

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

ВЕСТНИК ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА

УПРАВЛЕНИЕ,
ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА
И ИНФОРМАТИКА

TOMSK STATE UNIVERSITY
JOURNAL OF CONTROL AND COMPUTER SCIENCE

Научный журнал

2025

№ 71

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

16+

Учредитель – Национальный исследовательский Томский государственный университет (НИ ТГУ)

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

Замятин Александр Владимирович – гл. редактор, проф., д-р техн. наук, НИ ТГУ. Тел. +73822529496
Смагин Валерий Иванович – зам. гл. редактора, проф., д-р техн. наук, НИ ТГУ. Тел. +73822529485
Нежелская Людмила Алексеевна – ответственный секретарь, проф., д-р физ.-мат. наук, НИ ТГУ. Тел. +73822529485
E-mail: vestnik_uvti@mail.tsu.ru

Воробейчиков Сергей Эрикович – д-р физ.-мат. наук, НИ ТГУ
Вишневский Владимир Миронович – проф., д-р техн. наук, Институт проблем управления РАН (г. Москва)
Гордов Евгений Петрович – проф., д-р физ. мат. наук, Институт мониторинга климатических и экологических систем СО РАН (г. Томск)
Кошкин Геннадий Михайлович – проф., д-р физ.-мат. наук, НИ ТГУ
Костюк Юрий Леонидович – проф., д-р техн. наук, НИ ТГУ
Матросова Анжела Юрьевна – проф., д-р техн. наук, НИ ТГУ
Назаров Анатолий Андреевич – проф., д-р техн. наук, НИ ТГУ
Пашинская Татьяна Юрьевна – д-р физ.-мат. наук, НИ ТГУ
Перепелкин Евгений Александрович – проф., д-р техн. наук, Санкт-Петербургский государственный университет аэрокосмического приборостроения (г. Санкт-Петербург)
Пчелинцев Евгений Анатольевич – д-р физ.-мат. наук, НИ ТГУ
Самуйлов Константин Евгеньевич – проф., д-р техн. наук, РУДН (г. Москва)
Семенкин Евгений Станиславович – проф., д-р техн. наук, СибГУ им. акад. М.Ф. Решетнева (г. Красноярск)
Сушенко Сергей Петрович – проф., д-р техн. наук, НИ ТГУ
Фархадов Маис Паша оглы – д-р техн. наук, Институт проблем управления РАН (г. Москва)
Харин Юрий Семенович – академик НАНБ, проф., д-р физ.-мат. наук, БГУ, Республика Беларусь (г. Минск)
Цициашвили Гурами Шалвович – проф., д-р физ.-мат. наук, ИПМ ДВО РАН (г. Владивосток)

РЕДАКЦИОННЫЙ СОВЕТ

Александр Николаевич Дудин – проф., д-р физ.-мат. наук, БГУ, Минск, Республика Беларусь
Ана Роза Ковали (Ana Rosa Cavalli) – проф., д-р. философии (PhD), University VII Paris, Франция
Енцо Орсингер (Enco Orsinger) – проф., д-р. философии (PhD), University of Rome, Италия
Паоло Принетто (Paolo Prinetto) – проф. Politechnic Institute, Torino, Италия
Нина Владимировна Евтушенко – проф., д-р. техн. наук, ИСП РАН им. Иванникова В.П., Москва, Россия
Ервант Зориан (Yervant Zorian) – д-р. философии (PhD), Synopsys, Mountain View, США

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).

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

Зубов Н.Е., Лапин А.В., Рябченко В.Н. Формулы обращения квадратных матриц, разбитых на прямоугольные блоки, и их применение в модальном управлении	4
Мансимов К.Б. Необходимые условия оптимальности первого и второго порядков в обыкновенных динамических системах с запаздыванием при многоточечных функциональных ограничениях типа равенств и неравенств	17
Сорокина Н.В., Шашихин В.Н. Управление нелинейными системами с хаотической динамикой	27

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

Ишkulova A.P., Lyutov A.G. Моделирование процесса поляризации электродов при контроле и управлении процессом нанесения гальванического покрытия	39
Яблонский В.О. Математическая модель и анализ влияния режимных параметров гидроциклона на показатели процесса разделения вязкопластических сред двухстадийной флотацией	48

ОБРАБОТКА ИНФОРМАЦИИ

Abdushukurov A.A., Bozorov S.B. On analysis of survival function estimators with its identifiability with model of right random censorship	58
Воробьева Г.Р., Воробьев А.В., Орлов Г.О. Мультиосевая модель геофизических данных: концепция, формализация, реализация	71
Kitaeva A.V., Yu Cao. Shortages of perishables control for a stochastic inventory system in retail through dynamic pricing	82
Литовченко М.И., Костюк Ю.Л. Быстрая реализация алгоритма невыпуклого полосового слияния триангуляции Делоне	93
Рудько И.М. Применение УПС-фильтра при решении задачи обнаружения в режиме активной локации	103
Смагин В.И., Ким К.С. Идентификация в дискретных системах со случайными и интервальными параметрами с использованием EKF	112

ИНФОРМАТИКА И ПРОГРАММИРОВАНИЕ

Ушенина И.В. Реализация на ПЛИС и сравнительный анализ вычислителей сигмоида, работающих с полным диапазоном аргумента с учетом симметрии	120
---	-----

ПРОЕКТИРОВАНИЕ И ДИАГНОСТИКА ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

Матросова А.Ю., Савенкова М.М. Проверка эквивалентности частично построенных комбинационных схем	130
--	-----

CONTENTS

CONTROL OF DYNAMICAL SYSTEMS

Zubov N.E., Lapin A.V., Ryabchenko V.N. Formulas for inverting square matrices divided into rectangular blocks and their application in modal control	4
Mansimov K.B. Necessary conditions for first- and second-order optimality in ordinary dynamic systems with delay under multipoint functional constraints of the equalities and inequalities type	17
Sorokina N.V., Shashihin V.N. Control of nonlinear systems with chaotic dynamics	27

MATHEMATICAL MODELING

Ishkulova A.R., Lyutov A.G. Modeling of the electrode polarization process in monitoring and controlling the process of applying galvanic coating	39
Yablonsky V.O. Mathematical model and analysis of the effect of the regime parameters of the hydrocyclone on the indicators of the separation process of viscoplastic media by two-stage flotation	48

DATA PROCESSING

Abdushukurov A.A., Bozorov S.B. On analysis of survival function estimators with its identifiability with model of right random censorship	58
Vorobeva G.R., Vorobev A.V., Orlov G.O. Multi-axis model of geophysical data: concept, formalization, implementation	71
Kitaeva A.V., Yu Cao. Shortages of perishables control for a stochastic inventory system in retail through dynamic pricing	82
Litovchenko M.I., Kostyuk Yu.L. Fast implementation of a non-convex Stripe Merging algorithm for Delaunay triangulation	93
Rudko I.M. Using the TOS-filter in solving the detection problem in active location mode	103
Smagin V.I., Kim K.S. Identification in discrete systems with random and interval parameters using EKF	112

INFORMATICS AND PROGRAMMING

Ushenina I.V. FPGA implementation and comparative analysis of sigmoid calculators processing the full argument range in view of symmetry	120
--	-----

DESIGNING AND DIAGNOSTICS OF COMPUTER SYSTEMS

Matrosova A.Yu., Savenkova M.M. Partial equivalence checking for combinational circuits	130
---	-----

УПРАВЛЕНИЕ ДИНАМИЧЕСКИМИ СИСТЕМАМИ

CONTROL OF DYNAMICAL SYSTEMS

Научная статья
УДК 681.5:517
doi: 10.17223/19988605/71/1

Формулы обращения квадратных матриц, разбитых на прямоугольные блоки, и их применение в модальном управлении

Николай Евгеньевич Зубов¹, Алексей Владимирович Лапин², Владимир Николаевич Рябченко³

^{1,2,3} *Московский государственный технический университет им. Н.Э. Баумана, Москва, Россия*

¹ *nik.zubov@gmail.com*

² *AlexeyPoeme@yandex.ru*

³ *ryabchenko.vn@mail.ru*

Аннотация. Получены новые формулы обращения квадратных матриц, разбитых по строкам или столбцам на прямоугольные блоки. Наряду с формулами Фробениуса, где диагональные блоки являются квадратными, новые формулы, построенные с помощью матричных аннуляторов, позволяют упростить обращение блочной матрицы большой размерности, заменяя его обращением двух матриц меньшей размерности. Формулы пригодны для обращения матриц, записанных как в числовом (вещественном или комплексном), так и в аналитическом (символьном) виде. Для определенного широкого класса линейных стационарных динамических систем с использованием новых формул блочного обращения получены компактные аналитические алгоритмы вычисления матриц обратной связи при решении задач управления и оценки компонент вектора состояния. Эти алгоритмы являются упрощением обобщенных формул Басса–Гура и Аккермана в прямом и дуальном вариантах. Приведены примеры обращения матриц, разбитых на прямоугольные блоки: числовой матрицы с комплексными элементами, а также символьной матрицы. Решена задача модального управления пространственным движением самолета с использованием предложенного упрощения обобщенной формулы Аккермана, за счет удобной параметризации, не влияющей на расположение полюсов, выполнено обнуление требуемых компонент матрицы регулятора.

Ключевые слова: блочная матрица; прямоугольный блок; блочное обращение, матричный аннулятор; модальное управление; обобщенная формула Басса–Гура; обобщенная формула Аккермана.

Для цитирования: Зубов Н.Е., Лапин А.В., Рябченко В.Н. Формулы обращения квадратных матриц, разбитых на прямоугольные блоки, и их применение в модальном управлении // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2025. № 71. С. 4–16. doi: 10.17223/19988605/71/1

Original article
doi: 10.17223/19988605/71/1

Formulas for inverting square matrices divided into rectangular blocks and their application in modal control

Nikolay E. Zubov¹, Alexey V. Lapin², Vladimir N. Ryabchenko³

^{1,2,3} *Bauman Moscow State Technical University, Moscow, Russian Federation*

¹ *nik.zubov@gmail.com*

² *AlexeyPoeme@yandex.ru*

³ *ryabchenko.vn@mail.ru*

Abstract. New formulas for inverting square matrices divided into rectangular blocks by rows or columns are obtained. Along with Frobenius formulas, where diagonal blocks are square, the new formulas constructed using matrix zero divisors make it possible to simplify inverting a block matrix of large dimensions by inverting two matrices of

smaller dimensions. The formulas are applicable for inverting matrices written both in numerical (real or complex) and analytical (symbolic) form. For a certain wide class of linear time-invariant dynamic systems, compact analytical algorithms for calculating feedback matrices at solving control problems and evaluating components of the state vector are obtained using new formulas of block inversion. These algorithms are to simplify the generalized formulas of Bass – Gura and Ackermann both in direct and in dual versions. Examples of inverting some matrices divided into rectangular blocks are given for both a numerical matrix with complex-valued elements and a symbolic matrix. The problem of modal control of an aircraft spatial motion is solved using the proposed simplification of the generalized Ackermann formula and zeroing the required components of the controller matrix due to convenient parameterization, which does not affect the poles location.

Keywords: block matrix; rectangular block; block inversion; matrix zero divisor; modal control; generalized Bass–Gura formula; generalized Ackermann formula.

For citation: Zubov, N.E., Lapin, A.V., Ryabchenko, V.N. (2025) Formulas for inverting square matrices divided into rectangular blocks and their application in modal control. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 71. pp. 4–16. doi: 10.17223/19988605/71/1

Введение

Математический аппарат блочного представления квадратных матриц [1] находит широкое применение в прикладных научно-технических задачах для линейных стационарных систем (ЛСС), например в задачах управления [2], наблюдения [3], мониторинга [4] и реконфигурации [5]. Наиболее успешно используется разбиение на квадратные блоки [6]. Для таких блочных матриц используются формулы обращения Фробениуса [7]. Однако возникает потребность в разбиении квадратной матрицы на прямоугольные блоки и нахождении обратной матрицы при таком разбиении [8, 9]. Известные подходы в этом направлении существенно ограничены. Так, например, в [10. С. 148] приведена формула, позволяющая обратить квадратную матрицу $\mathbf{A} \in \mathfrak{R}^{n \times n}$, где $\mathfrak{R} = \mathbb{R}$ (вещественные числа) или $\mathfrak{R} = \mathbb{C}$ (комплексные числа), если эта матрица разбита на прямоугольные блоки следующим образом:

$$\mathbf{A} = [\mathbf{A}_1 \mid \mathbf{A}_2], \quad \mathbf{A}_1 \in \mathfrak{R}^{n \times n_1}, \quad \mathbf{A}_2 \in \mathfrak{R}^{n \times n_2}, \quad n = n_1 + n_2. \quad (1)$$

При условии, что выполняется равенство

$$\mathbf{A}_1^T \mathbf{A}_2 = \mathbf{0}_{n_1 \times n_2}, \quad (2)$$

где $\mathbf{0}_{n_1 \times n_2}$ – нулевая матрица размерности $n_1 \times n_2$, имеет место формула обращения [10]

$$\mathbf{A}^{-1} = [\mathbf{A}_1 \mid \mathbf{A}_2]^{-1} = \begin{bmatrix} (\mathbf{A}_1^* \mathbf{A}_1)^{-1} \mathbf{A}_1^* \\ (\mathbf{A}_2^* \mathbf{A}_2)^{-1} \mathbf{A}_2^* \end{bmatrix}. \quad (3)$$

Здесь верхним индексом $*$ обозначена операция эрмитова сопряжения [11] для случая $\mathfrak{R} = \mathbb{C}$. При $\mathfrak{R} = \mathbb{R}$ формулу (3) можно упростить до следующей:

$$\mathbf{A}^{-1} = [\mathbf{A}_1 \mid \mathbf{A}_2]^{-1} = \begin{bmatrix} (\mathbf{A}_1^T \mathbf{A}_1)^{-1} \mathbf{A}_1^T \\ (\mathbf{A}_2^T \mathbf{A}_2)^{-1} \mathbf{A}_2^T \end{bmatrix}.$$

Формулы обращения матрицы (1), сходные по структуре формуле (3), были получены на основе приведения к каноническому базису сначала в [12], а затем (через два года) в [13]. В настоящей работе представлена универсальная формула обращения матрицы (1), также подобная по структуре формуле (3), но полученная на основе техники аннуляторов (делителей нуля).

1. Теоретический результат

Справедлива следующая теорема.

Теорема 1. Пусть задана квадратная неособенная матрица $\mathbf{A} \in \mathfrak{R}^{n \times n}$, разбитая горизонтально на прямоугольные блоки (1)

$$\mathbf{A} = [\mathbf{A}_1 \mid \mathbf{A}_2], \quad \mathbf{A}_1 \in \mathfrak{R}^{n \times n_1}, \quad \mathbf{A}_2 \in \mathfrak{R}^{n \times n_2}, \quad n = n_1 + n_2,$$

и для каждого из блоков найдены левые аннуляторы максимального ранга (АМР) [14], т.е. любые матрицы $\overline{\mathbf{A}}_1^L$ и $\overline{\mathbf{A}}_2^L$, удовлетворяющие условиям

$$\overline{\mathbf{A}}_1^L \mathbf{A}_1 = \mathbf{0}_{n_2 \times n_1}, \quad \text{rank } \overline{\mathbf{A}}_1^L = n_2; \quad (4)$$

$$\overline{\mathbf{A}}_2^L \mathbf{A}_2 = \mathbf{0}_{n_1 \times n_2}, \quad \text{rank } \overline{\mathbf{A}}_2^L = n_1. \quad (5)$$

Тогда формула для обратной матрицы имеет вид:

$$\mathbf{A}^{-1} = [\mathbf{A}_1 \mid \mathbf{A}_2]^{-1} = \begin{bmatrix} (\overline{\mathbf{A}}_2^L \mathbf{A}_1)^{-1} \overline{\mathbf{A}}_2^L \\ (\overline{\mathbf{A}}_1^L \mathbf{A}_2)^{-1} \overline{\mathbf{A}}_1^L \end{bmatrix}. \quad (6)$$

Доказательство. Докажем обратимость произведений $\overline{\mathbf{A}}_2^L \mathbf{A}_1$ и $\overline{\mathbf{A}}_1^L \mathbf{A}_2$. Предположим, что матрица $\overline{\mathbf{A}}_2^L \mathbf{A}_1$ вырожденная, т.е. у нее найдется линейно зависимая комбинация строк:

$$\boldsymbol{\alpha}^T \overline{\mathbf{A}}_2^L \mathbf{A}_1 = \mathbf{0}_{1 \times n_1},$$

где $\boldsymbol{\alpha}^T \in \mathfrak{R}^{1 \times n_1}$ – ненулевой вектор. Тогда вектор $\boldsymbol{\alpha}^T \overline{\mathbf{A}}_2^L$ окажется левым аннулятором матрицы $\mathbf{A}$, поскольку

$$\boldsymbol{\alpha}^T \overline{\mathbf{A}}_2^L \underbrace{[\mathbf{A}_1 \mid \mathbf{A}_2]}_{\mathbf{A}} = [\mathbf{0}_{1 \times n_1} \mid \mathbf{0}_{1 \times n_2}] = \mathbf{0}_{1 \times n}.$$

Но у квадратной неособенной матрицы $\mathbf{A}$ не может быть аннуляторов. Полученное противоречие доказывает обратимость матрицы $\overline{\mathbf{A}}_2^L \mathbf{A}_1$. Аналогично доказывается обратимость матрицы $\overline{\mathbf{A}}_1^L \mathbf{A}_2$.

Далее для доказательства теоремы в силу невырожденности матрицы $\mathbf{A}$ достаточно показать, что независимо от конкретного выбора левых АМР $\overline{\mathbf{A}}_1^L$ и $\overline{\mathbf{A}}_2^L$ выполняется тождество [14]

$$\mathbf{A}^{-1} \mathbf{A} = \mathbf{I}_n, \quad (7)$$

где $\mathbf{I}_n$ – единичная матрица порядка n .

Общие выражения $\boldsymbol{\Omega}_{L1} \overline{\mathbf{A}}_1^L$ и $\boldsymbol{\Omega}_{L2} \overline{\mathbf{A}}_2^L$ для левых АМР, удовлетворяющих условиям (4) и (5) соответственно, записываются через произвольные квадратные неособенные матрицы $\boldsymbol{\Omega}_{L1}$ и $\boldsymbol{\Omega}_{L2}$ подходящей размерности [14].

Выполним умножение матриц (6) и (1):

$$\begin{aligned} \mathbf{A}^{-1} \mathbf{A} &= \begin{bmatrix} (\boldsymbol{\Omega}_{L2} \overline{\mathbf{A}}_2^L \mathbf{A}_1)^{-1} \boldsymbol{\Omega}_{L2} \overline{\mathbf{A}}_2^L \\ (\boldsymbol{\Omega}_{L1} \overline{\mathbf{A}}_1^L \mathbf{A}_2)^{-1} \boldsymbol{\Omega}_{L1} \overline{\mathbf{A}}_1^L \end{bmatrix} [\mathbf{A}_1 \mid \mathbf{A}_2] = \\ &= \begin{bmatrix} (\overline{\mathbf{A}}_2^L \mathbf{A}_1)^{-1} \overline{\mathbf{A}}_2^L \mathbf{A}_1 & (\overline{\mathbf{A}}_2^L \mathbf{A}_1)^{-1} \overline{\mathbf{A}}_2^L \mathbf{A}_2 \\ (\overline{\mathbf{A}}_1^L \mathbf{A}_2)^{-1} \overline{\mathbf{A}}_1^L \mathbf{A}_1 & (\overline{\mathbf{A}}_1^L \mathbf{A}_2)^{-1} \overline{\mathbf{A}}_1^L \mathbf{A}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{n_1} & \mathbf{0}_{n_1 \times n_2} \\ \mathbf{0}_{n_2 \times n_1} & \mathbf{I}_{n_2} \end{bmatrix} = \mathbf{I}_n. \end{aligned}$$

При формировании нулевых блоков учтены свойства (4) и (5) левых АМР. Таким образом, тождество (7) доказано. *Доказательство завершено.*

В силу симметрии справедливо также другое утверждение.

Теорема 2. Пусть задана квадратная неособенная матрица $\mathbf{B} \in \mathfrak{R}^{n \times n}$, разбитая вертикально на прямоугольные блоки

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{bmatrix}, \quad \mathbf{B}_1 \in \mathfrak{R}^{n_1 \times n}, \quad \mathbf{B}_2 \in \mathfrak{R}^{n_2 \times n}, \quad n = n_1 + n_2,$$

и для каждого из блоков найдены правые АМР [15], т.е. любые матрицы $\overline{\mathbf{B}}_1^R$ и $\overline{\mathbf{B}}_2^R$, удовлетворяющие условиям

$$\begin{aligned}\mathbf{B}_1 \overline{\mathbf{B}}_1^R &= \mathbf{0}_{n_1 \times n_2}, \quad \text{rank } \overline{\mathbf{B}}_1^R = n_2; \\ \mathbf{B}_2 \overline{\mathbf{B}}_2^R &= \mathbf{0}_{n_2 \times n_1}, \quad \text{rank } \overline{\mathbf{B}}_2^R = n_1.\end{aligned}$$

Тогда формула для обратной матрицы имеет вид:

$$\mathbf{B}^{-1} = \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{bmatrix}^{-1} = \begin{bmatrix} \overline{\mathbf{B}}_2^R (\mathbf{B}_1 \overline{\mathbf{B}}_2^R)^{-1} & \vdots & \overline{\mathbf{B}}_1^R (\mathbf{B}_2 \overline{\mathbf{B}}_1^R)^{-1} \end{bmatrix}. \quad (8)$$

2. Применение блочного обращения в модальном управлении

Обращение матриц, состоящих из двух и более прямоугольных блоков, используется в обобщенных формулах Басса–Гура [8] и Аккермана [9]. При этом теорема 1 и формула (6) актуальны для прямого варианта (задачи модального управления) [16], а теорема 2 и формула (8) – для дуального варианта (задачи модального наблюдения) [17].

В качестве применения рассмотрим использование блочного обращения (6) и (8) для модификации (упрощения) обобщенных формул Аккермана [9].

Пусть задана ЛСС с матрицами состояния $\mathbf{A} \in \mathbb{R}^{n \times n}$ и управления $\mathbf{B} \in \mathbb{R}^{n \times m}$, для которой справедливы равенства:

$$n = 2m, \quad \det \underbrace{\begin{bmatrix} \mathbf{B} & \mathbf{AB} \end{bmatrix}}_{\mathbf{U}} \neq 0.$$

Требуется определить матрицу регулятора $\mathbf{K} \in \mathbb{R}^{m \times n}$, обеспечивающую матрице замкнутой ЛСС «объект–регулятор» $\mathbf{A} - \mathbf{BK}$ желаемый характеристический полином $p_B^*(\lambda)$.

