^*), \quad T^* > 0. \quad (5)$$

где $p(\tau_k | T^*)$ – плотность вероятности, определенная формулами (2), (3), (4) в которых $\tau = \tau_k$ (τ_k – измерения), T^* – переменная величина ($T^* > 0$). С учетом (3), (4) выражение (5) примет вид:

$$L(T^* | \tau_1, \tau_2, \dots, \tau_n) = L_1(T^* | \tau_1, \tau_2, \dots, \tau_n) L_2(T^* | \tau_1, \tau_2, \dots, \tau_n);$$

$$L_1(T^* | \tau_1, \tau_2, \dots, \tau_n) = \prod_{k: \tau_k < T^*}^n p_1(\tau_k | T^*), \quad L_2(T^* | \tau_1, \tau_2, \dots, \tau_n) = \prod_{k: \tau_k > T^*}^n p_2(\tau_k | T^*). \quad (6)$$

Упорядочим величины τ_k , $k = \overline{1, n}$, по возрастанию: $0 < \tau^{(1)} < \tau^{(2)} < \dots < \tau^{(n)} < \infty$. Тогда (6) переписывается в виде:

$$\begin{aligned}
L\left(T^* \mid \tau^{(1)}, \dots, \tau^{(n)}\right) &= \prod_{k=1}^n p_2\left(\tau^{(k)} \mid T^*\right), \quad 0 < T^* < \tau^{(1)}; \\
L\left(T^* \mid \tau^{(1)}, \dots, \tau^{(n)}\right) &= p_1\left(\tau^{(1)} \mid T^*\right) \prod_{k=2}^n p_2\left(\tau^{(k)} \mid T^*\right), \quad \tau^{(1)} \leq T^* < \tau^{(2)}; \\
L\left(T^* \mid \tau^{(1)}, \dots, \tau^{(n)}\right) &= \prod_{k=1}^2 p_1\left(\tau^{(k)} \mid T^*\right) \prod_{k=3}^n p_2\left(\tau^{(k)} \mid T^*\right), \quad \tau^{(2)} \leq T^* < \tau^{(3)}; \\
&\vdots \\
L\left(T^* \mid \tau^{(1)}, \dots, \tau^{(n)}\right) &= \prod_{k=1}^{n-1} p_1\left(\tau^{(k)} \mid T^*\right) p_2\left(\tau^{(n)} \mid T^*\right), \quad \tau^{(n-1)} \leq T^* < \tau^{(n)}; \\
&\vdots \\
L\left(T^* \mid \tau^{(1)}, \dots, \tau^{(n)}\right) &= \prod_{k=1}^n p_1\left(\tau^{(k)} \mid T^*\right), \quad \tau^{(n)} \leq T^* < \infty.
\end{aligned} \tag{7}$$

Теорема. Функция $p(\tau^{(k)} | T^*)$ переменной T^* ($T^* > 0$) достигает своего глобального максимума в точке $T^* = \tau^{(k)}$.

Доказательство. Обозначим $p_1(\tau^{(k)} | T^*)$ – плотность вероятности, определенную формулой (3) в которой $\tau = \tau^{(k)}$ ($\tau^{(k)}$ – измерение), T^* – переменная величина ($T^* > 0$). Тогда очевидно, что $dp_1(\tau^{(k)} | T^*)/dT^* < 0$ для $T^* \geq \tau^{(k)}$, т.е. функция $p_1(\tau^{(k)} | T^*)$ – убывающая функция переменной T^* и достигает своего глобального максимума в точке $T^* = \tau^{(k)}$, $k = \overline{1, n}$.

Рассмотрим функцию $p_2(\tau^{(k)} | T^*)$ – плотность вероятности, определенную формулой (4), в которой $\tau = \tau^{(k)}$ ($\tau^{(k)}$ – измерение), T^* – переменная величина, $0 < T^* \leq \tau^{(k)}$, $k = \overline{1, n}$. При этом $p_1(\tau^{(k)} | T^* = \tau^{(k)}) = p_2(\tau^{(k)} | T^* = \tau^{(k)})$, $k = \overline{1, n}$. Производная по T^* функции $p_2(\tau^{(k)} | T^*)$ выпишется в виде:

$$p'_2(\tau^{(k)} | T^*) = \frac{1}{(z_2 - z_1)(\alpha_1 + \alpha_2)(T^*)^2} \varphi(T^*), \quad 0 < T^* \leq \tau^{(k)}, \quad k = \overline{1, n}, \quad (8)$$

$$\begin{aligned} \varphi(T^*) = & (\alpha_1 + \alpha_2)(z_2 - z_1)[e^{-z_1\tau^{(k)}} + e^{-z_2\tau^{(k)}}] + z_2(\alpha_1 + \alpha_2 - z_1)(z_1T^* - 1)e^{z_1(T^* - \tau^{(k)})} - z_1(\alpha_1 + \alpha_2 - z_2) \times \\ & \times (z_2T^* - 1)e^{z_2(T^* - \tau^{(k)})} + z_1(\alpha_1 + \alpha_2 - z_2)[(\alpha_1 + \alpha_2 - z_1)T^* + 1]e^{-(\alpha_1 + \alpha_2 - z_1)T^*} e^{-z_1\tau^{(k)}} - \\ & - z_2(\alpha_1 + \alpha_2 - z_1)[(\alpha_1 + \alpha_2 - z_2)T^* + 1]e^{-(\alpha_1 + \alpha_2 - z_2)T^*} e^{-z_2\tau^{(k)}}. \end{aligned} \quad (9)$$

Знак производной (8) определяется знаком функции (9). Рассмотрим функцию $\varphi(T^*)$ как функцию двух переменных T^* и $\tau^{(k)}$, т.е. $\varphi(T^*) = \varphi(T^*, \tau^{(k)})$, $0 \leq T^* \leq \tau^{(k)}$, $\tau^{(k)} \geq 0$, $k = \overline{1, n}$. Имеем $\varphi(T^* = 0, \tau^{(k)}) = 0$.

Производная функции $\varphi(T^*, \tau^{(k)})$ по переменной T^* выпишется в виде:

$$\begin{aligned} \varphi'(T^*, \tau^{(k)}) = & \varphi_1(T^*)\varphi_2(T^*, \tau^{(k)}), \quad 0 \leq T^* \leq \tau^{(k)}, \quad \tau^{(k)} \geq 0, \quad k = \overline{1, n}, \\ \varphi_1(T^*) = & z_1z_2 - (\alpha_1 + \alpha_2 - z_1)(\alpha_1 + \alpha_2 - z_2)e^{-(\alpha_1 + \alpha_2)T^*}, \\ \varphi_2(T^*, \tau^{(k)}) = & (\alpha_1 + \alpha_2 - z_1)z_1e^{-z_1(\tau^{(k)} - T^*)} - (\alpha_1 + \alpha_2 - z_2)z_2e^{-(\tau^{(k)} - T^*)}. \end{aligned}$$

Нетрудно показать, что $\varphi_1(T^*) > 0$ для $0 \leq T^* \leq \tau^{(k)}$, $k = \overline{1, n}$; $\varphi_2(T^*, \tau^{(k)}) > 0$ для $0 \leq T^* \leq \tau^{(k)}$, $\tau^{(k)} \geq 0$, $k = \overline{1, n}$. Тогда $\varphi'(T^*, \tau^{(k)}) > 0$ для $0 \leq T^* \leq \tau^{(k)}$, $\tau^{(k)} \geq 0$, $k = \overline{1, n}$. Отсюда следует, что $\varphi(T^*)$ – возрастающая функция переменной T^* , $T^* > 0$. Последнее означает, что $p'_2(\tau^{(k)} | T^*) > 0$, $0 < T^* \leq \tau^{(k)}$, $k = \overline{1, n}$. Тогда $p_2(\tau^{(k)} | T^*)$ – возрастающая функция переменной T^* ($0 < T^* \leq \tau^{(k)}$, $k = \overline{1, n}$) и достигает своего глобального максимума в точке $T^* = \tau^{(k)}$. Объединяя два утверждения относительно глобальных максимумов функций $p_1(\tau^{(k)} | T^*)$ и $p_2(\tau^{(k)} | T^*)$, получаем утверждение теоремы. Теорема доказана.

Результат теоремы позволяет утверждать, что на отрезке $[0, \tau^{(1)}]$ изменения T^* функция правдоподобия (7) является возрастающей функцией и достигает своего локального максимума в точке $T^* = \tau^{(1)}$. На полуинтервале $[\tau^{(n)}, \infty)$ изменения T^* функция правдоподобия (7) является убывающей функцией и достигает своего локального максимума в точке $T^* = \tau^{(n)}$.

Таким образом, для отыскания глобального максимума функции правдоподобия (7) необходимо исследовать отрезок $[\tau^{(1)}, \tau^{(n)}]$ изменения переменной T^* ($\tau^{(1)} \leq T^* \leq \tau^{(n)}$).

Так как функция $p(\tau^{(k)} | T^*)$ ($T^* > 0$) в точке $T^* = \tau^{(k)}$, $k = \overline{1, n}$, достигает глобального максимума, то будем считать точку $T^* = \tau^{(k)}$, $k = \overline{1, n}$, точкой подозрительной на локальный максимум функции правдоподобия (7). Тогда алгоритм нахождения приближенной МП-оценки $\hat{T}^*$ параметра T^* будет выглядеть следующим образом: 1) вычисляются значения функции правдоподобия (7) в точках $T^* = \tau^{(k)}$, $k = \overline{1, n}$; 2) находится максимальное значение функции (7) на множестве этих точек; 3) в качестве приближенной МП-оценки параметра T^* выбирается $\hat{T}^*$, обеспечивающее максимальное значение функции (7) на предыдущем шаге алгоритма.

С учетом второго условия рекуррентности в [29] получено выражение для условной плотности вероятности $p(\tau | T)$, когда T – детерминированная величина:

$$p(\tau | T) = 0, \quad 0 \leq \tau < T; \quad p(\tau | T) = z_1e^{-z_1(\tau - T)}, \quad \tau > T, \quad z_1 = \lambda_1 + p\alpha_1. \quad (10)$$

Тогда аналогично предыдущему случаю, находим

$$p_1(\tau) = \frac{1}{T^*}(1 - e^{-z_1\tau}), \quad 0 \leq \tau < T^*; \quad p_2(\tau) = \frac{1}{T^*}e^{-z_1\tau}(e^{z_1T^*} - 1), \quad \tau \geq T^*. \quad (11)$$

Дальнейший ход изложения повторяет рассмотренный выше.

Наконец, для третьего условия рекуррентности в [29] получено выражение для $p(\tau|T)$ в виде (10), где $z_1 = \lambda_1 + q\alpha_1$. Тогда $p_1(\tau)$ и $p_2(\tau)$ сохраняют вид (11), и ход изложения сохраняется аналогично рассмотренному для первого условия рекуррентности.

С целью установления качества получаемых методом максимального правдоподобия оценок параметра T^* поставлены статистические эксперименты. Статистические эксперименты поставлены для рекуррентного наблюдаемого потока, для которого справедливо первое условие рекуррентности: $\lambda_1\lambda_2 - pq\alpha_1\alpha_2 = 0$.

Первый статистический эксперимент (установление стационарного режима). Отдельный j -й эксперимент ($j = 1, \dots, N$) заключается в следующем: 1) при заданных значениях параметров потока λ_i , $i = 1, 2$, α_1 , $\alpha_2 = \lambda_1\lambda_2/(pq\alpha_1)$, p , q , T^* и заданном времени моделирования Tm единиц времени (Tm – время наблюдения за потоком) осуществляется имитационное моделирование наблюдаемого потока [33]; выходом имитационной модели в отдельном j -м эксперименте является последовательность значений $\tau^{(1)}, \dots, \tau^{(n)}$; 2) численно реализуется процедура поиска глобального максимума функции правдоподобия (7) в точках, подозрительных на экстремум [34], т.е. находится значение оценки $\hat{T}_j^*$; 3) осуществляется повторение N раз шагов 1, 2. Результатом работы алгоритма является выборка $(\hat{T}_1^*, \dots, \hat{T}_N^*)$, на основании которой вычисляются

$$\hat{M}(\hat{T}^*) = (1/N) \sum_{j=1}^N \hat{T}_j^*, \quad \hat{V}(\hat{T}^*) = (1/N) \sum_{j=1}^N [\hat{T}_j^* - T^*]^2, \quad (12)$$

где T^* – известное из имитационной модели значение параметра. После этого время моделирования Tm увеличивается на ΔTm , и алгоритм переходит на выполнение первого шага отдельного j -го эксперимента. При проведении первого статистического эксперимента выбраны следующие параметры имитационной модели: $\lambda_1 = 3$, $\lambda_2 = 2$, $\alpha_1 = 0,5$, $\alpha_2 = 48$, $p = 0,5$, $q = 0,5$, $T^* = 1$. Ниже приведены результаты первого статистического эксперимента ($N = 100$; $Tm = 100, 200, \dots, 1\,000$) (табл. 1).

Таблица 1

Первый статистический эксперимент

Tm	100	200	300	400	500	600	700	800	900	1 000
$\hat{M}(\hat{T}^*)$	1,0043	1,0012	1,0028	1,0004	0,9989	1,0013	0,9978	1,0003	0,9973	1,0003
$\hat{V}(\hat{T}^*)$	0,0036	0,0021	0,0008	0,0006	0,0007	0,0006	0,0006	0,0004	0,0003	0,0003

Анализ численных результатов показывает, что в смысле введенного критерия (выборочная вариация оценки $\hat{T}^*$) стационарный режим устанавливается при $Tm \geq 400$ ед. времени. При этом отклонение значений $\hat{M}(\hat{T}^*)$ от истинного значения параметра $T^* = 1$ вполне удовлетворительное.

Второй статистический эксперимент (исследование влияния параметра T^* на качество оценок). Второй статистический эксперимент организован аналогично первому и поставлен при фиксированном времени моделирования $Tm = 500$ ед. времени, что соответствует времени установления стационарного режима, и при тех же значениях параметров потока, что и первый статистический эксперимент, за исключением значений T^* . Сначала второй статистический эксперимент реализуется для $T^* = 1$, затем для $T^* = 2, \dots$, затем для $T^* = 5$. Ниже приведены результаты второго статистического эксперимента ($N = 100$, $Tm = 500$) (табл. 2).

Таблица 2

Результаты второго статистического эксперимента

T^*	1	2	3	4	5
$\hat{M}(\hat{T}^*)$	0,9989	2,0005	3,0002	3,9825	5,0059
$\hat{V}(\hat{T}^*)$	0,0007	0,0028	0,0046	0,0079	0,0116

Анализ численных результатов показывает, что в смысле введенного критерия $\hat{V}(\hat{T}^*)$ увеличение параметра T^* отрицательно сказывается на качестве оценок $\hat{T}^*$, что является вполне естественным: увеличение параметра T^* приводит к увеличению числа потерянных событий исходного потока. Отклонение значений $\hat{M}(\hat{T}^*)$ от истинных значений параметра T^* вполне удовлетворительное.

3. Приближенная МП-оценка параметра T^* в особом случае

Рассматриваемый в настоящем разделе особый случай соответствует ситуации, когда $z_1 = z_2$, т.е. в (1) реализуется деление на ноль. Этот особый случай возможен тогда, когда для параметров потока выполняется соотношение: $(\lambda_1 - \lambda_2 + \alpha_1 - \alpha_2)^2 + 4\alpha_1\alpha_2(1-p)(1-q) = 0$. В свою очередь, выполнение последнего соотношения возможно в трех вариантах: 1) $\lambda_1 + \alpha_1 = \lambda_2 + \alpha_2$, $p = 1$, 2) $\lambda_1 + \alpha_1 = \lambda_2 + \alpha_2$, $q = 1$, 3) $\lambda_1 + \alpha_1 = \lambda_2 + \alpha_2$, $p = q = 1$.

Рассмотрим первый вариант. В [29] при $\lambda_1 + \alpha_1 = \lambda_2 + \alpha_2$, $p = 1$, $\delta = 0$ ($\lambda_1\lambda_2 - q\alpha_1\alpha_2 = 0$) приведена формула для $p(\tau|T)$:

$$p(\tau|T) = 0, \quad 0 \leq \tau < T, \\ p(\tau|T) = \left\{ \lambda_1 + \alpha_1 - \frac{\alpha_1\alpha_2(1-q)}{\alpha_1 + \alpha_2} \left[1 - \frac{\lambda_1 - \alpha_2}{\lambda_1 + \alpha_1} e^{-(\alpha_1 + \alpha_2)T} \right] \left[1 - (\lambda_1 + \alpha_1)(\tau - T) \right] \right\} e^{-(\lambda_1 + \alpha_1)(\tau - T)}, \quad \tau \geq T. \quad (13)$$

Тогда, используя (2), находим

$$p_1(\tau) = \frac{1}{T^*} \{ 1 - 2e^{-(\lambda_1 + \alpha_1)\tau} + e^{-(\alpha_1 + \alpha_2)\tau} \}, \quad 0 \leq \tau < T^*; \quad (14)$$

$$p_2(\tau) = \frac{1}{T^*} e^{-(\lambda_1 + \alpha_1)\tau} \left\{ -2 + e^{(\lambda_1 + \alpha_1)T^*} + e^{(\lambda_1 - \alpha_2)T^*} + \frac{(\lambda_1 + \alpha_1)(\lambda_1 - \alpha_2)}{(\alpha_1 + \alpha_2)} (\tau - T^*) \times \right. \\ \left. \times \left[e^{(\lambda_1 - \alpha_2)T^*} - e^{(\lambda_1 + \alpha_1)T^*} \right] \right\}, \quad \tau \geq T^*. \quad (15)$$

В точке $\tau = T^*$ имеем: 1) $p_1(T^*) = p_2(T^*)$, 2) $p'_1(T^*) \neq p'_2(T^*)$, т.е. плотность $p(\tau)$ в точке $\tau = T^*$ непрерывна, но имеет излом. Дальнейшее изложение идентично изложению раздела 2.

По-прежнему обозначим $p(\tau^{(k)}|T^*)$ – плотность вероятности, определенная формулами (2), (14), (15), в которых $\tau = \tau^{(k)}$ ($\tau^{(k)}$ – измерения), T^* – переменная величина ($T^* > 0$). Для введенной функции $p(\tau^{(k)}|T^*)$ переменной T^* справедлива теорема раздела 2 и вытекающая из теоремы процедура нахождения приближенной МП-оценки $\hat{T}^*$ параметра T^* по методу максимального правдоподобия.

Второй вариант соотношения параметров $\lambda_1 + \alpha_1 = \lambda_2 + \alpha_2$, $q = 1$ симметричен рассмотренному: нужно только в формуле (13) заменить q на p , что в конечном итоге приводит к формулам (14), (15).

Третий вариант соотношения параметров $\lambda_1 + \alpha_1 = \lambda_2 + \alpha_2$, $p = q = 1$, при котором $\lambda_1 - \alpha_2 = 0$, приводит к формулам

$$p_1(\tau) = \frac{1}{T^*} [1 - e^{-(\alpha_1 + \alpha_2)\tau}], \quad 0 \leq \tau < T^*, \\ p_2(\tau) = \frac{1}{T^*} e^{-(\alpha_1 + \alpha_2)\tau} [-1 + e^{(\alpha_1 + \alpha_2)T^*}], \quad \tau \geq T^*.$$

Для второго и третьего вариантов алгоритмы нахождения приближенной МП-оценки $\hat{T}^*$ параметра T^* по методу максимального правдоподобия идентичны алгоритму раздела 2.

С целью установления качества получаемых методом максимального правдоподобия (в особом случае) оценок параметра T^* поставлены статистические эксперименты. Статистические эксперименты поставлены для первого варианта соотношения параметров $\lambda_1 + \alpha_1 = \lambda_2 + \alpha_2$, $p = 1$.

Первый и второй статистические эксперименты описаны в разделе 2. В табл. 3, 4 приведены численные результаты для величин $\hat{M}(\hat{T}^*)$, $\hat{V}(\hat{T}^*)$, вычисленные по формулам (12). При проведении первого статистического эксперимента выбраны следующие параметры имитационной модели: $\lambda_1 = 3$, $\lambda_2 = 2$, $\alpha_1 = 3$, $\alpha_2 = \lambda_1 - \lambda_2 + \alpha_1 = 4$, $p = 1$, $q = \lambda_1 \lambda_2 / \alpha_1 \alpha_2 = 0,5$, $T^* = 1$ ($N = 100$; $Tm = 100, 200, \dots, 1\,000$).

Таблица 3

Результаты первого статистического эксперимента

Tm	100	200	300	400	500	600	700	800	900	1 000
$\hat{M}(\hat{T}^*)$	0,9947	1,0000	1,0017	1,0025	0,9986	0,9978	0,9985	0,9971	1,0004	1,0004
$\hat{V}(\hat{T}^*)$	0,0019	0,0007	0,0006	0,0004	0,0003	0,0003	0,0002	0,0002	0,0002	0,0002

Анализ численных результатов (как и в общем случае) показывает, что в смысле введенного критерия (выборочная вариация оценки T^*) стационарный режим устанавливается при $Tm \geq 500$ ед. времени. Отклонение выборочного среднего $\hat{M}(\hat{T}^*)$ от истинного значения параметра $T^* = 1$ вполне удовлетворительное.

Второй статистический эксперимент поставлен при фиксированном времени моделирования $Tm = 500$ ед. времени, что соответствует времени установления стационарного режима, и при тех же значениях параметров потока, что и первый статистический эксперимент, за исключением значений T^* . Сначала второй статистический эксперимент реализуется для $T^* = 1$, затем для $T^* = 2, \dots$, затем для $T^* = 5$ ($N = 100$).

Таблица 4

Результаты второго статистического эксперимента

T^*	1	2	3	4	5
$\hat{M}(\hat{T}^*)$	1,0018	1,9927	2,9974	3,9916	4,9965
$\hat{V}(\hat{T}^*)$	0,0003	0,0009	0,0029	0,0044	0,0085

Анализ численных результатов показывает (как и в общем случае), что в смысле введенного критерия $\hat{V}(\hat{T}^*)$ увеличение параметра T^* отрицательно сказывается на качестве оценок $\hat{T}^*$: увеличение параметра T^* приводит к увеличению числа потерянных событий исходного потока. Отклонение же значений $\hat{M}(\hat{T}^*)$ от истинных значений параметра T^* вполне удовлетворительное.

Заключение

По результатам проведенного исследования (для общего и особого случаев) можно сделать следующие выводы:

- 1) получен явный вид (7) функции правдоподобия для оценивания параметра T^* равномерного распределения длительности непродлевающегося случайного мертвого времени;
- 2) показано, что точка глобального максимума функции правдоподобия (7) принадлежит отрезку $[\tau^{(1)}, \tau^{(n)}]$;
- 3) предложена процедура отыскания точки, подозрительной на глобальный максимум функции правдоподобия (7); эта точка выбирается в качестве значения оценки параметра T^* ;
- 4) результаты имитационного моделирования показывают, что качество оценок в смысле введенного критерия (выборочная вариация оценки $\hat{T}^*$) вполне удовлетворительное; при этом смещение оценок относительно истинного значения параметра T^* невелико.

ЛИТЕРАТУРА

1. Cox D.R. The analysis of non-Markovian stochastic processes by the inclusion of supplementary variables // *Proceedings of the Cambridge Philosophical Society*. 1955. V. 51, is. 3. P. 433–441.
2. Kingman J.F.C. On doubly stochastic Poisson process // *Proceedings of the Cambridge Philosophical Society*. 1964. V. 60, is. 4. P. 923–930.
3. Башарин Г.П., Кокотушкин В.А., Наумов В.А. О методе эквивалентных замен расчета фрагментов сетей связи. Ч. 1 // *Известия АН СССР. Техническая кибернетика*. 1979. № 6. С. 92–99.
4. Башарин Г.П., Кокотушкин В.А., Наумов В.А. О методе эквивалентных замен расчёта фрагментов сетей связи. Ч. 2 // *Известия АН СССР. Техническая кибернетика*. 1980. № 1. С. 55–61.
5. Neuts M.F. A versatile Markovian point process // *Journal of Applied Probability*. 1979. V. 16, № 4. P. 764–779.
6. Lucantoni D.M. New results on single server with a bath Markovian arrival process // *Communications in Statistics Stochastic Models*. 1991. V. 7, № 1. P. 1–46.
7. Башарин Г.П., Самуйлов К.Е., Яркина Н.В., Гудкова Н.А. Новый этап развития математической теории телетрафика // *Автоматика и телемеханика*. 2009. № 12. С. 16–28.
8. Basharin G.P., Gaidamaka Y.V., Samouylov K.E. Mathematical Theory of Teletraffic and Its Application to the Analysis of Multi-service Communication of Next Generation Networks // *Automatic Control and Computer Sciences*. 2013. V. 47, № 2. P. 62–69.
9. Вишневский В.М., Ляхов А.И. Оценка пропускной способности локальной беспроводной сети при высокой нагрузке и помехах // *Автоматика и телемеханика*. 2001. № 8. С. 81–96.
10. Vishnevsky V.M., Larionov A.A., Smolnikov R.V. Optimization of topological structure of broadband wireless networks along the long traffic routes // *Distributed Computer and Communications Networks: Control, Computation, Communications : proc. of the 18th Int. Scientific Conf. (DCCN–2015) (Moscow, 19–22 October 2015)*. M. : ICS RAS, 2015. P. 27–35.
11. Вишневский В.М., Ларионов А.А. Открытая сеть массового обслуживания с коррелированными входными потоками для оценки производительности широкополосных беспроводных сетей // *Информационные технологии и математическое моделирование (ИТММ–2016) : материалы XV Междунар. конф. Катунь, 12–16 сентября 2016. Томск : Изд-во ТГУ, 2016. Ч. 1. С. 36–50.*
12. Апанасович В.В., Коляда А.А., Чернявский А.Ф. Статистический анализ случайных потоков в физическом эксперименте. Минск : Университетское, 1988. 256 с.
13. Баруча-Рид А.Т. Элементы теории марковских процессов и их приложения. М. : Наука, 1969. 512 с.
14. Вишневский В.М., Дудин А.Н., Клименок В.И. Стохастические системы с коррелированными потоками. Теория и применение в телекоммуникационных сетях. М. : Техносфера, 2018. 564 с.
15. Леонова М.А., Нежелская Л.А. Вероятность ошибки при оценивании состояний обобщенного асинхронного потока событий // *Вестник Томского государственного университета. Управление, вычислительная техника и информатика*. 2012. № 2 (19). С. 88–101.
16. Горцев А.М., Леонова М.А., Нежелская Л.А. Сравнение МП- и ММ-оценок длительности мертвого времени в обобщенном асинхронном потоке событий // *Вестник Томского государственного университета. Управление, вычислительная техника и информатика*. 2013. № 4 (25). С. 32–42.
17. Горцев А.М., Калягин А.А., Нежелская Л.А. Оценка максимального правдоподобия длительности мертвого времени в обобщенном полусинхронном потоке // *Вестник Томского государственного университета. Управление, вычислительная техника и информатика*. 2015. № 1 (30). С. 27–37.
18. Nezhe'skaya L. Probability density function for modulated MAP event flows with unextendable dead time // *Communications in Computer and Information Sciences*. 2015. V. 564. P. 141–151.
19. Nezhe'skaya L., Tumashkina D. Optimal state estimation of semi-synchronous event flow of the second order under its complete observability // *Communications in Computer and Information Sciences*. 2018. V. 912. P. 93–105.
20. Gortsev A.M., Solov'ev A.A. Joint probability density of interarrival interval of a flow of a physical events with unextendable dead time period // *Russian Physics Journal*. 2014. V. 57, is. 7. P. 973–983.
21. Gortsev A.M., Klimov I.S. Estimation of intensity of Poisson stream of event for conditions under which it is partially unobservable // *Telecommunications and Radio Engineering*. 1992. V. 47, is. 1. P. 33–38.
22. Васильева Л.А. Оценивание параметров дважды стохастического потока событий в условиях присутствия мертвого времени // *Вестник Томского государственного университета*. 2002. № S1-1. С. 9–13.
23. Горцев А.М., Завгородняя М.Е. Оценивание параметра непродлевающегося мертвого времени случайной длительности в пуассоновском потоке событий // *Вестник Томского государственного университета. Управление, вычислительная техника и информатика*. 2017. № 40. С. 32–40.
24. Глухова Е.В., Терпугов А.Ф. Оценка интенсивности пуассоновского потока событий при наличии продлевающегося мёртвого времени // *Известия вузов. Физика*. 1995. Т. 38, № 3. С. 22–31.
25. Горцев А.М., Зуевич В.Л. Оптимальная оценка параметров асинхронного дважды стохастического потока событий с произвольным числом состояний // *Вестник Томского государственного университета. Управление, вычислительная техника и информатика*. 2010. № 2 (11). С. 44–65.
26. Nezhe'skaya L.A., Pershina A.A. Estimate of the parameter of unextendable random dead time in a recurrent generalized asynchronous flow of physical events // *Russian Physics Journal*. 2020. V. 63, is. 1. P. 99–104.

27. Нежелская Л.А., Першина А.А. Процедура оценивания параметра равномерного распределения длительности непродлевающегося случайного мёртвого времени в рекуррентном обобщенном асинхронном потоке событий в особом случае // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2021. № 54. С. 65–73.
28. Нежелская Л.А., Першина А.А. Оценивание параметра непродлевающегося мёртвого времени случайной длительности в обобщенном асинхронном потоке событий // Информационные технологии и математическое моделирование (ИТММ-2019) : материалы XVIII Междунар. конф., 26–30 июня 2019 г. Томск : Изд-во НТЛ, 2019. С. 352–357.
29. Горцев А.М., Леонова М.А., Нежелская Л.А. Совместная плотность вероятностей длительности интервалов обобщенного асинхронного потока событий при непродлеваемом мертвом времени // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2012. № 4 (21). С. 14–25.
30. Горцев А.М., Нежелская Л.А. Оценивание параметров синхронного дважды стохастического потока событий методом моментов // Вестник Томского государственного университета. 2002. № S1-1. С. 24–29.
31. Шуленин В.П. Математическая статистика. Томск : Изд-во НТЛ, 2012. Ч. 1. 540 с.
32. Малинковский Ю.В. Теория вероятностей и математическая статистика. Гомель : ГТУ им. Ф. Скорины, 2004. Ч. 2: Математическая статистика. 146 с.
33. Лифшиц А.Л., Малыц Э.А. Статистическое моделирование систем массового обслуживания. М. : Сов. радио, 1978. 248 с.
34. Соболев И.М. Численные методы Монте-Карло. М. : Наука, 1973. 312 с.

Поступила в редакцию 6 февраля 2021 г.

Nezhel'skaya L.A, Pershina A.A. (2021) ESTIMATION OF THE UNIFORM DISTRIBUTION PARAMETER OF UNEXTENDABLE DEAD TIME DURATION IN A GENERALIZED RECURRENT ASYNCHRONOUS FLOW OF EVENTS BY THE MAXIMUM LIKELIHOOD METHOD. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika* [Tomsk State University Journal of Control and Computer Science] 55. pp. 53–64

DOI: 10.17223/19988605/55/7

A generalized recurrent asynchronous flow of events (a generalized MMPP flow) is considered, which is a common mathematical model of a flow of elementary particles, information flows of applications operating in telecommunication and information-computer communication networks, and belongs to the class of doubly stochastic flows of events. The functioning of the flow is considered under the conditions of a random unextendable dead time distributed according to the uniform law on the interval $[0, T^*]$. A special and common case are considered. The dead time parameter T^* is estimated using the likelihood method. The results of statistical experiments are presented (in special and common cases). An analysis of the numerical results shows that, in the sense of the introduced criterion, an increase in the parameter T^* negatively affects the quality of the estimates, which is quite natural: an increase in the parameter T^* leads to an increase in the number of lost events in the initial stream.

Based on the results of the study, the following conclusions can be drawn the results of simulation modeling show that the quality of the assessments in the sense of the introduced criterion (selective variation of the assessment) is quite satisfactory; in this case, the bias of the estimates relative to the true value of the parameter T^* does not exceed hundredth values.

Keywords: recurrent generalized asynchronous flow of events; unextendable random dead time; parameter estimation; method of moments.

NEZHEL'SKAYA Lyudmila Alekseevna (Doctor of Physical and Mathematical Sciences, Professor of Applied Mathematics Department of the Institute of Applied Mathematics and Computer Science, National Research Tomsk State University, Russian Federation).
E-mail: ludne@mail.tsu.ru

PERSHINA Anna Alexandrovna (Post-graduate Student of the Institute of Applied Mathematics and Computer Science, National Research Tomsk State University, Russian Federation).
E-mail: ann.shitina@gmail.com

REFERENCES

1. Cox, D.R. (1955) The analysis of non-Markovian stochastic processes by the inclusion of supplementary variables. *Proceedings of the Cambridge Philosophical Society*. 51(3). pp. 433–441. DOI: 10.1017/S0305004100030437
2. Kingman, J.F.C. (1964) On doubly stochastic Poisson process. *Proceedings of the Cambridge Philosophical Society*. 60(4). pp. 923–930. DOI: 10.1017/S030500410003838X
3. Basharin, G.P., Kokotushkin, V.A. & Naumov, V.A. (1979) O metode ekvivalentnykh zamen rascheta fragmentov setey svyazi. Ch. 1 [On the equivalent substitutions method for computing fragments of communication networks. Ch. 1]. *Izvestiya Akademii Nauk USSR. Tekhnicheskaya kibernetika*. 6. pp. 92–99.
4. Basharin, G.P., Kokotushkin, V.A. & Naumov, V.A. (1980) O metode ekvivalentnykh zamen rascheta fragmentov setey svyazi. Ch. 2 [On the equivalent substitutions method for computing fragments of communication networks. Ch.2]. *Izvestiya Akademii Nauk USSR. Tekhnicheskaya kibernetika*. 1. pp. 55–61.

5. Neuts, M.F. (1979) A versatile Markovian point process. *Journal of Applied Probability*. 16(4). pp. 764–779. DOI: 10.2307/3213143
6. Lucantoni, D.M. (1991) New results on the single server queue with a batch markovian arrival process. *Communications in Statistics Stochastic Models*. 7(1). pp. 1–46. DOI: 10.1080/15326349108807174
7. Basharin, G.P., Samouylov, K.E., Yarkina, N.V. & Gudkova N.A. (2009) A new stage in the development of the mathematical theory of teletraffic. *Automation and Telemekhanics*. 12. pp. 16–28. DOI: 10.1134/S0005117909120030
8. Basharin, G.P., Gaidamaka, Y.V. & Samouylov, K.E. (2013) Mathematical Theory of Teletraffic and Its Application to the Analysis of Multiservice Communication of Next Generation Networks. *Automatic Control and Computer Sciences*. 47(2). pp. 62–69. DOI: 10.3103/S0146411613020028
9. Vishnevsky, V.M. & Lyakhov A.I. (2001) A Wireless LAN under High Load and in Noise: Throughput Evaluation. *Automation and Remote Control*. 8. pp. 81–96. DOI: 10.1023/A:1010257612786
10. Vishnevsky, V.M., Larionov, A.A. & Smolnikov, R.V. (2015) Optimization of topological structure of broadband wireless networks along the long traffic routes. *Distributed Computer and Communication Networks: Control, Computation, Communications*. Proc. of the 18th International Conference (DCCN–2015). Moscow, October 19–22. Moscow: ICS RAS. pp. 27–35.
11. Vishnevsky, V.M. & Larionov, A.A. (2016) Otkrytaya set' massovogo obsluzhivaniya s korrelirovannymi vkhodnymi potokami dlya otsenki proizvoditel'nosti shirokopolosnykh besprovodnykh setey [Open queueing network with correlated input flows for estimating the performance of broadband wireless networks]. *Informatsionnye tekhnologii i matematicheskoe modelirovanie (ITMM–2016)* [Information Technologies and Mathematical Modelling (ITMM 2016)]. Proc. of the 15th International Conference. September 12–16. Tomsk. pp. 36–50.
12. Apanasovich, V.V., Kolyada, A.A. & Chernyavskiy, A.F. (1988) *Statisticheskii analiz sluchaynykh potokov v fizicheskom eksperimente* [The statistical analysis of series of random events in physical experiment]. Minsk: University Press.
13. Barucha-Rid, A.T. (1969) *Elementy teorii markovskikh protsessov i ikh prilozheniya* [Elements of the theory of Markov processes and their applications]. Translated from English by V.V. Kalashnikov. Moscow: Nauka.
14. Vishnevsky, V.M., Dudin, A.N. & Klimenok, V.N. (2018) *Stokhasticheskie sistemy s korrelirovannymi potokami. Teoriya i primeneniye v telekommunikatsionnykh setyakh* [Stochastic systems with correlated flows. Theory and application in telecommunication networks]. Moscow: Tekhnosfera.
15. Leonova, M.A. & Nezhelskaya, L.A. (2012) Error probability when evaluating states of a generalized asynchronous flow of events. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 2(19). pp. 88–101.
16. Gortsev, A.M., Leonova, M.A. & Nezhelskaya, L.A. (2013) Comparison of ML- and MM- estimates of the duration of dead time in a generalized asynchronous flow of events. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 4(25). pp. 32–42.
17. Gortsev, A.M., Kalyagina, A.A. & Nezhelskaya, L.A. (2015) Estimation of the maximum likelihood of the duration of dead time in a generalized semisynchronous flow. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 1(30). pp. 27–37.
18. Nezhelskaya, L. (2015) Probability density function for modulated map event flows with unextendable dead time. *Communications in Computer and Information Sciences*. 564. pp. 141–151.
19. Nezhelskaya, L. & Tumashkina, D. (2018) Optimal state estimation of semi-synchronous event flow of the second order under its complete observability. *Communications in Computer and Information Sciences*. 912. pp. 93–105.
20. Gortsev, A.M. & Soloviev, A.A. (2014) Joint probability density of interarrival interval of a flow of a physical events with unextendable dead time period. *Russian Physics Journal*. 57(7). pp. 973–983. DOI: 10.1007/s11182-014-0333-4
21. Gortsev, A.M. & Klimov, I.S. (1992) Estimation of intensity of Poisson stream of event for conditions under which it is partially unobservable. *Telecommunications and Radio Engineering*. 47(1). pp. 33–38.
22. Vasilieva, L.A. (2002) Otsenivanie parametrov dvazhdy stokhasticheskogo potoka sobytiy v usloviyakh prisutstviya mertvogo vremeni [The abstract of clause estimation of parameters twice-stochastic flow of events in conditions of presence of dead time]. *Vestnik Tomskogo gosudarstvennogo universiteta – Tomsk State University Journal*. s1-1. pp. 9–13.
23. Gortsev, A.M. & Zavgorodnyaya, M.E. (2017) Estimation of the parameter of unextendable dead time random duration in the Poisson flow of events. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravleniye, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 40. pp. 32–40. DOI: 10.17223/19988605/40/4
24. Glukhova, E.V. & Terpugov, A.F. (1995) Otsenka intensivnosti puassonovskogo potoka sobytiy pri nalichii prodlevayushchegosya mertvogo vremeni [Estimation of the intensity of the Poisson flow of events in the presence of prolonged dead time]. *Izvestiya vuzov. Fizika*. 38(3). pp. 22–31.
25. Gortsev, A.M. & Zuevich, V.L. (2010) Optimal estimation of parameters of an asynchronous doubly stochastic flow of events with an arbitrary number of states. *Vestnik Tomskogo gosudarstvennogo universiteta, Upravleniye, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 2(11). pp. 44–65.
26. Nezhelskaya, L.A. & Pershina, A.A. (2020) Estimation of the parameter of unextendable random dead time in a recurrent generalized asynchronous flow of physical events. *Russian Physics Journal*. 63(1). pp. 99–104.
27. Nezhelskaya, L.A. & Pershina, A.A. (2021) Estimation procedure of the uniform distribution parameter of unextendable dead time duration in a generalized recurrent asynchronous flow of events in special case. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 54. pp. 65–73. DOI: 10.17223/19988605/54/8

28. Nezhelskaya, L.A. & Pershina, A.A. (2019) Otsenivanie parametra neprodlevayushchegosya mertvogo vremeni sluchaynoy dlitel'nosti v obobshchennom asinkhronnom potoke sobytiy [Estimation of the unextendable dead time random duration parameter in a generalized asynchronous flow of events]. *Informatsionnye tekhnologii i matematicheskoe modelirovanie (ITMM-2019)* [Information Technologies and Mathematical Modelling (ITMM-2019)]. Proc. of the 18th International Conference. Tomsk, June 26–30. Tomsk: NTL. pp. 352–357.
29. Gortsev, A.M., Leonova, M.A. & Nezhelskaya, L.A. (2012) Joint probability density of the durations of the intervals of the generalized asynchronous flow of events with non-prolonging dead time. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 4(21). pp. 14–25.
30. Gortsev, A.M. & Nezhelskaya, L.A. (2002) Estimation of parameters of a synchronous doubly stochastic flow of events by the method of moments. *Vestnik Tomskogo gosudarstvennogo universiteta – Tomsk State University Journal*. S1-1. pp. 24–29.
31. Shulenin, V.P. (2011) *Matematicheskaya statistika* [Mathematical statistics]. Tomsk: NTL Publ.
32. Malinkovsky, Yu.V. (2004) *Teoriya veroyatnostey i matematicheskaya statistika* [Probability Theory and Mathematical Statistics]. Gomel: Francisk Skorina Gomel State University.
33. Lifshits, A.L. & Maltz, E.A. (1978) *Statisticheskoe modelirovanie sistem massovogo obsluzhivaniya* [Statistical modeling of system queuing]. Moscow: Sovetskoe radio.
34. Sobol, I.M. (1973) *Chislennyye metody Monte-Karlo* [Numerical methods of Monte Carlo]. Moscow: Nauka.

УДК 62-50

DOI: 10.17223/19988605/55/8

A.I. Rouban

**THE SENSITIVITY COEFFICIENTS FOR MULTIVARIATE DYNAMIC SYSTEMS
DESCRIBED BY ORDINARY DIFFERENCE EQUATIONS
WITH VARIABLE AND CONSTANT PARAMETERS**

The variational method of calculation of sensitivity coefficients connecting first variation of quality functionals with variations of variable and constant parameters for multivariate non-linear dynamic systems described by ordinary difference equations is developed. Sensitivity coefficients are components of sensitivity functionals and they are before a variations of variable and constant parameters. The base of calculation is the decision of corresponding difference conjugate equations for Lagrange's multipliers in the opposite direction of discrete time.

Keywords: variational method; sensitivity coefficient; difference equation; conjugate equation; Lagrange's multiplier.

The sensitivity functional connect the first variation of quality functional with variations of variable and constant parameters and the sensitivity coefficients (SC) are components of vector gradient from quality functional according to parameters.

The problem of calculation of SC for dynamic systems is principal in the analysis and syntheses of control laws, identification, optimization, stability [1–16]. The first-order sensitivity characteristics are mostly used. Later on we shall examine only SC of the first-order.

Consider a vector output $x(t)$ of dynamic object model under discrete time $t \in [0, 1, \dots, N + 1]$ implicitly depending on vectors parameters α and functional $I(\alpha)$ constructed on $x(t)$ under $t \in [0, 1, \dots, N + 1]$ and on α :

$$I(\alpha) = \sum_{t=0}^{N+1} f_0(x(t), \alpha, t).$$

SC with respect to constant parameters α are called a gradient of $I(\alpha)$ on α : $(dI(\alpha)/d\alpha)^T \equiv \nabla_{\alpha} I(\alpha)$. SC are a coefficients of single-line relationship between the first variation of functional $\delta_{\alpha} I(\alpha)$ and the variations $\delta\alpha$ of constant parameters α :

$$\delta_{\alpha} I(\alpha) = (\nabla_{\alpha} I(\alpha))^T \delta\alpha = \frac{dI(\alpha)}{d\alpha} \delta\alpha \equiv \sum_{j=1}^m \frac{\partial I(\alpha)}{\partial \alpha_j} \delta\alpha_j.$$

The direct method of SC calculation (by means of the differentiation of quality functional with respect to constant parameters) inevitably requires a solution of cumbersome sensitivity equations to sensitivity functions $W(t)$. $W(t)$ is the matrix of single-line relationship of the first variation of dynamic model output with parameter variations: $\delta x(t) = W(t) \delta\alpha$. For instance, for functional $I(\alpha)$ we have following SC vector (row vector):

$$\frac{dI(\alpha)}{d\alpha} = \sum_{t=0}^{N+1} \left[\frac{\partial f_0(x(t), \alpha, t)}{\partial x(t)} W(t) + \frac{\partial f_0(x(t), \alpha, t)}{\partial \alpha} \right].$$

For obtaining the matrix $W(t)$ it is necessary to decide a bulky system equations – sensitivity equations. The j -th column of matrix $W(t)$ is made of the sensitivity functions $dx(t)/d\alpha_j$ with respect to component α_j of vector α . They satisfy a vector equation (if x is a vector) resulting from dynamic model (for x) by derivation on a parameter α_j .

For variable parameters such method essentially becomes complicated and practically is not applicable.

At a choice of good initial constant parameters at identification of objects and also at consecutive calculation of control actions on object often apply a gradient algorithms. It appears that for calculation of components of a gradient from an optimized functional to required variables and constant parameters, it is convenient to apply the conjugate equations (in relation to the dynamic equations of object).

Variational method [6], ascending to Lagrange's, Hamilton's, Euler's memoirs, makes possible to simplify the process of determination of conjugate equations and formulas of account of SC. On the basis of this method it is an extension of quality functional by means of inclusion into it dynamic equations of object by means of Lagrange's multipliers and obtaining the first variation of extended functional on phase coordinates of object and on interesting parameters. Dynamic equations for Lagrange's multipliers are obtained due to set equal to a zero (in the first variation of extended functional) the functions before the first variations of phase coordinates. Given simplification first variation of extended functional brings at presence in the right part only parameter variations, i.e. it is got the sensitivity functional concerning parameters.

In difference from other papers devoted to calculation of SC in given paper the generalized difference models are used. Thus variables and constant parameters enter into the right parts of ordinary difference equations of dynamic object, in an indicator of quality of system work, in the measuring device model and initial values of phase coordinates depend on constant parameters. It is proved that both methods to calculation of SC (with use of Lagrange's functions or with use of sensitivity functions) yield the same result, but the first method it is essential more simple in the computing relation.

1. Problem statement

We suppose that the dynamic object is described by system of non-linear ordinary difference equations [13]

$$x(t+1) = f(x(t), \tilde{\alpha}(t), \alpha, t), \quad t = 0, 1, 2, \dots, N, \quad x(0) = x_0(\alpha). \quad (1)$$

Here: $\tilde{\alpha}(t)$, α are a vector-columns of interesting variable and constant parameters; x is a vector-column of phase coordinates; $f(\cdot)$ is known continuously differentiated limited vector-functions.

The quality of functioning of system it is characterised of functional

$$I(\tilde{\alpha}, \alpha) = \sum_{t=0}^N f_0(x(t), \tilde{\alpha}(t), \alpha, t) + f_0(x(N+1), \tilde{\alpha}(N+1), \alpha, N+1) \quad (2)$$

depending on $\tilde{\alpha}(t)$ and α . The conditions for function $f_0(\cdot)$, $I(\cdot)$ are the same as for $f(\cdot)$. With use of a functional (2) the optimization problem (in the theory of optimal control) are named as the Bolts's problem. From it as the individual variants follow: Lagrange's problem (when there is only the first group of the summands) and Mayer's problem (when there is only the last summand – function from phase coordinates at a finishing point).

With the purpose of simplification of appropriate deductions with preservation of a generality in all transformations (1), (2) there are two vectors of parameters $\tilde{\alpha}(t)$, α . If in the equations (1), (2) parameters are different then it is possible formally to unit them in two vectors $\tilde{\alpha}(t)$, α , to use obtained outcomes and then to make appropriate simplifications, taking into account a structure of a vectors $\tilde{\alpha}(t)$, α .

It shown also that the variation method without basic modifications allows to receive SC in relation to variable and constant parameters:

$$\begin{aligned} \delta I(\tilde{\alpha}(t), \alpha) &= \sum_{t=0}^{N+1} \frac{\partial I(\tilde{\alpha}(t), \alpha)}{\partial \tilde{\alpha}(t)} \delta \tilde{\alpha}(t) + \frac{\partial I(\tilde{\alpha}(t), \alpha)}{\partial \alpha} \delta \alpha. \\ \nabla_{\tilde{\alpha}(t)} I(\tilde{\alpha}, \alpha) &= \left(\frac{\partial I(\tilde{\alpha}, \alpha)}{\partial \tilde{\alpha}_1(t)} \quad \dots \quad \frac{\partial I(\tilde{\alpha}, \alpha)}{\partial \tilde{\alpha}_{m_1}(t)} \right)^T, \quad t = 0, 1, 2, \dots, N, N+1, \\ \nabla_{\alpha} I(\tilde{\alpha}, \alpha) &= \left(\frac{\partial I(\tilde{\alpha}, \alpha)}{\partial \alpha_1} \quad \dots \quad \frac{\partial I(\tilde{\alpha}, \alpha)}{\partial \alpha_{m_2}} \right)^T. \end{aligned} \quad (3)$$

By obtaining of results the obvious designations:

$$\begin{aligned} f(t) &\equiv f(x(t), \tilde{\alpha}(t), \alpha, t), \quad t = 0, 1, 2, \dots, N, \\ f_0(t) &\equiv f_0(x(t), \tilde{\alpha}(t), \alpha, t), \quad t = 0, 1, 2, \dots, N+1 \end{aligned} \quad (4)$$

are used.

The index t in functions $f(x(t), \tilde{\alpha}(t), \alpha, t)$ and $f_0(x(t), \tilde{\alpha}(t), \alpha, t)$ also reflects not only obvious dependence on step number, but also that the kind of functions from a step to a step can change.

Let's receive the conjugate equations for calculation of Lagrange's multipliers and on the basis of them formulas for calculation of SC.

2. Conjugate equations

To an initial indicator of optimality $I(\tilde{\alpha}, \alpha)$ the dynamic equations (1) (written down in the form of restrictions of type of equalities) by means of Lagrange's multipliers $\lambda(t)$ are added. The size of the expanded indicator of optimality always coincides with size initial functional on which judge an optimality of work of system. SC for both functionals coincide also – section 4 of given paper see. Then for the received expanded indicator the first variation leaves and the dynamic equations for $\lambda(t)$ from an additional condition that factors before variations of phase coordinates $\delta x(N+1), \dots, \delta x(1)$ addressed in a zero are worked out. Factors before variations of parameters $\delta \tilde{\alpha}(t), \delta \alpha$ in the first variation of the expanded indicator I represent required SC (3).

Complement a quality functional (2) by restrictions-equalities (1) by means of Lagrange's multipliers $\lambda(t)$, $t = 0, 1, 2, \dots, N+1$ (column vectors) and get the expended functional

$$\begin{aligned} I &= I(\tilde{\alpha}, \alpha) + \lambda^T(0)[-x(0) + x_0(\alpha)] + \sum_{t=0}^N \lambda^T(t+1)[-x(t+1) + f(t)] = \\ &= [-\lambda^T(N+1)x(N+1) + f_0(N+1)] + \\ &+ \sum_{t=0}^N [-\lambda^T(t)x(t) + \lambda^T(t+1)f(t) + f_0(t)] + \lambda^T(0)x_0(\alpha), \end{aligned} \quad (5)$$

which complies with $I(\tilde{\alpha}, \alpha)$ when (1) is fulfilled.

We calculate the first variation of extended functional, caused by a variation of phase coordinates, and also a variation of variables and constant parameters:

$$\delta I = \sum_{t=0}^{N+1} \frac{\partial I}{\partial x(t)} \delta x(t) + \sum_{t=0}^{N+1} \frac{\partial I}{\partial \tilde{\alpha}(t)} \delta \tilde{\alpha}(t) + \frac{\partial I}{\partial \alpha} \delta \alpha. \quad (6)$$

The factors standing in the formula (6) before variations of phase coordinates look like:

$$\begin{aligned} \frac{\partial I}{\partial x(N+1)} &= -\lambda^T(N+1) + \frac{\partial f_0(N+1)}{\partial x(N+1)}, \\ \frac{\partial I}{\partial x(t)} &= -\lambda^T(t) + \lambda^T(t+1) \frac{\partial f(t)}{\partial x(t)} + \frac{\partial f_0(t)}{\partial x(t)}, \quad t = N, N-1, \dots, 1, 0. \end{aligned} \quad (7)$$

We equate to their zero and it is received the conjugate equations for Lagrange's multipliers:

$$\begin{aligned} \lambda^T(N+1) &= \frac{\partial f_0(N+1)}{\partial x(N+1)}, \\ \lambda^T(t) &= \lambda^T(t+1) \frac{\partial f(t)}{\partial x(t)} + \frac{\partial f_0(t)}{\partial x(t)}, \quad t = N, N-1, \dots, 1, 0, \end{aligned} \quad (8)$$

These equations are decided in the opposite direction changes of an independent integer variable t .

3. Sensitivity coefficients

In the equation (6) SC for variables and constant parameters look like:

$$\begin{aligned} \frac{\partial I}{\partial \tilde{\alpha}(N+1)} &= \frac{\partial f_0(N+1)}{\partial \tilde{\alpha}(N+1)}, \\ \frac{\partial I}{\partial \tilde{\alpha}(t)} &= \lambda^T(t+1) \frac{\partial f(t)}{\partial \tilde{\alpha}(t)} + \frac{\partial f_0(t)}{\partial \tilde{\alpha}(t)}, \quad t = N, N-1, \dots, 1, 0, \\ \frac{\partial I}{\partial \alpha} &= \frac{\partial f_0(N+1)}{\partial \alpha} + \sum_{t=0}^N \left[\lambda^T(t+1) \frac{\partial f(t)}{\partial \alpha} + \frac{\partial f_0(t)}{\partial \alpha} \right] + \lambda^T(0) \frac{dx_0(\alpha)}{d\alpha}. \end{aligned} \quad (9)$$

This result is more common in relation to appropriate results of monograph [13].

4. Equivalence of sensitivity coefficient for initial (2) and expended (5) functionals

We take expanded functional, presented in an initial part of the formula (5):

$$I = I(\tilde{\alpha}, \alpha) + \lambda^T(0)[-x(0) + x_0(\alpha)] + \sum_{t=0}^N \lambda^T(t+1)[-x(t+1) + f(t)].$$

Before $\lambda^T(\cdot)$ in square brackets there are the dynamic equations of the object which has been written down in the form of the equation of equality type. Hence, values of functions in square brackets are always equal to zero.

Let's calculate from both parts of the previous equation derivatives in the beginning on a vector of constant parameters α :

$$\begin{aligned} \frac{\partial I}{\partial \alpha} &= \frac{\partial I(\tilde{\alpha}, \alpha)}{\partial \alpha} + \lambda^T(0) \left[-W_{\alpha}(0) + \frac{dx_0(\alpha)}{d\alpha} \right] + \\ &+ \sum_{t=0}^N \lambda^T(t+1) \left[-W_{\alpha}(t+1) + \frac{\partial f(t)}{\partial x(t)} W_{\alpha}(t) + \frac{\partial f(t)}{\partial \alpha} \right]. \end{aligned}$$

Before $\lambda^T(\cdot)$ now there are sensitivity equations for a matrix of sensitivity functions. These equations are written down as in the form of restriction of equality type. Values of functions in square brackets also are always equal to zero.

Hence, SC rather both for initial functional and for its expanded variant have identical values.

That the sensitivity equation had the specified appearance, it is necessary (1) to impose a condition of differentiability of $f(t)$ on phase coordinates and on considered parameters on the right member of equation of movement of dynamic object (1).

We receive the same result and for SC in relation to variable parameters. The sensitivity equations for each fixed value of argument of variable parameters $\tilde{\alpha}(j)$, $j = 0, 1, \dots, N+1$ look like:

$$\begin{aligned} W_{\tilde{\alpha}(j)}(t+1) &= \frac{\partial f(t)}{\partial x(t)} W_{\tilde{\alpha}(j)}(t), \quad t = j+1, j+2, \dots, N; \\ W_{\tilde{\alpha}(j)}(j+1) &= \frac{\partial f(j)}{\partial \tilde{\alpha}(j)}; \quad W_{\tilde{\alpha}(j)}(t) = 0, \quad t = 0, 1, \dots, j; \\ j &= 0, 1, \dots, N-1. \end{aligned}$$

At $j = N$

$$W_{\tilde{\alpha}(N)}(N+1) = \frac{\partial f(N)}{\partial \tilde{\alpha}(N)}; \quad W_{\tilde{\alpha}(N)}(t) = 0, \quad t = 0, 1, \dots, N.$$

5. Example

We suppose that directed by a problem (1), (2) there are only constant parameters. These parameters are initial values of phase coordinates in the equation (1). Then

$$\begin{aligned} f(t) &\equiv f(x(t), t), \quad t = 0, 1, 2, \dots, N, \quad x_0(\alpha) = \alpha, \\ f_0(t) &\equiv f_0(x(t), t), \quad t = 0, 1, 2, \dots, N+1. \end{aligned} \quad (10)$$

Let's calculate a vector of SC to parameters α . As $f(t)$ and $f_0(t)$ do not depend on α , that $\partial I / \partial \alpha \equiv dI / d\alpha = \lambda^T(0)$. From the decision of the conjugate equations (8) it is received SC:

$$\begin{aligned} \frac{dI}{d\alpha} = \lambda^T(0) &= \frac{\partial f_0(N+1)}{\partial x(N+1)} \frac{\partial f(N)}{\partial x(N)} \dots \frac{\partial f(0)}{\partial x(0)} + \frac{\partial f_0(N)}{\partial x(N)} \frac{\partial f(N-1)}{\partial x(N-1)} \dots \frac{\partial f(0)}{\partial x(0)} + \dots + \\ &+ \frac{\partial f_0(1)}{\partial x(1)} \frac{\partial f(0)}{\partial x(0)} + \frac{\partial f_0(0)}{\partial x(0)}. \end{aligned}$$

For check of correctness of the decision we calculate SC with use of sensitivity functions: $dx(t)/d\alpha \equiv W_\alpha(t)$:

$$\frac{dI(\alpha)}{d\alpha} = \sum_{t=0}^{N+1} \frac{\partial f_0(t)}{\partial x(t)} W_\alpha(t).$$

Sensitivity functions satisfy to the equations:

$$W_\alpha(t+1) = \frac{\partial f(t)}{\partial x(t)} W_\alpha(t), \quad t = 0, 1, 2, \dots, N, \quad W(0) = E.$$

Here E is an unitary matrix. The same result is received.

SC $dI/d\alpha$ (with use expanded functional and Lagrange's functions) and $dI(\alpha)/d\alpha$ (with use initial functional and sensitivity functions) coincide. With increase in dimension of a vector α the first approach has essential computing advantages.

6. The account of the measuring device model

At additional use of model of the measuring device it is necessary to make changes to problem statement:

$$\begin{aligned} f_0(t) &\equiv f_0(\eta(t), \tilde{\alpha}(t), \alpha, t), \quad t = 0, 1, 2, \dots, N+1; \\ \eta(t) &\equiv \eta(x(t), \tilde{\alpha}(t), \alpha, t), \quad t = 0, 1, 2, \dots, N+1. \end{aligned}$$

In the received results it is necessary to execute small replacements:

$$\begin{aligned} \frac{\partial f_0(t)}{\partial x(t)} &\text{ to replace by } \frac{\partial f_0(t)}{\partial \eta(t)} \frac{\partial \eta(t)}{\partial x(t)}; \\ \frac{\partial f_0(t)}{\partial \alpha(t)} &\text{ to replace by } \frac{\partial f_0(t)}{\partial \alpha(t)} + \frac{\partial f_0(t)}{\partial \eta(t)} \frac{\partial \eta(t)}{\partial \alpha(t)}; \\ \frac{\partial f_0(t)}{\partial \alpha} &\text{ to replace by } \frac{\partial f_0(t)}{\partial \alpha} + \frac{\partial f_0(t)}{\partial \eta(t)} \frac{\partial \eta(t)}{\partial \alpha}. \end{aligned}$$

Conclusion

Variational method is applicability for calculation of SC of multivariate non-linear dynamic systems described by ordinary difference equations. Variables and constant parameters are present at object model, at model of the measuring device and at generalized quality functional for system (the Bolts's problem).

In a basis of calculation of SC the decision of the difference equations of object model in a forward direction of time and obtained difference equations for Lagrange's multipliers in the opposite direction of time.

It is proved that both methods to calculation of SC (with use of Lagrange's functions or with use of sensitivity functions) yield the same result, but the first method it is essential more simple in the computing relation.

Variation method of calculation of SC allows to generalize it for objects described by vectorial difference equations with distributed memory for phase coordinates and variables parameters.

Results of present paper are applicable at design of high-precision systems and devices.

This paper continues research in [13, 16].

REFERENCES

1. Ostrovskiy, G.M. & Volin, Yu.M. (1967) *Methods of optimization of chemical reactors*. Moscow: Khimiya.
2. Rosenwasser, E.N. & Yusupov, R.M. (1969) *Sensitivity of Automatic Control Systems*. Leningrad: Energiya.
3. Krutyko, P.D. (1969) The decision of a identification problem by a sensitivity theory method. *News of Sciences Academy of the USSR. Technical Cybernetics*. 6. pp. 146–153.
4. Petrov, B.N. & Krutyko, P.D. (1970) Application of the sensitivity theory in automatic control problems. *News of Sciences Academy of the USSR. Technical Cybernetics*. 2. pp. 202–212.
5. Gorodetskiy, V.I., Zacharin, F.M., Rosenwasser, E.N. & Yusupov, R.M. (1971) *Methods of Sensitivity Theory in Automatic Control*. Leningrad: Energiya.
6. Bryson, A.E. & Ho, Ju-Chi. (1972) *Applied Theory of Optimal Control*. Moscow: Mir.
7. Speedy, C.B., Brown, R.F. & Goodwin, G.C. (1973) *Control Theory: Identification and Optimal Control*. Moscow: Mir.
8. Gekher, K. (1973) *Theory of Sensitivity and Tolerances of Electronic Circuits*. Moscow: Sovetskoe radio.
9. Ruban, A.I. (1975) *Nonlinear Dynamic Objects Identification on the Base of Sensitivity Algorithm*. Tomsk: Tomsk State University.
10. Rosenwasser, E.N. & Yusupov, R.M. (eds) (1977) *Cybernetics Problems. Sensitivity Theory and its Application*. Vol. 23. Moscow: Svyaz.
11. Voronov, A.A. (1979) *Stability, Controllability, Observability*. Moscow: Nauka.
12. Rosenwasser, E.N. & Yusupov, R.M. (1981) *Sensitivity of Control Systems*. Moscow: Nauka.
13. Ruban, A.I. (1982) *Identification and Sensitivity of Complex Systems*. Tomsk: Tomsk State University.
14. Haug, E.J., Choi, K.K. & Komkov, V. (1988) *Design Sensitivity Analysis of Structural Systems*. Moscow: Mir.
15. Afanasyev, V.N., Kolmanovskiy, V.B. & Nosov, V.R. (1998) *The Mathematical Theory of Designing of Control Systems*. Moscow: Vysshaya shkola.
16. Rouban, A.I. (2019) The sensitivity functionals in the Bolts's problem for multivariate dynamic systems described by integral equations with delay time. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 46. pp. 83–92. DOI: 10.17223/19988605/46/10

Received: February, 19, 2021

Rouban A.I. (2021) THE SENSITIVITY COEFFICIENTS FOR MULTIVARIATE DYNAMIC SYSTEMS DESCRIBED BY ORDINARY DIFFERENCE EQUATIONS WITH VARIABLE AND CONSTANT PARAMETERS. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika* [Tomsk State University Journal of Control and Computer Science]. 55. pp. 65–70

DOI: 10.17223/19988605/55/8

Рубан А.И. КОЭФФИЦИЕНТЫ ЧУВСТВИТЕЛЬНОСТИ ДЛЯ МНОГОМЕРНЫХ ДИНАМИЧЕСКИХ СИСТЕМ, ОПИСЫВАЕМЫХ ОБЫКНОВЕННЫМИ РАЗНОСТНЫМИ УРАВНЕНИЯМИ С ПЕРЕМЕННЫМИ И ПОСТОЯННЫМИ ПАРАМЕТРАМИ. *Вестник Томского государственного университета. Управление, вычислительная техника и информатика*. 2021. № 55. С. 65–70

Вариационный метод применен для расчета коэффициентов чувствительности, которые связывают первую вариацию функционалов качества работы систем с вариациями переменных и постоянных параметров для многомерных нелинейных динамических систем, описываемых обобщенными разностными уравнениями и обобщенным функционалом качества работы системы (функционалом Больца).

Ключевые слова: вариационный метод; коэффициент чувствительности; разностное уравнение; функционал качества работы системы; сопряженное уравнение; множитель Лагранжа.

ROUBAN Anatoly Ivanovich (Doktor of Technical Sciences, Professor of Computer Science Department of Institute of Space and Information Technologies, Siberian Federal University, Krasnoyarsk, Russian Federation).

E-mail: ai-rouban@mail.ru

УДК 519.234:621.391
DOI: 10.17223/19988605/55/9

И.М. Рудько

ОПТИМИЗАЦИЯ ТРАЕКТОРИИ ДВИЖЕНИЯ ОБЪЕКТА В КОНФЛИКТНОЙ АНИЗОТРОПНОЙ СРЕДЕ ПО ВЕРОЯТНОСТНОМУ КРИТЕРИЮ СКРЫТНОСТИ

Рассмотрена задача выбора траектории движения маневрирующего объекта и закона изменения его скорости при движении объекта в трехмерной анизотропной среде распространения сигнала, когда его пытаются обнаружить несколько наблюдателей, расположенных в заданном районе. Критерием выбора траектории объекта служит вероятность его обнаружения на всей траектории ни одним из наблюдателей. Предложен трехшаговый последовательный дискретный метод оптимизации этого критерия на основе принципа динамического программирования.

Ключевые слова: траектория движения объекта; вероятность обнаружения объекта на траектории; анизотропная среда; динамическое программирование.

Рассматриваемая задача относится к классу задач об управлении, получивших в отечественной литературе название задач управления подвижными объектами в конфликтной среде [1, 2]. Под конфликтной средой понимается совокупность объектов (они называются конфликтующими), сближение с которыми для управляемого объекта нежелательно в ходе выполнения им основной задачи. Минимизация негативного воздействия конфликтующих объектов на управляемый объект является целью управления. Эта цель достигается путем выбора маршрута его движения, параметров движения и (или) режимов работы технических средств. Задачи об оптимизации закона уклонения подвижного объекта от обнаружения рассматривались в ряде работ, отличающихся предположениями о характеристиках информационных полей, в которых происходит обнаружение [1–6].

В работах [3, 5, 6] рассматриваются задачи уклонения применительно к обработке гидроакустической информации. Как известно, распространение гидроакустического сигнала происходит в анизотропной среде, которая обладает контрастной структурой с чередованием зон акустической освещенности и зон акустической тени [7].

Чаще всего в качестве критерия решения задачи используется вероятностный критерий. Удобство его использования определяется тем, что он позволяет получать не только текущие локальные, но и интегральные оценки потенциальной возможности обнаружения объекта поиска по траектории, т.е. маршруту (в горизонтальной плоскости) и трассе (по глубине) следования.

В работах [5, 6] предлагается дискретный метод оптимизации этого критерия при решении задачи об оптимизации закона уклонения морского подводного объекта (МПО) от обнаружения несколькими независимыми стационарными наблюдателями (СГАС) и маневренными средствами (МС) на основе принципа динамического программирования (ДП) для изотропной [5] и анизотропной [6] среды. Считается, что время движения МПО ограничено известной величиной. В [5] рассмотрена аппроксимация путями графа, вершины которого равномерно покрывают район. При решении задачи ДП для каждого ребра графа ДП по известным параметрам (координаты (x, y) начала и конца ребра, глубина $h_{\text{ист}}$, скорость $V_{\text{ист}}$) вычисляется вероятность обнаружения $P_{\text{обн}i}$ для i -го наблюдателя, и затем находится интегральная вероятность обнаружения для независимых наблюдений [8]:

$$P_{\text{обн}} = 1 - \prod_{i=1}^L (1 - P_{\text{обн}i}), \quad (1)$$

где L – число наблюдателей. Таким образом, задача оптимизации решается с использованием 4-мерного вектора $[x, y, h_{\text{ист}}, V_{\text{ист}}]$.

В настоящей работе предлагается подход к решению задачи об оптимизации закона уклонения МПО от обнаружения несколькими СГАС и МС для случая, когда известно единое поле вероятности обнаружения для всего района, пересекаемого МПО. Вектор управлений МПО включает траекторию уклонения и закон изменения скорости V на траектории.

1. Оптимизация закона уклонения в изотропном поле

Зададим закон движения МПО от начальной точки A до точки назначения B за время T параметрическими уравнениями

$$x = x(t), y = y(t), 0 \leq t \leq T, \quad x(0) = x_A, y(0) = y_A.$$

Поскольку параметром является время, то тем самым задается также закон изменения вектора скорости $(\dot{x}(t), \dot{y}(t))$. Качество маневрирования объекта будем характеризовать вероятностью того, что за время прохождения объектом маршрута его не обнаружит ни один из наблюдателей, контролирующих район. Будем называть ее вероятностью необнаружения, обозначив через $P_{\text{необн}}$. Зависимость этой вероятности от закона движения будем определять в предположении, что в районе расположено N СГАС в пунктах с координатами (x_k, y_k) , $k = 1, \dots, N$. Оптимальный закон движения из начального пункта в конечный ищется из условия максимума $P_{\text{необн}}$ или, что то же самое, из условия минимума вероятности $P_{\text{обн}} = 1 - P_{\text{необн}}$ того, что при прохождении маршрута объект будет обнаружен хотя бы одним из наблюдателей.

Расстояние, на котором вероятность правильного обнаружения сигнала $P_{\text{обн}}$ равна установленному значению при соблюдении требований к вероятности ложных тревог $P_{\text{лт}}$, принято называть дальностью действия ГАС [9]. Данное определение в общем виде относится к однократному наблюдению (взгляду) при установленных зависимостях между временем накопления T_0 , отношением сигнал / помеха ρ и заданных $P_{\text{обн}}$ и $P_{\text{лт}}$.

Для случая многократных последовательных независимых наблюдений (вероятность наступления события обнаружения хотя бы один раз за серию наблюдений) оценки производятся по формуле интегральной (кумулятивной) вероятности (1).

Из основного уравнения гидролокации (ОУГ) [10] для однократного наблюдения в [5] получено выражение для вероятности правильного обнаружения как функции от дальности для изотропного поля:

$$P_{\text{обн}}(r) = 1 - F_n\left(\frac{h_0}{1 + \rho}\right), \quad (2)$$

где $F_n(\cdot)$ – функция χ^2 -распределения с n степенями свободы, ρ – отношение сигнал / помеха; h_0 – квантиль уровня $(1 - \alpha)$, $\alpha = P_{\text{лт}}$, n – число степеней свободы, которое в случае цифровой обработки определяется из условий $T = n\Delta t$, $\Delta t = 1/\Delta F$, Δt – интервал дискретизации по времени, ΔF – полоса пропускания входного фильтра системы обнаружения, а в случае аналоговой обработки определяется как $n = 2T\Delta F$.

При заданной вероятности ложной тревоги $P_{\text{лт}} = \alpha$ вероятность правильного обнаружения $P_{\text{обн}}$ возрастает с увеличением ρ , а ρ монотонно убывает с ростом расстояния r до цели. Задаваясь минимально обнаруживаемым сигналом, т.е. минимальным значением ρ_{min} , можно найти максимальную дальность обнаружения R_{max} . При $r \geq R_{\text{max}}$ цель не может быть обнаружена с заданной вероятностью α , а при $r \leq R_{\text{max}}$ обнаружение происходит с вероятностью, не меньшей $P_{\text{обн}}$ при заданном значении α . Пределом R_{max} является расстояние r , для которого $P_{\text{обн}} = P_{\text{лт}} = \alpha$, так как обнаружение с вероятностью меньшей, чем $P_{\text{лт}}$, теряет смысл.

Расстояние r можно условно разбить на три участка:

- $0 < r \leq R_{\text{min}}$, на котором объект будет обнаружен с заданной $P_{\text{обн}}$ за одно или, в худшем случае, за заданное число разовых наблюдений (1) – зона устойчивого обнаружения;
- $r \geq R_{\text{max}}$, на котором скрытность объекта максимальна ($P_{\text{обн}} \leq P_{\text{лт}}$) – зона необнаружения;
- $R_{\text{min}} \leq r \leq R_{\text{max}}$ – зона неустойчивого обнаружения.

Таким образом, объект следует считать видимым лишь вне зоны необнаружения, а задача оптимизации траектории движения имеет смысл лишь в зоне неустойчивого обнаружения.

Единое поле вероятности обнаружения для всех наблюдателей, контролирующих заданный район, $P_{\text{обн}}(x, y)$ может быть вычислено следующим образом. Пусть имеются два наблюдателя с шириной полосы частот ΔF_1 и ΔF_2 и со временами накопления T_1 и T_2 ($T_1 \neq T_2$) соответственно, в которых решение об обнаружении принимаются независимо. Для расчета вероятности совместного обнаружения хотя бы одним из двух наблюдателей, чтобы использовать формулу (1), нужно одинаковое время накопления, но $T_1 \neq T_2$. Определим наименьшее общее кратное – НОК(T_1, T_2) = T_0 . Тогда $T_0 - L_1 = T_0/T_1$ – число взглядов для 1-го наблюдателя, $L_2 = T_0/T_2$ – число взглядов для 2-го наблюдателя.

Для 1-го наблюдателя, считая, что цель неподвижна, согласно (1) за время T_0 получаем

$$P_{\text{необн1}}(T_0) = \prod_{i=1}^{L_1} P_{\text{необн1}}(T_1) = [P_{\text{необн1}}(T_1)]^{L_1}, \quad (3)$$

а для 2-го наблюдателя за время T_0

$$P_{\text{необн2}}(T_0) = \prod_{i=1}^{L_2} P_{\text{необн2}}(T_2) = [P_{\text{необн2}}(T_2)]^{L_2}.$$

Допустим, что у нас есть 3-й наблюдатель, который имеет время наблюдения T_2 и за время T_0 обеспечивает по формуле (3) такую же вероятность необнаружения $P_{\text{необн}}$, как и 1-й наблюдатель, т.е.

$$[P_{\text{необн3}}(T_2)]^{L_2} = [P_{\text{необн1}}(T_1)]^{L_1} \text{ и } P_{\text{необн3}}(T_2) = [P_{\text{необн1}}(T_1)]^{L_1/L_2} = [P_{\text{необн1}}(T_1)]^{T_2/T_1}.$$

Третий наблюдатель эквивалентен 1-му наблюдателю в том смысле, что рассчитанные для них по формуле (3) за время T_0 вероятности необнаружения $P_{\text{необн}}$ равны между собой. Так как время наблюдения – T_2 – для 3-го наблюдателя равно времени наблюдения для 2-го наблюдателя, то для определения интегральной вероятности для независимых наблюдателей можно применить формулу (1):

$$P_{\text{необн}(1+2)} = P_{\text{необн2}}(T_2)P_{\text{необн3}}(T_2) = P_{\text{необн2}}(T_2)[P_{\text{необн1}}(T_1)]^{T_2/T_1}. \quad (4)$$

Таким образом, вероятности необнаружения для двух СГАС с различными временами взглядов можно объединить, используя формулу (4), в $P_{\text{необн}}$ с единым временем взгляда.

Пусть СГАС, расположенные в районе, имеют координаты (x_k, y_k) . Тогда, используя формулу для вычисления $P_{\text{обн}}(d_i)$, приведенную в [5], можно рассчитать вероятность необнаружения для всех точек района (x, y) , т.е. сформировать поле (матрицу) вероятностей необнаружения $P_{\text{необн}}(x, y)$:

$$P_{\text{необн}}(x, y) = \sqrt{a^2 + b^2} \lim_{x \rightarrow \infty} \prod_{k=1}^K P_{\text{необнк}}(x_k, y_k)^{T_M/T_k}, \quad (5)$$

где K – число наблюдателей, $T_M = \max[T_k]$, T_k – длительность взгляда для k -го наблюдателя, (x_k, y_k) – координаты k -го наблюдателя, $P_{\text{необнк}}$ – вероятности необнаружения для k -го наблюдателя.

В качестве примера на рис. 1 приведены вероятности обнаружения (карта вероятностей) $P_{\text{обн}}(x, y) = 1 - P_{\text{необн}}(x, y)$ для трех независимых наблюдателей (СГАС), расположенных в заданном районе, в случае однократного наблюдения с использованием формулы (5).

На рис. 1 зона необнаружения выделена зеленым цветом, а зона обнаружения – коричневым. Границы зоны неустойчивого обнаружения определены из условий: $R_{\min} - P_{\text{обн}} = 0,5$, $R_{\max} - P_{\text{обн}} = 0,02$. Приведены также четыре возможных маршрута – 2 маршрута, обходящих районы СГАС, и 2 – проходящих между ними.

Шумность МПО зависит от скорости V , следовательно, карта вероятности (см. рис. 1) будет относительно координат k -ой СГАС изменяться: с ростом V граница зоны необнаружения будет расширяться, а с уменьшением V – сжиматься.

Учитывая, что обычно траектория объекта проходит на достаточно больших расстояниях от наблюдателя и $(1/r^2)$ является малой величиной, как показано в [11], можно разложить функцию зависимости изменения вероятности обнаружения $P_{\text{обн}}$ от дистанции r – $P_{\text{обн}}(r)$ в ряд по степеням $(1/r^2)$ и ограничиться несколькими членами разложения. В первом приближении имеем

$$P_{\text{обн}}(r) = \alpha + \frac{\gamma V^\mu}{\sigma_{\text{ш}}^2} \cdot \frac{(\chi_{1-\alpha,n}^2)^{\frac{n}{2}}}{2^{n/2} \Gamma(\frac{n}{2})} e^{-\frac{\chi_{1-\alpha,n}^2}{2}} \cdot \frac{1}{r^2}, \quad (6)$$

где $\chi_{1-\alpha,n}^2 - (1 - \alpha)$ – квантиль χ^2 -распределения с n степенями свободы, $\sigma_{\text{ш}}^2$ – дисперсия помехи, γV^μ – зависимость шумности МПО от его скорости V , γ и μ – некоторые размерные коэффициенты.