Обобщенная формула Аккермана [9] для расчета матрицы регулятора в рассматриваемой задаче имеет вид:

$$\mathbf{K} = \begin{bmatrix} \mathbf{P}_{B2} & \mathbf{P}_{B1} & \mathbf{I}_m \end{bmatrix} \begin{bmatrix} \mathbf{C}_B \\ \mathbf{C}_B \mathbf{A} \\ \mathbf{C}_B \mathbf{A}^2 \end{bmatrix}, \quad (9)$$

где блоки $\mathbf{P}_{B1} \in \mathbb{R}^{m \times m}$ и $\mathbf{P}_{B2} \in \mathbb{R}^{m \times m}$ таковы, что

$$\det(\lambda^2 \mathbf{I}_m + \lambda \mathbf{P}_{B1} + \mathbf{P}_{B2}) = p_B^*(\lambda). \quad (10)$$

В формуле (9) используется только нижняя блочная строка

$$\mathbf{C}_B = \begin{bmatrix} \mathbf{0}_{m \times m} & \mathbf{I}_m \end{bmatrix} \mathbf{U}^{-1} = (\overline{\mathbf{B}}^L \mathbf{AB})^{-1} \overline{\mathbf{B}}^L \quad (11)$$

обратной матрицы управляемости $\mathbf{U}^{-1}$, записанной согласно формуле (6):

$$\mathbf{U}^{-1} = \begin{bmatrix} (\overline{\mathbf{AB}}^L \mathbf{B})^{-1} \overline{\mathbf{AB}}^L \\ (\overline{\mathbf{B}}^L \mathbf{AB})^{-1} \overline{\mathbf{B}}^L \end{bmatrix}.$$

Поэтому нет необходимости полностью рассчитывать обратную матрицу $\mathbf{U}^{-1}$ порядка n . Достаточно обратить матрицу вдвое меньшего порядка m :

$$\mathbf{B}_1 = \overline{\mathbf{B}}^L \mathbf{AB}. \quad (12)$$

Пусть, кроме того, используется преобразование подобия [18]:

$$\mathbf{P}_{B1} = \mathbf{B}_1^{-1} \mathbf{P}_{B1}^* \mathbf{B}_1, \quad \mathbf{P}_{B2} = \mathbf{B}_1^{-1} \mathbf{P}_{B2}^* \mathbf{B}_1.$$

Тогда формула (9) для расчета регулятора после подстановки матрицы (11) с учетом обозначения (12) упрощается и принимает вид:

$$\mathbf{K} = \mathbf{B}_1^{-1} \begin{bmatrix} \mathbf{P}_{B2}^* & \vdots & \mathbf{P}_{B1}^* & \vdots & \mathbf{I}_m \end{bmatrix} \begin{bmatrix} \overline{\mathbf{B}}^L \\ \overline{\mathbf{B}}^L \mathbf{A} \\ \overline{\mathbf{B}}^L \mathbf{A}^2 \end{bmatrix}, \quad (13)$$

а в силу тождества (10) остается справедливым равенство полиномов

$$\det(\lambda^2 \mathbf{I}_m + \lambda \mathbf{P}_{B1}^* + \mathbf{P}_{B2}^*) = p_B^*(\lambda). \quad (14)$$

Модификация обобщенной дуальной формулы Аккермана. Пусть задана ЛСС с матрицами состояния $\mathbf{A} \in \mathbb{R}^{n \times n}$ и наблюдения $\mathbf{C} \in \mathbb{R}^{l \times n}$, для которой справедливы равенства

$$n = 2l, \quad \det \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \end{bmatrix} \neq 0.$$

Требуется определить матрицу наблюдателя $\mathbf{L} \in \mathbb{R}^{n \times l}$, обеспечивающую матрице замкнутой ЛСС «объект–наблюдатель» $\mathbf{A} - \mathbf{LC}$ желаемый характеристический полином $p_C^*(\lambda)$.

Обобщенная дуальная формула Аккермана [9] для расчета матрицы наблюдателя в рассматриваемой задаче имеет вид:

$$\mathbf{L} = \begin{bmatrix} \mathbf{B}_C & \vdots & \mathbf{AB}_C & \vdots & \mathbf{A}^2 \mathbf{B}_C \end{bmatrix} \begin{bmatrix} \mathbf{P}_{C2} \\ \mathbf{P}_{C1} \\ \mathbf{I}_l \end{bmatrix}, \quad (15)$$

где блоки $\mathbf{P}_{C1} \in \mathbb{R}^{l \times l}$ и $\mathbf{P}_{C2} \in \mathbb{R}^{l \times l}$ таковы, что

$$\det(\lambda^2 \mathbf{I}_l + \lambda \mathbf{P}_{C1} + \mathbf{P}_{C2}) = p_C^*(\lambda). \quad (16)$$

В формуле (15) используется *только правый блочный столбец*

$$\mathbf{B}_C = \mathbf{N}^{-1} \begin{bmatrix} \mathbf{0}_{l \times l} \\ \mathbf{I}_l \end{bmatrix} = \overline{\mathbf{C}}^R (\mathbf{CA} \overline{\mathbf{C}}^R)^{-1} \quad (17)$$

обратной матрицы наблюдаемости $\mathbf{N}^{-1}$, записанной согласно формуле (8):

$$\mathbf{N}^{-1} = \begin{bmatrix} \overline{\mathbf{C}}^R (\mathbf{CA} \overline{\mathbf{C}}^R)^{-1} & \vdots & \overline{\mathbf{C}}^R (\mathbf{CA} \overline{\mathbf{C}}^R)^{-1} \end{bmatrix}.$$

Поэтому нет необходимости полностью рассчитывать обратную матрицу $\mathbf{N}^{-1}$ порядка n . Достаточно обратить матрицу вдвое меньшего порядка l :

$$\mathbf{C}_1 = \mathbf{CA} \overline{\mathbf{C}}^R. \quad (18)$$

Пусть, кроме того, используется преобразование подобия [12]

$$\mathbf{P}_{C1} = \mathbf{C}_1 \mathbf{P}_{C1}^* \mathbf{C}_1^{-1}, \quad \mathbf{P}_{C2} = \mathbf{C}_1 \mathbf{P}_{C2}^* \mathbf{C}_1^{-1}.$$

Тогда формула (15) для расчета наблюдателя после подстановки матрицы (17) с учетом обозначения (18) упрощается и принимает вид:

$$\mathbf{L} = \begin{bmatrix} \overline{\mathbf{C}}^R & \vdots & \mathbf{A} \overline{\mathbf{C}}^R & \vdots & \mathbf{A}^2 \overline{\mathbf{C}}^R \end{bmatrix} \begin{bmatrix} \mathbf{P}_{C2}^* \\ \mathbf{P}_{C1}^* \\ \mathbf{I}_l \end{bmatrix} \mathbf{C}_1^{-1},$$

а в силу тождества (16) остается справедливым равенство полиномов

$$\det(\lambda^2 \mathbf{I}_l + \lambda \mathbf{P}_{C1}^* + \mathbf{P}_{C2}^*) = p_C^*(\lambda).$$

3. Числовой пример блочного обращения.

Требуется обратить числовую комплексную матрицу

$$\mathbf{A} = [\mathbf{A}_1 \mid \mathbf{A}_2] = \begin{bmatrix} i & 2 & -3 & 4i \\ 1 & 2 & 3 & 1 \\ 1+i & 2 & i-3 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \quad i^2 = -1. \quad (19)$$

Выполним обращение матрицы (19) по формуле (6). Для этого первоначально вычислим левые АМР $\overline{\mathbf{A}}_1^L$ и $\overline{\mathbf{A}}_2^L$, воспользовавшись, например, функцией null [19] программы MATLAB. Получим

$$\overline{\mathbf{A}}_1^L = \begin{bmatrix} -46,15 + 19,23i & -19,23 - 46,15i & 65,38 + 26,92i & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\overline{\mathbf{A}}_2^L = \begin{bmatrix} -15,19 + 32,78i & 51,83 + 33,02i & 64,29 + 23,19i & 15,01 + 4,55i \\ 1,1 + 14,68i & -33,42 + 3,85i & -0,29 - 10,26i & 92,43 + 2i \end{bmatrix}.$$

Подставив полученные АМР в формулу (6), окончательно будем иметь

$$\mathbf{A}^{-1} = \begin{bmatrix} (\overline{\mathbf{A}}_2^L \mathbf{A}_1)^{-1} \overline{\mathbf{A}}_2^L \\ (\overline{\mathbf{A}}_1^L \mathbf{A}_2)^{-1} \overline{\mathbf{A}}_1^L \end{bmatrix} = \begin{bmatrix} -1,6 - 2,2i & 2,2 - 0,6i & -0,6 + 2,8i & -10,4 + 4,2i \\ 0,4 + 1,3i & -0,8 - 0,1i & 0,9 - 1,2i & 5,1 - 0,3i \\ 0,4 - 0,2i & 0,2 + 0,4i & -0,6 - 0,2i & -0,4 - 1,8i \\ -0,4 + 0,2i & -0,2 - 0,4i & 0,6 + 0,2i & 1,4 + 1,8i \end{bmatrix}. \quad (20)$$

С помощью непосредственных вычислений можно убедиться, что матрица (20) действительно является обратной к матрице (19).

4. Аналитический пример блочного обращения

Требуется обратить символьную матрицу

$$\mathbf{A} = [\mathbf{A}_1 \mid \mathbf{A}_2] = \begin{bmatrix} 0 & 0 & 0 & 1 & a_{15} & 0 \\ 1 & 0 & 0 & a_{24} & a_{25} & a_{26} \\ 0 & 0 & 0 & 0 & 1 & a_{36} \\ 0 & 0 & 0 & 0 & a_{45} & a_{46} \\ 0 & 1 & 0 & a_{54} & a_{55} & a_{56} \\ 0 & 0 & 1 & a_{64} & a_{65} & a_{66} \end{bmatrix}. \quad (21)$$

Введем обозначения

$$\begin{bmatrix} \mathbf{a}_{14}^T \\ \mathbf{a}_{24}^T \end{bmatrix} = \begin{bmatrix} a_{15} & 0 \\ a_{25} & a_{26} \end{bmatrix}, \quad \mathbf{A}_{34} = \begin{bmatrix} 1 & a_{36} \\ a_{45} & a_{46} \end{bmatrix}, \quad \mathbf{a}_{43} = \begin{bmatrix} a_{54} \\ a_{64} \end{bmatrix}, \quad \mathbf{A}_{44} = \begin{bmatrix} a_{55} & a_{56} \\ a_{65} & a_{66} \end{bmatrix} \quad (22)$$

и запишем матрицу (21) в блочном виде:

$$\mathbf{A} = [\mathbf{A}_1 \mid \mathbf{A}_2], \quad (23)$$

где

$$\mathbf{A}_1 = \begin{bmatrix} 0 & \mathbf{0}_{1 \times 2} \\ 1 & \mathbf{0}_{1 \times 2} \\ \mathbf{0}_{2 \times 1} & \mathbf{0}_{2 \times 2} \\ \mathbf{0}_{2 \times 1} & \mathbf{I}_2 \end{bmatrix}, \quad \mathbf{A}_2 = \begin{bmatrix} 1 & \mathbf{a}_{14}^T \\ a_{24} & \mathbf{a}_{24}^T \\ \mathbf{0}_{2 \times 1} & \mathbf{A}_{34} \\ \mathbf{a}_{43} & \mathbf{A}_{44} \end{bmatrix}. \quad (24)$$

Определитель матрицы (23) равен $\det \mathbf{A} = -\det \mathbf{A}_{34}$.

При выполнении условия

$$\det \mathbf{A}_{34} \neq 0 \quad (25)$$

блочная матрица (23) обратима, а блоки $\mathbf{A}_1$ и $\mathbf{A}_2$ имеют полный ранг 3.

В силу условия (25) квадратные блочные матрицы

$$\bar{\mathbf{A}}_1 = \begin{bmatrix} 1 & \mathbf{0}_{1 \times 2} \\ \mathbf{0}_{2 \times 1} & \mathbf{I}_2 \end{bmatrix}, \quad \bar{\mathbf{A}}_2 = \begin{bmatrix} 1 & \mathbf{a}_{14}^T \\ \mathbf{0}_{2 \times 1} & \mathbf{A}_{34} \end{bmatrix} \quad (26)$$

являются неособенными. Поэтому левые АМР для матриц (24) соответственно могут быть записаны в виде:

$$\bar{\mathbf{A}}_1^{-L} = \begin{bmatrix} 1 & 0 & \mathbf{0}_{1 \times 2} & \mathbf{0}_{1 \times 2} \\ \mathbf{0}_{2 \times 1} & \mathbf{0}_{2 \times 1} & \mathbf{I}_2 & \mathbf{0}_{2 \times 2} \end{bmatrix}, \quad \bar{\mathbf{A}}_2^{-L} = \begin{bmatrix} -a_{22} & 1 & (a_{24}\mathbf{a}_{14}^T - \mathbf{a}_{24}^T)\mathbf{A}_{34}^{-1} & \mathbf{0}_{1 \times 2} \\ -\mathbf{a}_{43} & 0 & (\mathbf{a}_{43}\mathbf{a}_{14}^T - \mathbf{A}_{44})\mathbf{A}_{34}^{-1} & \mathbf{I}_2 \end{bmatrix}. \quad (27)$$

Для обращения блочной матрицы (23) порядка 6 по формуле (6) достаточно обратить матрицы (26) меньшего порядка 3:

$$\bar{\mathbf{A}}_2^{-L} \mathbf{A}_1 = \bar{\mathbf{A}}_1, \quad \bar{\mathbf{A}}_1^{-L} \mathbf{A}_2 = \bar{\mathbf{A}}_2.$$

Обратим эти блочно-треугольные матрицы по формулам Фробениуса [7]:

$$\bar{\mathbf{A}}_1^{-1} = \begin{bmatrix} 1 & \mathbf{0}_{1 \times 2} \\ \mathbf{0}_{2 \times 1} & \mathbf{I}_2 \end{bmatrix}, \quad \bar{\mathbf{A}}_2^{-1} = \begin{bmatrix} 1 & -\mathbf{a}_{14}^T \mathbf{A}_{34}^{-1} \\ \mathbf{0}_{2 \times 1} & \mathbf{A}_{34}^{-1} \end{bmatrix}. \quad (28)$$

Используя результаты (27), (28) и формулу (6), запишем искомую обратную матрицу в блочном виде:

$$\mathbf{A}^{-1} = \begin{bmatrix} \bar{\mathbf{A}}_1^{-1} \bar{\mathbf{A}}_2^{-L} \\ \bar{\mathbf{A}}_2^{-1} \bar{\mathbf{A}}_1^{-L} \end{bmatrix} = \begin{bmatrix} -a_{24} & 1 & (a_{24}\mathbf{a}_{14}^T - \mathbf{a}_{24}^T)\mathbf{A}_{34}^{-1} & \mathbf{0}_{1 \times 2} \\ -\mathbf{a}_{43} & \mathbf{0}_{2 \times 1} & (\mathbf{a}_{43}\mathbf{a}_{14}^T - \mathbf{A}_{44})\mathbf{A}_{34}^{-1} & \mathbf{I}_2 \\ 1 & 0 & -\mathbf{a}_{14}^T \mathbf{A}_{34}^{-1} & \mathbf{0}_{1 \times 2} \\ \mathbf{0}_{2 \times 1} & \mathbf{0}_{2 \times 1} & \mathbf{A}_{34}^{-1} & \mathbf{0}_{2 \times 2} \end{bmatrix}. \quad (29)$$

Подставив в матрицу (29) блоки (22), получим результат в скалярном виде:

$$\mathbf{A}^{-1} = \begin{bmatrix} -a_{24} & 1 & -\bar{a}_{25}/\tilde{a}_{46} & -\bar{a}_{26}/\tilde{a}_{46} & 0 & 0 \\ -a_{54} & 0 & -\bar{a}_{55}/\tilde{a}_{46} & -\bar{a}_{56}/\tilde{a}_{46} & 1 & 0 \\ -a_{64} & 0 & -\bar{a}_{65}/\tilde{a}_{46} & -\bar{a}_{66}/\tilde{a}_{46} & 0 & 1 \\ 1 & 0 & -a_{15}a_{46}/\tilde{a}_{46} & a_{15}a_{36}/\tilde{a}_{46} & 0 & 0 \\ 0 & 0 & a_{46}/\tilde{a}_{46} & -a_{36}/\tilde{a}_{46} & 0 & 0 \\ 0 & 0 & -a_{45}/\tilde{a}_{46} & 1/\tilde{a}_{46} & 0 & 0 \end{bmatrix}. \quad (30)$$

Здесь $\tilde{a}_{46} = a_{46} - a_{36}a_{45}$, а для номеров $i = 2, 5, 6$ введены обозначения

$$\tilde{a}_{i5} = a_{i5} - a_{15}a_{i4}, \quad \bar{a}_{i5} = \tilde{a}_{i5}a_{46} - a_{i6}a_{45}, \quad \bar{a}_{i6} = a_{i6} - a_{36}\tilde{a}_{i5}.$$

Выполним проверку прямого произведения матриц (21) и (30):

$$\mathbf{A}\mathbf{A}^{-1} = \begin{bmatrix} 0 & 0 & 0 & 1 & a_{15} & 0 \\ 1 & 0 & 0 & a_{24} & a_{25} & a_{26} \\ 0 & 0 & 0 & 0 & 1 & a_{36} \\ 0 & 0 & 0 & 0 & a_{45} & a_{46} \\ 0 & 1 & 0 & a_{54} & a_{55} & a_{56} \\ 0 & 0 & 1 & a_{64} & a_{65} & a_{66} \end{bmatrix} \begin{bmatrix} -a_{24} & 1 & -\bar{a}_{25}/\tilde{a}_{46} & -\bar{a}_{26}/\tilde{a}_{46} & 0 & 0 \\ -a_{54} & 0 & -\bar{a}_{55}/\tilde{a}_{46} & -\bar{a}_{56}/\tilde{a}_{46} & 1 & 0 \\ -a_{64} & 0 & -\bar{a}_{65}/\tilde{a}_{46} & -\bar{a}_{66}/\tilde{a}_{46} & 0 & 1 \\ 1 & 0 & -a_{15}a_{46}/\tilde{a}_{46} & a_{15}a_{36}/\tilde{a}_{46} & 0 & 0 \\ 0 & 0 & a_{46}/\tilde{a}_{46} & -a_{36}/\tilde{a}_{46} & 0 & 0 \\ 0 & 0 & -a_{45}/\tilde{a}_{46} & 1/\tilde{a}_{46} & 0 & 0 \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & (\tilde{a}_{25}a_{46} - a_{26}a_{45} - \bar{a}_{25})/\tilde{a}_{46} & (a_{26} - a_{36}\tilde{a}_{25} - \bar{a}_{26})/\tilde{a}_{46} & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & (\tilde{a}_{55}a_{46} - a_{56}a_{45} - \bar{a}_{55})/\tilde{a}_{46} & (a_{56} - a_{36}\tilde{a}_{55} - \bar{a}_{56})/\tilde{a}_{46} & 1 & 0 \\ 0 & 0 & (\tilde{a}_{65}a_{46} - a_{66}a_{45} - \bar{a}_{65})/\tilde{a}_{46} & (a_{66} - a_{36}\tilde{a}_{65} - \bar{a}_{66})/\tilde{a}_{46} & 0 & 1 \end{bmatrix} = \mathbf{I}_6.$$

Выполним проверку обратного произведения матриц (30) и (21):

$$\mathbf{A}^{-1}\mathbf{A} = \begin{bmatrix} -a_{24} & 1 & -\bar{a}_{25}/\tilde{a}_{46} & -\bar{a}_{26}/\tilde{a}_{46} & 0 & 0 \\ -a_{54} & 0 & -\bar{a}_{55}/\tilde{a}_{46} & -\bar{a}_{56}/\tilde{a}_{46} & 1 & 0 \\ -a_{64} & 0 & -\bar{a}_{65}/\tilde{a}_{46} & -\bar{a}_{66}/\tilde{a}_{46} & 0 & 1 \\ 1 & 0 & -a_{15}a_{46}/\tilde{a}_{46} & a_{15}a_{36}/\tilde{a}_{46} & 0 & 0 \\ 0 & 0 & a_{46}/\tilde{a}_{46} & -a_{36}/\tilde{a}_{46} & 0 & 0 \\ 0 & 0 & -a_{45}/\tilde{a}_{46} & 1/\tilde{a}_{46} & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 1 & a_{15} & 0 \\ 1 & 0 & 0 & a_{24} & a_{25} & a_{26} \\ 0 & 0 & 0 & 0 & 1 & a_{36} \\ 0 & 0 & 0 & 0 & a_{45} & a_{46} \\ 0 & 1 & 0 & a_{54} & a_{55} & a_{56} \\ 0 & 0 & 1 & a_{64} & a_{65} & a_{66} \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 & \tilde{a}_{25} - (\bar{a}_{25} + \bar{a}_{26}a_{45})/\tilde{a}_{46} & a_{26} - (a_{36}\bar{a}_{25} + a_{46}\bar{a}_{26})/\tilde{a}_{46} \\ 0 & 1 & 0 & 0 & \tilde{a}_{55} - (\bar{a}_{55} + \bar{a}_{56}a_{45})/\tilde{a}_{46} & a_{56} - (a_{36}\bar{a}_{55} + a_{46}\bar{a}_{56})/\tilde{a}_{46} \\ 0 & 0 & 1 & 0 & \tilde{a}_{65} - (\bar{a}_{65} + \bar{a}_{66}a_{45})/\tilde{a}_{46} & a_{66} - (a_{36}\bar{a}_{65} + a_{46}\bar{a}_{66})/\tilde{a}_{46} \\ 0 & 0 & 0 & 1 & a_{15} - a_{15}(a_{46} - a_{36}a_{45})/\tilde{a}_{46} & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} = \mathbf{I}_6.$$

Проверка подтверждает правильность определения обратной матрицы.

5. Пример аналитического синтеза регулятора для ЛСС

Рассмотрим непрерывную ($t \in \mathbb{R}$) ЛСС шестого порядка $\mathbf{x}(t) \in \mathbb{R}^6$ с тремя управляющими входами $\mathbf{u}(t) \in \mathbb{R}^3$:

$$\underbrace{\begin{bmatrix} \dot{\gamma}(t) \\ \dot{\beta}(t) \\ \dot{\alpha}(t) \\ \dot{\omega}_x(t) \\ \dot{\omega}_y(t) \\ \dot{\omega}_z(t) \end{bmatrix}}_{\dot{\mathbf{x}}(t)} = \underbrace{\begin{bmatrix} 0 & 0 & 0 & 1 & a_{15} & 0 \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & 0 \\ 0 & a_{32} & a_{33} & a_{34} & 0 & 1 \\ 0 & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ 0 & a_{52} & 0 & a_{54} & a_{55} & a_{56} \\ 0 & 0 & a_{63} & a_{64} & a_{65} & a_{66} \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} \gamma(t) \\ \beta(t) \\ \alpha(t) \\ \omega_x(t) \\ \omega_y(t) \\ \omega_z(t) \end{bmatrix}}_{\mathbf{x}(t)} + \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ b_{41} & 0 & 0 \\ 0 & b_{52} & b_{53} \\ 0 & b_{62} & b_{63} \end{bmatrix}}_{\mathbf{B}} \underbrace{\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}}_{\mathbf{u}(t)}, \quad (31)$$

которая представляет собой линеаризованную математическую модель пространственного движения летательного аппарата (ЛА) самолетного типа [20]. Здесь α , β и γ – соответственно углы атаки, скольжения и крена; ω_x , ω_y и ω_z – компоненты вектора угловой скорости соответственно по крену, рысканью и тангажу; u_1 , u_2 и u_3 – управляющие воздействия; $\mathbf{A}$ и $\mathbf{B}$ – матрицы состояния и управления.

Требуется построить регулятор с матрицей обратной связи $\mathbf{K}$, обеспечивающий матрице замкнутой ЛСС

$$\mathbf{A}_B^* = \mathbf{A} - \mathbf{BK}$$

желаемый спектр

$$\Lambda_B^* = \{\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6\}.$$

Запишем матрицы $\mathbf{A}$ и $\mathbf{B}$ в блочном виде, разделив углы и угловые скорости:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathbf{0}_{3 \times 3} \\ \mathbf{B}_{21} \end{bmatrix}, \quad (32)$$

где

$$\mathbf{A}_{11} = \begin{bmatrix} 0 & 0 & 0 \\ a_{21} & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{bmatrix}, \quad \mathbf{A}_{12} = \begin{bmatrix} 1 & a_{15} & 0 \\ a_{24} & a_{25} & 0 \\ a_{34} & 0 & 1 \end{bmatrix},$$

$$\mathbf{A}_{21} = \begin{bmatrix} 0 & a_{42} & a_{43} \\ 0 & a_{52} & 0 \\ 0 & 0 & a_{63} \end{bmatrix}, \quad \mathbf{A}_{22} = \begin{bmatrix} a_{44} & a_{45} & a_{46} \\ a_{54} & a_{55} & a_{56} \\ a_{64} & a_{65} & a_{66} \end{bmatrix}, \quad \mathbf{B}_{21} = \begin{bmatrix} b_{41} & 0 & 0 \\ 0 & b_{52} & b_{53} \\ 0 & b_{62} & b_{63} \end{bmatrix}.$$

Первые два блочных столбца матрицы управляемости Калмана образуют квадратную матрицу

$$\mathbf{U} = [\mathbf{B} \mid \mathbf{AB}] = \begin{bmatrix} \mathbf{0}_{3 \times 3} & \mathbf{A}_{12} \\ \mathbf{I}_3 & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{B}_{21} & \mathbf{0}_{3 \times 3} \\ \mathbf{0}_{3 \times 3} & \mathbf{B}_{21} \end{bmatrix},$$

определитель которой равен

$$\det \mathbf{U} = -\det \mathbf{A}_{12} \det^2 \mathbf{B}_{21}.$$

Таким образом, для полной управляемости ЛСС (31) достаточно одновременного выполнения следующих условий:

$$a_{25} \neq a_{15}a_{24}, \quad b_{41} \neq 0, \quad b_{52}b_{63} \neq b_{53}b_{62}. \quad (33)$$

Определив левый АМР матрицы $\mathbf{B}$

$$\bar{\mathbf{B}}^L = [\mathbf{I}_3 \mid \mathbf{0}_{3 \times 3}],$$

а также матрицу

$$\mathbf{B}_1 = \bar{\mathbf{B}}^L \mathbf{AB} = \mathbf{A}_{12} \mathbf{B}_{21},$$

рассчитаем регулятор по упрощенной обобщенной формуле Аккермана (13):

$$\mathbf{K} = \mathbf{B}_1^{-1} \left(\mathbf{P}_{B2}^* \bar{\mathbf{B}}^L + \mathbf{P}_{B1}^* \bar{\mathbf{B}}^L \mathbf{A} + \bar{\mathbf{B}}^L \mathbf{A}^2 \right) = \mathbf{B}_{21}^{-1} \left[\mathbf{A}_{21} - \tilde{\mathbf{A}}_{21} \mid \mathbf{A}_{22} - \tilde{\mathbf{A}}_{22} \right], \quad (34)$$

где

$$\tilde{\mathbf{A}}_{21} = -\mathbf{A}_{12}^{-1} (\mathbf{A}_{11}^2 + \mathbf{P}_{B1}^* \mathbf{A}_{11} + \mathbf{P}_{B2}^*), \quad \tilde{\mathbf{A}}_{22} = -\mathbf{A}_{12}^{-1} (\mathbf{A}_{11} + \mathbf{P}_{B1}^*) \mathbf{A}_{12},$$

а блоки $\mathbf{P}_{B1}^* \in \mathbb{R}^{3 \times 3}$ и $\mathbf{P}_{B2}^* \in \mathbb{R}^{3 \times 3}$ согласно соотношению (14) таковы, что

$$\det (\lambda^2 \mathbf{I}_3 + \lambda \mathbf{P}_{B1}^* + \mathbf{P}_{B2}^*) = \prod_{i=1}^6 (\lambda - \lambda_i). \quad (35)$$

Конкретный вид блоков $\mathbf{P}_{B1}^*$ и $\mathbf{P}_{B2}^*$ выбирается не только из условия (35) размещения полюсов, но и для обеспечения специальных свойств матрицы регулятора $\mathbf{K}$ (уменьшение нормы [21], обнуление столбцов [22], разделение каналов управления [14] и др.). Например, для упрощения выражения (34) и попытки разделения бокового (крен–рысканье) и продольного (тангаж) движений ЛА блоки $\mathbf{P}_{B1}^*$ и $\mathbf{P}_{B2}^*$ можно искать в удовлетворяющем условию (35) виде:

$$\mathbf{P}_{B1}^* = \begin{bmatrix} p_x & 0 & 0 \\ p_3 & p_y & p_1 \\ p_2 & 0 & p_z \end{bmatrix}, \quad \mathbf{P}_{B2}^* = \begin{bmatrix} q_x & 0 & 0 \\ q_3 & q_y & q_1 \\ q_2 & 0 & q_z \end{bmatrix}, \quad (36)$$

где p_i и q_i ($i = \overline{1, 3}$) – произвольные параметры, а

$$p_x = -\lambda_1 - \lambda_4, \quad p_y = -\lambda_2 - \lambda_5, \quad p_z = -\lambda_3 - \lambda_6, \\ q_x = \lambda_1 \lambda_4, \quad q_y = \lambda_2 \lambda_5, \quad q_z = \lambda_3 \lambda_6$$

коэффициенты Виета, такие что

$$\lambda_{14} = \lambda^2 + p_x \lambda + q_x = (\lambda - \lambda_1)(\lambda - \lambda_4), \\ \lambda_{25} = \lambda^2 + p_y \lambda + q_y = (\lambda - \lambda_2)(\lambda - \lambda_5), \\ \lambda_{36} = \lambda^2 + p_z \lambda + q_z = (\lambda - \lambda_3)(\lambda - \lambda_6).$$

Отметим, что получить блоки $\mathbf{P}_{B1}^*$ и $\mathbf{P}_{B2}^*$ в виде (36) с помощью классического декомпозиционного подхода [23] затруднительно или невозможно ввиду требования о разложимости полинома $\lambda^2 \mathbf{I}_3 + \lambda \mathbf{P}_{B1}^* + \mathbf{P}_{B2}^*$ на матричные множители $(\lambda \mathbf{I}_3 - \Phi_{B0})(\lambda \mathbf{I}_3 - \Phi_{B1})$, где $\text{eig } \Phi_{B0} \cup \text{eig } \Phi_{B1} = \Lambda_B^*$.

Поставим задачу в блоках $\tilde{\mathbf{A}}_{21}$ и $\tilde{\mathbf{A}}_{22}$ регулятора (34) обнулить позиции, соответствующие расположению параметров p_i и q_i в матрицах $\mathbf{P}_{B1}^*$ и $\mathbf{P}_{B2}^*$:

$$\begin{aligned}\tilde{\mathbf{A}}_{22} &= -\mathbf{A}_{12}^{-1}(\mathbf{A}_{11} + \mathbf{P}_{B1}^*)\mathbf{A}_{12} = -\begin{bmatrix} p_x & -a_{15}(\tilde{p}_y - p_x) & 0 \\ 0 & \tilde{p}_y & 0 \\ 0 & \tilde{a}_{65} & \tilde{p}_z \end{bmatrix}, \\ \tilde{\mathbf{A}}_{21} &= -\mathbf{A}_{12}^{-1}(\mathbf{A}_{11}^2 + \mathbf{A}_{11}\mathbf{P}_{B1}^* + \mathbf{P}_{B2}^*) = -\begin{bmatrix} q_x & -a_{15}\tilde{q}_y & 0 \\ 0 & \tilde{q}_y & 0 \\ 0 & \tilde{a}_{62} & \tilde{q}_z \end{bmatrix},\end{aligned}\quad (37)$$

где $\tilde{a}_{25} = a_{25} - a_{15}a_{24}$ и

$$\begin{aligned}\tilde{p}_y &= p_y + a_{22}, & \tilde{q}_y &= q_y + a_{22}\tilde{p}_y, & \tilde{\tilde{q}}_y &= \tilde{q}_y / \tilde{a}_{25}, & \tilde{a}_{62} &= a_{32}(\tilde{p}_z + a_{22}) + a_{15}a_{34}\tilde{\tilde{q}}_y, \\ \tilde{p}_z &= p_z + a_{33}, & \tilde{q}_z &= q_z + a_{33}\tilde{p}_z, & \tilde{\tilde{q}}_z &= \tilde{q}_z + a_{23}a_{32}, & \tilde{a}_{65} &= a_{32}\tilde{a}_{25} + a_{15}a_{34}(\tilde{p}_y - \tilde{p}_z).\end{aligned}$$

Обнуление достигается при следующих значениях параметров:

$$\begin{aligned}p_1 &= -a_{23}, & p_2 &= a_{34}(p_x - \tilde{p}_z) - a_{24}a_{32}, & p_3 &= a_{24}(p_x - \tilde{p}_y) - a_{21}, \\ q_1 &= -a_{23}\tilde{p}_y, & q_2 &= a_{34}q_x - a_{21}a_{32}, & q_3 &= a_{24}q_x - a_{21}\tilde{p}_y.\end{aligned}$$

Выполним проверку характеристического полинома матрицы замкнутой ЛСС (32) с регулятором (34) и подстановкой (37):

$$\begin{aligned}\Delta &= \det(\underbrace{\lambda\mathbf{I}_6 - \mathbf{A}_B^*}_{\mathbf{A}_{B\lambda}^*}) = \det\left(\lambda\mathbf{I}_6 - \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \tilde{\mathbf{A}}_{21} & \tilde{\mathbf{A}}_{22} \end{bmatrix}\right) = \\ &= \det\begin{bmatrix} \lambda & 0 & 0 & -1 & -a_{15} & 0 \\ -a_{21} & \lambda - a_{22} & -a_{23} & -a_{24} & -a_{25} & 0 \\ 0 & -a_{32} & \lambda - a_{33} & -a_{34} & 0 & -1 \\ q_x & -a_{15}\tilde{q}_y & 0 & \lambda + p_x & -a_{15}(\tilde{p}_y - p_x) & 0 \\ 0 & \tilde{q}_y & 0 & 0 & \lambda + \tilde{p}_y & 0 \\ 0 & \tilde{a}_{62} & \tilde{q}_z & 0 & \tilde{a}_{65} & \lambda + \tilde{p}_z \end{bmatrix}.\end{aligned}$$

Рассчитаем представленный определитель по пятой строке:

$$\Delta = (\lambda + \tilde{p}_y) \underbrace{(\lambda\Delta_1 + \Delta_2)}_{\Delta_{12}} - \tilde{q}_y \underbrace{((\lambda + \tilde{p}_z)\Delta_3 - \Delta_4)}_{\Delta_{34}}, \quad (38)$$

где

$$\begin{aligned}\Delta_1 &= \det \mathbf{A}_{B\lambda}^*([2,3,4,6],[2,3,4,6]), & \Delta_2 &= \det \mathbf{A}_{B\lambda}^*([2,3,4,6],[1,2,3,6]), \\ \Delta_3 &= \det \mathbf{A}_{B\lambda}^*([1,2,3,4],[1,3,4,5]), & \Delta_4 &= \det \mathbf{A}_{B\lambda}^*([1,2,4,6],[1,3,4,5]),\end{aligned}\quad (39)$$

а запись вида $\mathbf{A}_{B\lambda}^*(rows,cols)$ обозначает матрицу, составленную из элементов матрицы $\mathbf{A}_{B\lambda}^*$, стоящих на пересечении строк из множества $rows$ и столбцов из множества $cols$. Далее рассчитаем отдельно каждый из определителей (39) и подставим в формулу (38):

$$\begin{aligned}\Delta_1 &= (\lambda + p_x)(\lambda - a_{22})\lambda_{36} + a_{15}\tilde{q}_y(a_{23}a_{34}(p_x - \tilde{p}_z) - a_{24}(\lambda_{36} + a_{23}a_{32})), \\ \Delta_2 &= q_x(\lambda - a_{22})\lambda_{36} + a_{15}\tilde{q}_y(a_{23}a_{34}q_x - a_{21}(\lambda_{36} + a_{23}a_{32})), \\ \Delta_3 &= \lambda_{14}(a_{15}a_{23}a_{34} - \tilde{a}_{25}(\lambda - a_{33})) - a_{15}(\lambda + \tilde{p}_y)(a_{23}a_{34}\lambda + (a_{24}\lambda + a_{21})(\lambda - a_{33})), \\ \Delta_4 &= \lambda_{14}(\tilde{q}_z\tilde{a}_{25} - a_{15}a_{23}a_{34}(\tilde{p}_y - \tilde{p}_z)) + a_{15}(a_{24}\lambda + a_{21})(\lambda + \tilde{p}_y)\tilde{q}_z,\end{aligned}$$

$$\Delta_{12} = \lambda_{14}(\lambda - a_{22})\lambda_{36} + a_{15}\tilde{q}_y(a_{23}a_{34}(\lambda(p_x - \tilde{p}_z) + q_x) - (a_{24}\lambda + a_{21})(\lambda_{36} + a_{23}a_{32})),$$

$$\Delta_{34} = a_{15}(\lambda + \tilde{p}_y)(a_{23}a_{34}(\lambda(p_x - \tilde{p}_z) + q_x) - (a_{24}\lambda + a_{21})(\lambda_{36} + a_{23}a_{32})) - \tilde{a}_{25}\lambda_{14}\lambda_{36}.$$

Окончательно получим

$$\det \mathbf{A}_{B\lambda}^* = \Delta = \lambda_{14}\lambda_{25}\lambda_{36} = \prod_{i=1}^6 (\lambda - \lambda_i),$$

откуда и следует совпадение спектров $\text{eig } \mathbf{A}_B^* = \Lambda_B^*$.

Заключение

На основе математического аппарата матричных аннуляторов получены и доказаны новые аналитические формулы обращения квадратных матриц, разбитых по строкам или столбцам на два прямоугольных блока. Полученные формулы не имеют ограничений в применимости в виде каких-либо соотношений между прямоугольными блоками. Показано, как эффективно использовать предлагаемые формулы обращения блочных матриц для расчета матриц обратной связи при решении задач управления и оценки компонент вектора состояния. Для некоторого класса линейных стационарных систем управления разработана модификация полученных ранее обобщенных формул Басса–Гура и Аккермана.

Список источников

1. Zhou H., Yang D., Lin Y., Dai Y., Chen Q. A Parallel Acceleration Technique based on Bordered Block Diagonal Matrix Reordering for Exponential Integrator Method // 2nd Int. Symp. of Electronics Design Automation. 2024. P. 94–99. doi: 10.1109/ISED62518.2024.10617631
2. Dogruer C.U., Yildirim B. Switching Optimal Modal Control of Linear Time-Invariant Model of a Structural System // 25th Int. Conf. on Syst. Theory, Control and Computing. 2021. P. 220–226. doi: 10.1109/ICSTCC52150.2021.9607184
3. Зубов Н.Е., Лапин А.В., Рябченко В.Н. Аналитический алгоритм построения орбитальной ориентации космического аппарата при неполном измерении компонент вектора состояния // Известия РАН. Теория и системы управления. 2019. № 6. С. 128–138. doi: 10.1134/S0002338819040176
4. Zybin E.Yu., Kosyanchuk V.V., Karpenko S.S. Quantitative Model-Free Method for Aircraft Control System Failure Detection // MATEC Web of Conf. 2017. V. 99. Art. 03011. P. 1–3. doi: 10.1051/mateconf/20179903011
5. Kosyanchuk V.V., Selvesyuk N.I., Kulchak A.M. Aircraft Control Law Reconfiguration // Aviation. 2015. V. 19, is. 1. P. 14–18. doi: 10.3846/16487788.2015.1015290
6. Qin Y., Feng G., Ren Y., Zhang X. Block-Diagonal Guided Symmetric Nonnegative Matrix Factorization // IEEE Transactions on Knowledge and Data Engineering. 2023 V. 35, is. 3. P. 2313–2325. doi: 10.1109/TKDE.2021.3113943
7. Gantmacher F.R. The Theory of Matrices. Providence : AMS Chelsea Publishing, 2000.
8. Lapin A.V., Zubov N.E. Generalization of Bass–Gura Formula for Linear Dynamic Systems with Vector Control // Herald of the Bauman Moscow State Technical University. Series Natural Sciences. 2020. V. 89, is. 2. P. 41–64. doi: 10.18698/1812-3368-2020-2-41-64.
9. Лапин А.В., Зубов Н.Е., Пролетарский А.В. Обобщение формулы Аккермана для некоторого класса многомерных динамических систем с векторным входом // Вестник МГТУ им. Н.Э. Баумана. Сер. Естественные науки. 2023. № 4 (109). С. 18–38. doi: 10.18698/1812-3368-2023-4-18-38
10. Bernstein D.S. Matrix Mathematics. Princeton, NJ : Princeton University Press, 2005.
11. Воеводин В.В., Кузнецов Ю.А. Матрицы и вычисления. М. : Наука, 1984.
12. Мисриханов М.Ш., Рябченко В.Н. Блочное решение алгебраических уравнений при расчетах электрических сетей и систем // Вестник Тамбовского государственного технического университета. 2004. № 7 (42). С. 29–42.
13. Горюнов С.В., Буков В.Н. Обращение и канонизация блочных матриц // Математические заметки. 2006. Т. 79, № 5. С. 662–673.
14. Зубов Н.Е., Микрин Е.А., Рябченко В.Н. Матричные методы в теории и практике систем автоматического управления летательных аппаратов. М. : Изд-во МГТУ им. Н.Э. Баумана, 2016.
15. Kosyanchuk V.V., Zheltov Yu S., Zybin E.Yu. Aircraft Flight Control System Fault Tolerance under Structural and Parametric Uncertainties // J. Phys.: Conf. Ser. 2021. V. 1864. P. 1–12. doi: 10.1088/1742-6596/1864/1/012005
16. Ling J., Ji H., Yu C., Wang S., Yu D., Jiang S. Design of UPFC Sub-Synchronous Oscillation Damping Controller Based on Modal Control // China Int. Conf. on Electricity Distribution. 2018. P. 1439–1443. doi: 10.1109/CICED.2018.8592538
17. Dobroskok N.A., Vtorov V.B. Robust Pole Assignment in a Problem of State Observer-Based Modal Control // IV Int. Conf. on Control in Technical Syst. 2021. P. 30–33. doi: 10.1109/CTS53513.2021.9562851
18. Zong Q., Tian K., Zuo Q., Wang W. A Similarity Measurement Model of Complex Polygon with Holes Based on Fuzzy Matrix and Similarity Transformation // 3rd Int. Conf. on Geology, Mapping and Remote Sensing. 2022. P. 249–252. doi: 10.1109/ICGMRS55602.2022.9849295

19. Kisananin B., Agarwal G.C. *Linear Control Systems: With Solved Problems and MATLAB Examples*. New York : Kluwer Academic / Plenum Publishers, 2002.
20. Буков В.Н. Адаптивные прогнозирующие системы управления полетом. М. : Наука, 1987.
21. Nie F., Wang H., Cai X., Huang H., Ding C. Robust Matrix Completion via Joint Schatten p-Norm and lp-Norm Minimization // *IEEE 12th Int. Conf. on Data Mining*. 2012. P. 566–574. doi: 10.1109/ICDM.2012.160
22. Lapin A.V., Zubov N.E. Parametric Synthesis of Modal Control with Output Feedback for Descent Module Attitude Stabilization // *2019 Int. Russian Automation Conf.* 2019. P. 1–6. doi: 10.1109/RusAutoCon.2019.8867744
23. Lapin A.V., Zubov N.E. Analytic Solution of the Problem of Stabilizing Orbital Orientation of a Spacecraft with Flywheel Engines // *AIP Conf. Proc.* 2021. V. 2318, is. 1. Art. 130009. P. 1–8. doi: 10.1063/5.0036155