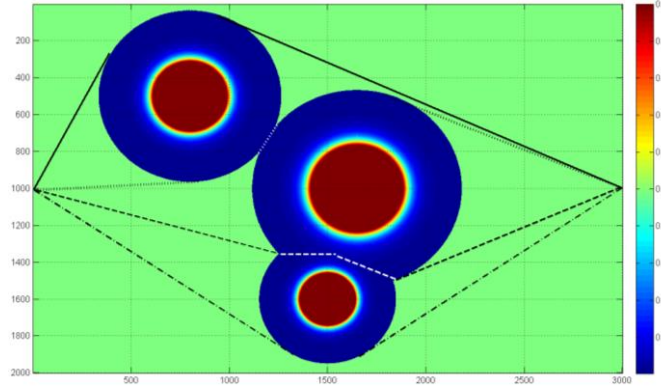


Рис. 1. Вероятности правильного обнаружения (карта вероятностей) $P_{\text{обн}}(x, y)$
Fig. 1. Probabilities of correct detection (probability map) $P_{\text{det}}(x, y)$

Рассмотрим прохождение МПО между двумя СГАС с перекрывающимися зонами наблюдения. Пусть расстояние между СГАС – R_0 . Тогда из (6) для 1-й и 2-й СГАС соответственно получаем

$$P_{\text{обн1}}(r) = \alpha_1 + V^\mu C_1 \frac{1}{r^2}, \quad P_{\text{обн2}}(r) = \alpha_2 + V^\mu C_2 \frac{1}{(R_0 - r)^2}. \quad (7)$$

В обобщенный параметр C_1 сведены все параметры ОУГ, относящиеся к первому СГАС, а в параметр C_2 – относящиеся ко второму СГАС. При выполнении условий $P_{\text{обн1}} = P_{\text{обн2}}$ и $\alpha_1 = \alpha_2$ параметры, относящиеся к МПО (шумность, зависящая от скорости V), сокращаются. Таким образом, положение $\min[P_{\text{обн}}(r)]$ **от скорости МПО не зависит**.

На рис. 2 приведен пример изменения вероятности обнаружения от скорости, где для различных скоростей V по формуле (5) рассчитана суммарная однократная вероятность обнаружения на линии, соединяющей два СГАС. Положение минимума интегральной вероятности (5) при изменении скорости не меняется, а величины $\min(P_{\text{обн}})$ растут с ростом V_k .

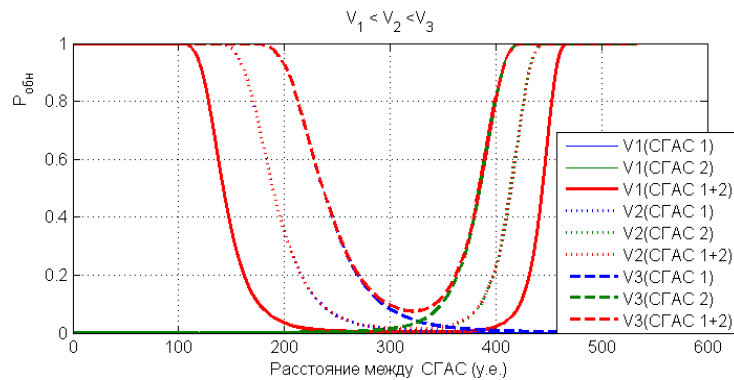


Рис. 2. Суммарная вероятность обнаружения $P_{\text{обн}}$ для двух СГАС
Fig. 2. The total probability of detection of P_{det} for two SSS

Таким образом, маршрут, проложенный между двумя СГАС с перекрывающимися полями обнаружения (см. рис. 1), не зависит от скорости V и должен проходить через точки карты вероятностей, обеспечивающие $\min(P_{\text{обн}}(x, y))$. Или, используя географическую терминологию, проходить по «дну долин» и через «перевалы» карты вероятностей.

Оптимизация маршрута по вероятностному критерию производится методом ДП [5]. В результате решения задачи получаем несколько синтезированных кусочно-прямолинейных субоптимальных маршрутов движения МПО (r_{Mi}), каждый из которых при фиксированной скорости V будет характеризоваться $P_{обн}(x, y)$ и временем перехода T_{Mi} .

Наличие ограничений на время перехода T_M требует проведения для каждого маршрута оптимизации по скорости V , которая также производится методом ДП.

2. Оптимизация закона уклонения в анизотропном поле

Предположение об изотропности гидроакустического поля справедливо лишь для однородной среды, в этом случае дальность обнаружения зависит только от расстояния между МПО и наблюдателем. В неоднородной среде дальность обнаружения зависит не только от расстояния между МПО и наблюдателем, но также от других параметров, таких как глубина и профиль дна, акустические свойства грунта, от гидрологических, определяемых видом вертикального распределения скорости звука (ВРСЗ), и метеорологических условий наблюдения, глубины погружения МПО, характеристик СГАС и глубины установки ее антенны и ряда других факторов [7]. Как результат влияния этих факторов круговая диаграмма, присущая изотропному полю, деформируется и зависит уже от четырех параметров: расстояния между МПО и наблюдателем (r), глубины погружения МПО ($h_{ист}$), азимута на МПО и глубины установки антенны СГАС ($h_{пр}$).

Кроме обратной квадратичной зависимости затухания сигнала от дистанции, для каждой фиксированной дистанции r в анизотропном поле необходимо учитывать также перераспределение энергии сигнала по глубине, вызванное непосредственно анизотропией поля [7] и зависящее от дистанции r и глубин источника $h_{ист}$ и приемника $h_{пр}$, – коэффициент анизотропии $\gamma(h_{пр}, r, h_{ист})$.

Для учета анизотропии поля в формулу (2) необходимо ввести значение коэффициента анизотропии $\gamma(h_{пр}, r, h_{ист})$. Тогда вероятность правильного обнаружения как функции от дальности (2) для анизотропного поля преобразуется к виду:

$$P_{обн}(h_{ист}, r, h_{пр}) = 1 - F_n \left(\frac{h_0}{1 + \rho_A} \right) = 1 - F_n \left(\frac{h_0}{1 + \rho \gamma(h_{пр}, r, h_{ист})} \right), \quad (8)$$

где ρ_A – отношение сигнал / помеха в анизотропном поле.

Задача расчета вероятности обнаружения в анизотропной среде хорошо разделяется на две независимые задачи: расчет анизотропного поля по заданной гидрологии и расчет вероятности обнаружения с учетом коэффициента анизотропии поля.

Такое разделение очень важно с учетом того, что гидрологические характеристики района остаются «квазистационарными» достаточно длительное время (от нескольких часов до многих суток), а сам расчет анизотропного поля требует очень большого объема вычислений. В то же время расчет вероятности обнаружения с учетом коэффициента анизотропии поля требует значительно меньшего объема вычислений и может проводиться в режиме online.

На рис. 3 в качестве примера приведены карты вероятности правильного обнаружения $P_{обн}(r, h_{ист}, h_{пр})$ для СГАС с антенной, расположенной на глубине $h_{пр}$, и двух заданных глубин – $h_{ист1}$ и $h_{ист2}$. Хорошо видно, что $P_{обн}$ зависит от глубины МПО и азимута на МПО. Следовательно, для анизотропного (слоисто-неоднородного) поля необходимо иметь набор карт вероятности $P_{обн}$ для дискретного ряда глубин ПЛ – $h_{ист1}, \dots, h_{истj}, \dots, h_{истJ}$. Таким образом, если для изотропного поля карты вероятности обнаружения $P_{обн}$ строятся в координатах (x, y) , то для анизотропного поля необходимо иметь набор из J карт, рассчитанных по формуле (8), в координатах (x, y) для ряда фиксированных глубин $h_{истj}$. Такой набор карт вероятности можно описывать трехмерной матрицей в координатах $(x, y, h_{ист})$ при фиксированных глубинах наблюдателей $h_{прi}$.

На рис. 4 приведены вероятности обнаружения $P_{обн}(x, y)$ в анизотропном поле для двух СГАС, расположенных в заданном районе, для заданной глубины $h_{ист}$. В качестве примера на этом же рисунке показан маршрут, который мог бы проложить оператор при условии минимизации вероятности обнаружения МПО при переходе их точки координатами $(0, 0)$ в точку $(500, 500)$.

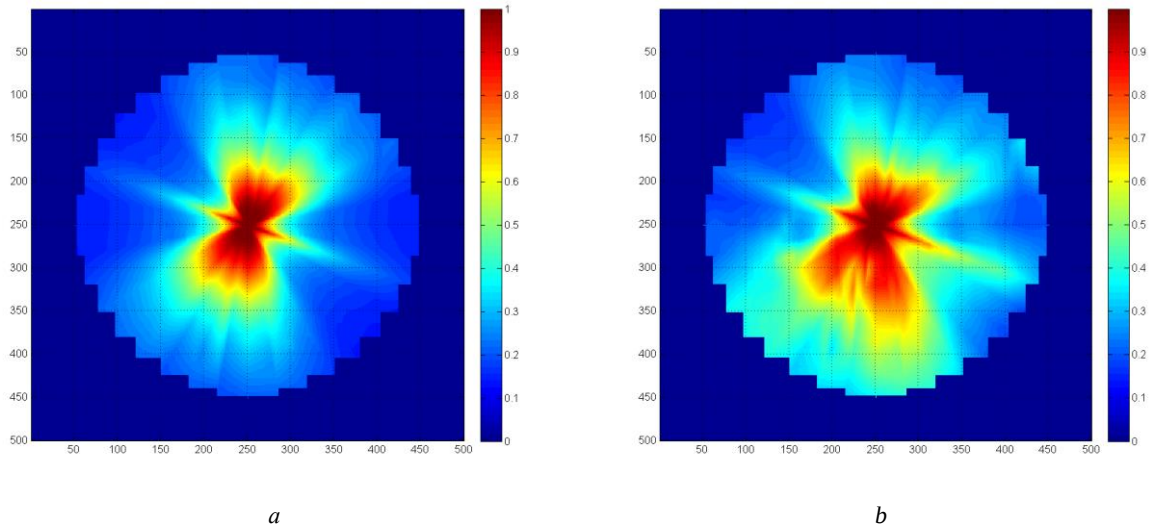


Рис. 3. Карты вероятности обнаружения $P_{\text{обн}}(x, y, h_{\text{ист}}, h_{\text{пр}})$: $a - h_{\text{ист}1}$, $b - h_{\text{ист}2}$
 Fig. 3. Probability maps of detection $P_{\text{det}}(x, y, h_{\text{orig}}, h_{\text{rec}})$: $a - h_{\text{orig}1}$, $b - h_{\text{orig}2}$

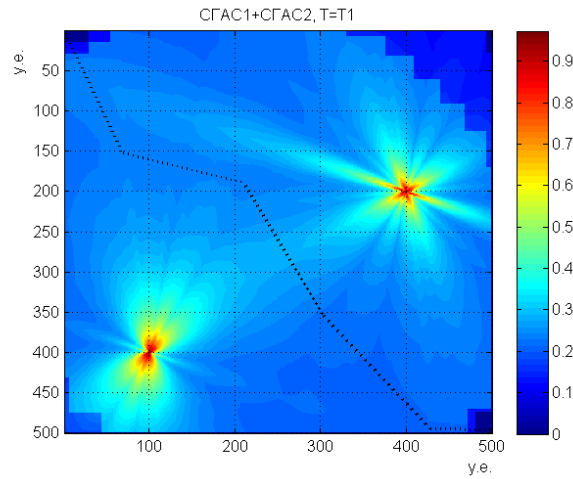


Рис. 4. Суммарное поле вероятности обнаружения для 1-го и 2-го наблюдателей по формуле (2)
 Fig. 4. The total field of probability of detection for the 1st and 2nd observer according to the formula (2)

Так же, как было показано выше для изотропного поля (см. рис. 1 и соответствующие формулы), рассмотрим прохождение МПО между двумя СГАС с перекрывающимися зонами наблюдения для анизотропного поля. Как следует из формулы (6), в ОУГ влияние анизотропии учитывается при расчете ρ_A – отношения сигнал / помеха в анизотропном поле. В свою очередь, ρ_A зависит от скорости МПО V и коэффициента анизотропии $\gamma(h_{\text{пр}}, r, h_{\text{ист}})$. Очевидно, что коэффициент анизотропии от скорости МПО не зависит и может быть учтен при расчете параметров C_1 и C_2 . Таким образом, в анизотропном поле положение $\min[P_{\text{обн}}(r)]$ также не зависит от скорости МПО.

На рис. 5, a для этого сечения приведены $P_{\text{обн}}$, рассчитанные при одинаковой скорости МПО для различных глубин $h_{\text{ист}}$, а также $P_{\text{обн}}$ для соответствующего изотропного поля. При изменении глубины МПО положение минимума вероятности $\min(P_{\text{обн}}(h_{\text{ист}}))$ практически не меняется, поэтому маршруты для различных глубин $h_{\text{ист}}$ при прохождении между двумя СГАС практически совпадают.

На рис. 5, b приведено (при фиксированной скорости МПО) сечения матрицы $P_{\text{обн}}(x, y, h_{\text{ист}})$ по линии (r) , соединяющей два СГАС, в координатах $(r, h_{\text{ист}})$, т.е. 2-мерная матрица, где по формуле (8) рассчитана интегральная вероятность обнаружения $P_{\text{обн}}(r, h_{\text{ист}})$. (Профиль дна окрашен в цвет, соответствующий $P_{\text{обн}} = 1$.)

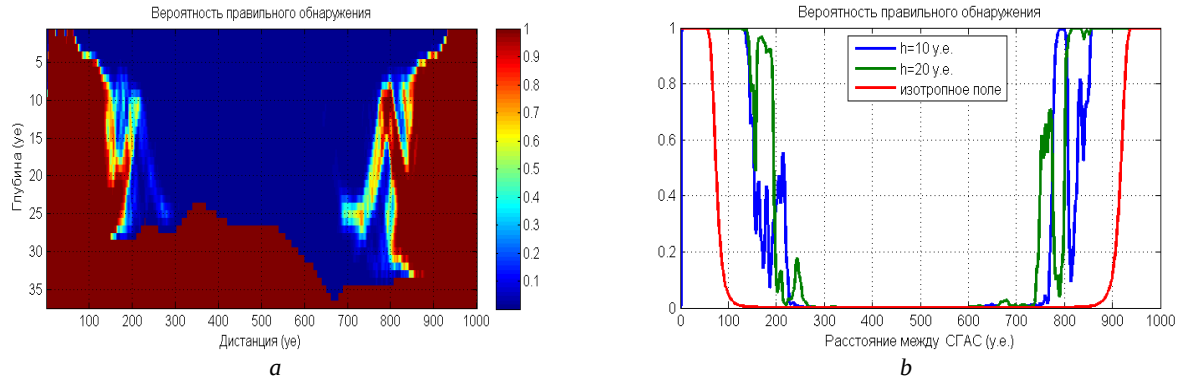


Рис. 5. Вероятности обнаружения $P_{\text{обн}}$ для различных глубин $h_{\text{ист}}$ при прохождении между двумя СГАС (а) и сечении матрицы $P_{\text{обн}}(x, y, h_{\text{ист}})$ по линии, соединяющей две СГАС (б)

Fig. 5. P_{det} detection probabilities for various depths of h_{orig} when passing between two SSS (a); the section of the matrix $P_{\text{det}}(x, y, h_{\text{orig}})$ along the line connecting the two SSS (b)

Рассмотрим сечение матрицы $P_{\text{обн}}(x, y, h_{\text{ист}})$ по маршруту, проходящему, как показано на рис. 4, между двух СГАС – $P_{\text{обн}}(r_M, h_{\text{ист}})$. На рис. 6 приведен пример такого сечения. Это сечение также можно рассматривать как карту вероятности обнаружения $P_{\text{обн}}$, но уже не в координатах (x, y) , а в координатах $(r_M, h_{\text{ист}})$, где r_M – маршрут, $h_{\text{ист}}$ – глубина, по которой аналогично построению маршрута методом ДП вычисляется набор возможных трасс, одна из которых приведена на рис. 6.

Для изотропного поля изменение скорости МПО при прохождении между двумя СГАС не приводит к изменению маршрута. При расчете вероятности $P_{\text{обн}}$ для анизотропного поля учитывается фактор аномалии, входящий, как один из параметров, в ОУГ [9]. При определении $\min[P_{\text{обн}}(r)]$ между двумя СГАС для анизотропного поля получаем уравнение, аналогичное (7), в котором параметры, характеризующие факторы аномалии, входят в параметры C_1 и C_2 соответствующих СГАС. Следовательно, положение $\min[P_{\text{обн}}(r)]$ для анизотропного поля также **не зависит от скорости МПО**.

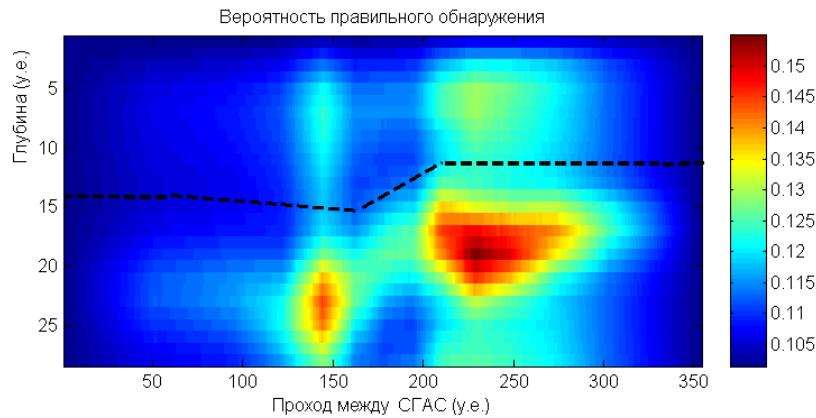


Рис. 6. Сечение матрицы $P_{\text{обн}}(x, y, h_{\text{ист}})$ по маршруту r_M – $P_{\text{обн}}(r_M, h_{\text{ист}})$
Fig.6. The cross section of the matrix $P_{\text{det}}(x, y, h_{\text{orig}})$ along the route r_M – $P_{\text{det}}(r_M, h_{\text{orig}})$

С учетом перебора по скоростям V_k можно построить 4-мерную карту (матрицу) $P_{\text{обн}}(x, y, h_{\text{ист}}, V)$, которая для заданного района может быть рассчитана заранее для фиксированного набора глубин $h_{\text{ист}j}$ и скоростей V_k МПО при известных координатах наблюдателей и их технических характеристиках, гидрофизических характеристиках района (ВРСЗ, карта глубин, метеорологические условия наблюдения и ряд других факторов).

Таким образом, рассматриваемый алгоритм построения оптимальной траектории с использованием матрицы вероятности (риска) $P_{\text{обн}}(x, y, h_{\text{ист}}, V)$ выполняется за три последовательных этапа:

1. Построение маршрута r_M по карте $P_{\text{обн}}(x, y)$ – сечению матрицы $P_{\text{обн}}(x, y, h_{\text{ист}}, V)$ при $h_{\text{ист}} = \text{const}$, $V_k = \text{const}$.

2. Построение для выбранного маршрута r_M трассы по $h_{ист}$ по карте $P_{обн}(r_M, h_{ист})$ – сечению матрицы $P_{обн}(x, y, h_{ист}, V)$ по r_M при $V_k = \text{const}$.

3. Расчет для выбранной траектории $T_{rh}(r_M, h_{ист})$ оптимальной скорости V_k по карте $P_{обн}(T_{rh}, V)$ – сечению матрицы $P_{обн}(x, y, h_{ист}, V)$ по T_{rh} .

Такая последовательная структура алгоритма позволяет существенно уменьшить объем вычислений при построении оптимальной траектории по сравнению, например, с работой [5], где оптимизация производится одновременно по четырем параметрам: $x, y, h_{ист}, V$. Общее число двумерных карт вероятности $P_{обн}(x, y)$ определяется размерами векторов глубин $h_{ист}$ и скоростей V , т.е. $(J \times K)$.

Кроме того, расчеты на первых двух этапах (расчет траектории) *желательно* проводить при оптимальных для наблюдателей условиях – при движении ПЛ с максимально допустимой скоростью в подводном звуковом канале (если он есть). Эти условия позволяют производить ДП при максимально узких «долинах» и «перевалах» карты вероятности и, следовательно, более точно рассчитать маршрут и трассу (см. рис. 4, 6).

Заключение

Важной особенностью алгоритма является то, что наиболее трудоемкие расчеты вероятностей для анизотропного поля вынесены на предварительный этап решения задачи оптимизации.

Рассматриваемый алгоритм может применяться для решения задачи оптимизации маршрута в режиме online и позволяет корректировать расчеты в случае появления новых наблюдателей.

Алгоритм ориентирован на условия известного взаимного положения и движения объектов.

Алгоритм построения оптимальной траектории универсален, поскольку он не зависит от количества средств обнаружения, их взаимного расположения, их параметров, при условии наличия моделей для расчетов вероятностей обнаружения объекта этими средствами.

ЛИТЕРАТУРА

1. Галяев А.А., Маслов Е.П., Рубинович Е.Я. Об одной задаче управления движением объекта в конфликтной среде // Известия РАН. Теория и системы управления. 2009. № 3. С. 134–140.
2. Галяев А.А., Маслов Е.П. Оптимизация законов уклонения подвижного объекта от обнаружения // Известия РАН. Теория и системы управления. 2010. № 4. С. 52–62.
3. Какалов В.А., Поленин В.И., Дымент А.Б., Сергеев В.А. Метод оптимизации боевого маневрирования подводной лодки по вероятностному критерию скрытности. // Морская радиоэлектроника. 2003. № 3 (6). С. 44–47.
4. Zabaranin M., Uryasev S., Pardalos P. Optimal Risk Path Algorithms // Cooperative Control and Optimization / eds. R. Murphey, P. Pardalos. Dordrecht : Kluwer Acad., 2002. Ch. 1. P. 273–298.
5. Добровидов А.В., Кулида Е.Л., Рудько И.М. Выбор траектории движения объекта в конфликтной среде // Проблемы управления. 2011. № 3. С. 64–75.
6. Добровидов А.В., Кулида Е.Л., Рудько И.М. Оптимизация траектории движения объекта по вероятностному критерию в режиме пассивной гидролокации в анизотропной среде // Проблемы управления. 2014. № 4. С. 31–37.
7. Сташкевич А.П. Акустика моря. Л. : Судостроение, 1966. 350 с.
8. Абчук В.А., Суздаль В.Г. Поиск объектов. М. : Сов. радио, 1977. 336 с.
9. Евтютов А.П., Колесников А.Е., Ляликов А.П. Справочник по гидроакустике. Л. : Судостроение, 1988. 552 с.
10. Burdic W.S. Acoustic system analysis. Englewood Cliffs, NJ : Prentice Hall, Inc., 1984. 392 p.
11. Сысоев Л.П. Критерий вероятности обнаружения на траектории в задаче управления движением объекта в конфликтной среде // Проблемы управления. 2010. № 6. С. 65–72.

Поступила в редакцию 12 августа 2020 г.

Rudko I.M. (2020) OPTIMIZATION OF THE TRAJECTORY OF THE OBJECT IN A CONFLICT ANISOTROPIC MEDIUM ACCORDING TO THE PROBABILISTIC CRITERION OF STEALTH. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika* [Tomsk State University Journal of Control and Computer Science]. 55. pp. 71–79

DOI: 10.17223/19988605/55/9

The problem of choosing the trajectory of a marine moving object (MMO) and the law of change in its velocity during the movement of an object in a three-dimensional anisotropic medium of signal propagation, when several stationary observers (SSS)

are located in a given area, are considered. The criterion for choosing the trajectory of an object is the probability of its non-detection on the entire trajectory by any of the observers.

An approach is proposed to solve the problem of optimizing the law of MMO evasion from detection by several SSSs for the case when a single detection probability field is known for the entire area intersected by MMO. The MMO control vector includes the evasion path and the law of variation of the velocity V on the path.

Relations are derived in the work that make it possible to calculate in a isotropic and anisotropic observation fields a single field (map) of the probability of detecting an object for several SSSs with different technical characteristics. It is shown that for an anisotropic field, the map of the probability of detection can be represented as a 4-dimensional matrix in the coordinates $(x, y, h_{\text{orig}}, V)$, where x, y are the coordinates of the surface of the region, h_{orig} is the depth of MMO, V is the MMO velocity.

It is shown that using the P_{det} detection probability map $(x, y, h_{\text{orig}}, V)$, when solving the problem of dynamic programming (DP) of the route, it is possible to replace the 4-dimensional vector $(x, y, h_{\text{orig}}, V)$, according to which optimization is carried out, on a three-step sequential discrete optimization method based on the principle of DP. In this case, at each of the 3 steps, optimization is performed only for two-dimensional vectors:

- Construction of the r_M route from the P_{det} map (x, y) - to the section of the P_{det} matrix $(x, y, h_{\text{orig}}, V)$ with $h_{\text{orig}} = \text{const}$, $V_k = \text{const}$.

- Construction for the chosen route r_M of the route according to h_{orig} on the P_{det} map (r_M, h_{orig}) - section of the P_{det} matrix $(x, y, h_{\text{orig}}, V)$ along r_M at $V_k = \text{const}$.

- Calculation for the chosen trajectory T_{rh} (r_M, h_{orig}) of the optimal speed V_k according to the P_{det} map (T_{rh}, V) - to the section of the P_{det} matrix $(x, y, h_{\text{orig}}, V)$ according to T_{rh} .

Such a consistent structure of the algorithm can significantly reduce the amount of computation when constructing the optimal trajectory.

Keywords: trajectory of the object; the probability of an object not being detected on the trajectory; anisotropic medium; dynamic programming.

RUDKO Igor Mikhailovich (Candidate of Technical Sciences, Senior Researcher, V.A.Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russian Federation).

E-mail: igor-rudko@mail.ru

REFERENCES

1. Galyaev, A.A., Maslov, E.P. & Rubinovich, E.Ya. (2009) On one problem of controlling the movement of an object in a conflict environment. *Izvestiya RAN. Teoriya i sistemy upravleniya – Journal of Computer and Systems Sciences*. 3. pp. 134–140.
2. Galyaev, A.A. & Maslov, E.P. (2010) Optimization of a mobile object evasion laws from detection. *Journal of Computer and Systems Sciences*. 4. pp. 52–62.
3. Kakalov, V. A., Polenin, V. I., Dymant, A.B. & Sergeev V.A. (2003) Metod optimizatsii boevogo manevrirovaniya podvodnoy lodki po veroyatnostnomu kriteriyu skrytnosti [A method for optimizing the combat maneuvering of a submarine by the probabilistic stealth criterion]. *Morskaya radioelektronika – Marine Radio electronics*. 3(6). pp. 44–47.
4. Zabaranin, M., Uryasev, S. & Pardalos, P. (2002) Optimal Risk Path Algorithms. In: Murphey, R. & Pardalos, P. (eds) *Cooperative Control and Optimizatoin*. Vol. 1. Dordrecht: Kluwer Acad. pp. 273–298.
5. Dobrovidov, A.V., Kulida, E.L. & Rudko, I.M. (2011) Vybora traektorii dvizheniya ob"ekta v konfliktnoy srede [The choice of the trajectory of the object in a conflict environment]. *Problemy upravleniya – Control Sciences*. 3. pp. 64–75.
6. Dobrovidov, A.V., Kulida, E.L. & Rudko, I.M. (2015) Path optimization for a moving object in an anisotropic environment using the probabilistic criterion in the passive sonar mode. *Automation and Remote Control*. 76(7). pp. 1271–1281. DOI: 10.1134/S0005117915070127
7. Stashkevich, A.P. (1966) *Akustika morya* [Acoustics of the Sea]. Leningrad: Sudostroenie.
8. Abchuk, V.A. & Suzdal, V.G. (1977) *Poisk ob"ektov* [Search for Objects]. Moscow: Sovetskoe radio.
9. Evtuyutov, A.P., Kolesnikov, A.E. & Lyalikov, A.P. (1988) *Spravochnik po gidroakustike* [Handbook of Hydroacoustics]. Leningrad: Sudostroenie.
10. Burdic, W.S. (1984) *Acoustic System Analysis*. Prentice. Hall, Inc. Englewood Cliffs.
11. Sysoev, L.P. (2010) Kriteriy veroyatnosti obnaruzheniya na traektorii v zadache upravleniya dvizheniem ob"ekta v konfliktnoy srede [The probability criterion for detection on the trajectory in the task of controlling the movement of an object in a conflict environment]. *Problemy upravleniya – Control Sciences*. 6. pp. 65–72.

УДК 519.95

DOI: 10.17223/19988605/55/10

Ф.Г. Фейзинов, Н.Б. Абаева

УСЛОВИЯ ОРТОГОНАЛЬНОСТИ ВХОДНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ОДНОГО КЛАССА ДВОИЧНЫХ 4D-НЕЛИНЕЙНЫХ МОДУЛЯРНЫХ ДИНАМИЧЕСКИХ СИСТЕМ

Приводится понятие ортогональной входной последовательности для двоичных 4D-модулярных динамических систем, заданных в виде двухзначного аналога полинома Вольтерры. Доказываются необходимые и достаточные условия ортогональности в совокупности входных последовательностей и необходимые и достаточные условия собственной ортогональности каждой входной последовательности. Предлагается методика построения ортогональной входной последовательности на базе коротких вспомогательных ортогональных последовательностей. Приводятся достаточные условия ортогональности вспомогательных последовательностей и формулы их определения.

Ключевые слова: 4D-нелинейные модулярные динамические системы; ортогональные входные последовательности; условия ортогональности.

Модулярные динамические системы (МДС), или последовательностные машины, являются одним из важных классов дискретных динамических систем [1–4]. Они находят широкое применение в различных областях науки и техники [1, 3–8]. Исследованы различные задачи теории и приложений одно- и многопараметрических МДС (nD -МДС) [8–15]. К ним относится и задача синтеза МДС. Для нахождения наилучшего решения задачи синтеза МДС используется метод, основанный на применении ортогональных входных последовательностей. Ортогональные входные последовательности строятся с учетом особенности конкретных классов МДС. В работах [3, 4, 16–18] в случаях $n \in \{1, 2, 3\}$ для некоторых нелинейных МДС (nD -НМДС) найдены соответствующие условия ортогональности входной последовательности и разработаны алгоритмы их построения. Одним из классов nD -НМДС является 4D-НМДС [19], который имеет более общую структуру, чем nD -НМДС, $n \in \{1, 2, 3\}$. В работе [20] рассмотрен вопрос нахождения наилучшего решения задачи синтеза, и для решения этой задачи отмечена необходимость нахождения условий ортогональности для входных последовательностей 4D-НМДС. А в данной работе рассматривается вопрос вывода условий ортогональности входных последовательностей двоичных 4D-НМДС.

1. Постановка задачи

Рассмотрим двоичную 4D-НМДС с фиксированной памятью n_0 , ограниченной связью $P = P_1 \times P_2 \times P_3$, степенью S , описываемую в виде двухзначного аналога полинома Вольтерры [20]:

$$y[n, c_1, c_2, c_3] = \sum_{i=1}^S \sum_{v=1}^{\lambda_i} \sum_{(\bar{j}, \bar{\sigma}, \bar{\rho}) \in L_{i,v}} \sum_{\bar{\tau} \in \Gamma_{i,v}} h_{i,v}[\bar{j}, \bar{\sigma}, \bar{\rho}, \bar{\tau}] \times$$

$$\times \prod_{(\alpha, \beta, \gamma) \in Q_{i,v}} \prod_{\xi_{\alpha, \beta, \gamma} = 1}^{m_{\alpha, \beta, \gamma}} v_{i,v}[n - \tau(\alpha, \beta, \gamma, \xi_{\alpha, \beta, \gamma}), c_1 + p_1(j_\alpha), c_2 + p_2(\sigma_\beta), c_3 + p_3(\rho_\gamma)], GF(2).$$

Здесь $n \in \{0, 1, 2, \dots\}$, $c_\alpha \in \{\dots, -1, 0, 1, \dots\}$, $\alpha = \overline{1, 3}$; $y[n, c_1, c_2, c_3]$ и $v_{i,v}[n, c_1, c_2, c_3]$ есть соответственно выходная и входная последовательности над полем $GF(2)$ (см.: [1, 5]). Присутствие записи $GF(2)$ в формуле (1) указывает, что эта формула выполняется над полем $GF(2)$, т.е. операции сложения и

умножения есть сложение и умножение по mod 2 (это не касается выражений $n - \tau(\alpha, \beta, \gamma, \xi_{\alpha, \beta, \gamma})$, $c_1 + p_1(j_\alpha)$, $c_2 + p_2(\sigma_\beta)$ и $c_3 + p_3(\rho_\gamma)$);

$P_\alpha = \{p_\alpha(1), \dots, p_\alpha(r_\alpha)\}$, $-\infty < p_\alpha(1) < \dots < p_\alpha(r_\alpha) < \infty$, $p_\alpha(j) \in \{\dots, -1, 0, 1, \dots\}$, $j = \overline{1, r_\alpha}$, $\alpha = \overline{1, 3}$;
 $\lambda_i = |F(i)|$ (через $|F(i)|$ обозначено число элементов множества $F(i)$), где

$$F(i) = \{(\ell_1, \ell_2, \ell_3, \bar{m}) \mid \bar{m} = (m_{1,1,1}, \dots, m_{1,1,\ell_3}, m_{1,2,1}, \dots, m_{1,2,\ell_3}, \dots, m_{1,\ell_2,1}, \dots, m_{1,\ell_2,\ell_3}),$$

$$\sum_{\alpha=1}^{\ell_1} \sum_{\beta=1}^{\ell_2} \sum_{\gamma=1}^{\ell_3} m_{\alpha, \beta, \gamma} = i; \quad m_{\alpha, \beta, \gamma} \in \{0, \dots, n_0 + 1\}, \quad \alpha = \overline{1, \ell_1}, \beta = \overline{1, \ell_2}, \gamma = \overline{1, \ell_3};$$

$$(\forall \alpha \in \{1, \dots, \ell_1\})(\exists \beta \in \{1, \dots, \ell_2\})(\exists \gamma \in \{1, \dots, \ell_3\}) \Rightarrow (m_{\alpha, \beta, \gamma} \neq 0),$$

$$(\forall \beta \in \{1, \dots, \ell_2\})(\exists \alpha \in \{1, \dots, \ell_1\})(\exists \gamma \in \{1, \dots, \ell_3\}) \Rightarrow (m_{\alpha, \beta, \gamma} \neq 0),$$

$$(\forall \gamma \in \{1, \dots, \ell_3\})(\exists \alpha \in \{1, \dots, \ell_1\})(\exists \beta \in \{1, \dots, \ell_2\}) \Rightarrow (m_{\alpha, \beta, \gamma} \neq 0); \quad \ell_\sigma \in \{1, \dots, r_\sigma\}, \sigma = \overline{1, 3};$$

$$L(\ell) = L_1(\ell_1) \times L_2(\ell_2) \times L_3(\ell_3),$$

где

$$L_2(\ell_2) = \{\bar{\sigma} = (\sigma_1, \dots, \sigma_{\ell_2}) \mid 1 \leq \sigma_1 < \dots < \sigma_{\ell_2} \leq r_2\}, \quad L_3(\ell_3) = \{\bar{\rho} = (\rho_1, \dots, \rho_{\ell_3}) \mid 1 \leq \rho_1 < \dots < \rho_{\ell_3} \leq r_3\};$$

$$h_{i,v}[\bar{j}, \bar{\sigma}, \bar{\rho}, \bar{\tau}] \equiv h_{i,(\ell_1, \ell_2, \ell_3, \bar{m})_v}[\bar{j}, \bar{\sigma}, \bar{\rho}, \bar{\tau}], \quad Q_{i,v} \equiv Q_0(i, (\ell_1, \ell_2, \ell_3, \bar{m})_v), \quad \Gamma_{i,v} \equiv \Gamma(\ell_1, \ell_2, \ell_3, \bar{m})_v,$$

$$L_{i,v} \equiv L((\ell_1, \ell_2, \ell_3)_v),$$

где

$$\Gamma(\ell_1, \ell_2, \ell_3, \bar{m}) = \times_{(\alpha, \beta, \gamma) \in Q_0(i, \ell_1, \ell_2, \ell_3, \bar{m})} \Gamma_1(m_{\alpha, \beta, \gamma})$$

$$\Gamma_1(m_{\alpha, \beta, \gamma}) = \{\bar{\tau}_{\alpha, \beta, \gamma} = (\tau(\alpha, \beta, \gamma, 1), \dots, \tau(\alpha, \beta, \gamma, m_{\alpha, \beta, \gamma})) \mid 0 \leq \tau(\alpha, \beta, \gamma, 1) < \dots < \tau(\alpha, \beta, \gamma, m_{\alpha, \beta, \gamma}) \leq n_0\};$$

$$Q_0(i, \ell_1, \ell_2, \ell_3, \bar{m}) = \{(\alpha, \beta, \gamma) \mid m_{\alpha, \beta, \gamma} \neq 0, \alpha = \overline{1, \ell_1}, \beta = \overline{1, \ell_2}, \gamma = \overline{1, \ell_3}\}.$$

Пусть $n \in [0, N] \equiv \{0, 1, \dots, N\}$, $c_\alpha \in [0, C_\alpha] \equiv \{0, 1, \dots, C_\alpha\}$, $\alpha = \overline{1, 3}$, и $V_0(i, v, \bar{j}, \bar{\sigma}, \bar{\rho}, \bar{\tau}_k)$ есть следующая $(N+1)(C_1+1)(C_2+1)(C_3+1) \times 1$ -мерная матрица:

$$V_0(i, v, \bar{j}, \bar{\sigma}, \bar{\rho}, \bar{\tau}_k) =$$

$$= \left\{ \prod_{(\alpha, \beta, \gamma) \in Q_{i,v}} \prod_{\xi_{\alpha, \beta, \gamma}=1}^{m_{\alpha, \beta, \gamma}} v_{i,v} [n - \tau^{(k)}(\alpha, \beta, \gamma, \xi_{\alpha, \beta, \gamma}), c_1 + p_1(j_\alpha), c_2 + p_2(\sigma_\beta), c_3 + p_3(\rho_\gamma)] \right\}. \quad (2)$$

В матрице (2) каждому набору $(n, c_1, c_2, c_3) \in [0, N] \times [0, C_1] \times [0, C_2] \times [0, C_3]$ соответствует одна строка.

На основе матрицы $V_0(i, v, \bar{j}, \bar{\sigma}, \bar{\rho}, \bar{\tau}_k)$ строятся матрицы

$$V_1(i, v, \bar{j}, \bar{\sigma}, \bar{\rho}) = (V_0(i, v, \bar{j}, \bar{\sigma}, \bar{\rho}, \bar{\tau}_1) \dots V_0(i, v, \bar{j}, \bar{\sigma}, \bar{\rho}, \bar{\tau}_{|\Gamma_{i,v}|})),$$

$$V_2(i, v) = (V_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho})_1) \dots V_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho})_{|L_{i,v}|})), \quad (3)$$

$$V_3(i) = (V_2(i, 1) \dots V_2(i, \lambda_i)), \quad V = (V_3(1) \dots V_3(S)).$$

Блочная матрица V как обыкновенная матрица имеет размерность $(N+1)(C_1+1)(C_2+1)(C_3+1) \times R$,

где $R = \sum_{i=1}^S C_{(n_0+1)r_1r_2r_3}^i$. Пусть

$$\{v_{i,v} [n, c_1, c_2, c_3] : n \in [0, N], c_\alpha \in [0, C_\alpha], \alpha = \overline{1, 3}, v \in \{1, \dots, \lambda_i\}, i \in \{1, \dots, S\}, \quad (4)$$

таковы, что матрица V , образованная из них по формулам (2), (3), удовлетворяет условию ортогональности

$$V^T V = \text{diag}\{d_{1,1}, \dots, d_{R,R}\}; \quad d_{\alpha, \alpha} > 0, \alpha = 1, \dots, R. \quad (5)$$

Тогда последовательности (4) называются ортогональными (в совокупности) входными последовательностями (ОВП) для 4D-НМДС (1).

2. Условия ортогональности для входной последовательности 4D-НМДС

Обозначим через $R_1(i, v, \bar{j}, \bar{\sigma}, \bar{\rho})$, $R_2(i, v)$, $R_3(i)$, R количество столбцов матрицы $V_1(i, v, \bar{j}, \bar{\sigma}, \bar{\rho})$, $V_2(i, v)$, $V_3(i)$, V соответственно. Ясно, что

$$R_1(i, v, \bar{j}, \bar{\sigma}, \bar{\rho}) = |\Gamma_{i,v}|, \quad R_2(i, v) = |\Gamma_{i,v}| \cdot |L_{i,v}|, \quad R_3(i) = \sum_{i=1}^{\lambda_i} R_2(i, v).$$

Теорема 1. Пусть имеют место формулы (2), (3). Для того чтобы входная последовательность (4) была ортогональной (в совокупности) входной последовательностью для 4D-НМДС (1), необходимо и достаточно, чтобы для всех $v \in \{1, \dots, \lambda_i\}$, $i \in \{1, \dots, S\}$, выполнялись соотношения

$$V_2(i, v)^T V_2(i, v) = \text{diag}\{d_{1,1}(2, i, v), \dots, d_{R_2(i,v), R_2(i,v)}(2, i, v)\}, \quad d_{\gamma,\gamma}(2, i, v) > 0, \quad \gamma = 1, \dots, R_2(i, v), \quad (6)$$

где $d_{\gamma,\gamma}(2, i, v)$ есть элемент матрицы $V_2(i, v)^T V_2(i, v)$, а для всех $v \in \{1, \dots, \lambda_i\}$, $i \in \{1, \dots, S\}$, $v' \in \{1, \dots, \lambda_{i'}\}$, $i' \in \{1, \dots, S\}$, $(i, v) \neq (i', v')$ выполнялось соотношение

$$V_2(i, v)^T V_2(i', v') = 0_{R_2(i,v) \times R_2(i',v')}, \quad (7)$$

где в (7) через $0_{R_2(i,v) \times R_2(i',v')}$ обозначена нулевая матрица с размерностью $R_2(i, v) \times R_2(i', v')$.

Доказательство. По определению матрицы V имеем:

$$V^T V = (V_3^T(\alpha) V_3(\beta)), \quad \alpha = \overline{1, S}, \quad \beta = \overline{1, S}. \quad (8)$$

По (8) для того, чтобы V удовлетворяла условию ортогональности, необходимо и достаточно, чтобы для всех $\alpha \in \{1, \dots, S\}$ выполнялись соотношения

$$V_3(\alpha)^T V_3(\alpha) = \text{diag}\{d_{1,1}(3, \alpha), \dots, d_{R_3(\alpha), R_3(\alpha)}(3, \alpha)\}, \quad d_{\gamma,\gamma}(3, \alpha) > 0, \quad \gamma = 1, \dots, R_3(\alpha), \quad (9)$$

где $d_{\gamma,\gamma}(3, \alpha)$ есть элемент матрицы $V_3(\alpha)^T V_3(\alpha)$, а для всех $\alpha \in \{1, \dots, S\}$, $\beta \in \{1, \dots, S\}$, $\alpha \neq \beta$ выполнялось соотношение

$$V_3(\alpha)^T V_3(\beta) = 0_{R_3(\alpha), R_3(\beta)}. \quad (10)$$

Из формулы $V_3(i) = (V_2(i, 1) \dots V_2(i, \lambda_i))$ имеем

$$V_3(i)^T V_3(i) = (V_2(i, \alpha)^T V_2(i, \beta)), \quad \alpha = \overline{1, \lambda_i}, \quad \beta = \overline{1, \lambda_i}. \quad (11)$$

Из (11) следует, что для того, чтобы $V_3(\alpha)$ удовлетворяла условию ортогональности (9), необходимо и достаточно, чтобы для всех $v \in \{1, \dots, \lambda_i\}$ выполнялись

$$V_2(i, v)^T V_2(i, v) = \text{diag}\{d_{1,1}(2, i, v), \dots, d_{R_2(i,v), R_2(i,v)}(2, i, v)\}, \quad d_{\gamma,\gamma}(2, i, v) > 0, \quad \gamma = 1, \dots, R_2(i, v), \quad (12)$$

где $d_{\gamma,\gamma}(2, i, v)$ есть элемент матрицы $V_2(i, v)^T V_2(i, v)$, а для всех $\alpha \in \{1, \dots, \lambda_i\}$, $\beta \in \{1, \dots, \lambda_i\}$, $\alpha \neq \beta$ выполнялось

$$V_2(i, \alpha)^T V_2(i, \beta) = 0_{R_2(i,\alpha), R_2(i,\beta)}. \quad (13)$$

Таким образом, учитывая соотношения (8)–(13), для удовлетворения матрицей V условия ортогональности (5) необходимыми и достаточными условиями являются условия (6), (7). Теорема доказана.

Соотношения (6) есть условие собственной ортогональности каждой последовательности (4), а соотношения (7) – взаимной ортогональности последовательностей (4) и последовательностей

$$\{v_{i',v'}[n, c_1, c_2, c_3] : n \in [0, N], \quad c_\alpha \in [0, C_\alpha], \alpha = \overline{1, 3}, \quad v' \in \{1, \dots, \lambda_{i'}\}, \quad i' \in \{1, \dots, S\}.$$

Теорема 2. Пусть $v \in \{1, \dots, \lambda_i\}$, $i \in \{1, \dots, S\}$. Для собственной ортогональности последовательностей $\{v_{i,v}[n, c_1, c_2, c_3] : n \in [0, N], c_\alpha \in [0, C_\alpha], \alpha = \overline{1, 3}\}$ необходимо и достаточно, чтобы для всех $(\bar{j}, \bar{\sigma}, \bar{\rho}) \in L_{i,v}$ выполнялись соотношения

$$V_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho}))^T V_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho})) = \text{diag}\{d_{1,1}(1), \dots, d_{R_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho})), R_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho}))}(1)\},$$

$$d_{\alpha, \alpha}(1) > 0, \alpha = 1, \dots, R_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho})),$$
(14)

где $d_{\alpha, \alpha}(1)$ есть элемент матрицы $V_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho}))^T \cdot V_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho}))$, а для всех $(\bar{j}, \bar{\sigma}, \bar{\rho}) \in L_{i,v}$, $(\bar{j}', \bar{\sigma}', \bar{\rho}') \in L_{i,v}$, $(\bar{j}, \bar{\sigma}, \bar{\rho}) \neq (\bar{j}', \bar{\sigma}', \bar{\rho}')$ выполнялось соотношение

$$V_1(i, v, \bar{j}, \bar{\sigma}, \bar{\rho})^T \cdot V_1(i, v, \bar{j}', \bar{\sigma}', \bar{\rho}') = 0_{R_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho})) \times V_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho}))}. \quad (15)$$

Доказательство. По формулам (3) можем записать матрицу $V_2(i, v)^T V_2(i, v)$ в следующем компактном виде:

$$V_2(i, v)^T V_2(i, v) = (V_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho}))_\alpha)^T V_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho}))_\beta), \alpha = 1, \overline{|L_{i,v}|}, \beta = 1, \overline{|L_{i,v}|}. \quad (16)$$

По (16) ясно, что в матрице $V_2(i, v)^T V_2(i, v)$ на главной диагонали стоят элементы

$$V_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho}))_\alpha^T \cdot V_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho}))_\alpha, \alpha = 1, \overline{|L_{i,v}|},$$

а элементы, стоящие вне главной диагонали, есть элементы

$$V_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho}))_\alpha^T \cdot V_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho}))_\beta, \alpha = 1, \overline{|L_{i,v}|}, \beta = 1, \overline{|L_{i,v}|}, \alpha \neq \beta.$$

Поэтому если для каждого $\alpha = \{1, \dots, |L_{i,v}|\}$ в матрице $V_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho}))_\alpha^T V_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho}))_\alpha$ все диагональные элементы ненулевые, тогда выполняется условие (14), а если для каждого $\alpha \in \{1, \dots, |L_{i,v}|\}$, $\beta \in \{1, \dots, |L_{i,v}|\}$, $\alpha \neq \beta$, матрица $V_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho}))_\alpha^T V_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho}))_\beta$ есть нулевая матрица, тогда выполняется условие (15). Таким образом, для удовлетворения соотношения (6) необходимо и достаточно, чтобы выполнялись условия (14) и (15). Теорема доказана.

Теорема 3. Пусть: I. а) Для всех $v \in \{1, \dots, \lambda_i\}$, $i \in \{1, \dots, S\}$ вспомогательная последовательность $\bar{v}_{i,v}[n, c_1, c_2, c_3]$ является $\{0, 1\}$ – последовательностью с периодом $T(i, v) + 1$, $A_1(i, v) + 1$, $A_2(i, v) + 1$ и $A_3(i, v) + 1$ соответственно по аргументам n , c_1 , c_2 и c_3 , где $T(i, v) + 1 < N$, $A_\alpha(i, v) + 1 < C_\alpha$, $\alpha = \overline{1, 3}$;

б) Для всех $v \in \{1, \dots, \lambda_i\}$, $i \in \{1, \dots, S\}$ справедливы соотношения

$$\bar{V}_2(i, v)^T \bar{V}_2(i, v) = \text{diag}\{d_{1,1}(2, i, v), \dots, d_{R_2(i, v), R_2(i, v)}(2, i, v)\},$$

$$d_{\gamma, \gamma}(2, i, v) > 0, \gamma = 1, \dots, R_2(i, v),$$
(17)

где $d_{\gamma, \gamma}(2, i, v)$ есть элемент матрицы $\bar{V}_2(i, v)^T \bar{V}_2(i, v)$, а матрица $\bar{V}_2(i, v)$ образована из последовательностей $\{\bar{v}_{i,v}[n, c_1, c_2, c_3] : n \in [0, T(i, v)], c_1 \in [0, A_1(i, v)], c_2 \in [0, A_2(i, v)], c_3 \in [0, A_3(i, v)]\}$ последовательно по формулам:

$$\bar{V}_0(i, v, \bar{j}, \bar{\sigma}, \bar{\rho}, \bar{\tau}_k) =$$

$$= \left\{ \prod_{(\alpha, \beta, \gamma) \in Q_{i,v}} \prod_{\xi_{\alpha, \beta, \gamma} = 1}^{m_{\alpha, \beta, \gamma}} \bar{v}_{i,v}[n - \tau^{(k)}(\alpha, \beta, \gamma, \xi_{\alpha, \beta, \gamma}), c_1 + p_1(j_\alpha), c_2 + p_2(\sigma_\beta), c_3 + p_3(\rho_\gamma)] \right\}. \quad (18)$$

$$\bar{V}_1(i, v, \bar{j}, \bar{\sigma}, \bar{\rho}) = (\bar{V}_0(i, v, \bar{j}, \bar{\sigma}, \bar{\rho}, \bar{\tau}_1) \dots \bar{V}_0(i, v, \bar{j}, \bar{\sigma}, \bar{\rho}, \bar{\tau}_{|L_{i,v}|})), \quad (19)$$

$$\bar{V}_2(i, v) = (\bar{V}_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho}))_1 \dots \bar{V}_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho}))_{|L_{i,v}|}). \quad (20)$$

II. а) Для всех $v \in \{1, \dots, \lambda_i\}$, $i \in \{1, \dots, S\}$ и

$$(n, c_1, c_2, c_3) \in [0, T(i, v)] \times [0, A_1(i, v)] \times [0, A_2(i, v)] \times [0, A_3(i, v)]$$

последовательность $\bar{v}'_{i,v}[n, c_1, c_2, c_3]$ определяется по формуле

$$\bar{v}'_{i,v}[n, c_1, c_2, c_3] = \begin{cases} \bar{v}_{i,v}[n, c_1, c_2, c_3], & \text{если } (n, c_1, c_2, c_3) \in F(i, v) \times \left(\times_{\alpha=1}^3 G_{\alpha}(i, v) \right), \\ 0, & \text{если } (n, c_1, c_2, c_3) \notin F(i, v) \times \left(\times_{\alpha=1}^3 G_{\alpha}(i, v) \right), \end{cases} \quad (21)$$

где

$$\begin{aligned} F(i, v) &= [N_1(i, v) - \tau(i, v), N_1(i, v) - \tau(i, v) + T(i, v)] \subset [0, T'] , \\ G_{\alpha}(i, v) &= [D_{\alpha}(i, v), D_{\alpha}(i, v) + A_{\alpha}(i, v)] \subset [0, C'_{\alpha}] , \alpha = \overline{1, 3}, \\ \tau(i, v) &= \begin{cases} \max\{m_{1,1,1}, \dots, m_{1,1,\ell_3}, m_{1,2,1}, \dots, m_{1,2,\ell_3}, \dots, m_{\ell_1,\ell_2,\ell_3}\} - 1, & \text{если } N_1(i, v) > 0, \\ 0, & \text{если } N_1(i, v) = 0. \end{cases} \end{aligned}$$

б) Для всех $v \in \{1, \dots, \lambda_i\}$, $i \in \{1, \dots, S\}$ натуральные числа $N_1(i, v)$, $D_1(i, v)$, $D_2(i, v)$, $D_3(i, v)$ и область $[0, T'] \times [0, C'_1] \times [0, C'_2] \times [0, C'_3]$ таковы, что для любых $v' \in \{1, \dots, \lambda_{i'}\}$, $i' \in \{1, \dots, S\}$, $\langle i, v \rangle \neq \langle i', v' \rangle$, выполняются соотношения $F(i, v) \cap F(i', v') = \emptyset$ или $G_{\alpha}(i, v) \cap G_{\alpha}(i', v') = \emptyset$, $\alpha = \overline{1, 3}$.

III. Для всех $v \in \{1, \dots, \lambda_i\}$, $i \in \{1, \dots, S\}$ последовательность $\bar{v}_{i,v}[n, c_1, c_2, c_3]$ есть периодическое продолжение последовательности $\bar{v}'_{i,v}[n, c_1, c_2, c_3]$ из области $[0, T'] \times [0, C'_1] \times [0, C'_2] \times [0, C'_3]$ в остальные части области $[0, N] \times [0, C_1] \times [0, C_2] \times [0, C_3]$ с периодом $T(i, v) + 1$, $A_1(i, v) + 1$, $A_2(i, v) + 1$ и $A_3(i, v) + 1$ соответственно по аргументам n , c_1 , c_2 и c_3 . Тогда $\bar{v}_{i,v}[n, c_1, c_2, c_3]$, $v \in \{1, \dots, \lambda_i\}$, $i \in \{1, \dots, S\}$, есть ортогональные входные последовательности для 4D-НМДС (1).

С учетом условия (17) и $T(i, v) + 1 < N$, $A_{\alpha}(i, v) + 1 < C_{\alpha}$, $\alpha = \overline{1, 3}$, для каждого $v \in \{1, \dots, \lambda_i\}$, $i \in \{1, \dots, S\}$ вспомогательной последовательностью $\bar{v}_{i,v}[n, c_1, c_2, c_3]$ называются короткие вспомогательные ортогональные последовательности.

Условие I теоремы 3 есть условие независимости ортогональных входных последовательностей, условие II есть условие разделения ортогональных входных последовательностей друг от друга по временной области и пространственной области, а условие III – условие периодичности ортогональной входной последовательности.

Для построения ортогональных входных последовательностей $\bar{v}_{i,v}[n, c_1, c_2, c_3]$, $v \in \{1, \dots, \lambda_i\}$, $i \in \{1, \dots, S\}$, можем использовать методику построения ортогональных входных последовательностей, состоящую из следующих этапов:

1. Построение коротких ортогональных вспомогательных последовательностей $\bar{v}_{i,v}[n, c_1, c_2, c_3]$, $v \in \{1, \dots, \lambda_i\}$, $i \in \{1, \dots, S\}$, в соответствие с условием I теоремы 3 в отдельности, т.е. независимо от $\bar{v}_{i',v'}[n, c_1, c_2, c_3]$, $v' \in \{1, \dots, \lambda_{i'}\}$, $i' \in \{1, \dots, S\}$, $\langle i, v \rangle \neq \langle i', v' \rangle$;

2. В соответствии с условием II теоремы 3, разделяя область определения ортогональных последовательностей $\bar{v}_{i,v}[n, c_1, c_2, c_3]$, $v \in \{1, \dots, \lambda_i\}$, $i \in \{1, \dots, S\}$, по аргументу n , или c_1 , или c_2 , или c_3 , или по двум, или трем, или четырем аргументам, по формуле (21) строятся последовательности $\bar{v}'_{i,v}[n, c_1, c_2, c_3]$, $v \in \{1, \dots, \lambda_i\}$, $i \in \{1, \dots, S\}$, в области $[0, T'] \times [0, C'_1] \times [0, C'_2] \times [0, C'_3]$, где

$$[0, T'] \times [0, C'_1] \times [0, C'_2] \times [0, C'_3] \subset [0, N] \times [0, C_1] \times [0, C_2] \times [0, C_3];$$

3. В соответствии с условием III теоремы 3 с продолжением последовательностей $\bar{v}'_{i,v}[n, c_1, c_2, c_3]$, $v \in \{1, \dots, \lambda_i\}$, $i \in \{1, \dots, S\}$, из области $[0, T'] \times [0, C'_1] \times [0, C'_2] \times [0, C'_3]$ с периодом $T' + 1$, $C'_1 + 1$, $C'_2 + 1$ и $C'_3 + 1$ соответственно по аргументам n , c_1 , c_2 и c_3 в остальных частях области $[0, N] \times [0, C_1] \times [0, C_2] \times [0, C_3]$ строятся соответственно ортогональные входные последовательности $\bar{v}_{i,v}[n, c_1, c_2, c_3]$, $v \in \{1, \dots, \lambda_i\}$, $i \in \{1, \dots, S\}$.

Таким образом для построения ОВП $v_{i,v}[n, c_1, c_2, c_3]$, $v \in \{1, \dots, \lambda_i\}$, $i \in \{1, \dots, S\}$, сначала нужно выполнить построение ортогональных последовательностей $\bar{v}_{i,v}[n, c_1, c_2, c_3]$, $v \in \{1, \dots, \lambda_i\}$, $i \in \{1, \dots, S\}$.

3. Условия ортогональности коротких вспомогательных последовательности

Легко можно доказать следующую теорему.

Теорема 4. Пусть имеет место обозначение (18)–(20). Для того чтобы выполнялось условие ортогональности (17) необходимо и достаточно, чтобы:

а) для каждой четверки (n, c_1, c_2, c_3) в соответствующей ей строке в не более одной из матриц $\bar{V}_1(i, v, \bar{j}, \bar{\sigma}, \bar{\rho})$, $(\bar{j}, \bar{\sigma}, \bar{\rho}) \in L_{i,v}$, содержалось бы не более одного ненулевого элемента;

б) для каждого $(\bar{j}, \bar{\sigma}, \bar{\rho}) \in L_{i,v}$ все столбцы матрицы $\bar{V}_1(i, v, \bar{j}, \bar{\sigma}, \bar{\rho})$ содержали хотя бы один ненулевой элемент.

Через $\theta(i, v)$ обозначим количество ненулевых компонентов вектора

$$\bar{m}_{i,v} = (m_{1,1,1}, \dots, m_{1,1,\ell_3}, m_{1,2,1}, \dots, m_{1,2,\ell_3}, \dots, m_{1,\ell_2,1}, \dots, m_{1,\ell_2,\ell_3}).$$

Ясно, что $\theta(i, v) = |Q_{i,v}|$. Пусть последовательность ненулевых компонентов вектора $\bar{m}_{i,v}$ есть следующая последовательность:

$$\begin{aligned} & m_{1,\xi_{1,1},\eta(1,\xi_{1,1},1)}, \dots, m_{1,\xi_{1,1},\eta(1,\xi_{1,1},\pi(1,\xi_{1,1}))}, \dots, m_{1,\xi_{1,\phi(1)},\eta(1,\xi_{1,\phi(1)},1)}, \dots, m_{1,\xi_{1,\phi(1)},\eta(1,\xi_{1,\phi(1)},\pi(1,\xi_{1,\phi(1)}))}, \dots \\ & \dots, m_{2,\xi_{2,1},\eta(2,\xi_{2,1},1)}, \dots, m_{2,\xi_{2,1},\eta(2,\xi_{2,1},\pi(2,\xi_{2,1}))}, \dots, m_{2,\xi_{2,\phi(2)},\eta(2,\xi_{2,\phi(2)},1)}, \dots, m_{2,\xi_{2,\phi(2)},\eta(2,\xi_{2,\phi(2)},\pi(2,\xi_{2,\phi(2)}))}, \dots \\ & \dots, m_{\ell_1,\xi_{\ell_1,1},\eta(\ell_1,\xi_{\ell_1,1},1)}, \dots, m_{\ell_1,\xi_{\ell_1,1},\eta(\ell_1,\xi_{\ell_1,1},\pi(\ell_1,\xi_{\ell_1,1}))}, \dots, m_{\ell_1,\xi_{\ell_1,\phi(\ell_1)},\eta(\ell_1,\xi_{\ell_1,\phi(\ell_1)},1)}, \dots, m_{\ell_1,\xi_{\ell_1,\phi(\ell_1)},\eta(\ell_1,\xi_{\ell_1,\phi(\ell_1)},\pi(\ell_1,\xi_{\ell_1,\phi(\ell_1)}))}. \end{aligned} \quad (22)$$

Здесь для каждого $\alpha \in \{1, \dots, \ell_1\}$ считается, что $\phi(\alpha) \leq \ell_2$ и $1 \leq \xi_{\alpha,1} < \dots < \xi_{\alpha,\phi(\alpha)} \leq \ell_2$, а для каждого $\alpha \in \{1, \dots, \ell_1\}$ и $\beta \in \{\xi_{\alpha,1}, \dots, \xi_{\alpha,\phi(\alpha)}\}$ считается, что $\pi(\alpha, \beta) \leq \ell_3$ и $1 \leq \eta(\alpha, \beta, 1) < \dots < \eta(\alpha, \beta, \pi(\alpha, \beta)) \leq \ell_3$.

Ясно, что $\sum_{\alpha=1}^{\ell_1} \sum_{\beta=1}^{\phi(\alpha)} \eta(\alpha, \beta, \pi(\alpha, \beta)) = \theta(i, v)$.

Посмотрим следующие множества, построенные на основе первых, вторых и третьих индексов элементов последовательности (22):

$$\begin{aligned} \Omega(i, v) = & \{(1, \xi_{1,1}, \eta(1, \xi_{1,1}, 1)), \dots, (1, \xi_{1,1}, \eta(1, \xi_{1,1}, \pi(1, \xi_{1,1}))), \dots, (1, \xi_{1,\phi(1)}, \eta(1, \xi_{1,\phi(1)}, 1)), \dots \\ & \dots, (1, \xi_{1,\phi(1)}, \eta(1, \xi_{1,\phi(1)}, \pi(1, \xi_{1,\phi(1)}))), (2, \xi_{2,1}, \eta(2, \xi_{2,1}, 1)), \dots, (2, \xi_{2,1}, \eta(2, \xi_{2,1}, \pi(2, \xi_{2,1}))), \dots \\ & \dots, (2, \xi_{2,\phi(2)}, \eta(2, \xi_{2,\phi(2)}, 1)), \dots, (2, \xi_{2,\phi(2)}, \eta(2, \xi_{2,\phi(2)}, \pi(2, \xi_{2,\phi(2)}))), \dots \\ & \dots, (\ell_1, \xi_{\ell_1,1}, \eta(\ell_1, \xi_{\ell_1,1}, 1)), \dots, (\ell_1, \xi_{\ell_1,1}, \eta(\ell_1, \xi_{\ell_1,1}, \pi(\ell_1, \xi_{\ell_1,1}))), \dots \\ & \dots, (\ell_1, \xi_{\ell_1,\phi(\ell_1)}, \eta(\ell_1, \xi_{\ell_1,\phi(\ell_1)}, 1)), \dots, (\ell_1, \xi_{\ell_1,\phi(\ell_1)}, \eta(\ell_1, \xi_{\ell_1,\phi(\ell_1)}, \pi(\ell_1, \xi_{\ell_1,\phi(\ell_1)})))\}, \\ \Omega_1(i, v) = & \{(j_1, \sigma_{\xi_{1,1}}, \rho_{\eta(1, \xi_{1,1}, 1)}), \dots, (j_1, \sigma_{\xi_{1,1}}, \rho_{\eta(1, \xi_{1,1}, \pi(1, \xi_{1,1}))}), \dots, (j_1, \sigma_{\xi_{1,\phi(1)}}, \rho_{\eta(1, \xi_{1,\phi(1)}, 1)}), \dots \\ & \dots, (j_1, \sigma_{\xi_{1,\phi(1)}}, \rho_{\eta(1, \xi_{1,\phi(1)}, \pi(1, \xi_{1,\phi(1)}))}), (j_2, \sigma_{\xi_{2,1}}, \rho_{\eta(2, \xi_{2,1}, 1)}), \dots, (j_2, \sigma_{\xi_{2,1}}, \rho_{\eta(2, \xi_{2,1}, \pi(2, \xi_{2,1}))}), \dots \\ & \dots, (j_2, \sigma_{\xi_{2,\phi(2)}}, \rho_{\eta(2, \xi_{2,\phi(2)}, 1)}), \dots, (j_2, \sigma_{\xi_{2,\phi(2)}}, \rho_{\eta(2, \xi_{2,\phi(2)}, \pi(2, \xi_{2,\phi(2)}))}), \dots, (j_{\ell_1}, \sigma_{\xi_{\ell_1,1}}, \rho_{\eta(\ell_1, \xi_{\ell_1,1}, 1)}), \dots \\ & \dots, (j_{\ell_1}, \sigma_{\xi_{\ell_1,1}}, \rho_{\eta(\ell_1, \xi_{\ell_1,1}, \pi(\ell_1, \xi_{\ell_1,1}))}), \dots, (j_{\ell_1}, \sigma_{\xi_{\ell_1,\phi(\ell_1)}}, \rho_{\eta(\ell_1, \xi_{\ell_1,\phi(\ell_1)}, 1)}), \dots, (j_{\ell_1}, \sigma_{\xi_{\ell_1,\phi(\ell_1)}}, \rho_{\eta(\ell_1, \xi_{\ell_1,\phi(\ell_1)}, \pi(\ell_1, \xi_{\ell_1,\phi(\ell_1)}))})\}. \end{aligned}$$

Пусть для фиксированных (i, v) удовлетворяются:

1. $A_1(i, v)$, $A_2(i, v)$ и $A_3(i, v)$ есть какие-либо натуральные числа и множества E , $M_1, \dots, M_{\theta(i, v)}$ образованы из элементов множества $[0, A_1(i, v)] \times [0, A_2(i, v)] \times [0, A_3(i, v)]$. Числа $A_1(i, v)$, $A_2(i, v)$ и $A_3(i, v)$ и эти множества таковы, что:

Теорема 5. Пусть для фиксированных (i, v) удовлетворяются условия (1)–(4) и $[0, J(i, v)] \times [0, A_1(i, v)] \times [0, A_2(i, v)] \times [0, A_3(i, v)]$ есть область определения последовательностей (функций) $\bar{v}_{i,v}[n, c_1, c_2, c_3]$. Если элементы множества $\{J_\mu + 1 | \mu = 1, \dots, \theta(i, v)\}$ взаимно простые числа, тогда матрица $\bar{V}_2(i, v)$, определяемая по (18)–(20), удовлетворяет условиям ортогональности (17).

Доказательство. Рассмотрим произвольное $(\bar{j}, \bar{\sigma}, \bar{\rho}) \in L_{i,v}$. Согласно п. «г» из условия 1 существует такая тройка $(c_1, c_2, c_3) \in [0, A_1(i, v)] \times [0, A_2(i, v)] \times [0, A_3(i, v)]$, при котором для всех $(\alpha, \beta, \gamma) \in \Omega(i, v)$ тройка $(c_1 + p_1(j_\alpha), c_2 + p_2(\sigma_\beta), c_3 + p_3(\rho_\gamma))$ входит во множество M_μ , где

$$\mu = \sum_{\alpha_1=1}^{\alpha-1} \sum_{\beta_1=1}^{\phi(\alpha_1)} \eta(\alpha_1, \beta_1, \pi(\alpha_1, \beta)) + \sum_{\beta_1=1}^{\beta-1} \eta(\alpha, \beta_1, \pi(\alpha, \beta)) + \gamma,$$

и для всех $(\alpha, \beta, \gamma) \in (\{1, \dots, r_1\} \times \{1, \dots, r_2\} \times \{1, \dots, r_3\}) \setminus \Omega_1(i, v)$ тройка $(c_1 + p_1(\alpha), c_2 + p_2(\beta), c_3 + p_3(\gamma))$ входит в множество E . Поскольку элементы множества $\{J_\mu + 1 | \mu = 1, \dots, \theta(i, v)\}$ – взаимно простые числа, то по лемме 3.6 [3, С. 59–60] матрица, образованная из строк $\bar{V}_1(i, v, \bar{j}, \bar{\sigma}, \bar{\rho})$, соответствующих набору (n, c_1, c_2, c_3) , $n \in [0, \Pi(i, v)]$, ортогональна. Таким образом, все столбцы матрицы $\bar{V}_1(i, v, \bar{j}, \bar{\sigma}, \bar{\rho})$ содержат не менее чем один ненулевой элемент.

Докажем, что любая строка матрицы $\bar{V}_1(i, v, \bar{j}, \bar{\sigma}, \bar{\rho})$ содержит не более чем один ненулевой элемент. Для этого рассмотрим произвольные $(c_1, c_2, c_3) \in [0, A_1(i, v)] \times [0, A_2(i, v)] \times [0, A_3(i, v)]$. По п. «а» условия 1 возможно два случая:

1. $\Lambda(i, v) \leq \theta(i, v)$. В этом случае для каждого $n \in [0, \Pi(i, v)]$ строка, соответствующая набору (n, c_1, c_2, c_3) , содержит только нулевые элементы.
2. $\Lambda(i, v) = \theta(i, v)$. В этом случае, поскольку элементы множества $\{J_\mu + 1 | \mu = 1, \dots, \theta(i, v)\}$ взаимно простые числа, матрица, образованная из строк матрицы $\bar{V}_1(i, v, \bar{j}, \bar{\sigma}, \bar{\rho})$, соответствующих набору (n, c_1, c_2, c_3) , $n \in [0, J(i, v)]$, ортогональна по лемме 3.6 [Там же]. Таким образом, любая строка матрицы $\bar{V}_1(i, v, \bar{j}, \bar{\sigma}, \bar{\rho})$ содержит не более чем один ненулевой элемент.

Теперь докажем, что если строка, соответствующая какому-либо произвольному набору (n, c_1, c_2, c_3) , в матрице $\bar{V}_1(i, v, \bar{j}, \bar{\sigma}, \bar{\rho})$ содержит ненулевой элемент, то строка, соответствующая набору (n, c_1, c_2, c_3) , в любой матрице из множества

$$\{\bar{V}_1(i, v, \bar{j}', \bar{\sigma}', \bar{\rho}') | (\bar{j}', \bar{\sigma}', \bar{\rho}') \in L_{i,v}, (\bar{j}', \bar{\sigma}', \bar{\rho}') \neq (\bar{j}, \bar{\sigma}, \bar{\rho})\}$$

содержит только нулевые элементы. Предположим противное. Пусть существует такой набор $(n', c'_1, c'_2, c'_3) \in [0, J(i, v)] \times [0, A_1(i, v)] \times [0, A_2(i, v)] \times [0, A_3(i, v)]$ и тройки $(\bar{j}', \bar{\sigma}', \bar{\rho}')$ и $(\bar{j}, \bar{\sigma}, \bar{\rho})$, для которых строки, соответствующие тройке (n', c'_1, c'_2, c'_3) матрицы $\bar{V}_1(i, v, \bar{j}, \bar{\sigma}, \bar{\rho})$ и матрицы $\bar{V}_1(i, v, \bar{j}', \bar{\sigma}', \bar{\rho}')$, содержат ненулевые элементы, где $(\bar{j}', \bar{\sigma}', \bar{\rho}') \neq (\bar{j}, \bar{\sigma}, \bar{\rho})$. Не умаляя общности, предположим, что $\bar{j} = \bar{j}'$, $\bar{\sigma} = \bar{\sigma}'$, а $\bar{\rho} \neq \bar{\rho}'$. По определению $\bar{\rho} = (\rho_1, \dots, \rho_{\ell_3})$, $\bar{\rho}' = (\rho'_1, \dots, \rho'_{\ell_3})$. Не умаляя общности, предположим, что существует такое $f \in \{1, \dots, \ell_3\}$, при котором $\rho_f \neq \rho'_f$ и $\rho_g = \rho'_g$, $g \neq f$.

Рассмотрим следующие множества с $\theta(i, v)$ элементом

$$\Sigma = \{(c_1 + p_1(j_\alpha), c_2 + p_2(\sigma_\beta), c_3 + p_3(\rho_\gamma)) | (\alpha, \beta, \gamma) \in Q_{i,v}\},$$

$$\Sigma' = \{(c_1 + p_1(j'_\alpha), c_2 + p_2(\sigma'_\beta), c_3 + p_3(\rho'_\gamma)) | (\alpha, \beta, \gamma) \in Q_{i,v}\}.$$

Все клетки (элементы) множества Σ и Σ' по одному входят в соответствующих множества во $M_1, \dots, M_{\theta(i, v)}$. Однако, кроме элементов $(c_1 + p_1(j_\alpha), c_2 + p_2(\sigma_\beta), c_3 + p_3(\rho_f))$, $\alpha = \overline{1, \ell_1}$, $\beta = \overline{1, \ell_2}$, все остальные элементы множества Σ входят и в множество Σ' . А кроме элементов

$(c_1 + p_1(j_\alpha), c_2 + p_2(\sigma_\beta), c_3 + p_3(\rho'_f))$, $\alpha = \overline{1, \ell_1}$, $\beta = \overline{1, \ell_2}$, все остальные элементы множества Σ' входят и в множество Σ . Поэтому количество клеток, входящих в множества $M_1, \dots, M_{\theta(i,v)}$, больше, чем число $\theta(i, v)$. Отсюда вытекает, что $\Lambda(i, v) > \theta(i, v)$. А это противоречит п. «б» условия 2. Таким образом, полностью удовлетворяются условия теоремы 4. Поэтому матрица $\bar{V}_2(i, v)$, определяемая по (18)–(20), удовлетворяет условиям ортогональности (17). Теорема доказана.

Заключение

В работе для двоичных 4D-модулярных динамических систем, заданных в виде двухзначного аналога полинома Вольтерры, приведено понятие ортогональной входной последовательности. Доказана теорема о необходимом и достаточном условиях ортогональности в совокупности входных последовательностей двоичных 4D-НМДС (1). Доказана теорема о необходимом и достаточном условиях собственной ортогональности каждой входной последовательности двоичных 4D-НМДС (1). Предложена методика построения ортогональной входной последовательности на базе коротких вспомогательных ортогональных последовательностей. Приведены достаточные условия ортогональности коротких вспомогательных последовательностей и формулы их определения.

ЛИТЕРАТУРА

1. Фараджев Р.Г. Линейные последовательностные машины. М.: Сов. радио, 1975. 248 с.
2. Блюмин С.Л., Фараджев Р.Г. Линейные клеточные машины: подход пространства состояний (обзор) // Автоматика и телемеханика. 1982. № 2. С. 125–163.
3. Фараджев Р.Г., Фейзи́ев Ф.Г. Методы и алгоритмы решения задачи квадратичной оптимизации для двоичных последовательностных машин. Баку: Элм, 1996. 180 с.
4. Фейзи́ев Ф.Г., Фараджева М.Р. Модулярные последовательностные машины: основные результаты по теории и приложению. Баку: Элм, 2006. 234 с.
5. Блейхут Р. Теория и практика кодов, контролирующих ошибки: пер. с англ. М.: Мир, 1986. 576 с.
6. Блюмин С.Л., Корнеев А.М. Дискретное моделирование систем автоматизации и управления. Липецк: Липецкий экологический гуманитар. ин-т, 2005. 124 с.
7. Nagiyev A.T., Feyziyev F.G. The sequential cellular-machining model of the continuous objects with distributing parameters // Seminarberichte, Fachbereich Mathematic. 2001. Bd 71. P. 31–43.
8. Скобелев В.В. Автоматы на алгебраических структурах (обзор) // Известия Саратовского университета. Новая сер. Сер. Математика. Механика. Информатика. 2013. Т. 13, вып. 2, ч. 2. С. 58–66.
9. Фараджев Р.Г., Нагиев А.Т., Гусейнов И.Н. Критерии диагностируемости билинейных последовательностных машин // Доклады РАН. 1998. Т. 361, № 5. С. 606–607.
10. Сперанский Д.В. Эксперименты с нечеткими автоматами // Автоматика и телемеханика. 2015. № 2. С. 107–124.
11. Сперанский Д.В. Эксперименты с нестационарными билинейными автоматами // Автоматика и телемеханика. 2015. № 9. С. 161–174.
12. Сперанский Д.В. Тестирование нечетких линейных автоматов // Известия Саратовского университета. Новая сер. Сер. Математика. Механика. Информатика. 2019. Т. 19, вып. 2. С. 233–240.
13. Hacı Y. Optimal control problem for processes with multiparametric binary linear difference equation system // Applied and computational mathematics. 2009. V. 8, № 2. P. 263–269.
14. Hacı Y., Özen K. Terminal optimal control problem for processes represented by nonlinear multi-parametric binary dynamical system // Control and cybernetics. 2009. V. 38, № 3. P. 625–633.
15. Hacı Y., Candan M., Or A. On the Principle of Optimality for Linear Stochastic Dynamical System // International Journal in Foundations of Computer Science and Technology. 2016. V. 6, № 1. P. 57–63.
16. Байбатшаев М.Ш., Попков Ю.С. Об одной задаче квадратичной оптимизации двоичных нелинейных последовательностных машин // Автоматика и телемеханика. 1978. № 12. С. 37–47.
17. Фараджев Р.Г., Фейзи́ев Ф.Г. К задаче квадратичной оптимизации для двоичных многомерных нелинейных последовательно-клеточных машин // Автоматика и телемеханика. 1996. № 5. С. 104–119.
18. Ф.Г.Фейзи́ев, З.А.Самедова. Условия ортогональности для входных последовательностей двоичных 3D-нелинейных модулярных динамических систем // Известия НАН Азербайджана. Сер. физ.-техн. и мат. наук. Информатика и проблемы управления. 2010. Т. 30, № 3. С. 115–124.
19. Фейзи́ев Ф.Г., Абаева Н.Б. Полиномиальное соотношение для представления полной реакции одного класса двоичных 4D-модулярных динамических систем // Вестник Пермского университета. Сер. Математика. Механика. Информатика. 2019. Вып. 2 (45). С. 46–54.