References

1. Zhou, H., Yang, D., Lin, Y., Dai, Y. & Chen, Q. (2024) A Parallel Acceleration Technique based on Bordered Block Diagonal Matrix Reordering for Exponential Integrator Method. *2nd Int. Symp. of Electronics Design Automation*. pp. 94–99. DOI: 10.1109/ISED62518.2024.10617631
2. Dogruer, C.U. & Yildirim, B. (2021) Switching Optimal Modal Control of Linear Time-Invariant Model of a Structural System. *25th Int. Conf. on Syst. Theory, Control and Computing*. pp. 220–226. DOI: 10.1109/ICSTCC52150.2021.9607184.
3. Zubov, N.E., Lapin, A.V. & Ryabchenko, V.N. (2019) Analytical Algorithm for Constructing the Orbital Orientation of a Spacecraft with an Incomplete Measurement of the State Vector Components. *Journal of Computer and Systems Sciences International*. 58(6). pp. 969–979. DOI: 10.1134/S1064230719040178
4. Zybin, E.Yu., Kosyanchuk, V.V. & Karpenko, S.S. (2017) Quantitative Model-Free Method for Aircraft Control System Failure Detection. *MATEC Web of Conf.* 99(03011). pp. 1–3. DOI: 10.1051/mateconf/20179903011
5. Kosyanchuk, V.V., Selvesyuk, N.I. & Kulchak, A.M. (2015) Aircraft Control Law Reconfiguration. *Aviation*. 19(1). pp. 14–18. DOI: 10.3846/16487788.2015.1015290
6. Qin Y., Feng G., Ren Y. & Zhang X. (2023) Block-Diagonal Guided Symmetric Nonnegative Matrix Factorization. *IEEE Transactions on Knowledge and Data Engineering*. 35(3). pp. 2313–2325. DOI: 10.1109/TKDE.2021.3113943
7. Gantmacher, F.R. (2000) *The Theory of Matrices*. Providence: AMS Chelsea Publishing.
8. Lapin, A.V. & Zubov, N.E. (2020) Generalization of Bass – Gura Formula for Linear Dynamic Systems with Vector Control. *Vestnik MGTU im. N.E. Bauman. Ser. Estestvennye nauki – Herald of the Bauman Moscow State Technical University, Series Natural Sciences*. 89(2). pp. 41–64. DOI: 10.18698/1812-3368-2020-2-41-64
9. Lapin, A.V., Zubov, N.E. & Proletarskii, A.V. (2023) Generalization of Ackermann formula for a certain class of multidimensional dynamic systems with vector input. *Vestnik MGTU im. N.E. Bauman. Ser. Estestvennye nauki – Herald of the Bauman Moscow State Technical University, Series Natural Sciences*. 4(109). pp. 18–38. DOI: 10.18698/1812-3368-2023-4-18-38
10. Bernstein, D.S. (2005) *Matrix Mathematics*. Princeton University Press.
11. Voevodin, V.V. & Kuznetsov, Yu.A. (1984) *Matritsy i vychisleniya* [Matrices and Calculations]. Moscow: Nauka.
12. Misrikhanov, M.Sh. & Ryabchenko, V.N (2004) Blochnoe reshenie algebraicheskikh uravneniy pri raschetakh elektricheskikh setey i sistem [Block solution of algebraic equations in calculations of electrical networks and systems]. *Vestnik Tambovskogo gosudarstvennogo tekhnicheskogo universiteta*. 7(42). pp. 29–42.
13. Goryunov, S.V. & Bukov, V.N. (2006) Obrashchenie i kanonizatsiya blochnykh matrits [Inversion and canonization of block matrices]. *Matematicheskie zametki*. 79(5). pp. 662–673.
14. Zubov, N.E., Mikrin, E.A. & Ryabchenko, V.N. (2016) *Matrichnye metody v teorii i praktike sistem avtomaticheskogo upravleniya letatel'nykh apparatov* [Matrix methods in theory and practice of flying vehicles automatic control systems]. Moscow: Bauman MSTU.
15. Kosyanchuk, V.V. Zheltov, S.Yu. & Zybin, E.Yu. (2021) Aircraft Flight Control System Fault Tolerance under Structural and Parametric Uncertainties. *Journal of Physics: Conference Series*. 1864. pp. 1–12. DOI: 10.1088/1742-6596/1864/1/012005
16. Ling, J., Ji H., Yu, C., Wang, S., Yu, D. & Jiang, S. (2018) Design of UPFC Sub-Synchronous Oscillation Damping Controller Based on Modal Control. *China International Conference on Electricity Distribution*. pp. 1439–1443. DOI: 10.1109/CICED.2018.8592538.
17. Dobroskok, N.A. & Vtorov, V.B. (2021) Robust Pole Assignment in a Problem of State Observer-Based Modal Control. *IV International Conference on Control in Technical Systems*. pp. 30–33. DOI: 10.1109/CTS53513.2021.9562851.
18. Zong, Q., Tian, K., Zuo, Q. & Wang, W. (2022) A Similarity Measurement Model of Complex Polygon with Holes Based on Fuzzy Matrix and Similarity Transformation. *3rd Int. Conf. on Geology, Mapping and Remote Sensing*. pp. 249–252. DOI: 10.1109/ICGMRS55602.2022.9849295
19. Kisananin, B. & Agarwal, G.C. (2002) *Linear Control Systems: With Solved Problems and MATLAB Examples*. New York: Kluwer Academic/Plenum Publishers.
20. Bukov, V.N. (1987) *Adaptivnye prognoziruyushchie sistemy upravleniya poletom* [Adaptive predictive flight control systems]. Moscow: Nauka.
21. Nie, F., Wang, H., Cai, X., Huang, H. & Ding, C. (2012) Robust Matrix Completion via Joint Schatten p-Norm and lp-Norm Minimization. *IEEE 12th International Conference on Data Mining*. pp. 566–574. DOI: 10.1109/ICDM.2012.160.
22. Lapin, A.V. & Zubov, N.E. (2019) Parametric Synthesis of Modal Control with Output Feedback for Descent Module Attitude Stabilization. *2019 International Russian Automation Conference*. pp. 1–6. DOI: 10.1109/RusAutoCon.2019.8867744

23. Lapin, A.V. & Zubov, N.E. (2021) Analytic Solution of the Problem of Stabilizing Orbital Orientation of a Spacecraft with Flywheel Engines. *AIP Conference Proceedings*. 2318(1). Art. 130009. pp. 1–8. DOI: 10.1063/5.0036155

Информация об авторах:

Зубов Николай Евгеньевич – профессор, доктор технических наук, декан факультета «Ракетно-космическая техника», профессор кафедры «Системы автоматического управления» Московского государственного технического университета им. Н.Э. Баумана (Москва, Россия). E-mail: nik.zubov@gmail.com

Лапин Алексей Владимирович – кандидат технических наук, доцент кафедры «Системы автоматического управления» Московского государственного технического университета им. Н. Э. Баумана (Москва, Россия). E-mail: AlexeyPoeme@yandex.ru

Рябченко Владимир Николаевич – доцент, доктор технических наук, профессор кафедры «Системы автоматического управления» Московского государственного технического университета им. Н. Э. Баумана (Москва, Россия). E-mail: ryabchenko.vn@mail.ru

Вклад авторов: все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

Information about the authors:

Zubov Nikolay E. (Professor, Doctor of Technical Sciences, Dean of Rocket and Space Techniques Faculty, Professor of Department of Automatic Control Systems at Bauman Moscow State Technical University, Moscow, Russian Federation). E-mail: nik.zubov@gmail.com

Lapin Alexey V. (Candidate of Technical Sciences, Associate Professor of Department of Automatic Control Systems at Bauman Moscow State Technical University, Moscow, Russian Federation). E-mail: AlexeyPoeme@yandex.ru

Ryabchenko Vladimir N. (Associate Professor, Doctor of Technical Sciences, Professor of Department of Automatic Control Systems at Bauman Moscow State Technical University, Moscow, Russian Federation). E-mail: ryabchenko.vn@mail.ru

Contribution of the authors: the authors contributed equally to this article. The authors declare no conflicts of interests.

Поступила в редакцию 02.11.2024; принята к публикации 02.06.2025

Received 02.11.2024; accepted for publication 02.06.2025

*Посвящается светлой памяти члена-корреспондента
Национальной академии наук Республики Беларусь,
профессора Фаины Михайловны Кирилловой*

Научная статья
УДК 517.952; 62-50
doi: 10.17223/19988605/71/2

Необходимые условия оптимальности первого и второго порядков в обыкновенных динамических системах с запаздыванием при многоточечных функциональных ограничениях типа равенств и неравенств

Камиль Байрамали оглы Мансимов

*Бакинский государственный университет, Баку, Азербайджан
Институт систем управления Министерства науки и образования Азербайджана, Баку, Азербайджан
kamilbmansimov@gmail.com*

Аннотация. Рассматривается одна задача оптимального управления, описываемая системой дифференциальных уравнений с запаздывающим аргументом при наличии многоточечных функциональных ограничений типа равенств и неравенств. Во введении дается обзор работ, примыкающих к теме статьи. Первый раздел посвящен постановке задачи оптимального управления. Во втором разделе доказано необходимое условие оптимальности в форме аналога уравнения Эйлера. В третьем разделе сформулировано общее необходимое условие оптимальности второго порядка в терминах вторых вариаций функционалов, задающих ограничения и критерий качества. Из него получено необходимое условие оптимальности второго порядка, явно выраженное через параметры рассматриваемой задачи оптимального управления и носящее конструктивный характер. Установлен аналог условия Лежандра–Клебша. Изучен случай терминального критерия качества при наличии терминальных функциональных ограничений типа равенств и неравенств. Все необходимые условия оптимальности первого и второго порядков установлены без предположений о нормальности.

Ключевые слова: система дифференциальных уравнений с запаздыванием; допустимое управление; необходимое условие оптимальности; аналог уравнения Эйлера; условие Лежандра–Клебша; классический экстремаль.

Для цитирования: Мансимов К.Б. Необходимые условия оптимальности первого и второго порядков в обыкновенных динамических системах с запаздыванием при многоточечных функциональных ограничениях типа равенств и неравенств // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2025. № 71. С. 17–26. doi: 10.17223/19988605/71/2

Original article
doi: 10.17223/19988605/71/2

Necessary conditions for first- and second-order optimality in ordinary dynamic systems with delay under multipoint functional constraints of the equalities and inequalities type

Kamil B. Mansimov

*Baku State University, Baku, Azerbaijan
Institute of Control Systems of the Ministry of Science and Education of Azerbaijan, Baku, Azerbaijan
kamilbmansimov@gmail.com*

Abstract. The article considers one optimal control problem described by a system of differential equations with a retarded argument in the presence of multipoint functional constraints such as equalities and inequalities. The introduction provides an overview of works related to the topic of the article. The first section is devoted to the formulation of the optimal control problem. In the second section, the necessary condition for optimality is proved in the form

of an analogue of the Euler equation. In the third section, the general necessary condition for second-order optimality is formulated in terms of second variations of functionals defining constraints and quality criteria. From it, a necessary condition for second-order optimality was obtained, explicitly expressed through the parameters of the optimal control problem under consideration, and having a constructive nature. An analogue of the Legendre-Clebsch condition was established. Next, we study the case of a terminal quality criterion in the presence of terminal functional equations of the type of equalities and inequalities. All necessary conditions for first- and second-order optimality are established without assumptions of normality.

Keywords: system of differential equations with delay; admissible control; necessary condition for optimality analogue of the Euler equation; Legendre-Clebsch condition; classical extremal.

For citation: Mansimov, K.B. (2025) Necessary conditions for first- and second-order optimality in ordinary dynamic systems with delay under multipoint functional constraints of the equalities and inequalities type. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 71. pp. 17–26. doi: 10.17223/19988605/71/2

Введение

При выводе необходимых условий оптимальности второго порядка в различных задачах оптимального управления, описываемых обыкновенными дифференциальными уравнениями, в случае наличия фазовых или же функциональных ограничений на фазовую траекторию возникли существенные трудности (см., напр., обзор [1]).

Первые необходимые условия оптимальности второго порядка в общей задаче математического программирования в Банаховом пространстве при наличии ограничений и в некоторых задачах оптимального управления без дополнительных предположений о нормальности были получены в работах [2–6].

В дальнейшем выводу условий оптимальности второго порядка в задачах управления, описываемых дифференциальными уравнениями, при наличии фазовых или же функциональных ограничений были посвящены работы [7–15] и др. К настоящему времени теория необходимых условий оптимальности второго порядка, в частности теория особых управлений, достаточно полно развита также в случае задач управления, описываемых обыкновенными дифференциальными уравнениями, без запаздывания.

1. Постановка задачи

Пусть $[t_0, t_1]$ – заданный отрезок. Рассмотрим задачу о минимуме многоточечного функционала

$$S_0(u) = \varphi_0(x(T_1), x(T_2), \dots, x(T_k)) \quad (1)$$

при следующих ограничениях:

$$S_v(u) = \varphi_v(x(T_1), x(T_2), \dots, x(T_k)) \leq 0, v = \overline{1, p}, \quad (2)$$

$$S_v(u) = \varphi_v(x(T_1), x(T_2), \dots, x(T_k)) = 0, v = \overline{p+1, q}, \quad (3)$$

$$u(t) \in U \subset R^r, t \in [t_0, t_1], \quad (4)$$

$$\dot{x}(t) = f(t, x(t), x(t(h)), u(t_k)), t \in [t_0, t_1], \quad (5)$$

$$x(t) = x^0(t), t \in E_{t_0} = [h(t_0), t_0]. \quad (6)$$

Здесь U – заданное непустое, открытое и ограниченное множество, $f(t, x, y, u)$ – заданная n -мерная вектор-функция, непрерывная по совокупности переменных вместе с частными производными по (x, y, u) до второго порядка включительно, $\varphi_v(z_1, z_2, \dots, z_k), v = \overline{0, q}$ – заданные дважды непрерывно дифференцируемые скалярные функции, $h(t) < t$ – заданная непрерывно дифференцируемая скалярная функция, причем $\dot{h}(t) > 0$, $x^0(t)$ – заданная непрерывная на E_{t_0} начальная вектор-функция, а $u(t)$ – r -мерный измеримый и ограниченный вектор управляющих воздействий.

Предполагается, что каждой управляющей функции $u(t)$ соответствует единственное абсолютно непрерывное решение $x(t)$ системы (5), (6), определенное на отрезке $[t_0, t_1]$.

Управление $u(t)$ назовем допустимым управлением, если соответствующее решение $x(t)$ задачи (5), (6) удовлетворяет ограничениям (2), (3).

Допустимое управление $u(t)$, являющееся решением поставленной задачи, назовем оптимальным управлением, а соответствующий процесс $(u(t), x(t))$ – оптимальным процессом

2. Необходимое условие оптимальности первого порядка

Допустим, что $(u(t), x(t))$ – оптимальный процесс в задаче (1)–(5). Для простоты дальнейших изложений предположим, что для всех $v = \overline{1, p}$ $S_v(u) = 0$, и введем обозначения

$$\begin{aligned} H(t, x, y, u, \psi_v) &= \psi_v' f(t, x, y, u), \\ y(t) &= x(h(t)), \quad H_x^v[t] \equiv H_x(t, x(t), y(t), u(t), \psi_v(t)), \\ H_{xx}^v[t] &\equiv H_{xx}(t, x(t), y(t), u(t), \psi_v(t)), \\ H_u^v[t] &\equiv H_u(t, x(t), y(t), u(t), \psi_v(t)), \\ H_{uu}^v[t] &\equiv H_{uu}(t, x(t), y(t), u(t), \psi_v(t)). \end{aligned}$$

Здесь $\psi_v(t)$ – измеримая и ограниченная n -мерная вектор-функция, являющаяся решением интегрального уравнения

$$\begin{aligned} \psi_v(t) &= \int_t^{t_1} [f_x'[\tau] \psi_v(\tau) + \dot{r}(\tau) f_y'[\tau] \psi_v(r(\tau))] d\tau - \sum_{j=1}^k \alpha_j(t) \frac{\partial \varphi_v(x(T_1), x(T_2), \dots, x(T_k))}{\partial z_j}, \\ \psi_v(t) &= 0, t > t_1, \end{aligned} \quad (7)$$

где $\alpha_j(t)$ – характеристическая функция отрезка $[t_0, T_j]$, $r(t)$ – функция, обратная к $h(t)$, а штрих (') означает знак транспонирования. Учитывая введенные обозначения и то, что $\psi_v(t)$ является решением системы (7), доказывается, что первая и вторая вариации функционала $S_v(u)$ в «точке» $u(t)$ имеют следующий вид:

$$\begin{aligned} \delta^1 S_v(u : \delta u) &= - \int_{t_0}^{t_1} H_u^{(v)'}[t] \delta u(t) dt, \quad v = \overline{0, q}, \\ \delta^2 S_v(u : \delta u) &= \sum_{i,j=1}^k \delta x'(T_i) \frac{\partial^2 \varphi(x(T_1), x(T_2), \dots, x(T_k))}{\partial z_i \partial z_j} \delta x(T_j) - \\ &- \int_{t_0}^{t_1} [\delta x'(t) H_{xx}^{(v)}[t] \delta x(t) + \delta x'(t) H_{xy}^{(v)}[t] \delta y(t) + \delta y'(t) H_{yx}^{(v)}[t] \delta x(t) + \\ &+ \delta y'(t) H_{yy}^{(v)}[t] \delta y(t) + \delta u'(t) H_{ux}^{(v)}[t] \delta x(t) + \delta u'(t) H_{uy}^{(v)}[t] \delta y(t) + \\ &+ \delta x'(t) H_{xu}^{(v)}[t] \delta u(t) + \delta y'(t) H_{yu}^{(v)}[t] \delta u(t) + \delta u'(t) H_{uu}^{(v)}[t] \delta u(t)] dt, \quad v = \overline{0, q}. \end{aligned} \quad (9)$$

Здесь $\delta u(t) \in R^r$, $t \in [t_0, t_1]$ – произвольная измеримая и ограниченная вектор-функция (вариация управления), а $\delta x(t)$ – вариация траектории, являющаяся решением уравнения в вариациях

$$\delta \dot{x}(t) = f_x[t] \delta x(t) + f_y[t] \delta y(t) + f_u[t] \delta u(t), \quad (10)$$

$$\delta x(t) = 0, \quad t \in E_{t_0}. \quad (11)$$

Используя (8) по схеме, например из [6, 16], доказываем, что для оптимальности допустимого управления $u(t)$ необходимо существование вектора $\lambda = (\lambda_0, \lambda_1, \dots, \lambda_q) \in R^{q+1}$ такого, что

$$\lambda_v \geq 0, v = \overline{0, p}, \|\lambda\| = \sum_{v=0}^q |\lambda_v| = 1, \quad (12)$$

$$\sum_{v=0}^q \lambda_v \delta^1 S_v(u; \delta u) = 0 \quad (13)$$

для всех $\delta u(t) \in R^r, t \in [t_0, t_1]$.

В силу представления (8) условие оптимальности (13) превращается в следующее условие оптимальности (аналог уравнения Эйлера [7, 8, 17]):

$$H_u^{(\lambda)}[\theta] = 0, \quad (14)$$

для всех $\theta \in [t_0, t_1]$, где $H^{(\lambda)} = \sum_{v=0}^q \lambda_v H^{(v)}, \psi_\lambda(t) = \sum_{v=0}^q \psi_v(t)$, а $\theta \in [t_0, t_1]$ – произвольная правильная точка управления $u(t)$ (см., напр.: [18]).

Необходимое условие оптимальности (14), являясь условием оптимальности первого порядка, может выделить достаточно большое число допустимых управлений, подозрительных на оптимальность. Поэтому возникает проблема поиска дополнительных критериев оптимальности (условия оптимальности второго порядка) для отсева неоптимальных управлений, удовлетворяющих условию оптимальности (14).

3. Необходимое условие оптимальности второго порядка

Через $K(u; \delta u)$ обозначим множество критических вариаций управления $u(t)$ в задаче (1)–(6):

$$K(u; \delta u) = \left\{ \delta u; \delta^1 S_v(u; \delta u) \leq 0, v = \overline{0, p}, \delta^1 S_v(u; \delta u) = 0, v = \overline{p+1, q} \right\}. \quad (15)$$

Заметим, что понятие множества критических вариаций (15) было введено в [19] (см. также [20]).

Пусть $A(u)$ – множество всех векторов $\lambda \in R^{q+1}$, удовлетворяющих соотношениям (12), (13).

Приведем необходимое условие оптимальности второго порядка в рассматриваемой задаче.

По схеме из [21] без фактических изменений в рассуждениях доказываем

Теорема 1. Для оптимальности допустимого управления $u(t)$ задаче (1)–(6) необходимо, чтобы соотношение

$$\max_{\lambda \in A(u)} \sum_{v=0}^q \lambda_v \delta^2 S_v(u; \delta u) \geq 0 \quad (16)$$

выполнялось для всех $\delta u \in K(u; \delta u)$.

Неравенство (16) – неявный критерий оптимальности второго порядка, и поэтому его проверка является очень затруднительной.

Опираясь на условие оптимальности (16), получим конструктивно проверяемые необходимые условия оптимальности второго порядка, выраженные через параметры рассматриваемой задачи.

Решение уравнения в вариациях по аналогии с [22] может быть представлено в виде:

$$\delta x(t) = \int_{t_0}^t F(t, \tau) f_u[\tau] d\tau, \quad (17)$$

где $F(t, \tau)$ – $(n \times n)$ -матричная функция (аналог функции Коши), являющаяся решением задачи

$$\begin{aligned} F_\tau(t, \tau) &= -F(t, \tau) f_x[\tau] - \dot{r}(\tau) F(t, r(\tau)) f_y[r(\tau)], \tau < t, \\ F(t, t) &= E, F(t, \tau) = 0, \tau > t. \end{aligned}$$

Здесь E – $(n \times n)$ -единичная матрица.

Из (17) следует, что

$$\delta x(h(t)) = \delta y(t) = \int_{t_0}^t F(h(t), \tau) f_u[\tau] d\tau, \quad (18)$$

$$\delta x(T_i) = \delta y(t) = \int_{t_0}^{t_i} \alpha_i(t) F(T_i, t) f_u[t] \delta u(t) dt, \quad (19)$$

где $\alpha_i(t)$ – характеристическая функция отрезка $[t_0, T_i]$.