20. Фейзи́ев Ф.Г., Абаева Н.Б. Задача оптимального синтеза двоичных 4D-нелинейных модулярных динамических систем // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2020. № 53. С. 102–109.

Поступила в редакцию 23 декабря 2020 г.

Feyziyev F.G., Abayeva N.B. (2021) THE CONDITIONS OF ORTHOGONALITY OF THE INPUT SEQUENCES OF ONE CLASS OF BINARY 4D-NONLINEAR MODULAR DYNAMIC SYSTEMS. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika* [Tomsk State University Journal of Control and Computer Science]. 55. pp. 80–90

DOI: 10.17223/19988605/55/10

Binary 4D-nonlinear modular dynamic system (4D-NMDS) with fixed memory n_0 , limited connection $P_1 \times P_2 \times P_3$, with the degree S , described by two-valued analogue of Volterra's polynomial, is considered

$$y[n, c_1, c_2, c_3] = \sum_{i=1}^S \sum_{v=1}^{\lambda_i} \sum_{(\bar{j}, \bar{\sigma}, \bar{\rho}) \in L_{i,v}} \sum_{\bar{\tau} \in \Gamma_{i,v}} h_{i,v}[\bar{j}, \bar{\sigma}, \bar{\rho}, \bar{\tau}] \times \prod_{(\alpha, \beta, \gamma) \in Q_{i,v}} \prod_{\xi_{\alpha, \beta, \gamma}=1}^{m_{\alpha, \beta, \gamma}} v_{i,v}[n - \tau(\alpha, \beta, \gamma, \xi_{\alpha, \beta, \gamma}), c_1 + p_1(j_\alpha), c_2 + p_2(\sigma_\beta), c_3 + p_3(\rho_\gamma)], GF(2). \quad (1)$$

Here $n \in [0, N] \equiv \{0, 1, \dots, N\}$, $c_\alpha \in [0, C_\alpha] \equiv \{0, 1, \dots, C_\alpha\}$, $\alpha = \overline{1, 3}$; $P_\alpha = \{p_\alpha(1), \dots, p_\alpha(r_\alpha)\}$, $-\infty < p_\alpha(1) < \dots < p_\alpha(r_\alpha) < \infty$, $p_\alpha(j) \in \{\dots, -1, 0, 1, \dots\}$, $j = 1, \dots, r_\alpha$, $\alpha = \overline{1, 3}$. On the base of binary input sequences

$$\{v_{i,v}[n, c_1, c_2, c_3] : n \in [0, N], c_1 \in [0, C_1], c_2 \in [0, C_2], c_3 \in [0, C_3]\}, \quad v = 1, \dots, \lambda_i, \quad i = 1, \dots, S, \quad (2)$$

the matrices $V_0(i, v, \bar{j}, \bar{\sigma}, \bar{\rho}, \bar{\tau})$, $V_1(i, v, \bar{j}, \bar{\sigma}, \bar{\rho})$, $V_2(i, v)$, $V_3(i)$, V are constructed sequentially. Let a number of columns of this matrices be

$R_1(i, v, \bar{j}, \bar{\sigma}, \bar{\rho})$, $R_2(i, v)$, $R_3(i)$, R respectively. All these matrices have $(N+1)(C_1+1)(C_2+1)(C_3+1)$ rows. Here $R = \sum_{i=1}^S C_{(n_0+1)r_1r_2r_3}^i$.

If sequences (2) are such, that satisfied the conditions $V^T V = \text{diag}\{d_{1,1}, \dots, d_{R,R}\}$, $d_{\alpha,\alpha} > 0$, $\alpha = 1, \dots, R$, then sequences (2) are called the orthogonal input sequences for 4D-NMDS (1).

In order for the input sequence (2) to be an orthogonal input sequence for the 4D-LMDS (1), it is necessary and sufficient that for all $v \in \{1, \dots, \lambda_i\}$, $i \in \{1, \dots, S\}$ the relations hold

$$V_2(i, v)^T V_2(i, v) = \text{diag}\{d_{1,1}(2, i, v), \dots, d_{R_2(i, v), R_2(i, v)}(2, i, v)\}, \quad d_{\gamma, \gamma}(2, i, v) > 0, \quad \gamma = 1, \dots, R_2(i, v), \quad (3)$$

where $d_{\gamma, \gamma}(2, i, v)$ are elements of matrix $V_2(i, v)^T V_2(i, v)$, and for all $v \in \{1, \dots, \lambda_i\}$, $i \in \{1, \dots, S\}$, $v' \in \{1, \dots, \lambda_{i'}\}$, $i' \in \{1, \dots, S\}$, $(i, v) \neq (i', v')$ the ratio $V_2(i, v)^T V_2(i', v') = 0_{R_2(i, v) \times R_2(i', v')}$ was satisfied.

For own orthogonality of sequences $\{v_{i,v}[n, c_1, c_2, c_3] : n \in [0, N]\}$, $c_\alpha \in [0, C_\alpha]$, $\alpha = \overline{1, 3}$, it is necessary and sufficient that for all $(\bar{j}, \bar{\sigma}, \bar{\rho}) \in L_{i,v}$ the ratio

$$V_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho}))^T V_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho})) = \text{diag}\{d_{1,1}(1), \dots, d_{R_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho}))}, R_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho}))}(1)\}, \\ d_{\alpha, \alpha}(1) > 0, \quad \alpha = 1, \dots, R_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho})),$$

hold, and for all $(\bar{j}, \bar{\sigma}, \bar{\rho}) \in L_{i,v}$, $(\bar{j}', \bar{\sigma}', \bar{\rho}') \in L_{i,v}$, $(\bar{j}, \bar{\sigma}, \bar{\rho}) \neq (\bar{j}', \bar{\sigma}', \bar{\rho}')$ the ratio

$$V_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho}))^T \cdot V_1(i, v, (\bar{j}', \bar{\sigma}', \bar{\rho}')) = 0_{R_1(i, v, (\bar{j}, \bar{\sigma}, \bar{\rho})) \times R_1(i, v, (\bar{j}', \bar{\sigma}', \bar{\rho}'))}$$

was satisfied. The sufficient conditions for orthogonality of the input sequence for 4D-NMDS (1) are given. The technique to construct an orthogonal input sequences based on auxiliary orthogonal sequences is proposed. The sufficient conditions for orthogonality of short auxiliary sequences and formulas for their definition are given.

Keywords: 4D-nonlinear modular dynamic system; orthogonal input sequences; conditions of orthogonality..

FEYZIYEV Fikrat Gulali (Doctor of Physical and Mathematical Sciences, Professor, Sumgait State University, Sumgait, Azerbaijan). E-mail: FeyziyevFG@mail.ru

ABAYEVA Nigar Bahram (Dissertant of Ph.D in Mathematics, Sumgait State University, Sumgait, Azerbaijan). E-mail: abayeveldar404@gmail.com

REFERENCES

1. Faradzhev, R.G. (1975) *Lineynye posledovatel'nostnye mashiny* [Linear sequential machines]. Moscow: Sovetskoe radio.
2. Blyumin, S.L. & Faradzhev, R.G. (1982) Lineynye kletochnye mashiny: podkhod prostranstva sostoyaniy (obzor) [Linear cellular machine: The approach of the state space (review)]. *Avtomatika i telemekhanika—Automation and Remote Control*. 2. pp. 125–163.
3. Faradzhev, R.G. & Feyziev, F.G. (1996) *Metody i algoritmy resheniya zadachi kvadrachnoy optimizatsii dlya dvoichnykh posledovatel'nostnykh mashin* [Methods and algorithms for solving quadratic optimization problem for binary sequential machines]. Baku: Elm.
4. Feyziev, F.G. & Faradzheva, M.R. (2006) *Modulyarnye posledovatel'nostnye mashiny: osnovnye rezul'taty po teorii i prilozheniyu* [Modular sequential machine: The main results of the theory and application]. Baku: Elm.
5. Blahut, R. (1986) *Teoriya i praktika kodov, kontroliruyushchikh oshibki* [Theory and Practice of Error Control Codes]. Translated from English. Moscow: Mir.
6. Blyumin, S.L. & Korneev, A.M. (2005) *Diskretnoe modelirovanie sistem avtomatizatsii i upravleniya* [Discrete modeling automation and control systems]. Lipetsk: Lipetsk Ecological and Humanitarian Institute.
7. Nagiev, A.T. & Feyziyev, F.G. (2001) The sequential cellular-machining model of the continuous objects with distributing parameters. *Seminarberichte, Fachbereich Mathematic*. 71. pp. 31–43.
8. Skobelev, V.V. (2013) Automata on algebraic structures. *Izvestiya Saratovskogo universiteta. Novaya seriyaya. Seriya Matematika. Mechanica. Informatika – Izvestiya of Saratov University. Mathematics. Mechanics. Informatics*. 13(2-2). pp. 58–66. DOI: 10.18500/1816-9791-2013-13-2-2-58-66
9. Faradzhev, R.G., Nagiev, A.T. & Guseynov, I.N. (1998) Kriterii diagnostiruемости bilineynykh posledovatel'nostnykh mashin [The criteria of diagnosability for bilinear sequential machines]. *Doklady RAN*. 361(5). pp. 606–607.
10. Speranskiy, D.V. (2015) Eksperimenty s nechetkimi avtomatami [Experiments with fuzzy finite state machines]. *Avtomatika i telemekhanika—Automation and Remote Control*. 2. pp. 107–124.
11. Speranskii, D.V. (2015) Eksperimenty s nestatsionarnymi bilineynymi avtomatami [Experiments with nonstationary bilinear finite state machines]. *Avtomatika i telemekhanika—Automation and Remote Control*. 9. pp. 161–174.
12. Speranskii, D.V. (2019) Fuzzy linear automata testing. *Izvestiya Saratovskogo universiteta. Novaya seriyaya. Seriya Matematika. Mechanica. Informatika – Izvestiya of Saratov University. Mathematics. Mechanics. Informatics*. 19(2). pp. 233–240. DOI: 10.18500/1816-9791-2019-19-2-233-240
13. Haci, Y. (2009) Optimal control problem for processes with multiparametric binary linear difference equation system. *Applied and Computational Mathematics*. 8(2). pp. 263–269.
14. Haci, H. & Özen, K. (2009) Terminal optimal control problem for processes represented by nonlinear multiparametric binary dynamical system. *Control and Cybernetics*. 38(3). pp. 625–633.
15. Haci, Y., Candan, M. & Or, A. (2016) On the Principle of Optimality for Linear Stochastic Dynamical System. *International Journal in Foundations of Computer Science and Technology*. 6(1). pp. 57–63. DOI: 10.5121/ijfcs.2016.6105. 57
16. Baybatshaev, M.Sh. & Popkov, Yu.S. (1978) Ob odnoy zadache kvadrachnoy optimizatsii dvoichnykh nelineynykh posledovatel'nostnykh mashin [On one quadratic optimization problem for binary nonlinear sequential machines]. *Avtomatika i telemekhanika – Automation and Remote Control*. 12. pp. 37–47.
17. Faradzhev, R.G. & Feyziyev, F.G. (1996) K zadache kvadrachnoy optimizatsii dlya dvoichnykh mnogomernykh nelineynykh posledovatel'nostno-kletochnykh mashin [To the quadratic optimization problem for binary many-dimensional nonlinear sequential cellular machines]. *Avtomatika i telemekhanika – Automation and Remote Control*. 5. pp. 104–119.
18. Feyziyev, F.G. & Samedova, Z.A. (2010) Usloviya ortogonal'nosti dlya vkhodnykh posledovatel'nostey dvoichnykh 3D-nelineynykh modulyarnykh dinamicheskikh sistem [The conditions of orthogonality for input sequences for the binary 3D-nonlinear modular dynamical system]. *Izvestiya NAN Azerbaydzhana. Ser. fiz.-tekhn. i mat. nauk. Informatika i problemy upravleniya*. 30(3). pp. 115–124.
19. Feyziyev, F.G. & Abaeva, N.B. (2019) The polynomial ratio for description of full reaction of one classes binary 4D-multidimensional modular dynamic systems. *Vestnik Permskogo universiteta. Ser. Matematika. Mekhanika. Informatika – Bulletin of Perm University. Mathematics. Mechanics. Computer Science*. 2(45). pp. 46–54. DOI: 10.17072/1993-0550-2019-2-46-54
20. Feyziyev, F.G. & Abayeva, N.B. (2020) The problem of optimal synthesis of binary 4D- nonlinear modular dynamic systems. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 53. pp. 102–109. DOI: 10.17223/19988605/53/10

УДК 519.24

DOI: 10.17223/19988605/55/11

В.П. Шуленин

АСИМТОТИЧЕСКИЕ СВОЙСТВА И РОБАСТНОСТЬ ОЦЕНОК УРЕЗАННЫХ ВАРИАНТОВ СТАНДАРТНОГО ОТКЛОНЕНИЯ И СРЕДНЕГО АБСОЛЮТНЫХ ОТКЛОНЕНИЙ

Изучаются свойства робастных оценок масштабного параметра, который характеризует «разброс» случайной величины. Предложенные оценки асимптотически нормально распределены, имеют ограниченные функции влияния и, следовательно, в отличие от оценки стандартного отклонения, «защищены» от наличия выбросов в выборке. Рассматриваемые оценки вычисляются на основе упорядоченной статистики, из которой предварительно удаляется часть наблюдений. Предложен адаптивный вариант оценок, основанный на использовании выборочных оценок функционалов, характеризующих степень «затянутости хвостов» распределений. Приводятся результаты сравнения оценок масштабного параметра в условиях различных моделей наблюдений, в частности в рамках гауссовской модели с масштабным засорением.

Ключевые слова: масштабный параметр; робастные оценки; выбросы; функция влияния; адаптивные оценки.

При обработке результатов измерений $X_1, \dots, X_n$, полученных в эксперименте путем наблюдений за признаком X изучаемого объекта, традиционно вычисляются по статистическим данным $X_1, \dots, X_n$ оценки математического ожидания $M(X)$ и дисперсии $D(X)$ (или стандартного отклонения $[D(X)]^{1/2}$) и коэффициента корреляции, что вполне оправдано для гауссовской модели наблюдений. Однако оценки этих числовых характеристик изучаемой случайной величины X подвержены сильному влиянию грубых ошибок (выбросов), которые обычно присутствуют в реальных данных $X_1, \dots, X_n$ эксперимента [1–5]. Для больших совокупностей данных, особенно при их автоматизированной обработке на ЭВМ, тщательный анализ данных с целью обнаружения «выпадающих» наблюдений и их корректировки трудно осуществим. В таких ситуациях используются робастные процедуры [5–8], которые, с одной стороны, оказываются нечувствительными к наличию умеренного засорения данных грубыми ошибками, а с другой стороны, ведут себя «достаточно хорошо» при идеальных условиях нормальности или какого-либо другого предположения о типе распределения данных.

В настоящее время разработано большое количество робастных оценок параметра положения в качестве альтернативы оценке математического ожидания, которые имеют ограниченные функции влияния и «защищены» от наличия выбросов в выборке. Многие из этих оценок принадлежат к общим классам M-, L-, R- и MD-оценок, которые хорошо изучены и описаны в литературе (см., напр.: [6–12]).

Менее изученными являются робастные оценки масштабного параметра, который используется в качестве «меры», характеризующей степень разброса случайной величины, и определяется в виде функционала $S(F)$ от функции распределения F наблюдений. Общие требования, предъявляемые к таким функционалам, сформулированы в работах [2, 4].

Традиционно используемые на практике оценки масштабного параметра, такие как оценка $\hat{S}_1(0)$ стандартного отклонения $S_1(F)$, и оценка $\hat{S}_2(0)$ среднего абсолютных отклонений $S_2(F)$, имеют неограниченные функции влияния, и они очень чувствительны к наличию выбросов в выборке (см., напр.: [6–8]). В данной работе предложены урезанные варианты этих оценок $\hat{S}_1(\alpha)$ и $\hat{S}_2(\alpha)$, $0 \leq \alpha < 1/2$, которые вычисляются не по исходной выборке $X_1, \dots, X_n$, а на основе упорядоченной статистики $X_{(1)}, \dots, X_{(n)}$, из которой предварительно удаляются $[\alpha n]$ наименьших и наибольших поряд-

ковых статистик. Получены асимптотические формулы для вычисления характеристик качества оценок, основанные на «стандартизованных дисперсиях» оценок. Показано, что эти характеристики существенно зависят от параметра α . На практике это приводит к дополнительным усилиям по выбору данного параметра. По этой причине в работе предложены адаптивные варианты оценок, для которых параметр $\hat{\alpha}(X_1, \dots, X_n)$ доопределяется на основе информации, содержащейся в исходной выборке. Приводятся результаты сравнения оценок масштабного параметра в условиях различных моделей наблюдений.

Полученные результаты, приводят к следующему выводу. Предложенная адаптивная оценка $\hat{S}_1(\hat{\alpha})$ стандартного отклонения имеет ограниченную функцию влияния, она «защищена» от наличия выбросов в выборке, имеет высокую эффективность в условиях гауссовской модели и может быть рекомендована к использованию на практике для оценивания масштабного параметра при умеренных объемах выборки и при возможных отклонениях от модели, вызванных наличием выбросов в выборке.

1. Общий подход к построению оценок масштабного параметра

Введем необходимые понятия и обозначения. Пусть X – изучаемая случайная величина с функцией распределения (ф.р.) $F(x)$, $x \in R^1$, которая абсолютно непрерывна, имеет плотность $f(x)$, $x \in R^1$, и симметрична относительно точки θ_x , т.е. $F \in \mathfrak{S}_{S|0}$, где $\mathfrak{S}_{S|0} = \{F : F(x) = 1 - F(2\theta_x - x), \forall x \in R^1\}$. Везде ниже обратную (квантильную) функцию для ф.р. $F(x)$ будем обозначать через $F^{-1}(t) = \inf\{x : F(x) \geq t\}$, $0 \leq t \leq 1$. Масштабный параметр функции распределения F используется в качестве меры, характеризующей степень разброса случайной величины (с.в.) X с ф.р. F . Рассмотрим такие меры, которые могут быть представлены в виде функционала $S(F)$, $F \in \mathfrak{S}$, заданного на множестве допустимых распределений $\mathfrak{S}$ в условиях эксперимента, связанного с изучением с.в. X по статистическим данным $X_1, \dots, X_n$, полученным в серии n независимых и повторных наблюдений над с.в. X . Выборочная оценка $\hat{S}(X_1, \dots, X_n)$ функционала $S(F)$, $F \in \mathfrak{S}$, построенная методом подстановки, записывается в виде: $\hat{S}(X_1, \dots, X_n) = S(F_n)$, где $F_n(x)$ – эмпирическая функция распределения, построенная по выборке $X_1, \dots, X_n$. Везде ниже порядковые статистики исходной выборки $X_1, \dots, X_n$ будем обозначать через $X_{(1)}, \dots, X_{(n)}$. Общие требования, которым должен удовлетворять функционал $S(F)$, описывающий разброс случайной величины X , сформулированы в работах [2, 4]. Для формулировки этих требований отметим, что о случайных величинах X_1 и X_2 с функциями распределений F_1 и F_2 говорят, что с.в. X_2 стохастически больше чем с.в. X_1 (при этом используют обозначение в виде $F_1 <_{st} F_2$), если выполняется неравенство $P(X_1 > x) \leq P(X_2 > x)$, $\forall x \in R^1$. Отметим также, что $F_1 <_{st} F_2 \Rightarrow F_1(x) \geq F_2(x)$, $\forall x \in R^1$ и $F_1^{-1}(t) \leq F_2^{-1}(t)$, $0 \leq t \leq 1$. Далее, разброс с.в. X относительно θ_x (масштабный параметр с.в. X) определяют в терминах «расстояния» X от θ_x , т.е. с помощью величины $|X - \theta_x|$, при этом говорят, что с.в. X_1 имеет больший разброс относительно θ_{x_1} , чем с.в. X относительно θ_x , если с.в. $|X_1 - \theta_{x_1}|$ стохастически больше с.в. $|X - \theta_x|$. Функционал $S(F)$ определяет масштабный параметр ф.р. F , если он удовлетворяет следующим условиям.

Условие 1. Монотонность функционала $S(F)$ относительно стохастического возрастания распределений, т. е. выполняется выражение $S(F_1) \leq S(F_2)$ для $F_1 <_{st} F_2$, где F_1 и F_2 – функции распределения вероятностей случайных величин $|X_1 - \theta_{x_1}|$ и $|X_2 - \theta_x|$.

Условие 2. Оценка $\hat{S}(X_1, \dots, X_n) = S(F_n)$ функционала $S(F)$ является эквивариантной относительно линейных преобразований наблюдений $X_1, \dots, X_n$, т.е. для нее выполняется равенство $\hat{S}(aX_1 + b, \dots, aX_n + b) = |a| \hat{S}(X_1, \dots, X_n)$.

В литературе описаны различные подходы к построению «мер», характеризующих масштабный параметр (см., напр.: [6, 7]). Мы рассмотрим такие функционалы $S(F)$, которые построены с помощью отклонений каждого члена генеральной совокупности от некоторого «центрального» (типичного значения) с.в. X с ф.р. F . Обычно в качестве такого значения используется параметр положения, определенный функционалом $T(F)$, либо в виде среднего $T_1(F) = \int x dF(x) = M(X)$, либо в виде медианы $T_2(F) = F^{-1}(1/2) = MED(X)$. Обозначим через F_1 ф.р. случайной величины $|X - T(F)|$. Рассмотрим функционалы вида

$$S(F) = \left\{ \int_0^1 [F_1^{-1}(t)]^\gamma dV(t) \right\}^{1/\gamma}, \quad (1)$$

где $V(t)$ – некоторая функция распределения на $[0, 1]$ и $\gamma > 0$. Например, если в (1) в качестве параметра положения выбрать среднее значение $T_1(F)$, и положить $V(t) = t$, $0 \leq t \leq 1$, то при $\gamma = 1$ получим среднее абсолютных отклонений $S_2(F) = M\{|X - M(X)|\}$, оценка которого запишется в виде: $\hat{S}_2(0) = n^{-1} \sum |X_i - \bar{X}|$. При $\gamma = 2$ будем иметь стандартное отклонение $S_1(F) = \{M(X - M(X))^2\}^{1/2}$, оценка которого запишется в виде: $\hat{S}_1(0) = \{n^{-1} \sum (X_i - \bar{X})^2\}^{1/2}$. Если же положить $V(t) = t / (1 - \alpha)$, $0 \leq t \leq 1 - \alpha$, $0 \leq \alpha < 1/2$, то получим α -урезанные варианты $S_i(F, \alpha)$, $i = 1, 2$, $0 \leq \alpha < 1/2$, указанных мер масштабного параметра, оценки которых вычисляются на основе порядковых статистик $X_{(1)}, \dots, X_{(n)}$ (см. ниже).

Отметим еще раз, что для выбранного функционала $S(F)$, $F \in \mathfrak{F}$, описывающего масштабный параметр с.в. X с ф.р. F , его оценка строится методом подстановки и записывается в виде $\hat{S}(X_1, \dots, X_n) = S(F_n)$. Асимптотическая нормальность таких оценок масштабного параметра изучается методом Мизеса [9, 10] с использованием теоремы Слуцкого, центральной предельной теоремы и разложения вида:

$$S(F_n) = S(F) + \frac{1}{n} \sum_{i=1}^n IF(X_i; F, S) + o_p(n^{-1/2}). \quad (2)$$

Использование разложения (2) позволяет доказать асимптотическую нормальность оценки $\hat{S} = S(F_n)$, где $IF(x; F, S)$ – функция влияния Хампеля оценки $\hat{S} = S(F_n)$ функционала $S(F)$, $F \in \mathfrak{F}$ (см., напр.: [6, 13–15]). Функция влияния Хампеля определяется в виде:

$$IF(x; F, S) = \lim_{\lambda \rightarrow 0+} \frac{S[(1-\lambda)F + \lambda\Delta_x] - S(F)}{\lambda}, \quad 0 \leq \lambda \leq 1, \quad x \in R^1, \quad (3)$$

для тех $x \in R^1$, при которых предел существует. Здесь Δ_x обозначает «вырожденную функцию распределения» в точке $x \in R^1$, которая определяется в точке $x \in R^1$ через единичную функцию Хевисайда $C(y - x) = \begin{cases} 1, & y \geq x, \\ 0, & y < x \end{cases}$ в виде: $F_x(y) = C(y - x)$, $(y, x) \in R^1$. Асимптотическая дисперсия

$\sqrt{n}\hat{S}$ -оценки обозначается через $\sigma^2(F, \hat{S})$ и вычисляется по формуле

$$\sigma^2(F, \hat{S}) = \int_{-\infty}^{\infty} IF^2(x; F, S) dF(x). \quad (4)$$

Для сравнения различных оценок масштабного параметра при заданной ф.р. F будем использовать понятие асимптотической относительной эффективности, определенное через обратное отношение стандартизованных асимптотических дисперсий. Асимптотическую эффективность оценки $\hat{S}_1$ относительно $\hat{S}_2$ при заданной ф.р. F обозначим через $AO\mathfrak{E}_F(\hat{S}_1; \hat{S}_2)$ и, следуя работам [3, 7], определим в виде:

$$AO\mathfrak{D}_F(\hat{S}_1 : \hat{S}_2) = \frac{\tilde{\sigma}^2(F, \hat{S}_2)}{\tilde{\sigma}^2(F, \hat{S}_1)}, \quad (5)$$

где $\tilde{\sigma}^2(F, \hat{S})$ – стандартизованная дисперсия $\sqrt{n}\hat{S}$ -оценки, равная отношению асимптотической дисперсии к квадрату оцениваемого функционала, т.е.

$$\tilde{\sigma}^2(F, \hat{S}) = \sigma^2(F, \hat{S}) / S^2(F). \quad (6)$$

2. Урезанные варианты стандартного отклонения и среднего абсолютных отклонений

Пусть имеется последовательность $X_1, \dots, X_n$ н.о.р. случайных величин с непрерывной симметричной ф.р. $F(x)$, $x \in \mathbb{R}^1$, т.е. $F \in \mathfrak{S}_{\text{с}0}$, и везде ниже, согласно второму условию эквивариантности оценок, полагаем без потери общности точку симметрии $\theta = M(X) = 0$. Обозначим ф.р. случайной величины $Y = |X_1|$ через $G(x)$, а ф.р. случайной величины $Z = X_1^2$ обозначим через $H(x)$. Определим оценку α -урезанного стандартного отклонения в виде:

$$\hat{S}_1(\alpha) = \left\{ \frac{1}{n - [\alpha n]} \sum_{i=1}^{n - [\alpha n]} Z_{(i)} \right\}^{1/2}, \quad 0 \leq \alpha < 1/2, \quad (7)$$

где $Z_{(i)}$ – i -я порядковая статистика для преобразованной выборки $Z_1, \dots, Z_n$, $Z_i = X_i^2$, $i = 1, \dots, n$. Функционал, соответствующий этой оценке, может быть определен, либо через ф.р. $H(x)$ в виде:

$$S_1(H, \alpha) = \left\{ \frac{1}{1 - \alpha} \int_0^{H^{-1}(1-\alpha)} z dH(z) \right\}^{1/2}, \quad 0 \leq \alpha < 1/2, \quad (8)$$

либо через исходную ф.р. F в виде:

$$S_1(F, \alpha) = \left\{ \frac{1}{1 - \alpha} \int_{F^{-1}(\alpha/2)}^{F^{-1}(1-\alpha/2)} x^2 dF(x) \right\}^{1/2}, \quad 0 \leq \alpha < 1/2. \quad (9)$$

При $\alpha = 0$ оценка (7) превращается в обычную оценку стандартного отклонения

$$\hat{S}_1(0) = \left\{ n^{-1} \sum_{i=1}^n X_i^2 \right\}^{1/2}. \quad (10)$$

Оценку α -урезанного среднего абсолютных отклонений определим в виде:

$$\hat{S}_2(\alpha) = \frac{1}{n - [\alpha n]} \sum_{i=1}^{n - [\alpha n]} Y_{(i)}, \quad 0 \leq \alpha < 1/2, \quad (11)$$

где $Y_{(i)}$ – i -я порядковая статистика для преобразованной выборки $Y_1, \dots, Y_n$, $Y_i = |X_i|$, $i = 1, \dots, n$. Функционал, соответствующий этой оценке, имеет вид:

$$S_2(G, \alpha) = \frac{1}{1 - \alpha} \int_0^{1-\alpha} G^{-1}(t) dt, \quad 0 \leq \alpha < 1/2. \quad (12)$$

При $\alpha = 0$ оценка (11) превращается в обычную оценку среднего абсолютных отклонений

$$\hat{S}_2(0) = n^{-1} \sum_{i=1}^n |X_i|. \quad (13)$$

3. Асимптотические свойства оценок

Асимптотическая нормальность оценок (7) и (11) масштабного параметра изучается методом Мизеса [9, 10] с использованием теоремы Слуцкого, центральной предельной теоремы и теории

U-статистик на основе разложения (2) (см.: [8, 11, 16–19], где доказывалась асимптотическая нормальность оценок функционалов, построенных методом подстановки). Для сравнения рассмотренных оценок масштабного параметра, определенных функционалами (9) и (12) с использованием формулы (5) следует прежде всего вычислить их функции влияния (3), асимптотические дисперсии (4) и стандартизованные дисперсии, определенные в (6). Для удобства записи введем следующие обозначения:

$$I_0(F, \alpha) = 2 \int_0^{1-(\alpha/2)} \xi(t) dt, \quad I_1(F, \alpha) = \int_{\alpha/2}^{1-(\alpha/2)} \xi^2(t) dt, \quad I_2(F, \alpha) = \int_{\alpha/2}^{1-(\alpha/2)} \xi^4(t) dt, \\ \text{где } \xi(t) = F^{-1}(t), \quad 0 \leq t \leq 1. \quad (14)$$

Можно убедиться, что функция влияния $IF(u; H, \hat{S}_1(\alpha))$, $0 \leq u < \infty$, для $\hat{S}_1(\alpha)$ -оценки вида (7) записывается в виде:

$$IF(u; H, \hat{S}_1(\alpha)) = \frac{1}{2(1-\alpha)S_1(H, \alpha)} \begin{cases} u - b(\alpha), & 0 \leq u \leq H^{-1}(1-\alpha) \\ H^{-1}(1-\alpha) - b(\alpha), & H^{-1}(1-\alpha) < u < \infty \end{cases}, \\ b(\alpha) = \int_0^{H^{-1}(1-\alpha)} z dH(z) + \alpha H^{-1}(1-\alpha). \quad (15)$$

Используя приведенную формулу (15) и формулы (4) и (6) получаем, с учетом (14), выражение для вычисления стандартизованной дисперсии $\sqrt{n}\hat{S}_1(\alpha)$ -оценки в виде:

$$\tilde{\sigma}^2(F, \hat{S}_1(\alpha)) = \{I_2(F, \alpha) - I_1^2(F, \alpha) - 2\alpha\xi^2(1-\alpha/2)I_1(F, \alpha) + \alpha(1-\alpha)\xi^4(1-\alpha/2)\} / 4I_1^2(F, \alpha). \quad (16)$$

Далее, функция влияния $IF(y; G, \hat{S}_2(\alpha))$, $0 \leq y < \infty$ для $\hat{S}_2(\alpha)$ -оценки вида (11) записывается в виде:

$$IF(y; G, \hat{S}_2(\alpha)) = \frac{1}{(1-\alpha)} \begin{cases} y - c(\alpha), & 0 \leq y \leq G^{-1}(1-\alpha) \\ G^{-1}(1-\alpha) - c(\alpha), & G^{-1}(1-\alpha) < y < \infty \end{cases}, \quad c(\alpha) = \int_0^{G^{-1}(1-\alpha)} y dG(y) + \alpha G^{-1}(1-\alpha). \quad (17)$$

Используя приведенную формулу (17) и формулы (4) и (6) получаем, с учетом (14), выражение для вычисления стандартизованной дисперсии $\sqrt{n}\hat{S}_2(\alpha)$ -оценки в виде:

$$\tilde{\sigma}^2(F, \hat{S}_2(\alpha)) = \{I_1(F, \alpha) - I_0^2(F, \alpha) - 2\alpha\xi^2(1-\alpha/2)I_0(F, \alpha) + \alpha(1-\alpha)\xi^4(1-\alpha/2)\} / I_0^2(F, \alpha). \quad (18)$$

Отметим, что при $\alpha = 0$ имеем выражения для стандартизованных дисперсий $\sqrt{n}\hat{S}_1(0)$ - и $\sqrt{n}\hat{S}_2(0)$ -оценок в виде

$$\tilde{\sigma}^2(F, \hat{S}_1(0)) = \frac{\int x^4 dF(x) - \left(\int x^2 dF(x)\right)^2}{4\left(\int x^2 dF(x)\right)^2}, \quad \tilde{\sigma}^2(F, \hat{S}_2(0)) = \frac{\int x^2 dF(x) - \left(\int |x| dF(x)\right)^2}{\left(\int |x| dF(x)\right)^2}.$$

Замечание 1. Отметим, что приведенные функции влияния $\hat{S}_1(\alpha)$ -оценки и $\hat{S}_2(\alpha)$ -оценки (см. формулы (15) и (17)) являются *ограниченными* функциями при $0 < \alpha < 1/2$, следовательно, эти оценки «защищены» от наличия выбросов в выборке. Отметим также, что при изучении $\hat{S}_1(\alpha)$ - и $\hat{S}_2(\alpha)$ -оценок мы рассмотрели случай симметричных распределений при условии, что точка симметрии θ известна. Если она неизвестна, то для построения оценок масштабного параметра следует использовать оценки точки симметрии. Отметим, что при достаточно общих условиях асимптотические свойства $\hat{S}_1(\alpha)$ - и $\hat{S}_2(\alpha)$ -оценок не зависят от типа используемой оценки θ (см., напр.: [2, 3]).

Пример 1. Рассмотрим нормальную модель с засорением в виде: $\mathfrak{Z}_{\varepsilon, \tau}(\Phi) = \{F : F(x) = \Phi_{\varepsilon, \tau}(x)\}$, где $\Phi_{\varepsilon, \tau}(x) = (1-\varepsilon)\Phi(x) + \varepsilon\Phi(x/\tau)$, $0 \leq \varepsilon \leq 1$, $\tau \geq 1$. Обозначим через x^* квантиль уровня $1-(\alpha/2)$ для ф.р. $\Phi_{\varepsilon, \tau}(x)$, т.е. $x^* = \Phi_{\varepsilon, \tau}^{-1}(1-\alpha/2)$. С учетом этих обозначений выражения для I_0 , I_1 , I_2 из (14) для $F \in \mathfrak{Z}_{\varepsilon, \tau}(\Phi)$ записываются в виде:

$$\begin{aligned}
 I_0 &= (1-\varepsilon)[\sqrt{2/\pi} - 2\phi(x^*)] + \varepsilon\tau[\sqrt{2/\pi} - 2\phi(x^*/\tau)], \\
 I_1 &= (1-\varepsilon)[2\Phi(x^*) - 1 - 2x^*\phi(x^*)] + \varepsilon\tau^2[2\Phi(x^*/\tau) - 1 - 2\phi(x^*/\tau)x^*/\tau], \\
 I_2 &= (1-\varepsilon)[3\{2\Phi(x^*) - 1 - 2x^*\phi(x^*)\} - 2(x^*)^3\phi(x^*)] + \\
 &+ \varepsilon\tau^4[3\{2\Phi(x^*/\tau) - 1 - 2\phi(x^*/\tau)x^*/\tau\} - 2\phi(x^*/\tau)(x^*)^3/\tau],
 \end{aligned}$$

где Φ и ϕ обозначают соответственно функцию распределения и плотность стандартного нормального распределения и $x^* = \Phi_{\varepsilon,\tau}^{-1}(1-\alpha/2)$ – квантиль уровня $1-(\alpha/2)$ для ф.р. $\Phi_{\varepsilon,\tau}(x)$. Используя приведенные формулы, получаем выражения для стандартизованных дисперсий оценок $\hat{S}_1(\alpha)$ и $\hat{S}_2(\alpha)$ в виде:

$$\tilde{\sigma}^2(\Phi_{\varepsilon,\tau}, \hat{S}_1(\alpha)) = \{I_2 - I_1^2 - 2\alpha(x^*)^2 I_1 + \alpha(1-\alpha)(x^*)^4\} / 4I_1^2, \quad (19)$$

$$\tilde{\sigma}^2(\Phi_{\varepsilon,\tau}, \hat{S}_2(\alpha)) = \{I_1^2 - I_0^2 - 2\alpha(x^*)^2 I_0 + \alpha(1-\alpha)(x^*)^2\} / I_0^2. \quad (20)$$

В частности, при $\alpha \rightarrow 0$ получаем $I_0 = \sqrt{2/\pi}[1 + \varepsilon(\tau - 1)]$, $I_1 = 1 + \varepsilon(\tau^2 - 1)$, $I_2 = 3[1 + \varepsilon(\tau^4 - 1)]$.

Отметим, что для нормального распределения (при $\varepsilon=0$, или при $\tau=1$) имеем: $I_0 = \sqrt{2/\pi}$, $I_1 = 1$, $I_2 = 3$ и, следовательно,

$$\tilde{\sigma}^2(\Phi, \hat{S}_1(0)) = \frac{1}{4} \left(\frac{I_2 - I_1^2}{I_1^2} \right) = \frac{1}{4} \left(\frac{3-1}{1} \right) = 0,500, \quad \tilde{\sigma}^2(\Phi, \hat{S}_2(0)) = \frac{I_1^2 - I_0^2}{I_0^2} = \frac{\pi}{2} - 1 = 0,571.$$

Таким образом, асимптотическая относительная эффективность $\hat{S}_2(0)$ -оценки по отношению к $\hat{S}_1(0)$ для нормального распределения равна $AO\mathcal{E}_\Phi(\hat{S}_2(0) : \hat{S}_1(0)) = 0,500 / 0,572 = 0,876$, т.е. $\hat{S}_2(0)$ -оценка среднего абсолютных отклонений проигрывает $\hat{S}_1(0)$ -оценке стандартного отклонения по эффективности при нормальном распределении примерно 12%.

С учетом приведенных выражений стандартизованные дисперсии стандартного отклонения $\hat{S}_1(0)$ и среднего абсолютных отклонений $\hat{S}_2(0)$ для модели с засорением, т.е. для $F = \Phi_{\varepsilon,\tau}$, вычисляются по формулам

$$\begin{aligned}
 \tilde{\sigma}^2(\Phi_{\varepsilon,\tau}, \hat{S}_1(0)) &= \frac{3[1 + \varepsilon(\tau^4 - 1)] - [1 + \varepsilon(\tau^2 - 1)]}{4[1 + \varepsilon(\tau^2 - 1)]^2}, \\
 \tilde{\sigma}^2(\Phi_{\varepsilon,\tau}, \hat{S}_2(0)) &= \frac{\pi[1 + \varepsilon(\tau^2 - 1)] - 2[1 + \varepsilon(\tau - 1)]^2}{2[1 + \varepsilon(\tau - 1)]^2}.
 \end{aligned} \quad (21)$$

Согласно формулам (21), асимптотическая относительная эффективность $\hat{S}_2(0)$ -оценки по отношению к $\hat{S}_1(0)$ для супермодели $\mathfrak{T}_{\varepsilon,\tau}(\Phi)$ вычисляется по формуле

$$AO\mathcal{E}_F(\hat{S}_2(0) : \hat{S}_1(0)) = \frac{\{3[1 + \varepsilon(\tau^4 - 1)] - [1 + \varepsilon(\tau^2 - 1)]\}[1 + \varepsilon(\tau - 1)]^2}{2\{\pi[1 + \varepsilon(\tau^2 - 1)] - 2[1 + \varepsilon(\tau - 1)]^2\}[1 + \varepsilon(\tau^2 - 1)]^2}, \quad F \in \mathfrak{T}_{\varepsilon,\tau}(\Phi). \quad (22)$$

Из этой формулы также следует при $\varepsilon=0$, что $AO\mathcal{E}_\Phi(\hat{S}_2(0) : \hat{S}_1(0)) = 1/(\pi - 2) = 0,876$.

Приведем вычисленные по формуле (22) численные значения асимптотической относительной эффективности $\hat{S}_2(0)$ -оценки, $0 \leq \alpha < 1/2$, по отношению к $\hat{S}_1(0)$ для супермодели $\mathfrak{T}_{\varepsilon,\tau}(\Phi)$ (табл. 1).

Отметим, что значения эффективности $AO\mathcal{E}_{\Phi_{\varepsilon,\tau}}(\hat{S}_2(0) : \hat{S}_1(0))$, соответствующие большим значениям параметра τ и малым значениям ε , являются неожиданно высокими. По всей вероятности, это является следствием того факта, что функция влияния оценки $\hat{S}_1(0)$ не является ограниченной (см. также: [3]). Отметим также, что среднее абсолютных отклонений $\hat{S}_2(0)$, проигрывая лишь $\approx 12\%$ в эффективности стандартному отклонению $\hat{S}_1(0)$ при нормальном распределении, становится более эффективной при $\tau=3$ уже для $\varepsilon=0,002$.

Таблица 1

Значения $AO\mathfrak{E}_F(\hat{S}_2(0): \hat{S}_1(0))$ для супермодели $\mathfrak{S}_{\varepsilon, \tau}(\Phi)$

$\varepsilon \setminus \tau$	3	5	7	10	20
0,001	0,95	1,53	3,31	9,46	55,0
0,002	1,02	2,07	4,95	31,1	55,8
0,005	1,20	3,20	7,29	14,7	22,8
0,007	1,30	3,67	7,75	13,5	16,0
0,01	1,44	4,10	7,78	11,6	10,6
0,02	1,75	4,40	6,41	7,08	4,60
0,05	2,04	3,39	3,52	3,00	1,75
0,07	2,02	2,81	2,68	2,20	1,33
0,10	1,90	2,24	2,00	1,62	1,04
0,20	1,51	1,41	1,21	1,02	0,76

При умеренных значениях параметра τ ($\tau < 10$) и небольших значениях ε ($\alpha \leq 0,2$) среднее абсолютных отклонений $\hat{S}_2(0)$ предпочтительнее стандартного отклонения $\hat{S}_1(0)$.

Замечание 2. Заметим, что Эддингтон и Фишер придерживались противоположной точки зрения относительно преимуществ оценок $\hat{S}_1(0)$ и $\hat{S}_2(0)$ (см.: [7. С. 10]). Отметим, что $\hat{S}_1(0)$ и $\hat{S}_2(0)$ оценивают *разные* функционалы: $S_1(F) = \{M(X - M(X))^2\}^{1/2}$ и $S_2(F) = M\{|X - M(X)|\}$. И в условиях, например, нормальной модели наблюдений $N(0, \sigma^2)$ оценка $\hat{S}_1(0)$ сходится к σ , а оценка $\hat{S}_2(0)$ сходится к $\sigma\sqrt{2/\pi} \approx 0,80\sigma$. По этой причине, согласно [3, 7], при сравнении различных оценок масштаба используется не асимптотическая дисперсия, а *стандартизованная дисперсия* вида (6), и асимптотическая относительная эффективность определяется в виде (5).

Численные значения $AO\mathfrak{E}_F(\hat{S}_2(\alpha): \hat{S}_1(\alpha))$, для $F = \Phi_{\varepsilon, \tau}$, вычисленные по формулам (19) и (20), приведены в табл. (2).

Таблица 2

Значения $AO\mathfrak{E}_F(\hat{S}_2(\alpha): \hat{S}_1(\alpha))$, для $F = \Phi_{\varepsilon, \tau}$ при различных значениях параметра α

ε	$\tau \setminus \alpha$	$\alpha = 0,05$	$\alpha = 0,10$	$\alpha = 0,20$	$\alpha = 0,30$
$\varepsilon = 0,05$	$\varepsilon = 3$	0,928	0,892	0,876	0,872
	$\varepsilon = 5$	0,988	0,911	0,882	0,875
	$\varepsilon = 10$	1,089	0,929	0,885	0,887
$\varepsilon = 0,10$	$\varepsilon = 3$	1,032	0,930	0,890	0,879
	$\varepsilon = 5$	1,513	0,994	0,906	0,887
	$\varepsilon = 10$	3,401	1,104	0,920	0,907

Из приведенных данных следует, что при малых значениях параметра $\alpha \leq 0,10$, $\hat{S}_2(\alpha)$ -оценка предпочтительнее $\hat{S}_1(\alpha)$ -оценки лишь для распределений с «очень тяжелыми хвостами», в частности при $(\tau = 5, 10)$. В других случаях картина обратная, т.е. в рамках супермодели $\mathfrak{S}_{\varepsilon, \tau}(\Phi)$ урезанный вариант стандартного отклонения предпочтительнее.

4. Адаптивные оценки

Для построения $\hat{S}_1(\alpha)$ - и $\hat{S}_2(\alpha)$ -оценок на практике требуется задавать значения параметра α , $0 \leq \alpha < 1/2$, при изменении которого свойства этих оценок существенно меняются. Например, в условиях семейства распределений Стюдента с числом степеней свободы $r = 5$ при $\alpha = 0,10$ имеем $AO\mathfrak{E}_F(\hat{S}_1(0,10): \hat{S}_1(0)) = 2,35$, а при $\alpha = 0,20$, $AO\mathfrak{E}_F(\hat{S}_1(0,20): \hat{S}_1(0)) = 2,11$. Следуя работам [20, 21], выбор параметра α можно связать с поведением функционала $Q(F; \nu; \mu)$, характеризующего степень «тяжести хвостов» распределений, который определяется в виде:

$$Q(F; \nu, \mu) = \frac{(1/\nu) \left\{ \int_{1-\nu}^1 F^{-1}(t) dt - \int_0^\nu F^{-1}(t) dt \right\}}{(1/\mu) \left\{ \int_{1-\mu}^1 F^{-1}(t) dt - \int_0^\mu F^{-1}(t) dt \right\}}, \quad 0 < \nu < \mu \leq 0,5. \quad (23)$$

Оценка $Q(F_n)$ функционала $Q(F; \nu, \mu)$, построенная по выборке $X_1, \dots, X_n$ методом подстановки, записывается в виде:

$$Q(F_n; \nu, \mu) = \frac{m}{k} \left(\sum_{i=n-k+1}^n X_{(i)} - \sum_{i=1}^k X_{(i)} \right) / \left(\sum_{i=n-m+1}^n X_{(i)} - \sum_{i=1}^m X_{(i)} \right), \quad k = [\nu n], \quad m = [\mu n], \quad (24)$$

где $0 < \nu < \mu \leq 0,5$ и $X_{(1)}, \dots, X_{(n)}$ – порядковые статистики выборки $X_1, \dots, X_n$. Следуя работе [20],

везде ниже полагаем $\nu = 0,2$ и $\mu = 0,5$. Отметим, что $Q(F_n) \xrightarrow{p} Q(F; \nu, \mu)$ при $n \rightarrow \infty$.

Адаптивный параметр $\hat{\alpha}(X_1, \dots, X_n)$ определим в виде:

$$\hat{\alpha}(X_1, \dots, X_n) = \begin{cases} \alpha_1, & Q(F_n) \leq Q_1, \\ \frac{\alpha_2 - \alpha_1}{Q_2 - Q_1} \{Q(F_n) - Q_1\} + \alpha_1, & Q_1 < Q(F_n) < Q_2, \\ \alpha_2, & Q(F_n) \geq Q_2, \end{cases} \quad (25)$$

где параметры α_1 , α_2 , Q_1 и Q_2 задаются в соответствии с рассматриваемым типом супермодели, а выборочная оценка $Q(F_n)$ функционала $Q(F; \nu; \mu)$ определена в (24). В данной работе параметр $\hat{\alpha}(X_1, \dots, X_n)$ вычисляется по формуле (25) при следующих значениях параметров: $\alpha_1 = 0,05$, $\alpha_2 = 0,50$, $Q_1 = 1,75$, $Q_2 = 2,50$.

Для такого адаптивного выбора параметра $\hat{\alpha}(X_1, \dots, X_n)$ сравним адаптивные оценки $\hat{S}_1(\hat{\alpha})$ и $\hat{S}_2(\hat{\alpha})$ с другими оценками, для которых параметр α фиксирован. Сравнение проведено в условиях супермодели $\mathfrak{T}_{\varepsilon, \tau}(\Phi)$. Результаты приведены в табл. 3 в виде отношения стандартизованной дисперсии оценки при заданном распределении к минимальной стандартизованной дисперсии среди сравниваемых оценок.

Таблица 3

Отношения стандартизованных дисперсий оценок к минимальной среди сравниваемых оценок

ε	τ	$\hat{S}_1(0)$	$\hat{S}_1(0,1)$	$\hat{S}_1(\hat{\alpha})$	$\hat{S}_2(0)$	$\hat{S}_2(0,1)$	$\hat{S}_2(\hat{\alpha})$
0,05	3	2,25	1,02	1,00	1,20	1,15	1,12
	5	6,48	1,00	1,01	1,91	1,10	1,10
	10	13,5	1,00	1,09	4,49	1,08	1,17
0,10	3	2,37	1,00	1,00	1,25	1,08	1,06
	5	4,49	1,05	1,00	2,00	1,05	1,09
	10	6,07	1,13	1,00	3,74	1,02	1,10
0,20	3	1,67	1,05	1,00	1,11	1,00	1,01
	5	2,22	2,07	1,00	1,58	1,37	1,03
	10	2,30	7,26	1,00	2,26	3,15	1,00
0,30	3	1,26	1,25	1,03	1,00	1,08	1,06
	5	1,23	2,15	1,05	1,11	1,45	1,00
	10	1,26	3,55	1,00	1,48	2,70	1,03

Данные табл. 3 показывают, что в рамках супермодели $\mathfrak{T}_{\varepsilon, \tau}(\Phi)$ среди сравниваемых $\hat{S}_1(\alpha)$ - и $\hat{S}_2(\alpha)$ -оценок, предпочтение следует отдать адаптивному стандартному отклонению $\hat{S}_1(\hat{\alpha})$, для которого параметр $\hat{\alpha}(X_1, \dots, X_n)$ вычисляется по формуле (25) при следующих значениях параметров: $\alpha_1 = 0,05$, $\alpha_2 = 0,50$, $Q_1 = 1,75$, $Q_2 = 2,50$. Отметим, что при сравнении оценок использовались зна-

чения асимптотических стандартизованных дисперсий оценок, которые, как показали результаты моделирования для различных супермоделей, являются вполне удовлетворительной аппроксимацией стандартизованных дисперсий оценок при конечных объемах выборки $n \geq 40$. Исключением являются лишь некоторые распределения с очень «тяжелыми хвостами». Таким образом, адаптивная оценка $\hat{S}_1(\hat{\alpha})$ стандартного отклонения может быть рекомендована к использованию на практике для оценивания масштабного параметра при умеренных объемах выборки и при возможных отклонениях от гауссовской модели наблюдений, вызванных наличием выбросов в выборке.

5. Результаты сравнения оценок при конечных объемах выборки

Приведем результаты сравнения оценок масштабного параметра при конечных объемах выборки, полученные методом статистических испытаний. В эксперимент были включены оценка $\hat{S}_1(0)$ стандартного отклонения, оценка $\hat{S}_2(0)$ среднего абсолютных отклонений, средняя разность Джини $\hat{\Delta}_n = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n |X_i - X_j|$ и медиана абсолютных разностей $\hat{S}_4 = \text{med}\{|X_i - X_j|, 1 \leq i < j \leq n\}$ [5, 8].

В эксперименте вычислялись выборочные значения средних и дисперсий оценок и значения их стандартизованных дисперсий при различных объемах выборки n и различных статистических моделях исходного распределения. В качестве модели было взято гауссовское распределение с масштабным засорением, т.е. $F \in \mathfrak{Z}_{\varepsilon, \tau}(\Phi)$. Получение псевдовыборки $X_1, \dots, X_n$ для этой модели осуществлялось следующим образом. Датчиком случайных чисел $N(0, 1)$ генерировалась выборка объема $(n-k)$, оставшиеся k элементов выборки генерировались датчиком $N(0, \tau^2)$, при этом число k задавалось в соответствии с пропорцией засорения ε . В проведенных экспериментах объем выборки n менялся от 5 до 40, число экспериментов M изменялось от 1 000 до 10 000. В табл. 4 приведены значения стандартизованных дисперсий оценок при $n=20$ и $M=1000$ для супермодели $\mathfrak{Z}_{\varepsilon, \tau}(\Phi)$ при различных значениях τ и k . Отметим, что первая строка таблицы ($n \rightarrow \infty$) содержит значения асимптотических стандартизованных дисперсий оценок. В строке ($n = \infty / n = 20$) приведены отношения асимптотических значений к значениям, полученным экспериментально при $n = 20$.

Таблица 4

Стандартизованные дисперсии оценок при $n = 20$ и $M = 1\,000$ для супермодели $\mathfrak{Z}_{\varepsilon, \tau}(\Phi)$

Ф.р. $\Phi_{\varepsilon, \tau}$	k	$\hat{S}_1(0)$	$\hat{S}_2(0)$	$\hat{\Delta}_n$	$\hat{S}_4$
$F = \Phi$	$n \rightarrow \infty$	0,500	0,571	0,511	0,579
	$k = 0$	0,508	0,586	0,520	0,630
	$n = \infty / n = 20$	0,984	0,974	0,982	0,919
$\tau = 3$	$k = 2$	1,168	0,742	0,797	0,719
	$k = 4$	1,169	0,805	0,888	0,814
$\tau = 5$	$k = 2$	2,295	1,113	1,254	0,723
	$k = 4$	1,795	1,171	1,289	0,851
$\tau = 10$	$k = 2$	3,859	2,279	2,202	0,722
	$k = 4$	2,331	1,947	1,839	0,851

Из табл. 4 видно (см. строку $n = \infty / n = 20$), что асимптотика является вполне приемлемой аппроксимацией для объема выборки $n > 20$. Медиана абсолютных разностей $\hat{S}_4$ обладает преимуществом перед другими оценками при масштабном засорении. Ее стандартизованная дисперсия меняется лишь незначительно, в отличие от других оценок, при увеличении параметра τ . Это объясняется ограниченной функцией влияния оценки $\hat{S}_4$ (см.: [5]). Для распределений, «близких по знятнутости хвостов» к нормальному распределению, преимуществом обладает среднее абсолютных отклонений $\hat{S}_2(0)$.

Заключение

В работе изучены различные оценки масштабного параметра, характеризующего «разброс» случайной величины. Показано, что традиционно используемые на практике оценки масштабного параметра, такие как оценка $\hat{S}_1(0)$ стандартного отклонения $S_1(F)$ и оценка $\hat{S}_2(0)$ среднего абсолютных отклонений $S_2(F)$, имеют неограниченные функции влияния и очень чувствительны к наличию выбросов в выборке. В работе предложены урезанные варианты этих оценок $\hat{S}_1(\alpha)$ и $\hat{S}_2(\alpha)$, $0 \leq \alpha < 1/2$, которые вычисляются не по исходной выборке $X_1, \dots, X_n$, а на основе упорядоченной статистики $X_{(1)}, \dots, X_{(n)}$, из которой предварительно удаляются $[\alpha n]$ наименьших и наибольших порядковых статистик. Эти оценки «защищены» от наличия выбросов в выборке, они имеют ограниченные функции влияния, и их характеристики существенно зависят от параметра α , что на практике приводит к дополнительным усилиям по выбору этого параметра. Для устранения этой трудности в работе предложены адаптивные варианты оценок, для которых параметр $\hat{\alpha}(X_1, \dots, X_n)$ определяется на основе исходной выборке $X_1, \dots, X_n$. Получены асимптотические формулы для вычисления характеристик качества предложенных оценок масштабного параметра. Приведены результаты сравнения оценок в условиях гауссовской модели с масштабным засорением.

Полученные результаты приводят к следующему выводу. В тех случаях, когда нет уверенности, что исходное распределение гауссовское, либо выборка может содержать грубые ошибки (выбросы), для оценки масштабного параметра целесообразнее использовать адаптивное стандартное отклонение $\hat{S}_1(\hat{\alpha})$ либо оценку $\hat{S}_4$ медианы абсолютных разностей. Эти оценки имеют ограниченные функции влияния и, следовательно, «защищены» от наличия выбросов в выборке, и они предпочтительнее по эффективности перед другими рассмотренными оценками в условиях различных моделей наблюдений.

ЛИТЕРАТУРА

1. Bickel P.J., Lehmann E.L. Measures of location and scale // Proc. Prague Symp. Asymptotic Statist. 1973. Prague Charles Univ. 1974. V. 1. P. 25–36.
2. Bickel P.J., Lehmann E.L. Descriptive Statistics for nonparametric models. I. Introduction. II. Location // Ann. Statist. 1975. V. 3. P. 1038–1044; 1045–1069.
3. Bickel P.J., Lehmann E.L. Descriptive statistics for nonparametric models. III. Dispersion // Ann. Statist. 1976. V. 4. № 6. P. 1139–1158.
4. Bickel P.J., Lehmann E.L. Descriptive statistics for nonparametric models. IV. Spread // Contributions to Statistics. Hajek Memorial Volume / ed. by J. Jureckova. Prague : Academia, 1979. P. 33–40.
5. Шуленин В.П. Робастные альтернативы стандартному отклонению при обработке данных физических экспериментов // Известия вузов. Физика. 2016. Т. 59, № 6. С. 62–69.
6. Хампель Ф., Рончетти Э., Рауссей П., Штаэль В. Робастность в статистике. Подход на основе функций влияния. М. : Мир, 1989. 512 с.
7. Хьюбер П. Робастность в статистике. М. : Мир. 1984. 304 с.
8. Шуленин В.П. Робастные методы математической статистики. Томск : Изд-во НТЛ, 2016. 260 с.
9. Serfling R.J. Approximation Theorems of Mathematical Statistics. New York : Wiley, 1980. 371 p.
10. Serfling R.J. Generalized L-M-R-statistics // Ann. Statist. 1984. V. 12. P. 76–86.
11. Tarasenko F.P., Shulenin V.P. Connection of MD-estimates with classes of robust estimates of location parameter // 12th Prague Conf. Inf. Theory, Stat. Decision Functions, Random Processes. Prague, 1994. P. 220–223.
12. Jureckova J. M-L-R-estimators // Handbook of Statistics / eds. P.R. Krishnaiah, P. Sen. Elsevier Science Publishers, 1984. V. 4. P. 463–485.
13. Hampel F.R. The influence curve and its role in robust estimation // J. Amer. Statist. Assoc. 1974. V. 69, № 346. P. 383–393.
14. Hampel F.R. Contribution to the theory of robust estimation : Ph. D. diss. Berkeley : Univ. California, 1968. 103 p.
15. Hampel F.R. A general qualitative definition of robustness // Ann. Math. Statist. 1971. V. 42. P. 1887–1896.
16. Janssen P., Serfling R., Veraverbeke M. Asymptotic normality for a general class of statistical functions and applications to measures of spread // Ann. Statist. 1984. V. 12, № 4. P. 1369–1379.
17. Janssen P., Serfling R., Veraverbeke M. Asymptotic normality of U-statistics based on trimmed samples // J. Statist. Planning and Inference. 1987. V. 16. P. 63–74.

18. Шуленин В.П. Асимптотические свойства GL и U- статистик // Вестник Томского государственного университета. 2004. Приложение № 9 (11). С. 184–190.
19. Shulenin V.P. Asymptotic Properties of the Trimmed GL-and U-Statistics // PRAGUE STOCHASTICS'98 : Abstracts 6th Prague Symposium on Asymptotic Statistics, Prague, 1998, August 23–28. P. 84.
20. Hogg R.V. Adaptive robust procedures: a partial review and some suggestions for future applications an theory // J. Amer. Statist. Assoc. 1974. V. 69. P. 909–923.
21. Шуленин В.П. Свойства адаптивных оценок Ходжеса–Лемана в асимптотике и при конечных объемах выборки // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2010. № 2 (11). С. 96–112.

Поступила в редакцию 5 июня 2020 г.

Shulenin V.P. (2021) ASYMPTOTIC PROPERTIES AND ROBUSTNESS OF TRIMMED VERSIONS ESTIMATES OF STANDARD DEVIATION AND MEAN ABSOLUTE DEVIATIONS. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika*. [Tomsk State University Journal of Control and Computer Science]. 55. pp. 91–102

DOI: 10.17223/19988605/55/11

In the work, various estimates of the scale parameter characterizing the “spread” of a random variable are studied. It is shown that traditionally used in practice estimates of the scale parameter, such as a sample estimate of the standard deviation, and an estimate of the average absolute deviations, have unlimited influence functions, and they are very sensitive to the presence of outliers in the sample. The paper proposes truncated versions of these estimates and, which are calculated not from the initial sample, but based on ordered statistics, from which the smallest and largest ordinal statistics are previously removed. These estimates are “protected” from the presence of outliers in the sample, they have limited influence functions and their characteristics depend significantly on the parameter α , which in practice leads to additional efforts to select this parameter. In the work, adaptive versions of these estimates are proposed, for which the parameter is determined based on the initial sample. The results of comparing estimates under the conditions of various observation models, in particular, under the conditions of the Gaussian model with large-scale contaminating, are presented. The results obtained lead to the following conclusion. In cases where there is no certainty that the initial distribution is Gaussian, or the sample may contain gross errors (outliers), it is more advisable to use the adaptive standard deviation or the estimate in the form of an estimate of the median of the absolute differences. These estimates have limited influence functions and, therefore, are “protected” from the presence of outliers in the sample, and they are preferable in terms of efficiency over the other estimates considered under the conditions of various observation models.

Keywords: scale parameter; robust estimates; outliers; influence function; adaptive estimates.

SHULENIN Valery Petrovich (Candidate of Technical Sciences, National Research Tomsk State University, Tomsk, Russian Federation).

E-mail: shulenin-vp@rambler.ru

REFERENCES

1. Bickel, P.J. & Lehmann, E.L. (1974) Measures of location and scale. In: *Proc. Prague Symp. Asymptotic Statist.* 1973. Vol. 1. Prague: Charles Univ. pp. 25–36.
2. Bickel, P.J. & Lehmann, E.L. (1975) Descriptive Statistics for nonparametric models. I. Introduction. II. Location. *Ann. Statist.* 3, pp. 1038–1044; pp. 1045–1069.
3. Bickel, P.J. & Lehmann, E.L. (1976) Descriptive statistics for nonparametric models. III. Dispersion. *Ann. Statist.* 4(6). pp. 1139–1158.
4. Bickel, P.J. & Lehmann, E.L. (1979) Descriptive statistics for nonparametric models. IV. Spread. In: Jureckova, J. (ed.) *Contributions to Statistics. Hajek Memorial Volume*. Prague: Academia. pp. 33–40.
5. Shulenin, V.P. (2016) Robust alternatives to standard deviation in physical data processing experiments. *Izvestiya vuzov. Fizika – Russian Physics Journal*. 59(6). pp. 62–69.
6. Hampel, F., Ronchetti, E., Raussey, P. & Stael, V. (1989) *Robastnost' v statistike. Podkhod na osnove funktsiy vliyaniya* [Robustness in statistics. Function-based approach influence]. Translated from English. Moscow: Mir.
7. Huber, P. (1984) *Robastnost' v statistke* [Robustness in Statistics]. Translated from English. Moscow: Mir.
8. Shulenin, V.P. (2016) *Robastnye metody matematicheskoy statistiki* [Robust Methods of Mathematical Statistics]. Tomsk: NTL.
9. Serfling, R.J. (1980) *Approximation Theorems of Mathematical Statistics*. New York: Wiley.
10. Serfling, R.J. (1984) Generalized L-M-R-statistics. *Ann. Statist.* 12. pp. 76–86.
11. Tarasenko, F.P. & Shulenin, V.P. (1994) Connection of MD-estimates with classes of robust estimates of location parameter. *12th Prague Conf. Inf. Theory, Stat. Decision Functions, Random Processes*. Prague. pp. 220–223.
12. Jureckova, J. (1984) M-L-R-estimators. In: Krishnaiah, P.R. & Sen, P. (eds) *Handbook of Statistics*. Vol. 4. Elsevier Science Publishers. pp. 463–485.

13. Hampel, F.R. (1974) The influence curve and its role in robust estimation. *Journal of American Statist. Association.* 69(346). pp. 383–393.
14. Hampel, F.R. (1968) *Contribution to the Theory of Robust Estimation*. Ph. D. Diss. Berkeley, Univ. California.
15. Hampel, F.R. (1971) A general qualitative definition of robustness. *Annals of Mathematical Statistics.* 42. pp. 1887–1896. DOI: 10.1214/aoms/1177693054
16. Janssen, P., Serfling, R. & Veraverbeke, M. (1984) Asymptotic normality for a general class of statistical functions and applications to measures of spread. *Annals of Statistics.* 12(4). pp. 1369–1379. DOI: 10.1214/aos/1176346797
17. Janssen, P., Serfling, R. & Veraverbeke, M. (1987) Asymptotic normality of U-statistics based on trimmed samples. *Journal of Statist. Planning and Inference.* 16. pp. 63–74. DOI: 10.1016/0378-3758(87)90056-5
18. Shulenin, V.P. (2004) Asymptotic properties of GL and U-statistics. *Vestnik Tomskogo gosudarstvennogo universiteta – Tomsk State University Journal.* 9(11). Application. pp. 184–190.
19. Shulenin, V.P. (1998) Asymptotic Properties of the Trimmed GL-and U-Statistics. *6th Prague Symposium on Asymptotic Statistics.* Prague. August 23–28. “Prague Stochastics'98”. Abstracts. p. 84.
20. Hogg, R.V. (1974) Adaptive robust procedures: A partial review and some suggestions for future applications an theory. *Journal of American Statist. Association.* 69. pp. 909–923.
21. Shulenin, V.P. (2010) Asymptotic and nonasymptotic properties of Hodges - Lehmann adaptive estimators. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science.* 2(11). pp. 96–112.

ИНФОРМАТИКА И ПРОГРАММИРОВАНИЕ

УДК 621.396.62
DOI: 10.17223/19988605/55/12

Д.В. Шехалев

СТРУКТУРА МАТРИЧНОГО УМНОЖИТЕЛЯ ДЛЯ ДЕКОДЕРОВ QC-LDPC КОДА С МИНИМАЛЬНЫМ РЕСУРСОМ ПАМЯТИ ПЛИС

Работа выполнена в рамках программы повышения конкурентоспособности ТГУ, НИР 8.2.37.2020.

Рассмотрена архитектура построения декодера QC-LDPC кодов на основе алгоритма распространения доверия для кода с изменениями скорости кодирования и размера блока. Показаны архитектуры матричного умножителя для высокопроизводительного декодирования с переменным размером блока. Предложено решение задачи оптимизации архитектуры матричного умножителя по критерию занимаемого ресурса ПЛИС.

Ключевые слова: помехоустойчивое кодирование; ПЛИС; коды с минимальной плотностью проверок на четность; матричное умножение.

Развитие современных систем связи идет по пути роста скорости передачи данных и увеличения помехоустойчивости канала данных. Это приводит к тому, что все более широко внедряются современные методы помехоустойчивого кодирования, основанные на итерационном декодировании. Одними из таких кодов являются коды с минимальной проверкой на четность (LDPC). Коды задаются проверочной матрицей размера $m \times n$, где m – количество бит в блоке кодирования, n – количество проверочных бит [1. С. 851]. В качестве примера можно привести проверочную матрицу LDPC кода (6, 12) вида:

$$H = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}. \quad (1)$$

Наличие единицы в строке проверочной матрицы определяет участие бита данных в проверочном уравнении четности. Таким образом, проверочная матрица описывает связь символов блока кодирования друг с другом в n уравнениях четности. Скорость кодирования определяется по формуле

$$code\ rate = \frac{m - n}{m}.$$

Изменение размера блока или скорости кодирования кода осуществляется через изменение размеров и значений ячеек проверочной матрицы. Одним из основных алгоритмов декодирования таких кодов является алгоритм распространения доверия [1. С. 875]. Данный алгоритм обладает повышенными требованиями к вычислительной мощности модемного оборудования, так как ему необходимо большое количество итераций декодирования для получения хорошего энергетического выигрыша от кодирования.

В общем случае проверочная матрица может быть произвольной и не обладать структурой, что усложняет распараллеливание вычислений при декодировании. Но существует несколько семейств LDPC кодов, в которых эта проблема решена. Например, квазициклические коды с минимальной проверкой на четность (QC-LDPC) [2, 3].

Проверочная матрица таких кодов

$$H = \begin{bmatrix} H(1,1) & \dots & H(1,M) \\ \dots & \dots & \dots \\ H(N,1) & \dots & H(N,M) \end{bmatrix}$$

состоит из $M \times N$ базовых матриц $H(i, j)$ ($i = \overline{1, N}, j = \overline{1, M}$) размером $z \times z$ элементов. Базовые матрицы могут быть двух видов: единичная матрица с циклическим сдвигом или нулевая матрица.

Таким образом, проверочная матрица задается порождающей матрицей сдвигов, отрицательное значение элемента которой соответствует нулевой базовой матрице и означает, что биты данного блока не участвуют в вычислениях. Остальные значения задают циклический сдвиг блока бит соответствующего координатам элемента матрицы сдвигов. Скорость кодирования определяется аналогично LDPC кодам:

$$code\ rate = \frac{M - N}{M}, \quad (2)$$

а размер базовой матрицы определяет размер блока декодирования:

$$m = M \times z. \quad (3)$$

Единичная матрица – это квадратная матрица, элементы главной диагонали которой равны единице, а остальные равны нулю. Следовательно, при любом циклическом сдвиге базовой матрицы количество единиц в ряде и столбце матрицы будет равно единице. Это приведет к тому, что после подстановки базовых матриц со сдвигами, определенными порождающей матрицей, ряды и столбцы проверочной матрицы будут независимы друг от друга в пределах размера z базовых матриц; таким образом, эти элементы могут быть вычислены одновременно. А значит, такой код обладает хорошими возможностями по распараллеливанию вычислений.

При этом сохраняется строгая структура кода, не зависящая от размеров базовой матрицы z , определяющей размер блока кодирования по формуле (3), и от количества рядов порождающей матрицы сдвигов, определяющей скорость кодирования по формуле (2).

Эти особенности делают QC-LDPC коды удобными для высокоскоростных систем передачи данных и для систем с адаптивной системой кодирования и модуляции (АСМ). Поэтому коды именно этого семейства приняты в современных стандартах связи в качестве помехоустойчивых кодов для каналов данных [2, 3].

Реализация высокопроизводительных QC-LDPC кодов на ПЛИС требует использования большого количества вычислительных ресурсов ПЛИС, особенно одного из критических ресурсов – встроеной блочной памяти, используемой для хранения промежуточных результатов вычисления.

Количество промежуточных результатов зависит от используемого алгоритма декодирования, а производительность декодера – от скорости реализации матричного перемножения и степени распараллеливания алгоритма. Одним из популярных способов реализации высокопроизводительного декодирования является последовательная реализация алгоритма распространения доверия с минимизацией количества хранимых промежуточных результатов [4]. Несмотря на последовательные вычисления, высокая производительность такого декодера обеспечивается матричным умножителем, основанным на мультиплексорах и регистрах ПЛИС при работе с широким словом данных, представляющим собой все элементы, соответствующие базовой матрице. При использовании программного слова меньшего размера производительность такого декодера резко снижается.

В статье [4] рассмотрен стандарт 802.11n [5], в котором определены четыре порождающих матрицы и три размера базовых матриц 27/54/81 бит. Тогда как, например, в стандарте 5G NR [2] количество возможных размеров базовых матриц 51, от 2 до 384 бит, количество порождающих матриц 16.

При реализации декодера с возможностью изменения размера блока подобный матричный умножитель будет занимать большое количество логических ресурсов ПЛИС, так как должен быть рассчитан на использование базовой матрицы наибольшего размера. Таким образом, подобный декодер можно использовать только в режиме высокой производительности, при значительном ресурсе логики ПЛИС и малом ресурсе блочной памяти.

В статье рассмотрена альтернативная архитектура ядра декодера, основанная на параллельной реализации алгоритма распространения доверия, для реализации QC-LDPC декодеров с возможностью адаптации производительности с учетом отводимого под декодер ресурса ПЛИС. Рассмотрена версия подобного ядра для применения в системах с высокой пропускной способностью. Исследованы пути оптимизации использования ресурса ПЛИС для декодеров подобного вида. Предложено решение для экономии ресурсов ПЛИС для декодера QC-LDPC кодов с перестраиваемыми размерами блока и скорости кодирования.

1. Декодирование LDPC кодов по алгоритму распространения доверия

Алгоритм распространения доверия [1. С. 875; 6] основан на использовании представления проверочной матрицы в виде графа Таннера (рис. 1).

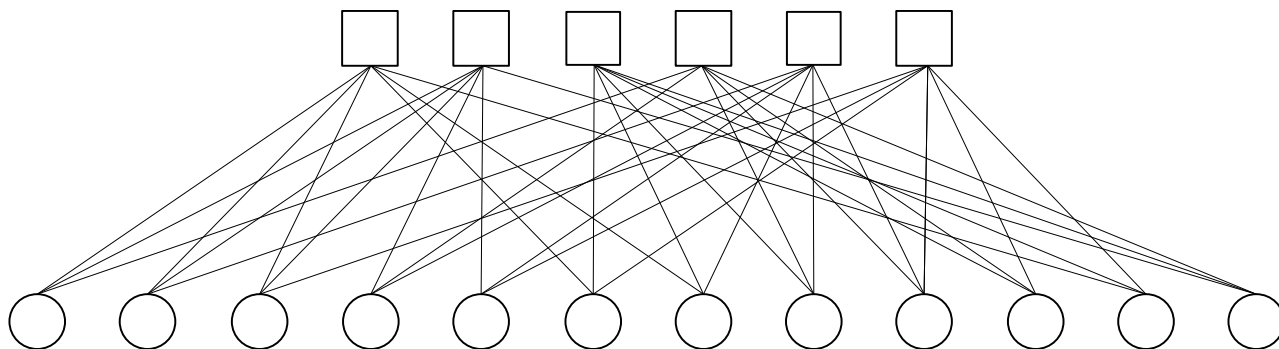


Рис. 1. Граф Таннера LDPC кода (6, 12) с проверочной матрицей по формуле (1)

Fig. 1. Tanner graph of LDPC code (6, 12) based upon check matrix equation (1)

В алгоритме вводится понятие узлов (node), кодовых узлов (cnode) и вертикальных узлов (vnode). Узел (node) представляет собой элемент памяти, в котором хранятся промежуточные результаты вычисления надежностей символов. Размер поля узлов в общем случае равен $m \times n$ и соответствует размеру проверочной матрицы. В операциях участвуют только те узлы, координатам которых в проверочной матрице соответствуют единицы.

Горизонтальные уравнения четности определяют собой кодовый узел (cnode) – результат вычисления совместной вероятности символов, входящих в уравнение четности. На основании совместной вероятности производится вычисление внешней вероятности каждого символа (extrinsic).

Вертикальное сложение узлов определяет вертикальный узел (vnode) – сумма исходной вероятности символа со всеми внешними вероятностями этого же символа, полученными на предыдущем шаге. На основании суммы вероятностей одного символа производится вычисление внешней вероятности для следующего шага.

Таким образом, каждая итерация алгоритма декодирования состоит из двух шагов:

1. Горизонтальный шаг: вычисление кодового узла и внешних вероятностей символа горизонтальных связей, обновление поля узлов.
2. Вертикальный шаг: вычисление вертикального узла, определение внешней вероятности вертикальных связей, обновление поля узлов.

Спустя заданное количество итераций декодирования работа алгоритма заканчивается на вертикальном шаге, на котором, после вычисления вертикального узла, принимается жесткое решение по символу.

В общем случае LDPC кодов, при произвольной проверочной матрице, вычисления производятся последовательно символ за символом. Это требует наличия быстрого, произвольного и атомарного доступа к памяти узлов. Поэтому при разработке такого декодера нельзя использовать какие-либо упорядоченные структуры при реализации генераторов адресов памяти узлов. Это увеличивает сложность реализации декодера на ПЛИС.

В отличие от LDPC кодов, QC-LDPC коды обладают четкой структурой, позволяющей значительно увеличить скорость декодирования, так как в этом случае узлы, соответствующие разным базовым матрицам, могут обрабатываться одновременно.

2. Типовая архитектура вычислительного ядра декодера QC-LDPC кодов на ПЛИС

В QC-LDPC кодах обработка строк и столбцов может выполняться одновременно, это позволяет значительно увеличить скорость декодирования, но требует организации памяти узлов с возможностью одновременного доступа к данным. При этом нужно помнить, что в памяти узлов размером $M \times N \times z$ данные должны храниться постоянно и передаваться между итерациями. При этом узлы, относящиеся к одной базовой матрице, при обработке на горизонтальном шаге необходимо умножить на матрицу сдвига при чтении и обратной записи.

Типовая архитектура вычислительного ядра QC-LDPC декодера с использованием естественного параллелизма для реализации на ПЛИС приведена на рис 2.

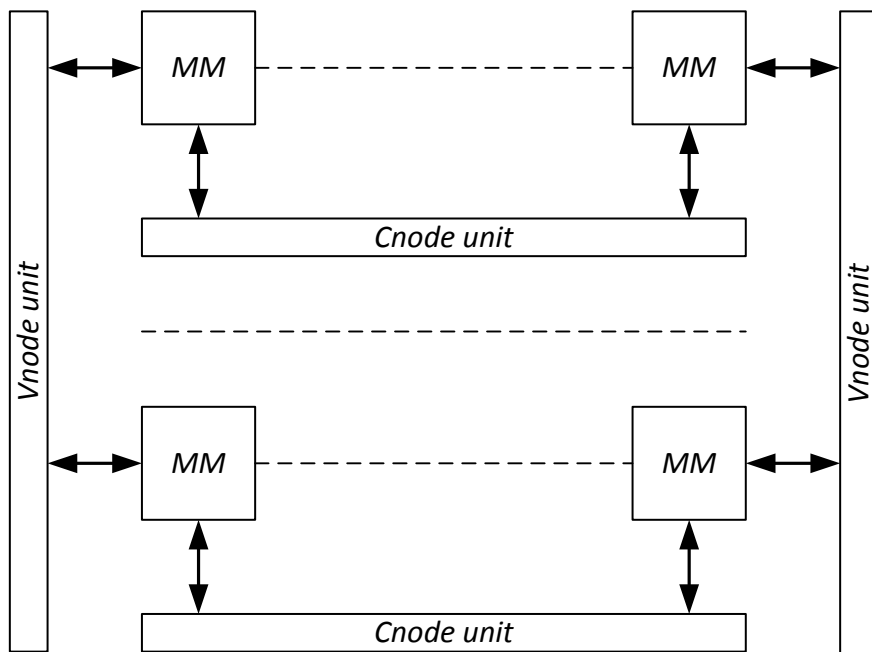


Рис. 2. Архитектура вычислительного ядра QC-LDPC декодера
Fig. 2. QC-LDPC decoder core architecture

На рис. 2 обозначены: *MM* – матричный множитель, *Vnode unit* – модуль операций вертикального шага, *Cnode unit* – модуль операций горизонтального шага. Из рисунка видно, что общее ускорение операций декодирования будет определяться количеством узлов, одновременно обрабатываемых на горизонтальном и вертикальном шагах. Итоговый результат ускорения зависит от количества доступных ядер обработки данных и может варьировать в широких пределах, исходя из доступного вычислительного ресурса ПЛИС.