Используя представления (17)–(19) получаем, что

$$\begin{aligned} & \sum_{i,j=1}^k \delta x'(T_i) \frac{\partial^2 \Phi_v(x(T_1), x(T_2), \dots, x(T_k))}{\partial z_i \partial z_j} \delta x(T_j) = \\ & = \int_{t_0}^{t_1} \int_{t_0}^{t_1} \sum_{i,j=1}^k \alpha_i(\tau) \alpha_j(s) \delta u'(\tau) f_u'[\tau] F'(T_i, \tau) \frac{\partial^2 \Phi_v(x(T_1), x(T_2), \dots, x(T_k))}{\partial z_i \partial z_j} F(T_j, s) f_u[s] \delta u(s) ds d\tau, \end{aligned} \quad (20)$$

$$\int_{t_0}^{t_1} \delta u'(t) H_{ux}^{(v)}[t] \delta x(t) dt = \int_{t_0}^{t_1} \int_t^{t_1} \delta u'(\tau) H_{ux}^{(v)}[\tau] F(\tau, t) f_u[t] \delta u(t) d\tau dt, \quad (21)$$

$$\int_{t_0}^{t_1} \delta u'(t) H_{uy}^{(v)}[t] \delta y(t) dt = \int_{t_0}^{t_1} \int_t^{t_1} \delta u'(\tau) H_{uy}^{(v)}[\tau] F(h(\tau), t) f_u[t] \delta u(t) d\tau dt. \quad (22)$$

Далее получаем, что

$$\begin{aligned} & \int_{t_0}^{t_1} [\delta u'(t) H_{xx}^{(v)}[t] \delta x(t) + \delta x'(t) H_{xy}^{(v)}[t] \delta y(t) + \delta y'(t) H_{yx}^{(v)}[t] \delta x(t) + \delta y'(t) H_{yy}^{(v)}[t] \delta y(t)] dt = \\ & = \int_{t_0}^{t_1} \int_{t_0}^{t_1} \delta u'(\tau) f_u'[\tau] \left[\int_{\max(\tau, s)}^{t_1} [F'(t, \tau) H_{xx}^{(v)}[t] F(t, s) + F'(t, \tau) H_{xy}^{(v)}[t] F(h(t), s) + \right. \\ & \left. + F'(h(t), \tau) H_{yx}^{(v)}[t] F(t, s) + F'(h(t), \tau) H_{yy}^{(v)}[t] F(h(t), \tau)] dt \right] f_u[s] \delta u(s) ds d\tau. \end{aligned} \quad (23)$$

Введем ряд обозначений. Пусть

$$\begin{aligned} M^{(v)}(\tau, s) &= - \sum_{i,j=1}^k \alpha_i(\tau) \alpha_j(s) F'(T_i, \tau) \frac{\partial^2 \Phi_v(x(T_1), x(T_2), \dots, x(T_k))}{\partial z_i \partial z_j} F(T_j, s) + \\ &+ \int_{\max(\tau, s)}^{t_1} [F'(t, \tau) H_{xx}^{(v)}[t] F(t, s) + F'(t, \tau) H_{xy}^{(v)}[t] F(h(t), s) + \\ &+ F'(h(t), \tau) H_{yx}^{(v)}[t] F(t, s) + F'(h(t), \tau) H_{yy}^{(v)}[t] F(h(t), \tau)] dt, \\ M^{(\lambda)}(\tau, s) &= \sum_{v=0}^q \lambda_v M^{(v)}(\tau, s). \end{aligned}$$

С учетом введенных обозначений и тождеств (20)–(23) вторая вариация (9) функционала $S_v(u)$ представляется в виде:

$$\begin{aligned} \delta^2 S_v(u; \delta u) &= - \int_{t_0}^{t_1} \int_{t_0}^{t_1} \delta u'(\tau) f_u'[\tau] M^{(v)}(\tau, s) f_u[s] \delta u(s) ds d\tau - \\ &- \int_{t_0}^{t_1} \int_t^{t_1} \delta u'(s) (H_{ux}^{(v)}[s] F(s, t) + H_{uy}^{(v)}[s] F(h(s), t)) f_u[t] \delta u(t) dt - \\ &- \int_{t_0}^{t_1} \delta u'(t) f_u[t] \left[\int_t^{t_1} (F'(s, t) H_{xu}^{(v)}[s] + F'(h(s), t) H_{yu}^{(v)}[s]) \delta u(s) ds \right] dt - \end{aligned}$$

$$-\int_{t_0}^{t_1} \delta u'(t) H_{uu}^{(v)}[t] \delta u(t) dt, v = \overline{0, q}. \quad (24)$$

Принимая во внимание формулу (24) в неравенстве (16), приходим к следующему утверждению.

Теорема 2. Для оптимальности допустимого управления $u(t)$ необходимо, чтобы для всех $\delta u(t) \in R^r, t \in [t_0, t_1]$, удовлетворяющих условиям

$$\int_{t_0}^{t_1} H_u^{(v)'}[t] \delta u(t) dt \geq 0, v = \overline{0, p}, \int_{t_0}^{t_1} H_u^{(v)'}[t] \delta u(t) dt = 0, v = \overline{p+1, q},$$

выполнялось неравенство

$$\begin{aligned} \min_{\lambda \in A(u)} & \left[\int_{t_0}^{t_1} \int_{t_0}^{t_1} \delta u'(\tau) f_u'[\tau] M^{(\lambda)}(\tau, s) f_u[s] \delta u(s) ds dt + \right. \\ & + \int_{t_0}^{t_1} \int_t^{t_1} \delta u'(s) \left(H_{ux}^{(\lambda)}[s] F(s, t) + H_{uy}^{(\lambda)}[s] F(h(s), t) \right) ds \Big] f_u[t] \delta u(t) dt + \\ & + \int_{t_0}^{t_1} \delta u'(t) f_u'[t] \left[\int_t^{t_1} \left(F'(s, t) H_{xu}^{(\lambda)}[s] + F'(h(s), t) H_{yu}^{(\lambda)}[s] \right) \delta u(s) ds \right] dt + \\ & + \int_{t_0}^{t_1} \delta u'(t) H_{uu}^{(\lambda)}[t] \delta u(t) dt \leq 0. \end{aligned} \quad (25)$$

Как видно, неравенство (25) является конструктивно проверяемым необходимым условием оптимальности второго порядка. Более того, из него, определяя тем или иным образом допустимую вариацию $\delta u \in K(u; \delta u)$ управления $u(t)$, несложно получить ряд легко проверяемых и удобных для практического использования необходимых условий оптимальности второго порядка.

Непосредственным следствием теоремы 2 является

Следствие 1. (Аналог условия Лежандра–Клебша). Для оптимальности допустимого управления $u(t)$ в задаче (1)–(6) необходимо, чтобы для всех $v \in R^r$ и $\theta \in [t_0, t_1]$ выполнялось неравенство

$$\min_{\lambda \in A(u)} v' H_{uu}^{(\lambda)}[\theta] v \leq 0. \quad (26)$$

Доказательство этого следствия проводится по аналогии со схемой доказательства соответствующего результата для случая системы уравнений без запаздывания с терминальными ограничениями из [6, 21] и поэтому не приводится.

Теперь введем обозначения

$$\begin{aligned} p^{(\lambda)}(\theta, e) &= \int_{\theta}^{t_1} \int_{\theta}^{t_1} e'(\tau) f_u'[\tau] M^{(\lambda)}(\tau, s) f_u[s] e(s) ds dt + \\ & + \int_{\theta}^{t_1} \delta u'(t) H_{uu}^{(\lambda)}[t] e(t) dt + \int_{\theta}^{t_1} \int_t^{t_1} e'(s) \left(H_{ux}^{(\lambda)}[s] F(s, t) + H_{uy}^{(\lambda)}[s] F(h(s), t) \right) ds \Big] f_u[t] e(t) dt + \\ & + \int_{t_0}^{t_1} e'(t) f_u'[t] \left[\int_t^{t_1} \left(F'(s, t) H_{xu}^{(\lambda)}[s] + F'(h(s), t) H_{yu}^{(\lambda)}[s] \right) e(s) ds \right] dt, \end{aligned} \quad (27)$$

$$\begin{aligned} R^{(\lambda)}(\theta) &= \int_{\theta}^{t_1} \int_{\theta}^{t_1} f_u'[\tau] M^{(\lambda)}(\tau, s) f_u[s] ds dt + \int_{\theta}^{t_1} H_{uu}^{(\lambda)}[t] dt + \int_{\theta}^{t_1} \int_t^{t_1} \left(H_{ux}^{(\lambda)}[s] F(s, t) + H_{uy}^{(\lambda)}[s] F(h(s), t) \right) ds \Big] f_u[t] dt + \\ & + \int_{t_0}^{t_1} f_u'[t] \left[\int_t^{t_1} \left(F'(s, t) H_{xu}^{(\lambda)}[s] + F'(h(s), t) H_{yu}^{(\lambda)}[s] \right) ds \right] dt, \end{aligned} \quad (28)$$

$$N^{(\lambda)}(\theta, e) = \int_{\theta}^{t_1} [M^{(\lambda)}(\theta, s)f_u[s] + F'(s, \theta)H_{xu}^{(\lambda)}[s] + F'(h(s), \theta)H_{yu}^{(\lambda)}[s]]e(s)ds, \quad (29)$$

$$L^{(\lambda)}(\theta) = \int_{\theta}^{t_1} [M^{(\lambda)}(\theta, s)f_u[s] + F'(s, \theta)H_{xu}^{(\lambda)}[s] + F'(h(s), \theta)H_{yu}^{(\lambda)}[s]]e(s)ds. \quad (30)$$

Имеют место следующие утверждения.

Теорема 3. Для оптимальности допустимого управления $u(t)$ необходимо, чтобы для всех $\theta \in [t_0, t_1]$ и $e(t) \in R^r, t \in [t_0, t_1]$, удовлетворяющих условиям

$$\int_{\theta}^{t_1} H_u^{(v)}[t]e(t)dt \geq 0, v = \overline{0, p}, \int_{\theta}^{t_1} H_u^{(v)}[t]e(t)dt = 0, v = \overline{p+1, q},$$

выполнялось неравенство

$$\min_{\lambda \in A(u)} p^{(\lambda)}(\theta, e) \leq 0. \quad (31)$$

Теорема 4. Для оптимальности допустимого управления $u(t)$ необходимо, чтобы неравенство

$$\min_{\lambda \in A(u)} v'R^{(\lambda)}(\theta)v \leq 0 \quad (32)$$

выполнялось для всех $v \in R^r$ и $\theta \in [t_0, t_1]$ таких, что

$$\int_{\theta}^{t_1} H_u^{(v)}[t]vdt \geq 0, v = \overline{0, p}, \int_{\theta}^{t_1} H_u^{(v)}[t]vdt = 0, v = \overline{p+1, q}.$$

Для доказательства теоремы 3 достаточно в (25) вариацию $\delta u(t)$ управления $u(t)$ определить по формуле

$$\delta u(t) = \begin{cases} 0, t \in [t_0, \theta), \\ e(t), t \in [\theta, t_1]. \end{cases}$$

Полагая $e(t) = v$, из теоремы 3 получаем утверждение теоремы 4. Ясно, что условие оптимальности (32) слабее, чем условие оптимальности (31), но его проверка в ряде случаев намного проще.

Рассмотрим один частный случай, позволяющий получить уравнение для определения функций $P^{(\lambda)}(\theta, e)$ и $R^{(\lambda)}(\theta)$. Пусть в задаче (1)–(6) функционалы $S_v(u)$ являются терминальными, т.е.

$$S_v(u) = \varphi_v(x(t_1)), v = \overline{0, q}. \quad (33)$$

В случае задачи (2)–(6), (33) формула для $M^{(\lambda)}(\tau, s)$ принимает следующий вид:

$$\begin{aligned} M^{(\lambda)}(\tau, s) = & -F'(t_1, \tau) \frac{\partial^2 \varphi_{\lambda}(x(t_1))}{\partial x^2} F(t_1, s) + \\ & + \int_{\max(\tau, s)}^{t_1} [F'(t, \tau)H_{xx}^{(\lambda)}[t]F(t, s) + F'(t, \tau)H_{xy}^{(\lambda)}[t]F(h(t), s) + \\ & + F'(h(t), \tau)H_{yx}^{(\lambda)}[t]F(t, s) + F'(h(t), \tau)H_{yy}^{(\lambda)}[t]F(h(t), \tau)]dt. \end{aligned} \quad (34)$$

Поэтому из (34) получаем, что

$$\frac{\partial M^{(\lambda)}(\theta, s)}{\partial \theta} = f'_x[\theta]M^{(\lambda)}(\theta, s) - f'_y[r(\theta)]\dot{r}(\theta)M^{(\lambda)}(r(\theta), s). \quad (35)$$

Принимая во внимание (35), из (29), (30) непосредственным дифференцированием получаем, что $N^{(\lambda)}(\theta, e)$ и $L^{(\lambda)}(\theta)$ являются решениями следующих задач:

$$\begin{aligned} \dot{N}^{(\lambda)}(\theta, e) = & -f'_x[\theta]N^{(\lambda)}(\theta, e) - \dot{r}(\theta)f'_y[r(\theta)]N^{(\lambda)}(r(\theta), e(\theta)) - \\ & - [H_{xu}^{(\lambda)}[\theta] + M^{(\lambda)}(\theta, \theta)f_u[\theta]]e(\theta), \quad N^{(\lambda)}(t_1, e) = 0, \end{aligned} \quad (36)$$

$$\begin{aligned}\dot{L}^{(\lambda)}(\theta) &= -f'_x[\theta]L^{(\lambda)}(\theta) - \dot{r}(\theta)f'_y[r(\theta)]L^{(\lambda)}(r(\theta)) - [H_{xu}^{(\lambda)}[\theta] + M^{(\lambda)}(\theta, \theta)f_u[\theta]], \\ L^{(\lambda)}(t_1) &= 0.\end{aligned}\quad (37)$$

Теперь с учетом (36), (37) из (27), (28) легко получить, что в случае задачи (2)–(6), (33) функции $P^{(\lambda)}(\theta, e)$ и $R^{(\lambda)}(\theta)$ являются решениями следующих задач Коши:

$$\begin{aligned}\dot{P}^{(\lambda)}(\theta, e) &= -e'(\theta)f'_u[\theta]N^{(\lambda)}(\theta, e) - N^{(\lambda)'}(\theta, e)f'_u[r(\theta)]e(\theta) - e'(\theta)H_{uu}^{(\lambda)}[\theta]e(\theta), \\ P^{(\lambda)}(t_1, e) &= 0,\end{aligned}\quad (38)$$

$$\dot{R}^{(\lambda)}(\theta) = -f'_u[\theta]R^{(\lambda)}(\theta) - R^{(\lambda)'}(\theta)f_u[\theta] - H_{uu}^{(\lambda)}[\theta], \quad R^{(\lambda)}(t_1) = 0. \quad (39)$$

Заключение

Рассматривается задача о минимуме многоточечного функционала, определенного на абсолютно непрерывных решениях обыкновенного дифференциального уравнения с переменным запаздыванием при наличии промежуточных функциональных ограничений типа равенств и неравенств на состояние системы. Допустимые управления входят в класс измеримых и ограниченных вектор-функций. Приведены первая и вторая вариации (в классическом смысле) критерия качества и функционалов, задающие ограничения. Доказаны аналог уравнения Эйлера и ряд общих необходимых условий оптимальности второго порядка, носящих конструктивный характер, позволяющих сузить множество классических экстремалей, подозрительных на оптимальность. В частности, получен аналог условия Лежандра–Клебша. Рассмотрен один частный случай.

Список источников

1. Гороховик В.В. Необходимые условия оптимальности высокого порядка для задачи управления с терминальными ограничениями : препринт ИМ АН БССР. Минск, 1982. № 1 (126). 50 с.
2. Левитин Е.С., Милютин А.А., Осмоловский Н.П. Об условиях локального минимума в задаче с ограничениями // Математическая экономика и функциональный анализ : сб. ст. М. : Наука, 1974. С. 139–202.
3. Левитин Е.С., Милютин А.А., Осмоловский Н.П. Условия высших порядков в гладких задачах на экстремум с ограничениями // Теоретические и прикладные вопросы оптимального управления : сб. ст. М. : Наука, 1985. С. 4–40.
4. Осмоловский Н.П. Условия второго порядка слабого локального минимума в задаче оптимального управления // Доклады Академии наук СССР. 1975. Т. 225, № 2. С. 259–262.
5. Дмитрук А.В. Квадратичные условия слабого минимума для особых режимов в задачах оптимального управления // Доклады Академии наук СССР. 1977. Т. 233, № 4. С. 523–526.
6. Гороховик В.В. Необходимые условия оптимальности в задачах оптимизации с векторным показателем качества // Проблемы оптимального управления : сб. ст. Минск : Наука и техника, 1981. С. 5–25.
7. Калинин А.И. К теории необходимых условий оптимальности второго порядка // Доклады Академии наук БССР. 1982. № 8. С. 677–680.
8. Калинин А.И. К проблеме особых управлений // Дифференциальные уравнения. 1985. № 3. С. 380–385.
9. Ащепков Л.Т., Васильев О.В. Исследование экстремалей Понтрягина методами нелинейного программирования // Вопросы оптимизации и устойчивости динамических систем : сб. ст. Иркутск : Изд-во Иркут. ун-та, 1983. С. 5–23.
10. Гороховик С.Я. Необходимые условия оптимальности в задаче с подвижным правым концом траектории // Дифференциальные уравнения. 1975. № 10. С. 1765–1773.
11. Колокольникова Г.А. Необходимое условие оптимальности второго порядка в задачах оптимального управления с подвижными концами. Иркутск : Иркут. гос. ун-т, 1990. Деп. в ВИНТИ АН СССР. № 1350.
12. Ащепков Л.Т. Оптимальное управление разрывными системами. Новосибирск : Наука, 1987. 226 с.
13. Гороховик В.В. Необходимые условия слабой эффективности в задаче управления с векторным показателем качества : препринт ИМ АН БССР. Минск, 1976. № 131 (3). 44 с.
14. Аграчев А.А. Необходимое условие оптимальности второго порядка в общем нелинейном случае // Математический сборник. 1977. Т. 102 (144), № 4. С. 551–568.
15. Warga J. Second order constition that strengthens Pontryagins maximum principle // Journ. Optim. Theory and Appl. 1978. V. 24 (3). P. 475–484.
16. Моисев Н.Н. Элементы теории оптимальных систем. М. : Наука, 1975. 526 с.
17. Габасов Р., Кириллова Ф.М. Особые оптимальные управления. М. : Либроком, 2011. 256 с.

18. Понтрягин Л.С., Болтянский В.Г., Гамкрелидзе Р.В., Мищенко Е.Ф. Математическая теория оптимальных процессов. М. : Наука, 1969. 384 с.
19. Дубовицкий А.Я., Милютин А.А. Задачи на экстремум при наличии ограничений // Журнал вычислительной математики и математической физики. 1965. № 3. С. 395–413.
20. Дубовицкий А.Я., Милютин А.А. Вторые вариации в задачах на экстремум с ограничениями // Доклады Академии наук СССР. 1965. Т. 160, № 1. С. 18–21.
21. Гороховик В.В. Необходимые условия слабой эффективности в задаче управления с векторным показателем качества : препринт ИМ АН БССР. Минск, 1976. № 13 (13). 44 с.
22. Габасов Р., Кириллова Ф.М. Оптимизация линейных систем. Минск : Изд-во Белорус. гос. ун-та, 1973. 256 с.

References

1. Gorokhovik, V.V. (1982) Neobkhodimye usloviya optimal'nosti vysokogo poryadka dlya zadachi upravleniya s terminal'nymi ogranicheniyami [Necessary high-order optimality conditions for a control problem with terminal constraints]. *Preprint IM AN BSSSR*. 1(126).
2. Levitin, E.S., Milyutin, A.A. & Osmolovsky, N.P. (1974) Ob usloviyakh lokal'nogo minimuma v zadache s ogranicheniyami [On the conditions for a local minimum in a problem with restrictions]. In: Mityagin, B.S. (ed.) *Matematicheskaya ekonomika i funktsional'nyy analiz* [Mathematical Economics and Functional Analysis]. Moscow: Nauka. pp. 139–202.
3. Levitin, E.S., Milyutin, A.A. & Osmolovsky N.P. (1985) Usloviya vysshikh poryadkov v gladkikh zadachakh na ekstremum s ogranicheniyami [Conditions of higher orders in smooth extremum problems with restrictions]. In: *Teoreticheskie i prikladnye voprosy optimal'nogo upravleniya* [Theoretical and Applied Issues of Optimal Control]. Moscow: Nauka. pp. 4–40.
4. Osmolovsky, N.P. (1975) Usloviya vtorogo poryadka slabogo lokal'nogo minimuma v zadache optimal'nogo upravleniya [Second-order conditions for a weak local minimum in an optimal control problem]. *Doklady Akademii nauk SSSR*. 225(2). pp. 259–262.
5. Dmitruk, A.V. (1977) Kvadratichnye usloviya slabogo minimuma dlya osobykh rezhimov v zadachakh optimal'nogo upravleniya [Quadratic conditions for a weak minimum for special modes in optimal control problems]. *Doklady Akademii nauk SSSR*. 233(4). pp. 523–526.
6. Gorokhovik, V.V. (1981) Neobkhodimye usloviya optimal'nosti v zadachakh optimizatsii s vektornym pokazatelem kachestva [Necessary conditions for optimality in optimization problems with a vector quality indicator]. In: Gabasov, R. & Kirillova, F.M. (eds) *Problemy optimal'nogo upravleniya* [Problems of Optimal Control]. Minsk: Nauka i tekhnika. pp. 5–25.
7. Kalinin, A.I. (1982) K teorii neobkhodimyykh usloviy optimal'nosti vtorogo poryadka [On the theory of necessary conditions for second-order optimality]. *Doklady Akademii nauk BSSR*. 8. pp. 677–680.
8. Kalinin, A.I. (1985) K probleme osobykh upravleniy [On the problem of singular controls]. *Differentsial'nye uravneniya*. 3. pp. 380–385.
9. Ashchepkov, L.T. & Vasilyev, O.V. (1983) Issledovanie ekstremaley Pontryagina metodami nelineynogo programmirovaniya [The study of Pontryagin extremals using nonlinear programming methods]. *Voprosy optimizatsii i ustoychivosti dinamicheskikh sistem* [Questions of optimization and stability of dynamic systems]. Irkutsk: Irkutsk State University. pp. 5–23.
10. Gorokhovik, S.Ya. (1975) Neobkhodimye usloviya optimal'nosti v zadache s podvizhnym pravym kontsom traektorii [Necessary conditions for optimality in a problem with a moving right end of a trajectory]. *Differentsial'nye uravneniya*. 10, pp. 1765–1773.
11. Kolokolnikova, G.A. (1990) *Neobkhodimoe uslovie optimal'nosti vtorogo poryadka v zadachakh optimal'nogo upravleniya s podvizhnymi kontsami* [A necessary condition for second-order optimality in optimal control problems with moving ends]. № 1350. Irkutsk: Irkutsk State University.
12. Ashchepkov, L.T. (1987) *Optimal'noe upravlenie razryvnymi sistemami* [Optimal Control of Bursting Systems]. Novosibirsk: Nauka.
13. Gorokhovik, V.V. (1976) Neobkhodimye usloviya slaboy effektivnosti v zadache upravleniya s vektornym pokazatelem kachestva [Necessary conditions for weak efficiency in a control problem with a vector quality indicator]. *Preprint IM AN BSSSR*. 131(3).
14. Agrachev, A.A. (1977) Neobkhodimoe uslovie optimal'nosti vtorogo poryadka v obshchem nelineynom sluchae [Necessary condition for second-order optimality in the general nonlinear case]. *Matematicheskii sbornik*. 102(144). no. 4. pp. 551–568.
15. Warga, J. (1978) Second order constitution that strengthens Pontryagins maximum principle. *Journal of Optimization Theory and Applications*. 24(3). pp. 475–484.
16. Moisey N.N. (1975) *Elementy teorii optimal'nykh sistem* [Elements of the Theory of Optimal Systems]. Moscow: Nauka.
17. Gabasov, R. & Kirillova, F.M. (2011) *Osobyie optimal'nye upravleniya* [Singular Optimal Controls]. Moscow: Librokom.
18. Pontryagin, L.S., Boltyansky, V.G., Gamkrelidze, R.V. & Mishchenko, E.F. (1969) *Matematicheskaya teoriya optimal'nykh protsessov* [Mathematical Theory of Optimal Processes]. Moscow: Nauka.
19. Dubovitsky, A.Ya. & Milyutin, A.A. (1965) Zadachi na ekstremum pri nalichii ogranicheniy [Extremum problems in the presence of restrictions]. *Zhurnal vychislitel'noy matematiki i matematicheskoy fiziki*. 3. pp. 395–413.
20. Dubovitsky, A.Ya. & Milyutin, A.A. (1965) Vtorye variatsii v zadachakh na ekstremum s ogranicheniyami [Second variations in extremum problems with restrictions]. *Doklady Akademii nauk SSSR*. 160(1). pp. 18–21.
21. Gorokhovik, V.V. (1976) Neobkhodimye usloviya slaboy effektivnosti v zadache upravleniya s vektornym pokazatelem kachestva [Necessary conditions for weak efficiency in a control problem with a vector quality indicator]. *Preprint IM AN BSSSR*. 13(13).
22. Gabasov, R. & Kirillova, F.M. (1973) *Optimizatsiya lineynykh sistem* [Optimization of Linear Systems]. Minsk: Belarus State University.

Информация об авторе:

Мансимов Камил Байрамали оглы – профессор, доктор физико-математических наук, заведующий лабораторией «Методы управления в сложных динамических системах» Института систем управления Министерства науки и образования Азербайджана (Баку, Азербайджан); заведующий кафедрой «Математическая кибернетика» Бакинского государственного университета (Баку, Азербайджан), E-mail: kamilbmansimov@gmail.com

Автор заявляет об отсутствии конфликта интересов.

Information about the authorss:

Mansimov Kamil B. (Professor, Doctor of Physical and Mathematical Sciences, Head of the Laboratory “Control Methods in Complex Dynamic Systems” of the Institute of Control Systems of the Ministry of Science and Education of Azerbaijan, Baku, Azerbaijan); Head Department of Mathematical Cybernetics, Baku State University (Baku, Azerbaijan), E-mail: kamilbmansimov@gmail.com

The author declares no conflicts of interests.

Поступила в редакцию 16.12.2024; принята к публикации 02.06.2025

Received 16.12.2024; accepted for publication 02.06.2025

Научная статья
УДК 681.5.013
doi: 10.17223/19988605/71/3

Управление нелинейными системами с хаотической динамикой

Наталья Владимировна Сорокина¹, Владимир Николаевич Шашихин²

^{1,2} Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия

¹ sorokina_nv@spbstu.ru

² shashihin@bk.ru

Аннотация. Рассматривается задача управления нелинейной системой с хаотической динамикой на основе централизованной структуры регулятора. Управление строится по принципу обратной связи, позволяющей реализовать в замкнутой системе заданный спектр характеристических показателей Ляпунова для подавления хаотической динамики. Рассмотрено использование предлагаемой методики синтеза на примере системы трех синхронных генераторов. Результаты численного эксперимента подтвердили подавление хаотических колебаний и обеспечение в замкнутой системе регулярного режима.

Ключевые слова: нелинейные системы; хаос; характеристические показатели Ляпунова.

Для цитирования: Сорокина Н.В., Шашихин В.Н. Управление нелинейными системами с хаотической динамикой // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2025. № 71. С. 27–38. doi: 10.17223/19988605/71/3

Original article
doi: 10.17223/19988605/71/3

Control of nonlinear systems with chaotic dynamics

Natalia V. Sorokina¹, Vladimir N. Shashihin²

^{1,2} Peter the Great Saint Petersburg Polytechnic University, Saint Petersburg, Russian Federation

¹ sorokina_nv@spbstu.ru

² shashihin@bk.ru

Abstract. The article considers the problem of controlling a nonlinear system with chaotic dynamics based on a centralized controller structure. Control is based on the principle of feedback, which allows the implementation of a given spectrum of Lyapunov exponents in a closed system to suppress chaotic dynamics. An example of the application of the proposed synthesis method to a system of three synchronous generators is considered. The results of the numerical experiment confirmed the suppression of the chaotic oscillations and the provision of a regular mode in the closed system.

Keywords: nonlinear systems; chaos; Lyapunov exponent.

For citation: Sorokina, N.V., Shashihin, V.N. (2025) Control of nonlinear systems with chaotic dynamics. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 71. pp. 27–38. doi: 10.17223/19988605/71/3

Введение

Хаотические режимы могут возникать во многих нелинейных динамических системах. Для одних систем хаотические режимы являются вредными – приводят к шумам в системах связи или вибрациям различных конструкций, для других – полезными, например для криптографических систем. Поэтому для одних систем необходимо подавить хаос и стабилизировать динамическую

систему, для других – создать хаос или усилить его. Вследствие этого одной из актуальных задач теории динамического хаоса является разработка методов управления хаосом [1].

Под стабилизацией динамической системы с хаотическим поведением подразумевается нахождение управляющих воздействий, позволяющих вывести систему из хаотического режима на регулярный.

Для стабилизации хаотических систем исторически первыми и наиболее активно развиваемыми в настоящее время являются следующие методы. Метод Отта–Гребоджи–Йорке (OGY-метод) [2] предназначен для стабилизации неустойчивой периодической траектории, вложенной в аттрактор хаотической системы, путем малых возмущений изменяемого параметра системы. Эти возмущения находятся с помощью дискретной модели системы, основанной на линеаризации отображения Пуанкаре.

Метод Пирагаса [3] обеспечивает стабилизацию периодической орбиты нелинейной системы путем построения обратной связи по фазовому вектору с запаздывающим аргументом. Этот метод чувствителен к выбору времени запаздывания в зависимости от периода цикла, который, как правило, неизвестен в хаотических системах, поэтому достичь требуемой сходимости можно в редких случаях.

Метод Магницкого [4] предназначен для локализации и стабилизации неустойчивых особых точек и периодических решений хаотических систем с помощью введения координатно-параметрической обратной связи в расширенном пространстве. К достоинствам метода относится то, что он позволяет найти неподвижные точки или периодические траектории хаотической системы, не требует информации о величине периода и положении искомого неустойчивого цикла в фазовом пространстве.

Также для управления хаотической динамикой используются традиционные подходы и методы автоматического управления. Например, в работе [5] стабилизация хаотической системы осуществляется с помощью обратной связи по отклонению, а для выбора коэффициентов регулятора используется критерий Рауса–Гурвица.

Настоящая работа посвящена решению задачи стабилизации нелинейных систем с хаотической динамикой. Процедура синтеза управления основана на формировании желаемого спектра характеристических показателей Ляпунова путем введения обратной связи по фазовому вектору. Коэффициенты обратной связи определяются на основе решения матричного уравнения Сильвестра.

1. Постановка задачи

1.1. Математическая модель нелинейной системы

Пусть возмущенные движения нелинейного динамического объекта описываются автономным векторным дифференциальным уравнением

$$\dot{x}(t) = F(x(t)) + Bu(t), \quad x(0) = x_0, \quad (1)$$

где $x(t) \in R^n$ – вектор состояния, $u(t) \in R^m$ – вектор управления, $m \leq n$, $F(x(t)) = (f_i(x(t)))_{i=1}^n$ – векторная функция, удовлетворяющая условиям существования решений уравнения (1), $f_i(x(t))$ – вещественные функции, определенные и непрерывные по совокупности аргументов в области

$$\Omega = \{(x, u) \mid \|x\| + \|u\| < \varrho, \varrho = \text{const} > 0\} \subset R^n \otimes R^m.$$

Векторная функция F обладает заданным классом гладкости по векторному аргументу $x(t)$ – $F \in C_x^r$, или удовлетворяет условию Липшица

$$\|F(x') - F(x'')\| \leq k \|x' - x''\|, \quad k > 0.$$

Кроме того, векторная функция F :

– обладает неустойчивостью по отношению к заданию начальных условий, т.е. имеется величина δ такая, что для некоторой точки $x \in R^n$ и $\varepsilon > 0$ существует точка $y \in R^n$, для которой расстояние

$$\text{dist}[x(t_0), y(t_0)] < \varepsilon, \text{ а } \text{dist}[x(t), y(t)] \geq \delta \text{ для } t > t_0;$$

– является топологически транзитивной, т.е. для любых двух открытых множеств N, M найдется такое l , что

$$F^l(N) \cap M \neq \emptyset;$$

– обладает элементом регулярности, или, иначе, плотностью периодических траекторий, – в любой окрестности любой точки фазового пространства существует по крайней мере одна и, следовательно, бесконечно много периодических траекторий.

Для траекторий $x(t, x_0)$ системы (1) может выполняться одна из трех возможностей:

– либо траектория $x(t, x_0) = x^0$ является точкой покоя или состоянием равновесия;
– либо траектория $x(t, x_0)$ соответствует периодическому решению, в этом случае существует такое число $T > 0$, что

$$x(t+T, x_0) \equiv x(t, x_0),$$

и траектория называется периодической, или замкнутой;

– либо для любых $t_1 \neq t_2$

$$x(t_1, x_0) \neq x(t_2, x_0).$$

Предположим, что вектор состояния системы (1) достаточно велик. В этом случае она может быть представлена в виде совокупности подсистем. Состояние i -й изолированной (невзаимодействующей) подсистемы определяется выражением

$$\dot{x}_i = g_i(t, x_i), \quad x_i(0) = x_{i0}, \quad g_i(t, 0) \equiv 0, \quad i = \overline{1, N}, \quad (2)$$

где $x_i \in \mathbb{R}^{n_i}$ – вектор состояния i -й подсистемы, $\sum_{i=1}^N n_i = n$, $g_i(t, x_i): \mathbb{R} \times \mathbb{R}^{n_i} \rightarrow \mathbb{R}^{n_i}$ – векторные функции,

определяющие состояние изолированных подсистем, N – количество подсистем в системе. Функции

$$h_i(t, x): \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^{n_i},$$

равные

$$h_i(t, x) = f_i(t, x) - g_i(t, x_i), \quad i = \overline{1, N}, \quad (3)$$

описывают взаимосвязи i -й подсистемы с другими подсистемами.

Поведение i -й взаимодействующей подсистемы может быть представлено уравнением

$$\dot{x}_i = g_i(t, x_i) + h_i(t, x), \quad i = \overline{1, N}. \quad (4)$$

Уравнение (3) описывает связи между изолированными подсистемами (2), а уравнение (4) – поведение крупномасштабной системы (1), представленной в виде взаимодействующих подсистем.

1.2. Характеристические показатели Ляпунова

Одна из особенностей нерегулярных режимов – неустойчивость траекторий, принадлежащих хаотическому (странному) аттрактору. Количественной мерой этой неустойчивости являются характеристические показатели, введенные первоначально Ляпуновым [6]. Формально характеристический показатель Ляпунова вводится следующим образом: характеристическим показателем функции $z(t)$ называется число (или символ $\pm\infty$), определенное следующим образом:

$$\lambda(z) \equiv \overline{\lim}_{t \rightarrow \infty} \left(t^{-1} \ln \|z(t)\| \right).$$

Характеристический показатель Ляпунова функции $z(t)$ есть результат сравнения скорости роста функции $z(t)$ при $t \rightarrow \infty$ с экспонентой $\exp\{\chi t\}$, для которой характеристический показатель равен χ . Среди всего набора характеристических показателей Ляпунова наиболее важен наибольший (старший) показатель $\chi_1 = \chi_{\max}$. Если $\chi_1 < 0$, то траектория нелинейной системы асимптотически устойчива; если $\chi_1 > 0$ – неустойчива. Набор характеристических показателей, упорядоченных по убыванию, $\chi_1 \geq \chi_2 \geq \dots \geq \chi_n$, называют спектром Ляпунова нелинейной динамической системы. В n -мерных системах сигнатура спектра Ляпунова (знаки характеристических показателей) может принимать следующий вид:

$$\underbrace{(-, -, -, \dots, -, -, -)}_n, \quad (5a)$$

$$(0, \underbrace{-, -, \dots, -, -, -}_{n-1}), \quad (5б)$$

$$(\underbrace{+, \dots, +}_s, 0, -, \dots, -), \quad (5в)$$

где формула (5a) определяет состояние равновесия, (5б) — предельный цикл, а (5в) — странный аттрактор.

1.3. Задача подавления хаоса

Задача подавления хаоса состоит в преобразовании нерегулярного режима системы (1), который характеризуется спектром Ляпунова (5в), в регулярный режим со спектром характеристических показателей (5a) или (5б), т.е. обеспечить аттрактор в виде особой точки или предельного цикла.

Для решения этой задачи будем искать управление в виде обратной связи по фазовому вектору нелинейной системы (1)

$$u(t) = Lx(t), \quad L \in \mathbb{R}^{m \times n},$$

которое обеспечит в замкнутой системе

$$\dot{x}(t) = F(x(t), Lx(t)), \quad x(0) = x_0,$$

спектр характеристических показателей Ляпунова

$$\sigma(F) = \{\chi_i(F), \quad i = \overline{1, n}\},$$

равный желаемому (требуемому) спектру, который определяется требуемым характером регулярного движения системы (1).

Таким образом, задача стабилизации сводится к подавлению хаотических колебаний путем приведения их к регулярным колебаниям (стабилизация предельного цикла) либо полному подавлению колебаний (стабилизация особой точки) путем введения обратной связи.

2. Синтез обратной связи

2.1. Топологическая эквивалентность

Решение задачи формирования необходимого спектра характеристических показателей Ляпунова основано на использовании их зависимости от собственных чисел матрицы Якоби. Изменение собственных чисел матрицы Якоби влечет за собой изменение характеристических показателей Ляпунова нелинейной системы. Задать желаемые собственные числа матрицы Якоби можно, например, с помощью синтеза модального управления на основе решения матричного алгебраического уравнения Сильвестра.

Правомочность такого подхода основывается на теореме о структурной устойчивости (грубости) нелинейных динамических систем, сформулированной в [7] и определяющей суть топологической эквивалентности нелинейной системы и гиперболической линеаризованной модели [8, 9]. Из теорем, приведенных в указанных работах, следует, что если линеаризованная система является гиперболической (не имеет чисто мнимых собственных чисел), то нелинейная система имеет устойчивое или неустойчивое многообразие, которые являются гладкими аналогами устойчивых или неустойчивых пространств линеаризованной системы. Иначе, нелинейная система и линеаризованная система имеют одинаковое количество особых точек и предельных циклов.

2.2. Синтез обратной связи для нелинейной системы

Алгоритм синтеза обратной связи для нелинейной крупномасштабной системы (1) включает в себя следующие шаги [10, 11].

При управлении спектром характеристических показателей Ляпунова используются собственные числа матриц Якоби $J(x)$, которые вычисляются в различных точках траектории нелинейной системы. Для этого фазовое пространство нелинейной системы разбивается на ячейки небольшого размера. Для каждой ячейки выбираются собственные числа матрицы Якоби замкнутой системы $\tilde{J}(x)$ в соответствии с решаемой задачей (5а) или (5б).

Далее вычисляется коэффициент обратной связи для каждой ячейки. Коэффициент обратной связи нелинейной системы определяется с учетом вероятности посещения траекторией ячеек фазового пространства системы (инвариантной меры динамической системы). Для вычисления вероятности посещения траекторией нелинейной системы ячеек фазового пространства p_i фазовое пространство делится на малые ячейки C_i и находится решение $x(t, x_0)$ динамической системы на достаточно большом отрезке времени. Затем для каждой ячейки определяется количество точек решения, попавших в нее, и находится вероятность попадания траектории в данную ячейку:

$$p_i = \frac{N_i}{N},$$

где N_i – количество точек, попавших в подмножество C_i , N – общее количество точек. Размер ячеек выбирается следующим образом:

$$h_j = \frac{1}{S(T) - S(T_0)} \sum_{k=S(T_0)}^{S(T)-1} |x_j(k+1) - x_j(k)|, \quad j = \overline{1, 4},$$

где T_0 – время начала расчета траектории нелинейной системы такое, что переходный процесс уже завершился, T – время окончания расчета траектории, $S(t)$ – номер шага, соответствующий текущему времени t . То есть длина стороны h_j ячейки, параллельной фазовой координате x_j , выбирается равной среднему по времени значению разности между координатами x_j следующей и предыдущей точек.

Требуемые собственные числа матрицы Якоби $\tilde{J}(x_i)$, соответствующие центру каждой ячейки, выбираются по формуле

$$\bar{v}(\tilde{J}(x_i)) = v(J(x_i)) + \alpha \operatorname{Re}(v(J(x_i))) + \beta \operatorname{Im}(v(J(x_i))),$$

где $v(J(x_i))$ – собственные числа матрицы Якоби исходной нелинейной системы, вычисленной в центре x_i ячейки C_i , α – коэффициент, влияющий на сдвиг собственных чисел по действительной оси; β – коэффициент, влияющий на сдвиг собственных чисел по мнимой оси. При решении задачи стабилизации нужно уменьшать собственные числа Якобиана, выбирая $\alpha < 0$ и $\beta < 0$.

Далее для каждой ячейки найдем коэффициенты обратной связи L_i , $i = \overline{1, N}$, которые обеспечат заданные характеристические показатели в замкнутой нелинейной системе. Для их вычисления предлагается использовать метод на основе решения матричного уравнения Сильвестра. Подробное описание метода приведено в разделе 2.3. Вычислив коэффициенты обратной связи для каждой ячейки, находим коэффициент обратной связи для нелинейной системы по формуле

$$L = \sum_{i=1}^N L_i p_i$$

и определяем характеристические показатели Ляпунова нелинейной системы (1), замкнутой управлением $u = Lx$.

Для задания в системе требуемого значения характеристического показателя Ляпунова вышеприведенные вычисления выполняются для нескольких наборов коэффициентов α и β , а затем определяется коэффициент обратной связи L^* , обеспечивающий желаемое значение характеристического показателя Ляпунова, по критерию

$$L^* = \arg \min \left\{ \varphi = \left(1 - \operatorname{sign} \left(\varepsilon - |\chi_{1i} - \chi_{1i}^*| \right) \right) \times \max_{j=1,r} \|L_j\| + \|L_i\| \mid \varphi < 2 \max_{j=1,r} \|L_j\| \right\}, \quad (6)$$

где $i \in \overline{1, r}$; r – количество коэффициентов α и β , χ_1^* – желаемый старший характеристический показатель Ляпунова нелинейной системы; χ_{1i} – старший характеристический показатель Ляпунова нелинейной системы с $u = L_i x$.

Для стабилизации системы (если желаемый аттрактор – неподвижная точка) характеристический показатель Ляпунова замкнутой системы должен быть меньше заданного отрицательного числа $\chi_{1i} < \overline{\chi_{1i}} < 0$. Критерий выбора L^* в этом случае имеет вид:

$$L^* = \arg \min \left\{ \varphi = \left(1 - \text{sign}(\overline{\chi_{1i}} - \chi_{1i}) \right) \times \max_{j=1, r} \|L_j\| + \|L_i\| \mid \varphi < 2 \max_{j=1, r} \|L_j\| \right\}.$$

При стабилизации системы (если желаемый аттрактор – предельный цикл) используется критерий (6), в котором желаемый старший характеристический показатель Ляпунова $\chi_1^* = 0$.

2.3. Синтез обратной связи для линеаризованной системы

Линеаризация системы

Пусть уравнение (1) описывает отклонения фазовых координат нелинейного объекта в окрестности частных решений x^S , соответствующих управляющим воздействиям u^S . Используя формулу Тейлора в предположении дифференцируемости компонент функции $F(x(t)) = (\phi_i(x(t)))_{i=1}^n$ в окрестности $\xi^S = (x^S, u^S)$, преобразуем уравнение (1) к квазилинейному виду:

$$\dot{x}(t) = A(\xi^S)x(t) + Bu(t) + f(\xi^S), \quad x(0) = x_0. \quad (7)$$

В уравнениях (7) $A(\xi^S)$ вычисляется в точке ξ^S по формуле

$$A(\xi^S) = \begin{bmatrix} \frac{\partial \phi_1}{\partial x_1} & \dots & \frac{\partial \phi_1}{\partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial \phi_n}{\partial x_1} & \dots & \frac{\partial \phi_n}{\partial x_n} \end{bmatrix}_{\substack{x(t)=x^S \\ u(t)=u^S}}. \quad (8)$$

Пусть для всех значений

$$\xi^S \in S(x^S, u^S, \rho) = \left\{ (x^S, u^S) : \|x - x^S\| + \|u - u^S\| \leq \rho, \rho > 0 \right\} \subset \mathbb{R}^n \otimes \mathbb{R}^m$$

выполняются оценки

$$\|f(\xi^S)\| \leq q \|\xi\|. \quad (9)$$

Если матрица Якоби вычислена по формуле (8) и выполняется условие (9), тогда уравнение (7) примет вид линеаризованной системы (или уравнения в вариациях):

$$\dot{y}(t) = Ay(t) + Bu(t), \quad (10)$$

где $y(t) \in \mathbb{R}^n$ – фазовый вектор линеаризованной системы.

Система (10) может быть использована для синтеза управления, стабилизирующего систему (1) в окрестности частного решения. Вещественные части собственных чисел матрицы Якоби определяют поведения траекторий исходной нелинейной системы.

Синтез централизованного управления

Рассматривается задача размещения полюсов системы, в которой определение параметров регулятора сводится к решению матричного уравнения Сильвестра. Для системы (10) необходимо найти стабилизирующий регулятор в виде обратной связи по вектору состояния

$$u(y(t)) = -Ly(t) \quad (11)$$

такой, что спектр замкнутой системы

$$\dot{y}(t) = (A - BL)y(t) = A_y y(t) \quad (12)$$

совпадает предписываемым спектром собственных значений A_y , задаваемым последовательностью $\mu = \{\mu_1, \dots, \mu_n\}$, $\operatorname{Re} \mu_i > 0$, $i = \overline{1, n}$:

$$\rho(A_y) = \rho(-W). \quad (13)$$

Здесь $W = \operatorname{diag}(\mu_i)_{i=1}^n \in \mathbb{R}^{n \times n}$ – матрица, на главной диагонали которой расположены числа $\mu_i > 0$.