Но даже такой организации вычислений может быть недостаточно для реализации высоких скоростей декодирования. А в случае изменения скорости кодирования оно может быть неэффективным, так как зарезервированные ресурсы ПЛИС под быстрое декодирование кодов с низкой скоростью кодирования могут не использоваться большую часть времени.

Подобная ситуация возникает при высоком отношении сигнал–шум в радиоканале. В таких условиях используют модуляции с высоким индексом и большой скоростью кодирования ($\geq 0,75$), так как это позволяет максимизировать пропускную способность системы связи для текущего состояния радиоканала, и в этом случае требуется высокая производительность декодера. При малом отношении сигнал–шум используются модуляции с низким индексом и коды с низкой скоростью кодирования, в таких условиях требования к пропускной способности декодера снижаются.

Таким образом, разумно введение еще одной степени параллелизма: на уровне узлов базовых матриц, т.е. обрабатывать больше одного узла, относящегося к одной базовой матрице, за один такт системной частоты. В этом случае избыточный ресурс, зарезервированный для декодирования кодов с низкой скоростью кодирования, будет задействован для вычислений при декодировании кодов с любой скоростью.

При обработке больше одного узла базовой матрицы за один такт возникает сложность организации умножения базовых матриц на матрицу сдвигов, так как элементы матрицы сдвигов задаются с кратностью один бит, а гранулярность слова данных становится больше. Типовое решение данной проблемы заключается в разделении памяти узлов, относящихся к базовой матрице, на несколько блоков одинакового размера, использовании индивидуального генератора адреса узла и бареловского сдвигателя. Традиционная архитектура матричного умножения представлена на рис. 3.

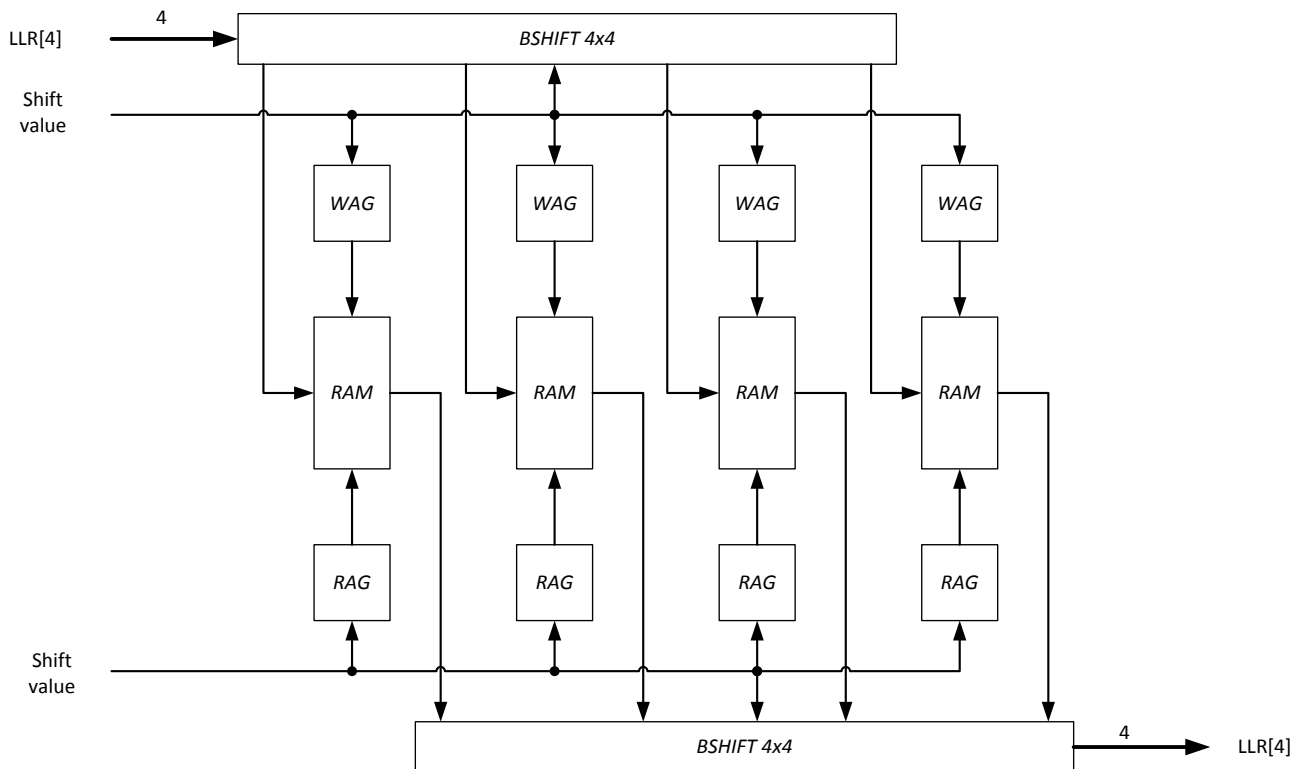


Рис. 3. Традиционный умножитель на матрицу сдвига для обработки 4 узлов за 1 такт частоты

Fig. 3. Classic shift matrix multiplier for 4 nodes per cycle processing

На рис. 3 обозначены: *BSHIFT* – бареловский сдвигатель, *RAG* / *WAG* – генератор адреса чтения / записи, *RAM* – массив памяти узлов. Данный умножитель позволяет выполнить умножение блока данных на матрицу сдвига с размером, кратным количеству данных. При этом умножение на матрицу сдвига происходит на лету, при чтении / записи памяти. Недостатками подобной схемы являются использование большого количества блоков памяти ПЛИС и неоптимальный режим их использования, так как типовой максимальный размер базовых блоков кодов составляет до 512 бит, а разрядность метрики узла 4–8 бит. Таким образом, требуемая емкость блока памяти составляет 2–4 килобита, а емкость блока памяти современных ПЛИС составляет 16 килобит. В результате большая часть па-

мнью не используется. Помимо этого, каждый блок памяти должен обладать собственным генератором адреса, что увеличивает сложность формирования адресов при изменении матрицы сдвига. В совокупности все это увеличивает требуемый ресурс ПЛИС для реализации умножителя. Достоинством схемы является минимальное количество тактов чтения / записи, необходимых для выполнения операции. Без учета латентности это время составит

$$T_{write}/T_{read} = \frac{z}{node_num},$$

где z – размер базовой матрицы, $node_num$ – количество узлов, обрабатываемых за 1 такт частоты работы ядра декодера

3. Предлагаемое решение

Для оптимизации ресурсов памяти ПЛИС предлагается архитектура матричного умножителя, представленная на рис. 4.

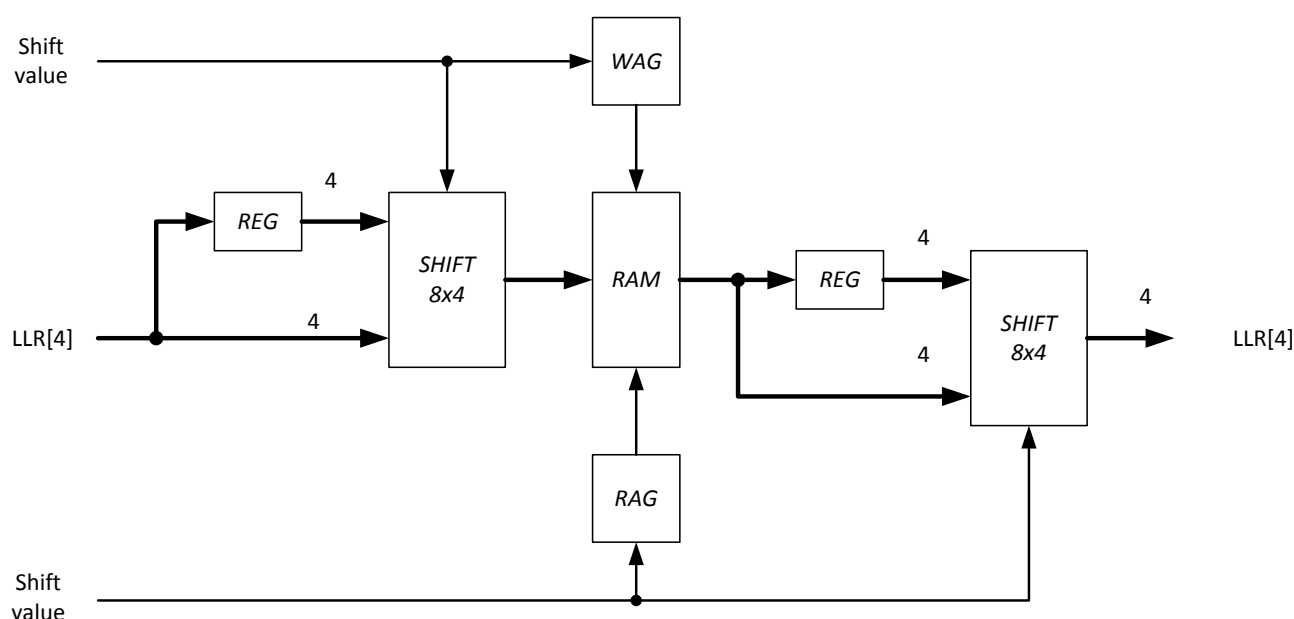


Рис. 4. Оптимизированный умножитель на матрицу сдвига
Fig. 4. Optimized shift matrix multiplier for 4 nodes per cycle processing

На рис. 4 обозначены: *REG* – регистр, *SHIFT* – регистр сдвига, *RAG* / *WAG* – генератор адреса чтения / записи, *RAM* – массив памяти узлов. Используется один блок памяти, хранящий несколько узлов в одной ячейке, по одному адресу. Для адресации используются один генератор адреса, дополнительный регистр и регистр сдвига. Предлагаемое решение занимает меньше ресурсов ПЛИС, чем традиционный умножитель. Но для формирования нужной последовательности узлов для обработки данных необходимо увеличить цикл чтения для предвыборки и цикл записи для записи последнего слова на один такт частоты относительно типовой схемы.

Таким образом, на каждом матричном умножении, при операциях чтения–записи, предлагаемая схема имеет задержку на два такта больше, чем типовой умножитель, но, благодаря оптимизации логических и трассировочных ресурсов ПЛИС, схема позволяет получить более высокие рабочие тактовые частоты и увеличить количество матричных умножителей для увеличения производительности декодера.

Помимо этого, так как в системе используется всего один генератор адреса, это упрощает формирование адресов при изменении матрицы сдвига и размера базовой матрицы, что делает предлагаемую схему умножителя удобной для декодеров с изменением параметров кодирования на лету.

4. Результаты сравнения традиционной и предлагаемой архитектур

Проведем сравнение архитектур по используемым ресурсам ПЛИС. В качестве примера рассмотрим базовый блок размером 1 024 бита, размер слова обработки 4 узла, размер метрики правдоподобия узла 4 бита.

Таблица 1

Сравнение ресурсов ПЛИС, требуемых для реализации классического и предлагаемого матричного умножителя, на логическом уровне

Составные модули	Традиционный умножитель	Предлагаемый умножитель
Память	4 блока памяти размером 256×4	1 блок памяти размером 256×16
Генератор адреса	8 генераторов адреса (счетчик по модулю 256 с установкой)	2 генератора адреса (счетчик по модулю 256 с установкой)
Сдвигатель 4x4	2 бареловских сдвигателя	2 сдвигателя
Регистры	0	2
Количество тактов на одну итерацию*	256	258

* – без учета латентности.

Для проверки предлагаемой реализации было разработано два варианта умножителя на матрицу сдвига на языке SystemVerilog. Результаты синтеза для ПЛИС Xilinx семейства Artix-7 в ПО Xilinx Vivado 2019.1.3 с настройками по умолчанию приведены в табл. 2, 3.

Таблица 2

Сравнение ресурсов ПЛИС Xilinx семейства Artix, требуемых для реализации классического и предлагаемого матричного умножителя

Ресурсы ПЛИС	Традиционный умножитель	Предлагаемый умножитель	Экономия ресурсов ПЛИС
RAMB18	4	1	75%
Slice LUT	364	167	54%
Slice REG	273	200	26%

Таблица 3

Сравнение показателей пиковой производительности реализаций классического и предлагаемого матричного умножителя

Показатели производительности	Традиционный умножитель	Предлагаемый умножитель	Выигрыш
Количество тактов на одну итерацию	261	263	–0,8%
Тактовая частота, МГц	262	281	8%
Производительность, млн умножений/с	1,007	1,07	6,3%

Из табл. 1–3 следует, что, несмотря на необходимость увеличения количества тактов, необходимых для выполнения матричного умножения, предложенное решение позволяет достичь более высокого значения тактовой частоты. Это позволяет нивелировать потенциальный проигрыш в производительности одного блока по тактовой частоте. С учетом меньшего количества использованных ресурсов ПЛИС для предлагаемого решения можно получить значительный выигрыш по производительности декодера или значительно уменьшить требуемый для декодера ресурс ПЛИС при сохранении той же производительности.

Заключение

В статье рассмотрена задача оптимизации декодера QC-LDPC кодов по критерию оптимального использования ресурсов ПЛИС. Предложена схема умножителя на матрицу сдвига для быстродействующих декодеров с поддержкой переменного размера блока и скорости кодирования. Показан выигрыш в требуемом ресурсе ПЛИС, достигаемый при использовании умножителя предлагаемой архитектуры.

ЛИТЕРАТУРА

1. Lin S., Costello D.J. *Error Control Coding*. 2nd ed. Upper Saddle River, NJ : Prentice-Hall, 2004. 1272 p.
2. ETSI 3rd Generation Partnership Project; Technical specification 5G; NR; multiplexing and channel coding (3GPP TS 38.212 version 15.8.0 Release 15). 2020. P. 18–25.
3. IEEE Standard for Air Interface for Broadband Wireless Access Systems // IEEE Std 802.16–2012. 2012. P. 1285–1289.
4. Mhaske S. et al. High-Throughput FPGA-Based QC-LDPC Decoder Architecture // 2015 IEEE 82nd Vehicular Technology Conference (VTC2015-Fall). Boston, MA : IEEE, 2015. P. 1–5.
5. IEEE Standard for Information technology. Telecommunications and information exchange between systems. Local and metropolitan area networks. Specific requirements // IEEE Std 802.11–2012. 2012. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. P. 2304–2306.
6. Jinghu C., Fossorier M.P.C. Near optimum universal belief propagation based decoding of low density parity check codes // IEEE Transactions on Communications. 2002. V. 50 (3). P. 406–414.
7. Chavet C., Coussy P. *Advanced Hardware Design for Error Correcting Codes*. Cham : Springer, 2015. IX, 192 p.

Поступила в редакцию 2 ноября 2020 г.

Shekhalev D.V. (2021) MATRIX MULTIPLIER ARCHITECTURE FOR QC-LDPC CODE WITH MINIMUM FPGA BLOCK RAM RESOURCE USING. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika* [Tomsk State University Journal of Control and Computer Science]. 55. pp. 103–111

DOI: 10.17223/19988605/55/12

Forward error correction coding based upon iterative decoding are widely used in modern communication systems. Low Density Parity Check codes (LDPC) is one of it. These codes have increased requirements for the processing performance of modem equipment, since they require a large number of decoding iterations to obtain a good coding gain. The LDPC check matrix can be arbitrary and vary with other encoding parameters in general. So, decoding such codes is performed sequentially symbol by symbol and does not allow high throughput from the decoder.

There are some classes of LDPC codes that have good capabilities for paralleling processing and changes in code characteristics on fly. Quasi-Cyclic Low Density Parity Check codes (QC-LDPC) is one of it. The check matrix of such codes is a shift matrix based upon base matrices array and has clear structure. It makes possible to increase the performance of the decoder due to its natural parallelism at the level of base matrices.

However, requirements of communication system performance can be so high that despite this feature, decoder performance may not be sufficient. An additional level of parallelism based upon elements of the base matrices is required. It's realized through a matrix multiplier unit based on block memory. The classic solution of it involves the use of a large amount of FPGA memory blocks used in a non-optimal mode. Each memory unit use an own write/read address generator. Multiplication by shift matrix is carried out due to barrel shifter. This architecture provides high throughput but requires a lot of FPGA resources. Especially, the block ram units which is one of most critical FPGA resources.

There is an alternative solution for architecture of matrix multiplier. The proposed solution is based upon one larger-width memory unit with single read/write address generator, bit shifter and registers for intermediate data storing. This solution has a similar bandwidth as classic one and requires a significantly smaller number of FPGA resources. Proposed solution shows reduction of FPGA block memory and LUT using up to 4 and 2 times respectively and obtain increased performance due to clock frequency increase up to 5–6% at same time. Moreover, the advantage of it is the ease of changing base matrix size and the shift value. Therefore, the proposed solution is recommended for constructing high throughput decoders with adaptive coding support.

Keywords: forward error correction; FPGA; LDPC; matrix multiply.

SHEKHALEV Denis Vladimirovich (Senior Engineer, Scientific and Educational Center “Engineering Center of Microwave Technic and Technology”, National Research Tomsk State University, Tomsk, Russian Federation).
E-mail: diod2003@list.ru

REFERENCES

1. Lin, S. & Costello, D.J. (2004) *Error Control Coding*. 2nd ed. Upper Saddle River, NJ: Prentice-Hall.
2. ETSI 3rd Generation Partnership Project. (2020) *Technical specification 5G; NR; multiplexing and channel coding (3GPP TS 38.212 version 15.8.0 Release 15)*. pp. 18–25.
3. *IEEE Standard for Air Interface for Broadband Wireless Access Systems (2012)*. IEEE Std 802.16-2012. pp. 1285–1289.
4. Mhaske, S. et al. (2015) High-Throughput FPGA-Based QC-LDPC Decoder Architecture. *2015 IEEE 82nd Vehicular Technology Conference (VTC2015-Fall)*. Boston, MA, USA: IEEE. pp. 1–5.

5. IEEE Standard for Information technology. (2012) Telecommunications and information exchange between systems. Local and metropolitan area networks. Specific requirements. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. *IEEE Std 802.11-2012*. pp. 2304–2306.
6. Jinghu, C. & Fossorier, M.P.C. (2002) Near optimum universal belief propagation based decoding of low density parity check codes. *IEEE Transactions on communications*. 50. pp. 23–31.
7. Chavet, C. & Coussy, P. (2015) *Advanced Hardware Design for Error Correcting Codes*. Heidelberg; New York; Dordrecht; London: Springer.

ПРОЕКТИРОВАНИЕ И ДИАГНОСТИКА ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

УДК 681.518.5

DOI: 10.17223/19988605/55/13

Ю.Ф. Мухопад, А.Ю. Мухопад, Д.Ц. Пунсык-Намжилов

ДИНАМИЧЕСКИЙ КОНТРОЛЬ УПРАВЛЯЮЩИХ АВТОМАТОВ СЛОЖНЫХ СИСТЕМ ДИСКРЕТНОЙ АВТОМАТИКИ

Рассмотрены способы контроля управляющих автоматов (УА) со структурной организацией Мура. По существующей парадигме снижение затрат на комбинационную схему переходов основано на сокращении числа состояний. В автоматах нового типа целенаправленно увеличено число состояний с введением адресного блока из (q) двухходовых элементов «И», одного элемента «ИЛИ» и RS -триггера. Контроль таких УА реализуется с помощью нового равновесного кода pCn из (p) трехразрядных групп с одной единицей в группе. Предложенный метод контроля УА проще методов с использованием других корректирующих кодов.

Ключевые слова: контроль; управляющие автоматы; кодирование состояний.

Самостоятельным направлением внедрения информационных технологий является разработка методологии синтеза информационно-управляющих систем мехатроники, робототехники, летательных аппаратов и оборонного комплекса. Такие информационно-управляющие системы (ИУС) эксплуатируются в экстремальных условиях механических и электромагнитных воздействий высокой интенсивности. Поэтому необходим динамический контроль правильности функционирования ИУС.

Сложные системы дискретной автоматики [1] с динамическим контролем функционирования начали широко применяться с начала 1950-х гг. [2]. К началу 1990-х гг. были созданы основы теории синтеза самоконтролируемых автоматов трудами Д.А. Поспелова, В.А. Горбатова, С.Х. Колдоуэлла, В.Н. Захарова, А.Д. Закревского, В.М. Глушкова, П.П. Пархоменко, Е.В. Слабакова, Е.С. Согомоняна, В.В. и Вл.В. Сапожниковых, Н.С. Щербакова и др. Обзор работ по синтезу УА с контролем приведен в работах [3, 4]. В статье рассматриваются методы контроля управляющих автоматов (УА), основанные на применении новых равновесных кодов в виде нескольких трехразрядных групп с одной единицей в группе. Двоичный эквивалент этого кода получен за счет замены каждой трехразрядной группы двумя двоичными разрядами. Этот новый двоичный позиционный код (ДПК) обеспечивает возможность создания самоконтролируемых управляющих автоматов с простой процедурой самоконтроля. Такая процедура является принципиально новой и более эффективной, что доказано результатами программной и аппаратной реализации управляющих автоматов нового типа.

1. Анализ управляющих автоматов

Математическая модель управляющих автоматов Мура имеет вид:

$$a(t+1) = F_1(a(t), \alpha_1 \dots \alpha_q); \quad A(t+1) = F_2(a(t)),$$

где $a(t)$, $a(t+1)$ – предыдущее и последующее состояния автомата; $A(t+1)$ – исполняемая команда; F_1 – комбинационная схема переходов; F_2 – схема формирования исполнительных команд; q – количество логических условий $\alpha_1 \dots \alpha_q$.

Контроль УА основан на применении кодов Хэмминга, Бергера, kCn и др. [5–11].

В работе [12] предложено разделять управляющие автоматы по степени сложности на сверхпростые (СП), простые (ПА), средней сложности (СА), автоматы сложные (АС), высоко сложные (ВС), особо сложные (ОС) и ультрасложные (УС).

В УА Мура схема переходов F_1 при программной реализации имеет объем (V), определяемый как $V = m_p 2^{m+q}$, где m_p – реальная разрядность БИС ПЗУ равная 4 или 8; m – разрядность кода состояния $a(t)$. При аппаратной реализации F_1 число входов ПЛМ равно $n \geq (m + q)$. Для УА типа ВС объем ПЗУ значителен. Поэтому контроль схемы переходов F_1 не является простым.

Оптимизация по существующей парадигме и методологии синтеза УА дает возможность снизить сложность комбинационных схем переходов только в 1,2–1,5 раза, причем чем сложнее УА, тем меньше значение коэффициента сложности.

Поэтому требуется концептуальное изменение существующей методологии синтеза УА с целью снижения сложности управляющих автоматов.

Контроль УА Мура с применением кодов Хэмминга, Бергера и других достаточно глубоко исследован. Для реализации самоконтролируемых УА с известными контролирующими кодами необходимо использовать большое количество логических элементов [5–9].

Эффективный способ контроля, основанный на применении кода kCn с фиксированным числом (k) разрядов в (n) –разрядах всего кода, представлен в работах [12, 13]. Для сокращения числа разрядов кода состояний используется конкатенация кодов: 2C5,2C5; 2C5,3C7; 3C7,3C7. В коде с двумя единицами в пяти разрядах (2C5) разрешенными являются 10 комбинаций из 32, а в коде (3C7) только 25 из 128. Однако комбинационные схемы принадлежности кода $a(t)$ к числу разрешенных кодов весьма сложны для (3C7) [14].

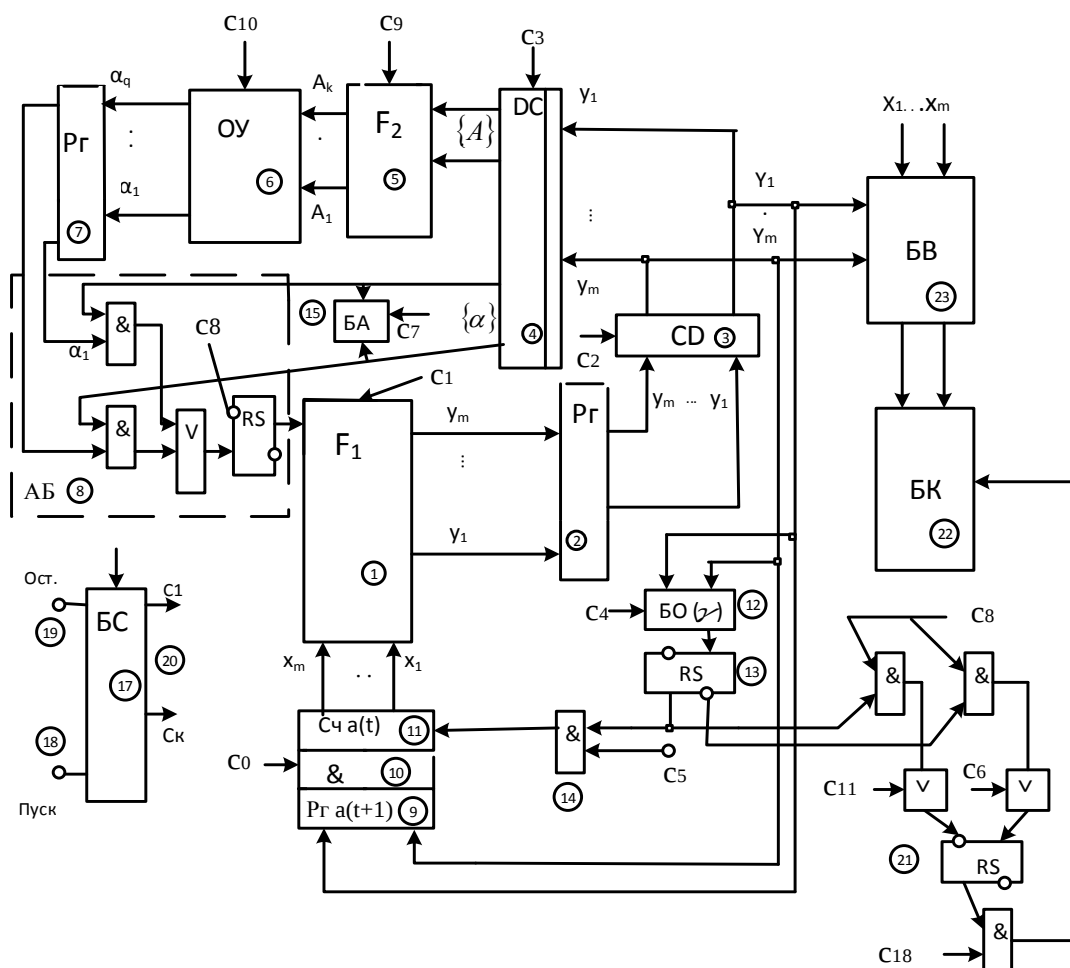


Рис. 1. Управляющий автомат Мухопад с блоками контроля 21-23

Fig. 1. Control automaton Mukhopad with control blocks 21-23

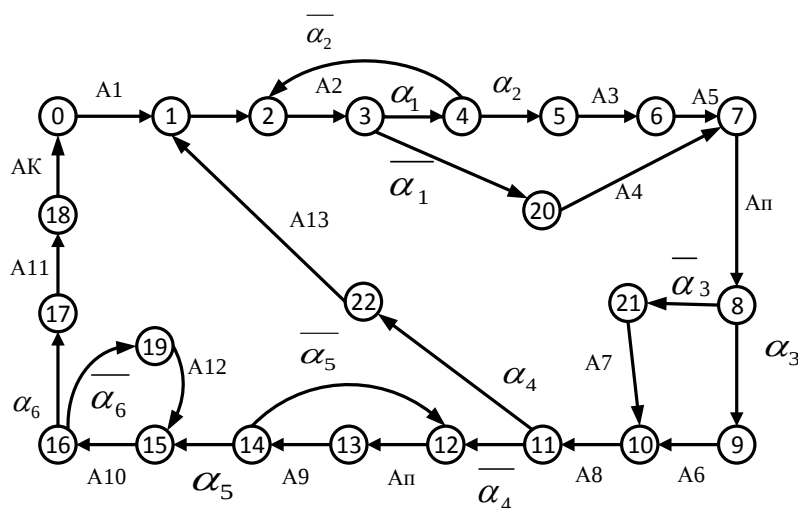


Рис. 3. Граф переходов нового управляющего автомата
Fig. 3. The transition graph of control automaton

Особенность этого варианта УА заключается в том, что только при одной комбинационной схеме $F_1(1)$ (а не при двух, как в УА Мили) этот автомат может работать как в режиме Мили, так и в режиме Мура, но с выбором одного логического условия для реализации схемы $F_1(1)$.

2. Контроль управляющих автоматов с особыми групповыми кодами

Наиболее эффективен контроль УА с равновесными кодами kCn , но схемы встроенного контроля сложны. В работе [17] был предложен особенный тип кода, также с фиксированным числом единиц, но состоящий из (p) трехразрядных групп. В каждой группе только одна единица. Этот код обозначен как pCn . В табл. 2 приведены двоичные позиционные коды (ДПК) и новые коды pCn (столбцы 1, 2) для номеров $0...7$. Двухразрядному коду ДПК «00» соответствует группа 001, коду «01» – 010, а коду «10» – 100.

Таблица 2

Двоичные непозиционные коды

№	ДПК	pCn	ДНК	ДПК
0	00.00	001.001	00.00	0
1	00.01	001.010	00.01	1
2	00.10	001.100	00.10	2
3	00.11	010.100	01.10	6
4	01.00	010.010	01.01	5
5	01.01	010.001	01.00	4
6	01.10	100.010	10.01	9
7	01.11	100.100	10.10	10
	1	2	3	4

Комбинации «11» в ДПК в коде pCn нет соответствия, но ее нельзя исключить из кода ДПК, поэтому при переходе к (pCn) осуществляется перенос «1» в более старшую группу кода pCn . В табл. 2 в столбцах 3, 4 представлен двоичный аналог кода pCn , полученный за счет обратного преобразования каждой трехразрядной группы в двоичный непозиционный код (ДНК). Этот код будет использован далее.

Контроль кода pCn производится для каждой i -й группы, $i = 1, \dots, p$. При этом определяется значение $z_i = a\bar{b}\bar{c} + \bar{a}b\bar{c} + \bar{a}\bar{b}c$. Здесь a, b, c – обозначения разрядов в группе. Если для всех групп значения $z_i = 1$, то фиксируется отсутствие ошибки. Но даже при правильных кодах каждой группы не исключена вероятность изменения значений сразу в двух или трех разрядах. После получения кода $a(t+1)$ необходимо восстановить предыдущий код $a(t)$, который обозначается как $a(t)_0$. Код $a(t)_0$ был

Операторы синхронизации

C_i	Семантика	C_i	Семантика
0	$a(t+1) \rightarrow a(t)$ Сч(11)	18	Запись $a(t)$ на Сч(11)
1	Опрос $F_1(1)$	19	Счит. $DC(4)$ без исполн.
2	Опрос шифратора $CD(3)$	20	Счит. $DC(4)$ с исп. A_i
3	Опрос $DC(4)$	21	Опрос схемы СВ(22)
4	Опрос БО (12) – γ	22	Вкл. УА(32)
5	Синхр. «И» (14)	23	«0» $\rightarrow$ Сч(11)
6	Уст. «0» RS (13)	24	$a(t)_0 \rightarrow Pr IV$
7	Опрос БА (15)	25	Счит. СК(30)
8	Уст «0» RS выбор α_j	26	$a(t+1) = a(i) + 1$
9	Вкл. F_2	27	$a(t+1) = a(t) - 1$
10	Перенос α_j на $Pr(7)$ от ОУ	28	$a(t)_0 \rightarrow < Pr IV > \rightarrow Pr I$
11	Запись $\gamma(t)$	29	$a(t) \rightarrow Pr II$
12	Перепись $\gamma(t)$ на $RS \gamma(t-1)$	30	Счит. Сх. Сп.24
13	Счит. $\gamma(t)$	31	Ошибка
14	Счит. $\gamma(t-1)$	32	$< Pr IV > \rightarrow Сч$
15	Уст. «1» RS синх. $DC(4)$	33	Вкл. Т
16	Уст. «0» RS синх. $DC(4)$	34	Оконч. программы
17	Счит. знач. RS синх. $DC(4)$		

Алгоритм контроля имеет два режима: Режим I соответствует $\gamma(t-1) = 1$, режим II соответствует $\gamma(t-1) = 0$.

Перед переходом в режим I или II выполняются следующие действия:

– опрашивается схема контроля СК(30), формирующая признак правильности кодов в каждой группе pCn кода состояния $a(t+1)$; сохраняется значение $\gamma(t-1)$ и $a(t)_0$.

Режим I. Находится $a(t)_1 = a(t+1) - 1$. Если $a(t)_1 \equiv a(t)_0$, то переход правильный и дается сигнал на исполнение очередного оператора действия $A_0 \dots A_k$, иначе фиксируется ошибка.

Режим II. Находится $a(t)_1$ через контрольное ПЗУ. В контрольном ПЗУ(К) количество констант равно числу состояний $a(t)$ вне счетчика, так как необходимо хранить значение $a(t)_1$ только для $\gamma(t-1) = 1$. Осуществляется переход к режиму I.

Сравнение кодов $a(t)_0$ и $a(t)_1$ реализуется схемой СхСр(24), построенной по известному принципу на двух регистрах со схемой поразрядного сложения по mod 2, выходы которой через элемент «ИЛИ» и RS-триггер формируют сигнал тождественности. В блоке БК(22) необходимо иметь конкретно реализованную схему контроля СхК на элементах алгебры логики, осуществляющую контроль для всех (p) групп с выдачей сигнала о правильности кода во всех группах. Функциональная структура схемы СхК тривиальна и поэтому не приводится.

Схема СК(30) представляет собой четырехуровневую восьмивходовую сеть с одним выходом 12 из двухразрядных схем неравнозначности (10) $z = \bar{a}\bar{b} + \bar{a}b$, разделенных схемами «И» с фиксацией промежуточных результатов на RS-триггерах (11) (рис. 5).

В блок контроля БК(22) включены схема сравнения СхСр(24) и схема контроля СхК(30). Управляющий автомат УА(32), схема сравнения СхСр(24) и схема контроля СхК(30) используются для определения наличия только одной единицы в коде после сложения по mod 2 кодов $a(t)_0$ и $a(t)_1$. В блок восстановления (рис. 6) включены вычитающий счетчик Сч(64) и контрольное ПЗУ(К)-34.

Реализация УА(32) по ГСА (см. рис. 4) осуществляется по структурной схеме [15] с новым определением состояний с переходами вне счетчика (табл. 5). По существующей классификации [6] блок УА(32) представляет собой простой управляющий автомат, так как $N = 41$, $m = 6$, $q = 6$.

Для сравнительного анализа аппаратных затрат рассмотрим наиболее простой способ контроля с кодом Бергера [9]. Для управляющего автомата (УА), построенного по граф-схеме алгоритма рис. 4, необходимо оценить затраты на устройство регистрации кодов со счетчиком числа «1», которые приблизительно равны затратам на блок восстановления рис. 6, но объем памяти V для рассматриваемого

примера с $m = 6$, $q = 6$ равен $V = m_p 2^{m+q} = 8 \times 2^{6+6} = 32$ кбит при реальной разрядности БИЗ ПЗУ (m_p) = 4 или 8. При использовании предлагаемого способа кодирования с управляющим автоматом объем памяти $W = m_p 2^{m+1} = 8 \times 2^{6+1} = 1$ кбит, т.е. выигрыш в объеме памяти в 32 раза. При реализации УА на ПЛИМ число ее входов (n) сократится в $n = (m + q)/(m + 1) = (6 + 6)/(6 + 1) = 1,7$ раза. Оба варианта (на ПЗУ и ПЛИМ) подтверждают эффективность предложенного способа кодирования. Другие коды (Хэмминга, kCn) не оцениваются, так как их средства контроля более сложны, чем для кода Бергера, следовательно, предлагаемый способ контроля будет еще более эффективным. Аналогичное заключение можно сделать и для УА с другими характеристиками m и q .