Задача нахождения матрицы L , определяющей «глубину» обратной связи по полному вектору состояния, сводится к решению матричного уравнения Сильвестра

$$AP + PW = BG \quad (14)$$

относительно матрицы $P \in \mathbb{R}^{n \times n}$ при произвольной матрице $G \in \mathbb{R}^{m \times n}$ и решению матричного уравнения

$$LP = G, \quad L = GP^{-1}. \quad (15)$$

Для динамической системы (12) условия существования решения задачи размещения полюсов и методика синтеза стабилизирующего управления содержатся в теореме, приведенной в [10]. Параметры регулятора (11), обеспечивающего выполнение условия (13) в замкнутой системе (12), определяются из соотношения (15), где матрица P – решение уравнения Сильвестра (14).

3. Исследование системы синхронных генераторов

3.1. Модель трехмашинной системы

Предлагаемая методика синтеза управлений для нелинейной крупномасштабной системы рассматривается на примере управления хаотическими колебаниями, возникающими в работе электроэнергетической системы, представленной в виде системы трех взаимосвязанных синхронных генераторов. Для анализа хаотических режимов электроэнергетической системы используется классическая модель синхронного генератора, которая позволяет провести качественный и количественный анализ, указывающий на нерегулярный характер отклонения угла ротора и частоты.

Уравнения математической модели трехмашинной электроэнергетической системы, имеющей неодинаковую инерционность роторов генераторов, входящих в нее, имеют вид [12]:

$$\begin{aligned} \frac{d\delta_1}{dt} &= \omega_1, \\ \frac{d\omega_1}{dt} &= -B_1 \sin\left(\left(1 + \frac{1}{\sqrt{2}}\right)\delta_1 + \frac{1}{\sqrt{2}}\delta_3\right) - C_{13} \sin(\delta_1 - \delta_3) + P_1, \end{aligned} \quad (16a)$$

$$\begin{aligned} \frac{d\delta_2}{dt} &= \omega_2, \\ \frac{d\omega_2}{dt} &= -B_2 \sin\left(\left(1 + \frac{1}{\sqrt{2}}\right)\delta_2 + \frac{1}{\sqrt{2}}\delta_3\right) - C_{21} \sin(\delta_2 - \delta_1) + P_2, \end{aligned} \quad (16b)$$

$$\begin{aligned} \frac{d\delta_3}{dt} &= \omega_3, \\ \frac{d\omega_3}{dt} &= -B_3 \sin\left(\left(1 + \frac{1}{\sqrt{2}}\right)\delta_1 + \frac{1}{\sqrt{2}}\delta_3\right) - C_{31} \sin(\delta_3 - \delta_1) + P_3, \end{aligned} \quad (16b)$$

где $\delta_1, \delta_2, \delta_3$ – отклонения угла поворота ротора генератора относительно вращающейся синхронно оси, $\omega_1, \omega_2, \omega_3$ – отклонение угловой частоты, P_1, P_2, P_3 – изменение мощности, выдаваемой в сеть генераторами. Исследования многомашинной системы проводились при следующих значениях параметров модели:

$$\begin{aligned} B_1 = \frac{P_1}{T_{j1}} = 1; C_{13} = \frac{P_{c13}}{T_{j1}} = 0,1; P_1 = \frac{\varepsilon_{01}}{T_{j1}} = 0,4; B_2 = \frac{P_2}{T_{j2}} = 1; C_{21} = \frac{P_{c21}}{T_{j2}} = 0,1; P_2 = \frac{\varepsilon_{02}}{T_{j2}} = 0,4; B_3 = \frac{P_3}{T_{j3}} = 1; \\ C_{31} = \frac{P_{c31}}{T_{j3}} = 0,1; P_3 = \frac{\varepsilon_{03}}{T_{j3}} = 0,3, \end{aligned}$$

где P_{c13} , P_{c21} , P_{c31} – синхронизирующие мощности между генераторами, ε_{01} , ε_{02} , ε_{03} – начальные значения мощностей, выдаваемые в сеть генераторами при возникновении возмущения сети.

Введя фазовый вектор

$$x(t) = (x_1(t) = \delta_1, x_2(t) = \omega_1, x_3(t) = \delta_2, x_4(t) = \omega_2, x_5(t) = \delta_3, x_6(t) = \omega_3)^T \in \mathbb{R}^6,$$

запишем систему (16) в виде:

$$\dot{x}(t) = F(x(t)),$$

где $F = [f_1(x(t)); f_2(x(t)); f_3(x(t)); f_4(x(t)); f_5(x(t)); f_6(x(t))]$,

$$f_1(x(t)) = x_2,$$

$$f_2(x(t)) = -B_1 \sin\left(\left(\frac{1}{\sqrt{2}} + 1\right)x_1 + \frac{1}{\sqrt{2}}x_5\right) - C_{13} \sin(x_1 - x_5) + P_1,$$

$$f_3(x(t)) = x_4,$$

$$f_4(x(t)) = -B_2 \sin\left(\left(\frac{1}{\sqrt{2}} + 1\right)x_3 + \frac{1}{\sqrt{2}}x_5\right) - C_{21} \sin(x_3 - x_1) + P_2,$$

$$f_5(x(t)) = x_6,$$

$$f_6(x(t)) = -B_3 \sin\left(\left(\frac{1}{\sqrt{2}} + 1\right)x_1 + \frac{1}{\sqrt{2}}x_5\right) - C_{31} \sin(x_5 - x_1) + P_3.$$

3.2. Исследование процессов в системе

Система без управления

Исследование системы (16) на наличие хаотических колебаний проводилось при начальных условиях

$$\delta_1(0) = 0,6; \omega_1(0) = 0,3; \delta_2(0) = 0,6; \omega_2(0) = 0,3; \delta_3(0) = 0,6; \omega_3(0) = 0,3.$$

На рис. 1, 2 представлены проекции фазового портрета системы (16) на плоскости x_1x_2 , x_3x_4 и x_5x_6 и временные диаграммы фазового вектора системы (16).

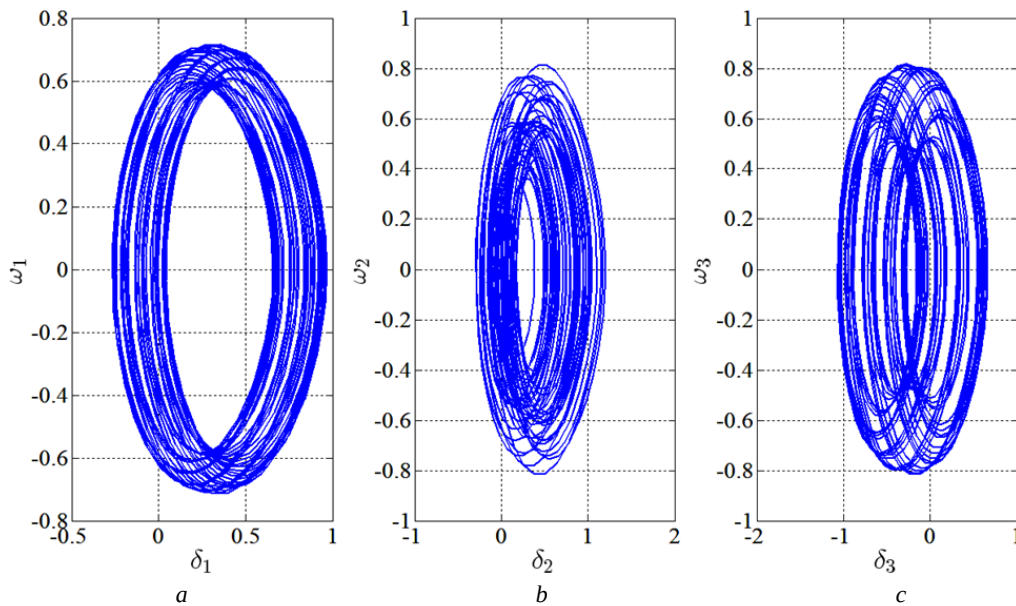


Рис. 1. Проекция фазового портрета системы (16): а – плоскость $x_1 = \delta_1$, $x_2 = \omega_1$;

б – плоскость $x_3 = \delta_2$, $x_4 = \omega_2$; в – плоскость $x_5 = \delta_3$, $x_6 = \omega_6$

Fig. 1. Projections of the phase portrait of the system (16): a) onto the plane $x_1 = \delta_1$, $x_2 = \omega_1$;

b) onto the plane $x_3 = \delta_2$, $x_4 = \omega_2$; c) onto the plane $x_5 = \delta_3$, $x_6 = \omega_6$

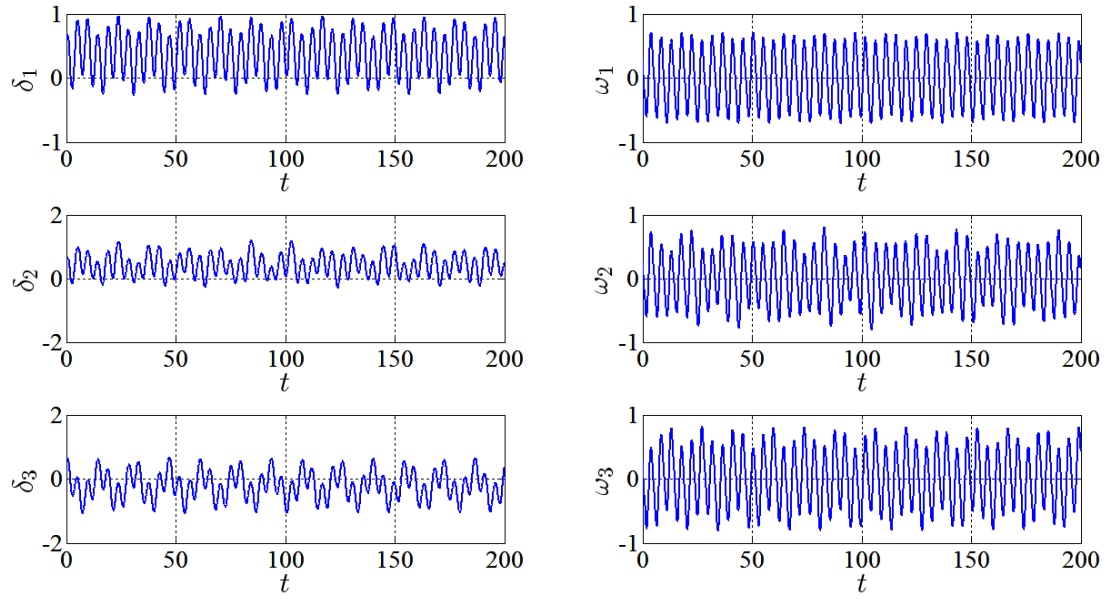


Рис.2. Временные диаграммы фазового вектора системы (16)

Fig.2. Spacetime diagrams of the phase vector of the system (16)

Матрица Якоби системы (16) имеет вид:

$$J = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ \partial f_2 / \partial x_1 & 0 & 0 & 0 & \partial f_2 / \partial x_5 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \partial f_4 / \partial x_1 & 0 & \partial f_4 / \partial x_3 & 0 & \partial f_4 / \partial x_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \partial f_6 / \partial x_1 & 0 & 0 & 0 & \partial f_6 / \partial x_5 & 0 \end{bmatrix}.$$

Одна из особых точек системы (16):

$$x_{\text{особ}} = (0,302; 0; 0,331; 0; -0,222; 0)^T.$$

Собственные числа матрицы Якоби в данной особой точке равны

$$\lambda_1 = i1,504, \lambda_2 = -i1,504, \lambda_3 = i1,291, \lambda_4 = -i1,291, \lambda_5 = i0,416, \lambda_6 = -i0,416.$$

При указанных значениях параметров и начальных условиях характеристические показатели Ляпунова системы (16) равны

$$\chi_1 = 0,014, \chi_2 = 0,003, \chi_3 = 0,009, \chi_4 = -0,011, \chi_5 = -0,003, \chi_6 = -0,012.$$

Спектр содержит положительные характеристические показатели Ляпунова, поэтому в системе (16) имеет место хаотический режим. Кроме того, на рис. 1 видно, что проекции траектории системы в фазовом пространстве являются странным аттрактором, что также свидетельствует о нерегулярном режиме.

Система с централизованным управлением

Введем в систему управление частотой каждого генератора; матрица B в системе (1) равна

$$B = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}^T,$$

и уравнения системы (16) с централизованным управлением (11) принимают вид:

$$\dot{x}(t) = F(x(t)) - BLx(t). \quad (17)$$

Коэффициенты обратной связи, вычисленные по методике синтеза централизованного регулятора с учетом (9) и (10), равны

$$L = \begin{pmatrix} -0,049 + i0,291 & -0,025 + i0,005 & -0,240 - i0,138 & 0,173 - i0,138 & 0,059 + i0,017 & 0,006 - i0,002 \\ 0,210 + i0,286 & 0,239 - i0,194 & -0,003 - i0,313 & -0,422 + i0,0004 & 0,114 - i0,008 & -0,033 - i0,015 \\ 0,024 + i0,168 & 0,328 + i0,050 & 0,130 - i0,178 & -0,320 - i0,178 & 0,045 + i0,027 & -0,270 - i0,006 \end{pmatrix}.$$

Спектр характеристических показателей Ляпунова системы (17) имеет вид:

$$\chi_1 = 0,0007, \chi_2 = 0,0007, \chi_3 = 0,003, \chi_4 = -0,0005, \chi_5 = -0,002, \chi_6 = -0,002.$$

Данный спектр свидетельствует о приведении системы к регулярному движению – предельному циклу, что подтверждается проекциями фазового портрета (рис. 3) и временными диаграммами (рис. 4).

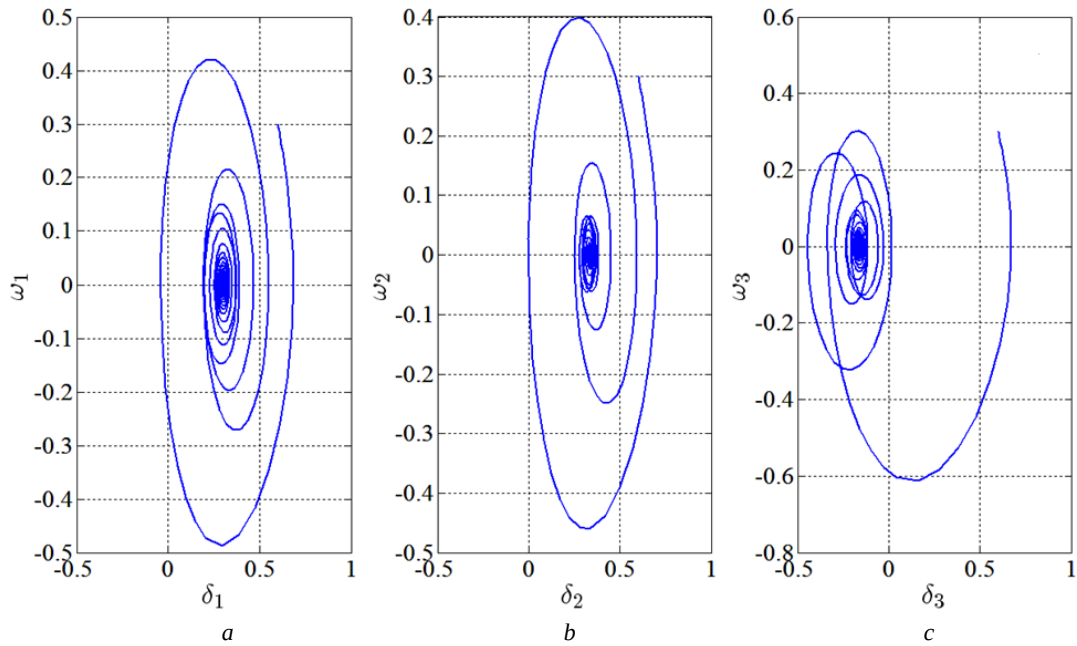


Рис. 3. Проекция фазового портрета системы (17): а – плоскость $x_1 = \delta_1, x_2 = \omega_1$;

б – плоскость $x_3 = \delta_2, x_4 = \omega_2$; с – плоскость $x_5 = \delta_3, x_6 = \omega_6$

Fig. 3. Projections of the phase portrait of the system (17): а) onto the plane $x_1 = \delta_1, x_2 = \omega_1$;

б) onto the plane $x_3 = \delta_2, x_4 = \omega_2$; с) onto the plane $x_5 = \delta_3, x_6 = \omega_6$

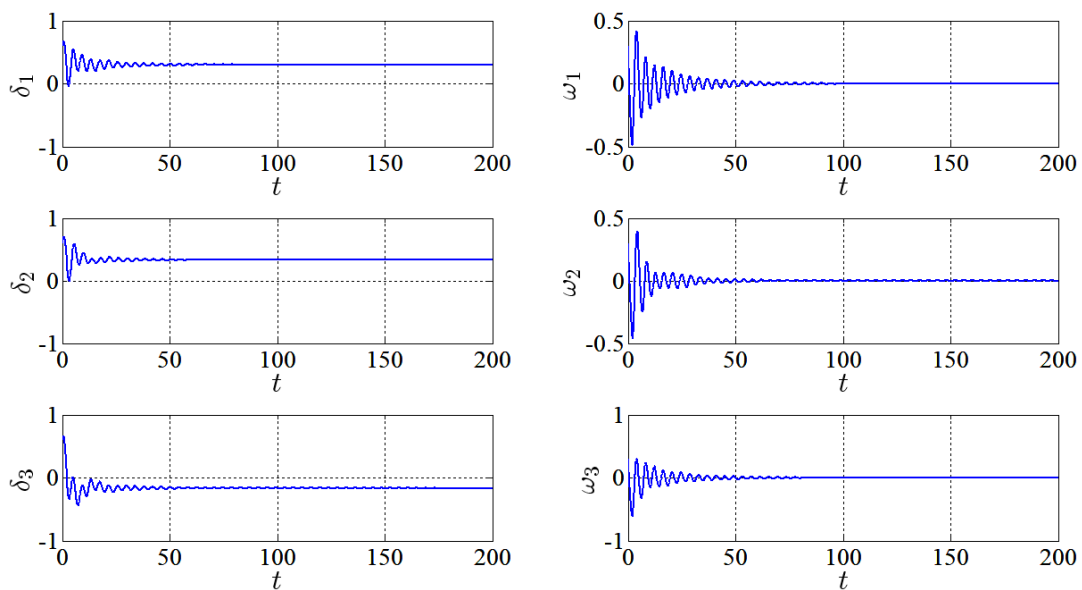


Рис.4. Временные диаграммы фазового вектора системы (17)

Fig.4. Spacetime diagrams of the phase vector of the system (17)

Заключение

В данной статье предложена методика синтеза стабилизирующего управления для нелинейных систем с хаотической динамикой. Параметры регулятора, обеспечивающего требуемый спектр характеристических показателей Ляпунова, определяются решением матричного уравнения Сильвестра и инвариантной мерой, рассчитанной на траекториях нелинейной системы.

Результаты исследования подтвердили подавление хаотических колебаний и обеспечение в замкнутой системе регулярного режима за счет формирования спектра с нулевыми и отрицательными характеристическими показателями Ляпунова.

Список источников

1. Андриевский Б.Р., Фрадков А.Л. Управление хаосом: методы и приложения. I. Методы // Автоматика и телемеханика. 2003. № 5. С. 3–45. doi: 10.1023/A:1023684619933
2. Ott E., Grebogi C., Yorke G. Controlling chaos // Physical Review Letter. 1990. V. 64. P. 1196–1199. doi: 10.1103/PhysRevLett.64.1196
3. Pyragas K. Continuous control of chaos by self-controlling feedback // Physical Letter. 1992. V. 170. P. 421–428. doi: 10.1016/0375-9601(92)90745-8
4. Магницкий Н.А., Сидоров С.В. Новые методы хаотической динамики. М. : Едиториал УРСС, 2004. 230 с.
5. Yamapi R., Chabi Orou J.B. Harmonic oscillations, stability and chaos control in a non-linear electromechanical system // Journal of Sound and Vibration. 2003. V. 259. P. 1253–1264. doi: 10.1006/jsvi.2002.5289
6. Ляпунов А.М. Общая задача об устойчивости движения. М. : Гостехиздат, 1950.
7. Андронов А.А., Понтрягин Л.С. Грубые системы // Доклады Академии наук СССР. 1937. Т. 14. С. 247–250.
8. Гробман Д.Г. Гомеоморфизм систем дифференциальных уравнений // Доклады Академии наук СССР, 1959. Т. 128. С. 880–881.
9. Оморов Р.О. Метод топологической грубости динамических систем: приложения к синергетическим системам // Научно-технический вестник информационных технологий, механики и оптики. 2020. Т. 20, № 2. С. 257–262. doi: 10.17586/2226-1494-2020-20-2-257-262
10. Shashihin V.N., Budnik S.V. Synthesis of control for nonlinear systems // Automatic Control and Computer Sciences. 2019. V. 53. P. 97–106. doi: 10.3103/S0146411619020068
11. Kozlov V.N., Shashikhin V.N. Synthesis of decentralized robust stabilizing control for the systems with parametric perturbations // Computing, Telecommunications and Control. 2020. V. 13. P. 49–60. doi: 10.18721/JCSTCS.13205
12. Рысеев Д.В., Федоров В.К. Современные проблемы динамики нелинейных энергосистем: электромеханический резонанс, энтропия, детерминированный хаос. Омск : Полиграф. центр КАН, 2012.