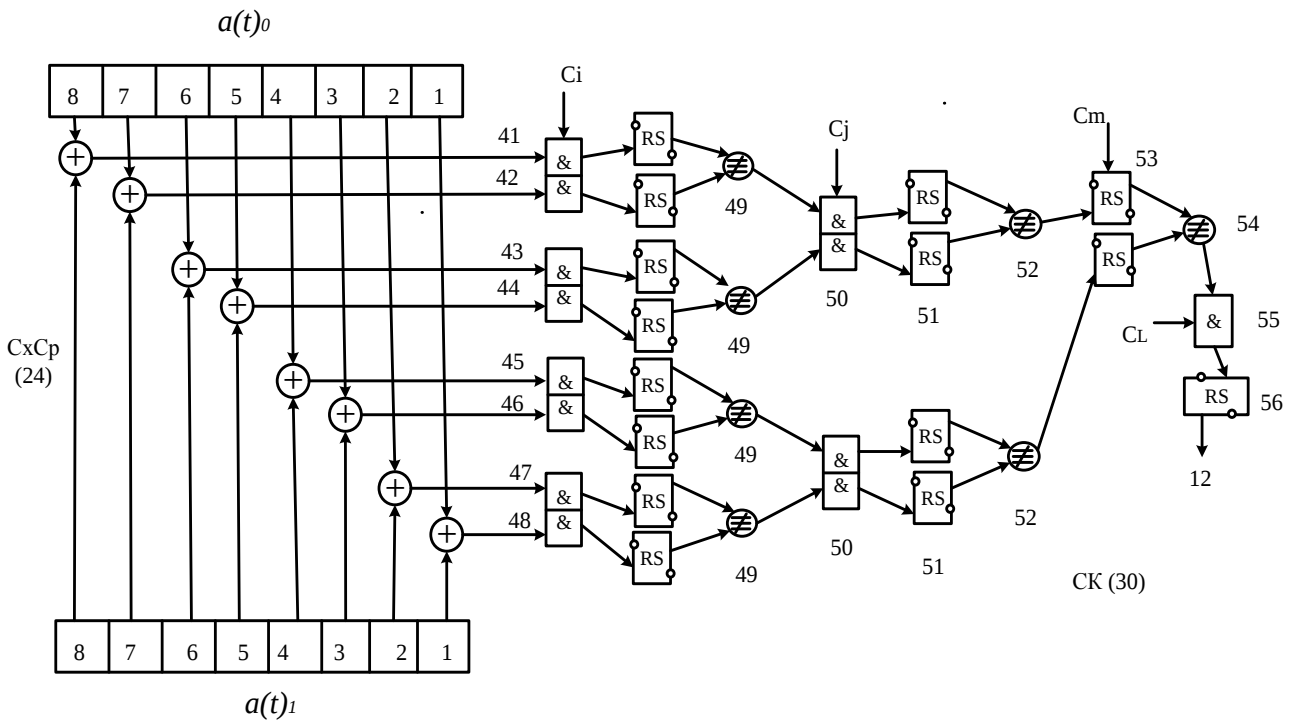


Рис. 5. Схема определения только одной «1» в логической сумме $(a(t)_0 \oplus a(t)_1)$
 Fig. 5. Scheme for determining only one «1» in the logical sum $(a(t)_0 \oplus a(t)_1)$

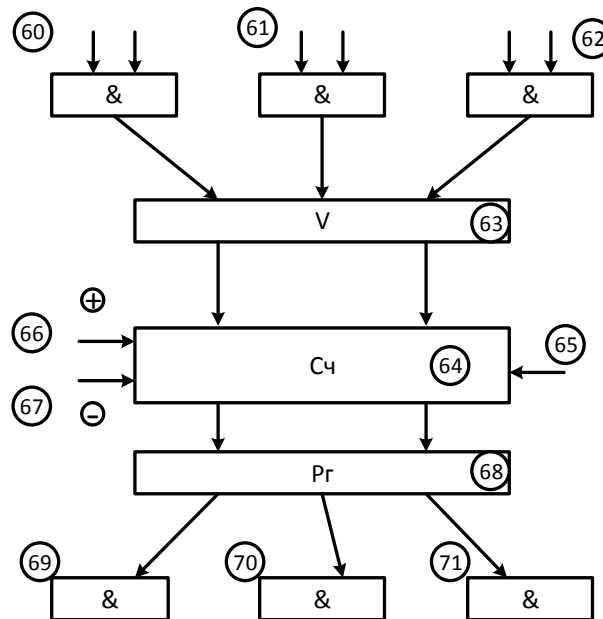


Рис. 6. Блок восстановления
 Fig. 6. Block recovery

Таблица 5

Переходы вне счетчика блока синхронизации

№	$a(t)$	α	$a(t+1)$	$N(t+1)$		№	$a(t)$	α	$a(t+1)$	$N(t+1)$	
1	7	$\overline{\alpha_2}$	31	011	111	7	39	–	2	000	010
2	15	–	40	110	000	8	15	–	37	100	101
3	40	–	33	100	001	9	18	α_4	29	011	101
4	6	$\overline{\alpha_1}$	37	100	101	10	30	–	16	010	000
5	37	–	35	100	011	11	25	α_1	15	001	111
6	33	$\overline{\alpha_6}$	33	100	110	12	28	–	3	000	011
						13	36	–	0	000	000

Примеры систем, в которых может использоваться предлагаемый метод, представлены в работах [17–20].

Заключение

Рассмотрены способы контроля управляющих автоматов со структурной организацией Мура с применением известных корректирующих кодов Хэмминга, Бергера, kSn и др. Показано, что сложность комбинационной схемы переходов самоконтролируемых автоматов связана с отсутствием адресной подсистемы. Адресная подсистема может быть введена в виде адресного блока из (q) двух-входовых элементов «И», одного элемента «ИЛИ» и RS-триггера. Такое решение позволяет создать управляющие автоматы с объемом комбинационной схемы переходов, в 2^{q-1} раз меньшей, чем объем той же схемы в УА Мура. По существующей парадигме снижение затрат на комбинационную схему переходов основано на сокращении числа состояний, декомпозиции граф-схем алгоритмов или графов переходов и последующей минимизации систем булевых функций. По известной методологии синтеза затраты на реализацию УА могут быть снижены в 1,2–1,5 раза, причем чем сложнее УА, тем ниже коэффициент снижения сложности. В автоматах нового типа с оригинальной структурной организацией целенаправленно увеличивается число состояний через разметку входов как операторов действия, так и логических условий. Объем комбинационной схемы переходов снижается в 2^{q-1} раз при программной реализации по сравнению с объемом той же схемы в УА Мура. Этот вариант реализации наиболее прост и поэтому именно такие управляющие автоматы контролируются встроенными средствами самоконтроля. Способ контроля основан на применении нового равновесного кода pSn из (p) трехразрядных групп только с одной «1» в группе без кода из всех нулей. Предлагаемый способ требует меньших затрат оборудования, чем варианты реализации средств контроля на основе известных контролирующих кодов.

ЛИТЕРАТУРА

- Новиков Д.А. Кибернетика // Проблемы управления. 2016. № 1. С. 73–81.
- Гаврилов М.А. Теория релейно-контактных схем. М. : Изд-во АН СССР, 1950. 143 с.
- Захаров В.Н., Поспелов Д.А., Хазацкий В.Е. Системы управления. М. : Энергия, 1972. 344 с.
- Соловьев В.В., Климович А. Логическое проектирование цифровых систем на основе ПЛИС. М. : Горячая линия – Телеком, 2008. 374 с.
- Сапожников В.В., Сапожников Вл.В., Ефанов Д.В. Коды Хэмминга в системах функционального контроля логических устройств. СПб. : Наука, 2018. 151 с.
- Мухопад Ю.Ф. Анализ структурных схем вычислителей по автоматным моделям // Труды междунар. конгресса IFAC : сб. Рига, 1974. Т. 5: Дискретные системы. С. 131–140.
- Баркалов А.А., Титаренко Л.А. Прикладная теория цифровых автоматов. Донецк : ДонНТУ, 2013. 320 с.
- Ульман Д.Д., Мотвани Р., Хопкрофт Д. Введение в теорию автоматов, языков и вычислений. М. : Вильямс, 2016. 528 с.
- Труды по теории синтеза и диагноза конечных автоматов и релейных устройств / под ред. В.В. Сапожникова и Вл.В. Сапожникова. СПб. : Элмор, 2009. 894 с.
- Кудрявцев В.Б., Алешин Ф.Б., Подколзин А.С. Теория автоматов / МГУ им. М. Ломоносова. М. : Юрайт, 2018. 320 с.
- Горбатов В.А. Семантическая теория проектирования автоматов. М. : Энергия, 1979. 263 с.

12. Мухопад А.Ю. Теория управляющих автоматов. Иркутск : ИрГУПС, 2018. 72 с.
13. Смолов В.Б., Чекмарев Ю.Д., Мухопад Ю.Ф. Использование системы геометрических кодов в ПЗУ // Приборостроение. 1971. Т. 14, № 6. С. 73–79.
14. Cook G., Anderson D., Metze G. Design of totally self checking Circuits for m-out-of-n Codes // IEEE Transact. 1973. V. 22, № 3. P. 255–263.
15. Патент на полезную модель № 183109. Российская Федерация, G06F 9/00, G05F 9/00. Управляющий автомат № 2018115113 / Ю.Ф. Мухопад, А.Ю. Мухопад, Д.Ц. Пунсык-Намжилов Заявл. 23.04.2018; опубл. 11.09.2018. Бюл. № 26.
16. Патент на полезную модель № 191742, G06F 9/00 Управляющий автомат с мажоритарно-резервируемой комбинационной схемой: № 2019104150 / Ю.Ф. Мухопад, А.Ю. Мухопад, Д.Ц. Пунсык-Намжилов. Заявл. 14.02.2019; опубл. 19.08.2019. Бюл. № 23.
17. Мухопад Ю.Ф. Микроэлектронные информационно-управляющие системы. Иркутск : ИрГУПС, 2004. 407 с.
18. Мухопад Ю.Ф., Мухопад А.Ю. Алгоритмические системы управления. Иркутск : ИрГУПС, 2018. 96 с.
19. Ефанов Д.В. Экспериментальные исследования кодов с суммированием единичных и одного взвешенного разрядов в системах функционального контроля // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2018. № 43. С. 81–94.
20. Филатов Д.А., Мухопад Ю.Ф. Системы управления турбомеханизмами // Научный вестник Новосибирского электротехнического института. 2013. Вып. 2. С. 17–22.

Поступила в редакцию 18 ноября 2020 г.

Mukhopad Yu.F., Mukhopad Au.Y., Punsik-Namzhilov D.C. (2021) DYNAMIC CONTROL OF CONTROL AUTOMATA OF COMPLEX DISCRETE AUTOMATION SYSTEMS. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika* [Tomsk State University Journal of Control and Computer Science]. 55. pp. 112–121

DOI: 10.17223/19988605/55/13

Methods for controlling control automata with Moore's structural organization and the use of well-known Hamming, Berger, kCn and other correction codes are considered. According to the existing paradigm, reducing the cost of a combinational transition scheme is based on reducing the number of States, decomposing graph schemes of algorithms or transition graphs, and then minimizing Boolean function systems. In the new type of automata proposed by the authors, the number of States is purposefully increased. The address subsystem is defined by introducing an address block of (q) two - input "And" elements, one "OR" element, and an RS trigger. Control of such control automata is implemented using a new equilibrium pCn code from (p) three-digit groups with one unit in the group. In the structure of the self-controlled automaton, an encoder (encoder) is introduced for translating the pCn code into a binary positional code. The proposed method for controlling control automata with pCn codes is much simpler than methods using other known correction codes. The binary equivalent of a three-digit group code is obtained by replacing each three-digit group with two binary digits. This new binary positional code makes it possible to create self-controlled control automata with a simple self-monitoring procedure. This procedure is fundamentally new and more efficient, which is proved by the results of software and hardware implementation of a new type of control automata.

Keywords: control; control automata; state coding.

MUKHOPAD Yuri Fedorovich (Doctor of Technical Science, Professor, Honored Worker of Science of the Russian Federation, Professor of the Department «Automation of Production Processes» Irkutsk State Transport University, Honored Professor Irkutsk State Transport University, Irkutsk, Russian Federation).
E-mail: bts48@mail.ru

MUKHOPAD Aleksandr Yuryevich (Doctor of Technical Science, Associate Professor, Professor of the Department «Automation of Production Processes», Irkutsk State Transport University, Irkutsk, Russian Federation).
E-mail: jcmg@mail.ru

PUNSIK-NAMZHILOV Daba Cyrenovich (Candidate of Technical Sciences, Associate Professor, Director of LLC «Sibproekt», Tomsk, Russian Federation).
E-mail: dablfff@mail.ru

REFERENCES

1. Novikov, D.A. (2016) Kibernetika [Cybernetics]. *Problemy upravleniya*. 1. pp. 73–81.
2. Gavrilov, M.A. (1950) *Teoriya releyno-kontaknykh skhem* [Theory of relay-contact schemes]. Moscow: USSR Academy of Sciences.
3. Zakharov, V.N., Pospelov, D.A. & Khazatsky, V.E. (1972) *Sistemy upravleniya* [Control Systems]. Moscow: Energia.

4. Solovyov, V. V. & Klimovich, A. (2008) *Logicheskoe proektirovanie tsifrovyykh sistem na osnove PLIS* [Logical design of digital systems based on FPGA]. Moscow: Goryachaya liniya – Telekom.
5. Sapozhnikov, V.V., Sapozhnikov, V.I. & Efanov, D.V. (2018) *Kody Khemminga v sistemakh funktsional'nogo kontrolya logicheskikh ustroystv* [Hamming Codes in systems of functional control of logical devices]. St. Petersburg: Nauka.
6. Mukhopad, Yu.F. (1974) Analiz strukturnykh skhem vychisliteley po avtomatnym modelyam [Analysis of structural schemes of calculators based on automatic models]. *Trudy mezhdunar. kongressa IFAS*. 5. pp. 131–140.
7. Barkalov, A.A. & Titarenko, L.A. (2013) *Prikladnaya teoriya tsifrovyykh avtomatov* [Applied theory of digital automata]. Donetsk: DonNTU.
8. Ullman, G.D., Motwani, R. & Hopcroft, D. (2016) *Vvedenie v teoriyu avtomatov, yazykov i vychisleniy* [Introduction to the Theory of Automata, Languages, and Computations]. Translated from English. Moscow: Vil'yams.
9. Sapozhnikov, V.V. & Sapozhnikov, V.I. (eds) (2009) *Trudy po teorii sinteza i diagnoza konechnyykh avtomatov i releynykh ustroystv* [Proceedings on the theory of synthesis and diagnosis of finite automata and relay devices]. St. Petersburg: Elmor.
10. Kudryavtsev, V.B., Aleshin, F.B. & Podkolzin, A.S. (2018) *Teoriya avtomatov* [Theory of Automata]. Moscow: Yurayt.
11. Gorbатов, V.A. (1979) *Semanticheskaya teoriya proektirovaniya avtomatov* [Semantic Theory of Automata Design]. Moscow: Energiya.
12. Mukhopad, A.Yu. (2018) *Teoriya upravlyayushchikh avtomatov* [Theory of Control Automata]. Irkutsk: Irkutsk State Transport University.
13. Smolov, V.B., Chekmarev, Yu.D. & Mukhopad, Yu.F. (1971) Using the system of geometric codes in ROM. *Priborostroenie*. 14(6). pp. 73–79.
14. Cook, G., Anderson, D. & Metze, G. (1973) Design of totally self-checking check circuits for m-out-of-n codes. *IEEE Transactions*. 22(3). pp. 255–263. DOI: 10.1109/T-C.1973.223705
15. Russian Federation. (2018) *The patent for useful model no. 183109. G06F 9/00, G05F 9/00. Control automaton No. 2018115113. Yu. F. Mukhopad, A.Yu. Mukhopad, D. C.-Punsik-Namzhilov Appl. 23.04.2018; publ. 11.09.2018. Bull. No. 26.*
16. Russian Federation. (2018) *The patent for useful model no 191742. G06F 9/00 Control automaton with the majority-redundant combinational circuit: no. 2019104150. Yu.F Mukhopad, A.Yu. Mukhopad, D. C. Puncik -Namzhilov. - Declared on 14.02.2019; published on 19.08.2019. Bull. no. 23.*
17. Mukhopad, Yu.F. (2004) *Mikroelektronnye informatsionno-upravlyayushchie sistemy* [Microelectronic information and control systems]. Irkutsk: Irkutsk State Transport University.
18. Mukhopad, Yu.F. & Mukhopad, A.Yu. (2018) *Algoritmicheskie sistemy upravleniya* [Algorithmic Control Systems]. Irkutsk: Irkutsk State Transport University.
19. Efanov, D.V. (2018) Experimental studies of one-bit and one weighted bit sum codes within concurrent error detection systems. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravleniye, vychislitel'naya tekhnika i informatika –Tomsk State University Journal of Control and Computer Science*. 43. p. 81–94. DOI: 10.17223/19988605/43/10
20. Filatov, D.A. & Mukhopad, Yu.F. (2013) *Sistemy upravleniya turbomekhanizmami* [Turbomechanism control Systems]. *Nauchnyy vestnik Novosibirskogo elektrotekh-nicheskogo instituta*. 2. pp. 17–22.

УДК 53.083.9

DOI: 10.17223/19988605/55/14

В.М. Соловьев, Д.В. Сперанский

АЛГОРИТМ МУРАВЬИНОЙ КОЛОНИИ СИНТЕЗА ТЕСТОВ ДЛЯ ЦИФРОВЫХ УСТРОЙСТВ

Описывается эволюционный алгоритм синтеза тестов для цифровых устройств (ЦУ), основанный на использовании алгоритма муравьиной колонии. Предлагаемый алгоритм ориентирован на диагностирование таких неисправностей ЦУ, которые обычно трудно обнаруживать и диагностировать другими широко практикуемыми методами.

Ключевые слова: цифровые устройства (ЦУ); контроль и диагностика ЦУ; алгоритм муравьиной колонии; синтез тестов.

Проблемы контроля и диагностирования цифровых устройств (ЦУ) относятся к числу актуальных, поскольку обеспечение высокой надежности функционирования ЦУ невозможно без их эффективного решения. Жесткие требования к надежности ЦУ диктуются постоянно возрастающей ответственностью функций, которые на них возлагаются. Это в полной мере относится, например, к ЦУ, используемым в системах управления опасными промышленными производствами (изготовление токсичных и взрывчатых веществ и т.п.), в системах обеспечения безопасности транспортных средств (авиация, железные дороги, морские суда) и т.д. Важным этапом обеспечения высокой надежности ЦУ является эффективная организация их тестового и функционального диагностирования [1].

Предлагаемая статья посвящена задаче синтеза тестов для ЦУ. Методы синтеза тестов можно условно разделить на две группы. К первой относятся сравнительно просто реализуемые методы, но вместе с тем получаемые ими тесты обнаруживают, как правило, сравнительно большую часть возможных неисправностей ЦУ, однако не все их множество. Примером может служить метод псевдослучайного синтеза тестов [1]. Ко второй группе относятся методы, более трудоемкие в реализации, но синтезируемые ими тесты способны обнаруживать не только «простые», но и «трудные» неисправности. Примером могут служить методы различающей функции и булевых производных [1]. Предлагаемый в статье метод относится именно к этой группе.

В последние два-три десятилетия получили широкое распространение так называемые эволюционные алгоритмы [2, 3], оказавшиеся эффективными для решения многих практических проблем. В статье описан метод синтеза тестов, базирующийся на идее одного из таких алгоритмов, алгоритме муравьиной колонии (АМК), разработанном М. Дориго [4, 5] и детально описанном в [6, 7].

1. Постановка задачи

Пусть задано ЦУ в виде структурной схемы и множество F его возможных неисправностей. Для простоты изложения далее предполагается, что ЦУ является комбинационным, а рассматриваемые неисправности – одиночными константными. Требуется разработать АМК для синтеза тестов, позволяющий обнаруживать все неисправности из множества F . Далее эту задачу будем именовать задачей синтеза тестов (ЗСТ).

Рассматриваемая задача наиболее близка к классической задаче коммивояжера (ЗК) [6], и аналогия с ней будет использована нами.

В ЗСТ моделью ЦУ является неориентированный граф $G = (V, E)$, где V – множество вершин, каждая из которых соответствует некоторому входному набору ЦУ. В множество V включаются

только те входные наборы ЦУ, которые обнаруживают неисправности из F . Если p – число входов, то $|F|$ может быть существенно меньше 2^p . Множество E образуют ребра между любой парой вершин графа G . Понятно, что граф $G = (V, E)$ является полным [8]. Заметим, что в ЗК используемый в качестве модели граф полным не является.

Каждому ребру ставится в соответствие число, называемое длиной (весом), о содержательном смысле которого будет сказано ниже. В терминах описанного графа решением ЗСТ является минимальный по длине путь (маршрут) в этом графе (сумма длин всех его ребер), которому соответствует тест, обнаруживающий все неисправности из F . Отметим, что если в ЗК ее решение есть путь, проходящий через все вершины графа, то в ЗСТ минимальный по длине «тестовый» путь может и не проходить через все вершины графа. При этом, как и в ЗК, каждая вершина «тестового» пути должна посещаться только один раз. В ЗСТ это требование объясняется тем, что повторение в тесте одного и того же входного набора не увеличивает числа обнаруживаемых неисправностей ЦУ.

Из сказанного выше вытекает, что хотя между ЗК и ЗСТ, несомненно, имеется аналогия, вместе с тем существующие между ними различия требуют существенной модификации АМК решения ЗК для возможности адаптации его к решению ЗСТ.

2. Описание алгоритма

Далее предполагается знакомство читателя с терминологией и понятиями из области муравьиных алгоритмов хотя бы в рамках небольшой по объему публикации [6]. Процесс построения теста каждым k -м муравьем (их число выбирается заранее) является пошаговым. Он начинается с включения в тест некоторого входного набора с последующим добавлением к «текущему» тесту $T_{\text{тек}}$ нового набора на каждом очередном шаге. С помощью $T_{\text{тек}}$ обнаруживается некоторое множество неисправностей ЦУ, которое исключается из первоначально заданного множества неисправностей F , и в результате остается множество F_k^H еще не обнаруженных неисправностей. Добавление к «текущему» тесту $T_{\text{тек}}$ нового набора продолжается до тех пор, пока множество F_k^H не окажется пустым.

Перейдем теперь к содержательному описанию предлагаемого АМК, предварительно введя некоторые обозначения, сопровождаемые комментариями.

$F(w)$ – множество неисправностей из F , обнаруживаемых с использованием входного слова w .

$F_{k,t}^H T_{\text{тек}}$ – множество еще не обнаруженных тестом $T_{\text{тек}}$ k -м муравьем неисправностей из F на итерации t ;

$d_{ij}(k,t) = |F_{k,t}^H(T_{\text{тек}}, i) \cap F_{k,t}^H(T_{\text{тек}}, j)|$ – число неисправностей, не обнаруживаемых k -м муравьем на итерации t как входным набором i , так и следующим за ним в текущем тесте входным набором j .

$$D_{ij} = \begin{cases} |F(i) \cap F(j)|, & \text{если } F(i) \neq F(j), \\ C - \text{большая константа,} & \text{если } F(i) = F(j), \end{cases}$$

D_{ij} – вес (длина) дуги (i, j) графа G).

Поясним содержательный смысл D_{ij} : чем больше множество «общих» неисправностей, обнаруживаемых наборами i и j , тем менее «привлекателен» для включения в тест после набора i входной набор j . Понятно, что вклад набора j в распознавание неисправностей из F в этом случае будет уменьшаться с ростом величины D_{ij} . Это хорошо согласуется с используемым в АМК для ЗК понятием видимости вершин графа [6]: чем больше значение D_{ij} , тем «дальше» вершина j от вершины i , т.е. тем хуже ее «видимость».

$\eta_{ij} = 1 / D_{ij}$ – видимость, являющаяся локальной статической информацией, которая выражает естественное желание включить в качестве очередного элемента искомого теста входной набор с наибольшим вкладом в обнаружение неисправностей множества F .

Отметим, что значения элементов D_{ij} матрицы D находятся применением логического моделирования ЦУ при наличии в нем всех неисправностей множества F .

Самоорганизация в колонии муравьев [6] есть результат взаимодействия четырех компонентов – случайности, многократности, а также положительной и отрицательной обратных связей. Что касается многократности, в АМК для ЗСТ она реализуется за счет итерационного процесса поиска теста, осуществляемого не одним, а группой муравьев. При этом каждый муравей этой группы решает ЗСТ независимо от других. В процессе выполнения итерации каждый муравей группы решает задачу синтеза теста до конца.

Упомянутый выше принцип положительной обратной связи в АМК реализуется в виде имитации предпочтения муравьем при выборе очередного ребра (входного набора) графа при построении тестового пути. Муравей может отдать предпочтение такому из альтернативных ребер, на котором концентрация феромона максимальна. Большая привлекательность для муравья такого ребра в качестве очередного звена тестового маршрута интуитивно связана с тем, что этот след феромона «укреплен» за счет перемещения по ребру многих муравьев, задействованных для решения той же задачи.

Условимся далее концентрацию (количество) феромона на ребре (i, j) на итерации t обозначать как $\tau_{ij}(t)$. Откладываемый на ребрах феромон позволяет «хорошим» тестовым маршрутам сохраняться в глобальной памяти колонии муравьев. Последующие итерации АМК могут привести к улучшению таких тестовых путей. Заметим, что концентрация феромона на ребрах должна изменяться после каждой итерации АМК, что отражает приобретение муравьями опыта в поиске теста.

Однако применение только положительной обратной связи может привести к ситуации, когда все муравьи двигаются одним и тем же субоптимальным (близким к оптимальному, но не оптимальным) путем. По аналогии с АМК для ЗК в нашем алгоритме будет использоваться отрицательная обратная связь, вызываемая испарением феромона. Этот процесс приводит к тому, что группа муравьев, решающая ЗСТ, одновременно исследует разные точки пространства решений и передает свой опыт через изменение ячеек глобальной памяти колонии муравьев. Такая память представляет собой совокупность значений концентрации феромона на ребрах графа G , накопленных в результате итераций АМК.

Компонент случайности в нашем АМК реализуется при переходе от одного набора синтезируемого теста к следующему. По аналогии с ЗК вероятность перехода k -го муравья из вершины i в вершину j на t -й итерации имеет следующий вид:

$$\begin{cases} P_{ij,k}(t) = \frac{[\tau_{ij}]^\alpha [\eta]^\beta}{\sum_{i \in J_{i,k}} [\tau_{il}(t)]^\alpha [\eta_{il}(t)]^\beta}, & \text{если } j \in J_{i,k}, \\ P_{ij,k}(t) = 0, & \text{если } j \notin J_{i,k}, \end{cases} \quad (1)$$

где $J_{i,k}$ – множество еще не использованных входных наборов ЦУ в синтезируемом тесте k -м муравьем, α и β – регулируемые параметры, задающие веса следа феромона и видимости при построении теста.

Понятно, что при $\alpha = 0$ концентрация феромона при выборе ребра не используется, т.е. выбор осуществляется только на основе видимости вершины j . Если же $\beta = 0$, то для выбора используется только концентрация феромона, что может привести к сваливанию тестовых путей к одному из субоптимальных.

Как и в АМК, для ЗК выбор очередного входного набора должен осуществляться по правилу рулетки. При этом каждый входной набор имеет в секторе рулетки площадь, пропорциональную вероятности (1).

По аналогии с ЗК формула корректировки концентрации феромона на ребре графа за счет его испарения (с параметром $\rho \in [0,1]$) имеет следующий вид:

$$\tau_{ij}(t+1) = (1-\rho)\tau_{ij}(t) + \Delta\tau_{ij}(t), \quad (2)$$

где $\Delta\tau_{ij}(t) = \sum_{k=1}^m \Delta\tau_{ij,k}(t)$, m – количество муравьев в группе (колонии), синтезирующей тест (m является регулируемым параметром).

Итак, в предложенном АМК для ЗСТ самоорганизация в колонии муравьев базируется на взаимодействии компонентов случайности, многократности, положительной и отрицательной обратных связей.

Заключение

Предложенный алгоритм расширяет область приложений эволюционных вычислений на решение важного класса задач контроля и диагностирования ЦУ. Качество получаемых с применением алгоритма муравьиной колонии результатов во многих других проблемах дает основание рассчитывать на эффективные результаты и в рассмотренной нами области.

ЛИТЕРАТУРА

1. Сперанский Д.В., Скобцов Ю.А., Скобцов В.Ю. Моделирование, тестирование и диагностика цифровых устройств. 2-е изд. М. : Нац. открытый ун-т ИНТУИТ, 2016. 535 с.
2. Курейчик В.В., Курейчик В.М., Родзин С.И. Теория эволюционных вычислений. М. : Физматлит, 2012. 260 с.
3. Скобцов Ю.А., Сперанский Д.В. Эволюционные вычисления : учеб. пособие. М. : Нац. открытый ун-т ИНТУИТ, 2015. 326 с.
4. Dorigo M., Maniezzo V., Colomi A. The Ant System: Optimization by a colony of cooperating objects // IEEE Trans. on Systems, Man and Cybernetics. Part B. 1996. № 26 (1). P. 29–41.
5. Dorigo M., Gambardella M.A. Ant colony systems: a cooperative learning approach to the traveling salesman problem objects // IEEE Trans. on Evolutionary Computation. 1997. № 1 (1). P. 53–66.
6. Штовба С.Д. Муравьиные алгоритмы // Экспонента Про. Математика в приложениях. 2003. № 4. С. 70–75.
7. Ермолаев С.Ю. Муравьиные алгоритмы оптимизации // Инфокоммуникационные технологии. 2008. Т. 6, № 1. С. 23–29.
8. Новиков Ф.А. Дискретная математика. 3-е изд. СПб. : Питер, 2017. 497 с.

Поступила в редакцию 3 сентября 2020 г.

Solovyev V.M., Speranskiy D.V. (2021) THE ANT COLONY ALGORITHM FOR TEST SYNTESIS FOR DIGITAL DEVICES. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika* [Tomsk State University Journal of Control and Computer Science]. 55. pp. 122–126

DOI: 10.17223/19988605/55/14

The problem of test synthesis for digital devices (DD) is studied. Test synthesis methods can be divided into two groups. The first includes methods that are relatively easy to implement, but the tests they synthesize are detecting significant part of all possible faults. An example is the pseudo-random test synthesis method. The second group includes methods that are more laborious in implementation, but are capable of detecting faults that are difficult to diagnose. These include methods, for example, of the distinguishing function and Boolean derivatives.

The proposed method belongs to this group. Nowadays, so-called evolutionary algorithms are widespread. These algorithms are very effective for many practical applications. The article proposes a method for test synthesis based on the ideas of evolutionary ant colony algorithms.

We briefly describe the problem under study. Let DD be determined in the form of a structural scheme and the set F of its possible faults. For simplicity, it is assumed that the DD is a combinational device, and the set F consists of single constant faults. We consider the problem of test synthesis (PTS) for DD, detecting out of faults from the set F .

In the PTS the mathematical model of the DD is the undirected graph $G = (V, E)$, where V is the set of vertices. Each of its vertices is some binary input of DD. Only the inputs that detect faults are included in the set V . The set E forms edges between any pair of vertices of the graph G . Each edge (i, j) is assigned to a some number (weight). Its meaning is to assess the feasibility of including the input j in the step-by-step process synthesis of test after the input i .

Considered PTS is the most close to the classic traveling salesman problem (TSP). The basis of the proposed ant colony algorithm for the PTS is the analogy with the ant colony algorithm for solving the TSP. The article presents formulas similar to those used in the ant colony algorithm for solving TSP to calculate the probabilities of transitions from one input symbol of the test to the next, correction of the pheromone concentration on the edges of the graph G etc. Note that in the proposed ant colony algorithm for solving an TSP the basic principle of self-organization is based on the interaction of components of randomness, multiplicity, positive and negative feedbacks.

Keywords: digital devices; control and diagnostics of DD; ant colony algorithm; test synthesis.

SOLOVYEV Vladimir Mikhailovich (Candidate of Technical Sciences, Associate Professor, Head of the Volga Regional Center for New Information Technologies of Saratov State University, Saratov, Russian Federation).

E-mail: ign1122@mail.ru

SPERANSKIY Dmitriy Vasilyevich (Professor, Doctor of Technical Sciences, Russian University of Transport (MIIT), Moscow, Russian Federation).

E-mail: speranskiy.dv@gmail.com

REFERENCES

1. Speranskiy, D.V., Skobtsov, Yu.A. & Skobtsov, V.Yu. (2016) *Modelirovanie, testirovanie i diagnostika tsifrovyykh ustroystv* [Modeling, testing and diagnostics of digital devices]. 2nd ed. Moscow: INTUIT.
2. Kureychik, V.V., Kureychik, V.M. & Rodzin, S.I. (2012) *Teoriya evolyutsionnykh vychisleniy* [Theory of Evolutionary Computations]. Moscow: Fizmatlit.
3. Skobtsov, Yu.A. & Speranskiy, D.V. (2015) *Evolutsionnye vychisleniya* [Evolutionary Computations]. Moscow: INTUIT.
4. Dorigo, M., Maniezzo, V. & Colomi, A. (1996) The Ant System: Optimization by a colony of cooperating objects. *IEEE Trans. On Systems, Man and Cybernetics*. Part B. 26(1). pp. 29–41.
5. Dorigo, M. & Gambardella, M.A. (1997) Ant colony systems: a cooperative learning approach to the traveling salesman problem objects. *IEEE Trans. on evolutionary computation*. 1(1). pp. 53–66.
6. Shtovba, S.D. (2003) Ant algorithms. *Exponenta Pro. Mathematics in Applications*. 4. pp. 70–75.
7. Ermolaev, S.Yu. (2008) Ant algorithms of optimization. *Infokommunikatsionnye tekhnologii*. 6(1). pp. 23–29.
8. Novikov, F.A. (2017) *Diskretnaya matematika* [Discrete mathematics]. 3rd. ed. St. Petersburg: Piter.

СВЕДЕНИЯ ОБ АВТОРАХ

АБАЕВА Нигяр Бахрам кызы – диссертант кафедры дифференциальных уравнений и оптимизации Сумгаитского государственного университета (Азербайджан). E-mail: abayeveldar404@gmail.com

АНДРИЕВСКАЯ Анна Андреевна – аспирант кафедры системного анализа и математического моделирования Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета. E-mail: anna.andriev@mail.tsu.ru

БРОКАРЕВ Иван Андреевич – аспирант, ассистент кафедры автоматизации технологических процессов Российского государственного университета нефти и газа (Национальный исследовательский университет) имени И.М. Губкина. (г. Москва). E-mail: brokarev.i@gubkin.ru

ВАСЬКОВСКИЙ Сергей Владимирович – кандидат технических наук, старший научный сотрудник Института проблем управления им. В.А.Трапезникова РАН (г. Москва). E-mail: v63v@yandex.ru

ВИК Ксения Васильевна – аспирант отделения информационных технологий Национального исследовательского Томского политехнического университета. E-mail: kvv11@tpu.ru

ВОРОБЬЕВ Андрей Владимирович – кандидат технических наук, доцент кафедры геоинформационных систем факультета информатики и робототехники Уфимского государственного авиационного технического университета. E-mail: geomagnet@list.ru

ГРИГОРЬЕВ Дмитрий Сергеевич – ассистент отделения информационных технологий Национального исследовательского Томского политехнического университета. E-mail: trygx@tpu.ru

ДРУКИ Алексей Алексеевич – кандидат технических наук, доцент отделения информационных технологий Национального исследовательского Томского политехнического университета. E-mail: druki@tpu.ru

ЗЕНКОВА Жанна Николаевна – доцент, кандидат физико-математических наук, доцент кафедры системного анализа и математического моделирования Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета; старший научный сотрудник Российской академии народного хозяйства и государственной службы при Президенте Российской Федерации (г. Москва). E-mail: zhanna.zenkova@mail.tsu.ru

КАЛАШНИКОВ Сергей Николаевич – доцент, доктор технических наук, профессор кафедры прикладных информационных технологий и программирования Сибирского государственного индустриального университета (г. Новокузнецк). E-mail: s.n.kalashnikov@yandex.ru

КОЛЕСНИКОВА Светлана Ивановна – доктор технических наук, профессор кафедры компьютерных технологий и программной инженерии института вычислительных систем и программирования Санкт-Петербургского государственного университета аэрокосмического приборостроения. E-mail: skolesnikova@yandex.ru

МАРТУСЕВИЧ Елена Владимировна – кандидат технических наук, доцент кафедры естественнонаучных дисциплин им. профессора В.М. Финкеля Сибирского государственного индустриального университета (г. Новокузнецк). E-mail: science_nvz@yandex.ru

МАРТУСЕВИЧ Ефим Александрович – младший научный сотрудник, аспирант кафедры прикладных информационных технологий и программирования Сибирского государственного индустриального университета (г. Новокузнецк). E-mail: program.pro666@yandex.ru

МУСОНИ Уильсон – аспирант кафедры системного анализа и математического моделирования Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета. E-mail: wmusoni@uok.ac.rw

МУХОПАД Александр Юрьевич – доцент, доктор технических наук, профессор кафедры автоматизации производственных процессов Иркутского государственного университета путей сообщения. E-mail: jcmg@mail.ru

МУХОПАД Юрий Федорович – профессор, доктор технических наук, заслуженный деятель науки РФ, заслуженный профессор Иркутского государственного университета путей сообщения, профессор кафедры автоматизации производственных процессов. E-mail: bts48@mail.ru

НЕЖЕЛЬСКАЯ Людмила Алексеевна – доцент, доктор физико-математических наук, профессор кафедры прикладной математики Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета. E-mail: ludne@mail.ru

ПАРАЕВ Юрий Иванович – профессор, доктор технических наук, заслуженный деятель науки РФ, заслуженный профессор Национального исследовательского Томского государственного университета, профессор кафедры прикладной математики Института прикладной математики и компьютерных наук. E-mail: paraev@mail.ru

ПЕРШИНА Анна Александровна – аспирант кафедры прикладной математики Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета. E-mail: ann.shitina@gmail.com

ПУНСЫК-НАМЖИЛОВ Даба Цыренович – доцент, кандидат технических наук, директор ООО «Сибпроект» (г. Томск). E-mail: dablutf@mail.ru

РУБАН Анатолий Иванович – профессор, доктор технических наук, заслуженный деятель науки РФ, заслуженный профессор Сибирского федерального университета, профессор кафедры информатики Института космических и информационных технологий (г. Красноярск). E-mail: ai-rouban@mail.ru

РУДЬКО Игорь Михайлович – кандидат технических наук, старший научный сотрудник Института проблем управления им. В.А.Трапезникова РАН (г. Москва). E-mail: igor-rudko@mail.ru

РЫБЕНКО Инна Анатольевна – доцент, доктор технических наук, заведующая кафедрой прикладных информационных технологий и программирования Сибирского государственного индустриального университета (г. Новокузнецк). E-mail: rybenkoi@mail.ru

СОЛОВЬЕВ Владимир Михайлович – доцент, кандидат технических наук, начальник Поволжского регионального центра новых информационных технологий Саратовского государственного университета. E-mail: svm@sgu.ru

СПЕРАНСКИЙ Дмитрий Васильевич – профессор, доктор технических наук, профессор кафедры систем управления транспортной инфраструктурой Российского университета транспорта (МИИТ) (г. Москва). E-mail: speranskiy.dv@gmail.com

СПИЦЫН Владимир Григорьевич – профессор, доктор технических наук, профессор отделения информационных технологий Национального исследовательского Томского политехнического университета. E-mail: spvg@tpu.ru

ФАРХАДОВ Маис Паша оглы – доктор технических наук, заведующий лабораторией Института проблем управления РАН (г. Москва). E-mail: mais@ipu.ru

ФЕЙЗИЕВ Фикрат Гюльали оглы – доктор физико-математических наук, профессор, заведующий кафедрой дифференциальных уравнений и оптимизации Сумгаитского государственного университета (Азербайджан). E-mail: FeyziyevFG@mail.ru

ЦВЕТНИЦКАЯ Светлана Александровна – кандидат технических наук, доцент кафедры прикладной математики Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета. E-mail: svetasa@sibmail.com

ШЕХАЛЕВ Денис Владимирович – ведущий инженер научно-образовательного центра «Инжиниринговый центр СВЧ Техники и Технологии» Национального исследовательского Томского государственного университета. E-mail: diod2003@list.ru

ШУЛЕНИН Валерий Петрович – доцент, кандидат технических наук, доцент Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета. E-mail: shulenin-vp@rambler.ru

Научный журнал

**ВЕСТНИК
ТОМСКОГО
ГОСУДАРСТВЕННОГО
УНИВЕРСИТЕТА**

**УПРАВЛЕНИЕ,
ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА
И ИНФОРМАТИКА**

**TOMSK STATE UNIVERSITY
JOURNAL OF CONTROL AND COMPUTER SCIENCE**

2021. № 55

Редактор Е.Г. Шумская
Оригинал-макет Е.Г. Шумской
Редакторы-переводчики: Г.М. Кошкин; В.Н. Горенинцева
Дизайн обложки Л.Д. Кривцовой

Подписано к печати 28.06.2021 г. Формат 60х84^{1/8}.
Гарнитура Times. Усл. печ. л. 15,6.
Тираж 250 экз. Заказ № 4698. Цена свободная.

Дата выхода в свет 29.06.2021 г.

Журнал отпечатан на полиграфическом оборудовании
Издательства Томского государственного университета
634050, г. Томск, Ленина, 36
Тел. 8(382-2)–52-98-49; 8(382-2)–52-96-75
Сайт: <http://publish.tsu.ru>; E-mail: rio.tsu@mail.ru