References

1. Andrievskiy, B.R. & Fradkov, A.L. (2003) Upravlenie khaosom: metody i prilozheniya. I. Metody [Chaos control: methods and applications. I. Methods]. *Avtomatika i telemekhanika – Automation and Remote Control*. 5. pp. 3–45. DOI: 10.1023/A:1023684619933
2. Ott, E., Grebogi, C. & Yorke, G. (1990) Controlling chaos. *Physical Review Letter*. 64. pp 1196–1199. DOI: 10.1103/PhysRevLett.64.1196
3. Pyragas, K. (1992) Continuous control of chaos by self-controlling feedback. *Physical Letter*. 170. pp. 421–428. DOI: 10.1016/0375-9601(92)90745-8
4. Magnitskiy, N.A. & Sidorov, S.V. (2004) *Novye metody khaoticheskoy dinamiki* [New Methods of Chaotic Dynamics] Moscow: Editorial URSS.
5. Yamapi, R. & Chabi Orou, J.B. (2003) Harmonic oscillations, stability and chaos control in a non-linear electromechanical system. *Journal of Sound and Vibration*. 259. pp. 1253–1264. DOI: 10.1006/jsvi.2002.5289
6. Lyapunov, A.M. (1950) *Obshchaya zadacha ob ustoychivosti dvizheniya* [The General Problem of the Stability of Motion]. Moscow: Gostekhizdat.
7. Andronov, A.A. & Pontryagin, L.S. (1937) Grubye sistemy [Rough systems]. *Doklady Akademii nauk SSSR*. 14. pp. 247–250.
8. Grobman, D.G. (1959) Gomeomorfizm sistem differentsial'nykh uravneniy [Homeomorphism of systems of differential equations]. *Doklady Akademii nauk SSSR*. 128. pp. 880–881.
9. Omorov, R.O. (2020) Metod topologicheskoy grubosti dinamicheskikh sistem: prilozheniya k sinergeticheskim sistemam [Method of topological roughness of dynamic systems: applications to synergetic systems]. *Nauchno-tekhnicheskij vestnik informatsionnykh tekhnologiy, mekhaniki i optiki*. 20(2). pp. 257–262. DOI: 10.17586/2226-1494-2020-20-2-257-262
10. Shashihin, V.N. & Budnik, S.V. (2019) Synthesis of control for nonlinear systems. *Automatic Control and Computer Sciences*. 53. pp. 97–106.
11. Kozlov, V.N. & Shashikhin, V.N. (2020) Synthesis of decentralized robust stabilizing control for the systems with parametric perturbations. *Computing, Telecommunications and Control*. 13. pp. 49–60. DOI: 10.18721/JCSTCS.13205

12. Ryseev, D.V. & Fyodorov, V.K. (2012) *Sovremennye problemy dinamiki nelineynykh energosistem: elektromekhanicheskiy rezonans, entropiya, determinirovanny khaos* [Modern Problems of Nonlinear Power Systems Dynamics: Electromechanical Resonance, Entropy, Deterministic Chaos]. Omsk: Poligraf. tsentr KAN.

Информация об авторах:

Сорокина Наталья Владимировна – старший преподаватель Высшей школы компьютерных технологий и информационных систем Института компьютерных наук и кибербезопасности Санкт-Петербургского политехнического университета Петра Великого (Санкт-Петербург, Россия). E-mail: sorokina_nv@spbstu.ru

Шашихин Владимир Николаевич – профессор, доктор технических наук, профессор Высшей школы компьютерных технологий и информационных систем Института компьютерных наук и кибербезопасности Санкт-Петербургского политехнического университета Петра Великого (Санкт-Петербург, Россия). E-mail: shashihin@bk.ru

Вклад авторов: все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

Information about the authors:

Sorokina Natalia V. (Senior Lecturer, Peter the Great Saint Petersburg Polytechnic University, Saint Petersburg, Russian Federation). E-mail: sorokina_nv@spbstu.ru

Shashikhin Vladimir N. (Doctor of Technical Sciences, Professor, Peter the Great Saint Petersburg Polytechnic University, Saint Petersburg, Russian Federation). E-mail: shashihin@bk.ru

Contribution of the authors: the authors contributed equally to this article. The authors declare no conflicts of interests.

Поступила в редакцию 27.02.2025; принята к публикации 02.06.2025

Received 27.02.2025; accepted for publication 02.06.2025

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

MATHEMATICAL MODELING

Научная статья

УДК 66.011:621.357

doi: 10.17223/19988605/71/4

Моделирование процесса поляризации электродов при контроле и управлении процессом нанесения гальванического покрытия**Алия Рифовна Ишкулова¹, Алексей Германович Лютов²**¹ Уфимский университет науки и технологий, Уфа, Россия² МИРЭА – Российский технологический университет, Москва, Россия¹ aliya.ishkulova@yandex.ru² lutov1@mail.ru

Аннотация. Рассматривается математическая модель процесса осаждения гальванического покрытия, отличительной особенностью которой является учет концентрационной поляризации. Представлены результаты исследований зависимости между поляризацией электродов, равномерностью распределения покрытия и параметрами процесса, такими как плотность тока, расстояние между электродами, скорость потока и температура электролита. Подтверждена гипотеза взаимосвязи равномерности покрытия от катодной поляризации. Установлено, что наилучшее распределение покрытия обеспечивается при более высоких значениях электродной поляризации.

Ключевые слова: математическое моделирование; гальваническое покрытие; электродная поляризация.

Для цитирования: Ишкулова А.Р., Лютов А.Г. Моделирование процесса поляризации электродов при контроле и управлении процессом нанесения гальванического покрытия // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2025. № 71. С. 39–47. doi: 10.17223/19988605/71/4

Original article

doi: 10.17223/19988605/71/4

Modeling of the electrode polarization process in monitoring and controlling the process of applying galvanic coating**Aliya R. Ishkulova¹, Alexey G. Lyutov²**¹ Ufa University of Science and Technology, Ufa, Russian Federation² MIREA – Russian Technological University, Moscow, Russian Federation¹ aliya.ishkulova@yandex.ru² lutov1@mail.ru

Abstract. The mathematical model of the process of deposition of galvanic coating is considered, the distinctive feature of which is the consideration of concentration polarization. The results of studies of the relationship between the polarization of the electrodes, the uniformity of the coating distribution and the process parameters, such as the current density, the distance between the electrodes, the flow rate and the temperature of the electrolyte are presented. The hypothesis of the relationship between the uniformity of the coating and the cathode polarization is confirmed. It is established that the best distribution of the coating is ensured at higher values of electrode polarization.

Keywords: mathematical modeling; galvanic coating; electrode polarization.

For citation: Ishkulova, A.R., Lyutov, A.G. (2025) Modeling of the electrode polarization process in monitoring and controlling the process of applying galvanic coating. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 71. pp. 39–47. doi: 10.17223/19988605/71/4

Введение

Гальванические покрытия (ГП) играют важную роль в различных сферах промышленности. Гальванический способ нанесения покрытий, основанный на осаждении чистых металлов или сплавов на поверхность изделий, открывает широкие возможности для улучшения их эксплуатационных характеристик. В частности, ГП значительно повышают твердость, износостойкость, коррозионную стойкость и декоративные свойства изделий. Однако получение покрытий с требуемыми свойствами и показателями качества представляет собой сложную задачу, зависящую от большого количества факторов. Помимо этого, сложность усугубляется тем, что в процессе электроосаждения невозможно измерить и определить качество покрытия, что делает невозможным своевременную корректировку режима электролиза. Таким образом, разработка метода косвенного определения или методов прогнозирования качества ГП в процессе нанесения покрытий остается актуальной задачей.

Один из основных показателей качества ГП – толщина покрытия, которая, в свою очередь, оказывает влияние и на защитные свойства покрытия, и на такие показатели качества, как твердость, износостойкость. Расчет параметров и длительности процесса осаждения покрытия для получения заданной толщины производится путем выполнения несложных расчетов. Однако распределение толщины покрытия на различных участках деталей может значительно отличаться: так, на более близких к анодам участках толщина будет больше, а на отдаленных участках и внутри отверстий, пазов толщина будет меньше, соответственно, не будет обладать требуемыми свойствами и обеспечивать необходимую защиту от коррозии. Помимо геометрической формы изделий на равномерность распределения толщины покрытий значительное влияние оказывают режим электролиза, состав и свойства электролита, наличие или отсутствие перемешивания электролита, расположение электродов в ванне.

1. Постановка гипотезы

Решению задачи улучшения равномерности ГП посвящен ряд работ, предлагающих различные методы управления параметрами процесса [1–7], а также управления расположением электродов в гальванической ванне [8–11]. Однако ни один из предложенных методов не предлагает измерения распределения толщины покрытия в процессе осаждения или косвенного определения равномерности покрытия, что затрудняет разработку эффективного алгоритма управления параметрами для улучшения качества ГП.

Во время осаждения покрытия можно контролировать и учесть при управлении такие параметры процесса, как температура, удельная проводимость, плотность и вязкость электролита, а также концентрация компонентов. Но изменения этих параметров в ходе нанесения гальванического покрытия происходят медленно, и они не могут в полной мере описать процесс распределения электрического потенциала на поверхности катода, от которого непосредственно зависит распределение толщины покрытия. Кроме того, они не могут в реальном времени отображать процессы, происходящие в электрохимической системе. Однако в процессе нанесения ГП может быть замерен такой параметр, как электродная поляризация, характеризующий процесс электрохимической реакции, протекающий на электродах, при подаче на них электрического тока [12].

Под действием электрического тока в электрохимической системе возникают необратимые процессы, вызывающие скачок потенциала на поверхности электрода. Данное явление называется поляризацией. Скачок потенциала происходит вследствие протекания процессов, которые приводят к переходу электронов и ионов осаждаемого металла через двойной электрический слой (ДЭС) и последующему его электровосстановлению на поверхности катода.

Скачок потенциала на поверхности катода возникает вследствие затруднений, возникающих в электрохимической системе в результате переноса носителей заряда через ДЭС, действия так назы-

ваемой поляризации разряда. Помимо этого, возникают затруднения, связанные с встраиванием в кристаллическую решетку адсорбированных атомов, разорвавшихся после разряда, данное явление получило название кристаллизационной поляризации, а также затруднения, возникающие вследствие замедления подвода из общего объема электролита частиц, разряжающихся на электроде, и низкой скорости реакции, предшествующей разряду иона, которые получили название диффузионной и реакционной поляризации соответственно. Суммарно диффузионная и реакционная поляризации вызывают изменение концентрации частиц вблизи электрода, по этой причине данные виды поляризации объединяют под общим определением концентрационной поляризации [13].

Влияние поляризационных процессов на кристаллическую структуру покрытий, соответственно, и на твердость, износостойкость, антикоррозионные свойства, установлена и доказана [14], но влияние поляризации на равномерность ГП до сих пор остается недостаточно изученным.

2. Математическое моделирование

На процесс электроосаждения покрытия оказывают влияние одновременно все виды поляризации, однако учет их всех является довольно сложной задачей и ввиду влияния множества стохастических факторов вряд ли может быть выражен математически. При этом, имея представление о концентрации электролита, его свойствах, температуре и приложенном на электроды токе, можно решить задачу учета влияния концентрационной поляризации на процесс осаждения ГП.

Концентрационная поляризация $\Delta\phi$ является сдвигом потенциала катода ϕ от его равновесного значения $\phi_{\text{равн}}$ при прохождении тока, который под действием электромагнитного поля приводит к миграции ионов металла из объема электролита к поверхности катода. В результате их концентрация у поверхности катода увеличивается, что приводит к затруднениям разряда ионов, соответственно, к отклонению потенциала электрода ϕ от равновесного значения $\phi_{\text{равн}}$:

$$\Delta\phi = \phi_{\text{равн}} - \phi.$$

Концентрационная поляризация $\Delta\phi$ описывается уравнением Нернста [12]. Для расчета $\Delta\phi$ в математической модели была применена форма уравнения Нернста, которая представляет собой соотношение между плотностью тока i и предельной диффузионной плотностью тока i_d [12]:

$$\Delta\phi = \frac{RT}{nF} \ln \left(1 - \frac{i}{i_d} \right),$$

где R – универсальная газовая постоянная, $R = 8,314 \text{ м}^2 \cdot \text{кг} / \text{с}^2 \cdot \text{К} \cdot \text{моль}$; T – температура, К; n – число электронов, переносимых через границу электрод–электролит при образовании одного атома металла; F – число Фарадея, $F = 96\,485 \text{ Кл/моль}$.

Концентрация ионов металла C_0 в математической модели учитывается путем расчета предельной диффузионной плотности тока i_d . Этот параметр характеризует максимальную плотность тока, при которой скорость доставки ионов металла из объема электролита к поверхности катода соответствует скорости их разряда. Расчет i_d производится по формуле [12]

$$i_d = \frac{nFDC_0}{\delta},$$

где D – коэффициент диффузии; δ – толщина диффузионного слоя, мкм.

Коэффициент диффузии D характеризует электрическую подвижность ионов металла и рассчитывается по уравнению Нерста–Эйнштейна [15]:

$$D = \frac{M_p RT}{nF^2},$$

здесь M_p – молярная электрическая проводимость, $\text{См} \cdot \text{м}^2 / \text{моль}$, которая вычисляется как

$$M_p = \frac{1000}{\rho(C_0/A)},$$

где A – атомная масса, осаждаемого металла, г/моль.

Толщина диффузионного слоя δ рассчитывается по формуле [16]

$$\delta(x, y, z) = 0,26 \sqrt{\frac{x_K(x, y, z)v}{U_0}} \sqrt[3]{\frac{D}{v}},$$

где $x_K(x, y, z)$ – ближайшее расстояние до края электрода; v – кинематическая вязкости электролита, Па·с; U_0 – скорость потока, см/с.

Распределение плотности тока $i(x, y, z)$ на поверхности детали за интервал времени τ рассчитывается по закону Ома:

$$i(x, y, z, \tau) = \frac{1}{\rho} \cdot \frac{\partial \varphi(x, y, z, \tau)}{\partial n}, \quad (1)$$

где ρ – удельное электросопротивление электролита, Ом·м; $\varphi(x, y, z, \tau)$ – потенциал электрического поля в точке (x, y, z) поверхности катода, В; n – направление внешней нормали к поверхности катода.

Распределение электрического потенциала в объеме электролита рассчитывается путем решения уравнения Лапласа при заданных величинах потенциалов электродов:

$$\frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} + \frac{\partial^2 \varphi}{\partial z^2} = 0, \quad (2)$$

где φ – электрический потенциал, В.

Граничными условиями для решения уравнения Лапласа (2) являются потенциалы электродов. В начальный момент времени они имеют следующий вид:

$$\varphi_a^1 = \varphi_a^0 - \Delta \varphi_a^0 \left(\frac{1}{\rho} \cdot \frac{\partial \varphi}{\partial n} \right) \Big|_{S_a} = U, \quad \varphi_k^1 = \varphi_k^0 + \Delta \varphi_k^0 \left(\frac{1}{\rho} \cdot \frac{\partial \varphi}{\partial n} \right) \Big|_{S_k} = 0, \quad (3)$$

где φ_a^1, φ_k^1 – анодный и катодный потенциал соответственно в начальный момент времени, В; U – напряжение в электролизере, В; $\Delta \varphi_a^0, \Delta \varphi_k^0$ – анодное и катодное перенапряжение в начальный момент времени, В; S_a, S_k – поверхности анодов и катодов соответственно, см².

Расчет напряжения U в электролизере производится по закону Ома при известных расстоянии между электродами l и удельной проводимости электролита ρ :

$$U = il\rho, \quad (4)$$

где i – рабочая плотность тока, А/дм².

Для каждого момента времени j производится перерасчет граничных условий на основании формул (3), которые принимают следующий вид:

$$\varphi_a^j = \varphi_a^{j-1} - \Delta \varphi_a^{j-1} \left(\frac{1}{\rho} \cdot \frac{\partial \varphi}{\partial n} \right) \Big|_{S_a}, \quad \varphi_k^j = \varphi_k^{j-1} + \Delta \varphi_k^{j-1} \left(\frac{1}{\rho} \cdot \frac{\partial \varphi}{\partial n} \right) \Big|_{S_k}. \quad (5)$$

На основании рассчитанного, с использованием уравнения Лапласа (2) распределения потенциала $\varphi(x, y, z, \tau)$ на поверхности катода при граничных условиях (3) для начального момента времени и (5) для последующих и определенной плотности тока $i(x, y, z, \tau)$ по уравнению (1), по закону Фарадея рассчитывается толщина осажденного покрытия $h_n(x, y, z, \tau)$ в некоторой точке поверхности катода (x, y, z) за интервал времени τ :

$$h(x, y, z, \tau) = \frac{i(x, y, z, \tau) \cdot \Theta \cdot \tau \cdot VT}{d}$$

где Θ – электрохимический эквивалент металла, мг/Кл, VT – выход металла по току, %, d – плотность металла, г/см³.

Оценка равномерности толщины покрытия на поверхности катода производится путем расчета коэффициента K_p , который определяется по формуле

$$K_p(t) = \frac{1}{n} \sum_{i=1}^n \frac{h_{\min}(t)}{h_i(x, y, z, t)},$$

где $h_{\min}$ – минимальная толщина покрытия; h_i – толщина покрытия в точке (x, y, z) поверхности катода, n – количество точек определения толщины покрытия на поверхности катода.

Проверка адекватности расчета поляризации предложенной в данной работе математической моделью производилась путем сравнения поляризационных кривых, представленных в справочниках [18, 19], с результатами расчета математической модели. Сравнение справочных и расчетных данных производилось для процесса электролиза хромового покрытия в электролите, содержащем 250 г/л CrO_3 и 2,5 г/л H_2SO_4 , при температурах электролита 18 и 60°C. Сравнительные справочные и расчетные поляризационные кривые представлены на рис. 1.

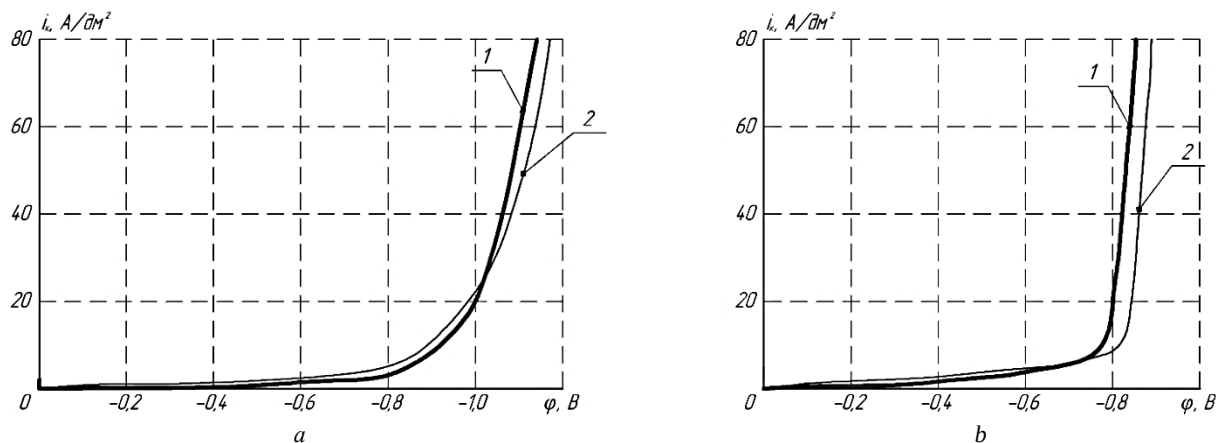


Рис. 1. Сравнительные справочные (1) и расчетные (2) поляризационные кривые электролиза хромового покрытия в электролите, содержащем 250 г/л CrO_3 и 2,5 г/л H_2SO_4 , при температурах электролита 18°C (а) и 60°C (б)
Fig. 1. Comparative reference (1) and calculated (2) polarization curves of electrolysis of a chromium coating in an electrolyte containing 250 g/l CrO_3 and 2.5 g/l H_2SO_4 at electrolyte temperatures of 18°C (a) and 60°C (b)

Адекватность математической модели была оценена путем вычисления при сопоставлении справочных и расчетных данных. Для эксперимента (а) этот показатель составил 10,2%, а для эксперимента (б) – 12,9%. Средняя квадратичная относительная погрешность для обоих экспериментов не превысила 15%. На основании этого можно сделать вывод, что результаты расчета математической модели соответствуют справочным данным.

Таким образом, предложенная в данной работе математическая модель может быть использована для проведения исследований и установления зависимостей равномерности покрытий от поляризации при различных режимах осаждения покрытий.

3. Экспериментальная часть

В данной работе проведено исследование зависимости равномерности ГП от величины катодной поляризации с использованием предложенной математической модели. В ходе исследования были проведены вычислительные эксперименты при различных значениях плотности тока, расстояния между электродами, скорости потока и температуры электролита.

Исследования производились на примере осаждения хромового покрытия в сульфатном электролите CrO_3 (250 г/л), H_2SO_4 (25 г/л) с удельной электропроводностью $0,0166 \text{ Ом}^{-1}\cdot\text{см}^{-1}$ и выходом металла по току 35% при следующих диапазонах изменения параметров электролиза:

- плотность тока – 45–55 А/дм^2 ;
- расстояние между электродами – 5–20 см;
- интенсивность перемешивания электролита – 5–50 см/с;
- температура электролита – 20–60°C.

Равномерность ГП определяется путем расчета коэффициента равномерности (13). Катодная поляризация рассчитывается по формуле (2). Оба параметра определялись на основе рассчитанного распределения плотности тока на поверхности катода, которое было получено путем решения уравнения Лапласа (8).

Результаты проведенных исследований представлены на рис. 2–5.

Графики, представленные на рис. 2–5, демонстрируют зависимости равномерности ГП и катодной поляризации от плотности тока, расстояния между электродами, скорости потока и температуры электролита. Кроме того, сопоставление графиков подтверждает взаимосвязь между равномерностью покрытия и катодной поляризацией.

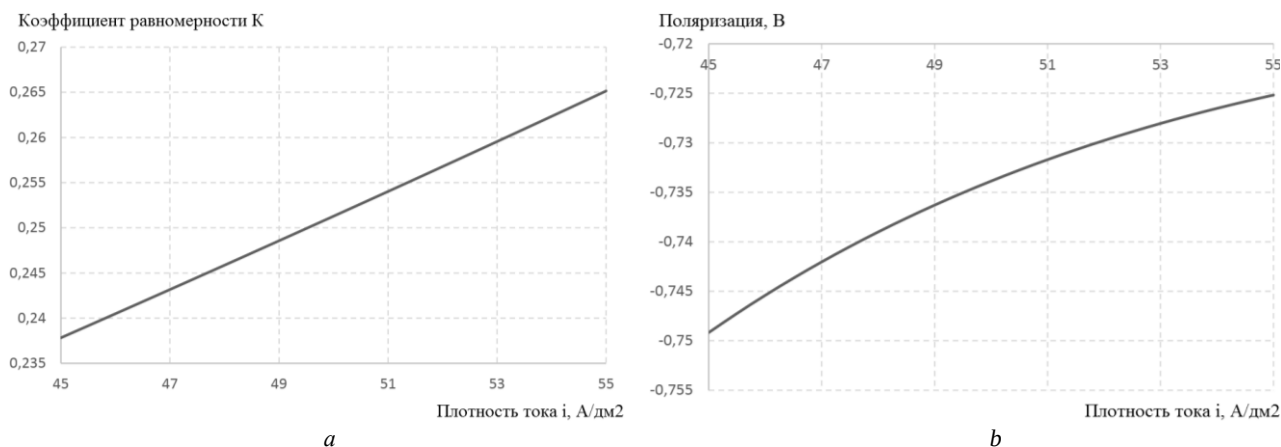


Рис. 2. Зависимости равномерности покрытия (a) и катодной поляризации (b) от плотности тока
Fig. 2. Dependences of coating uniformity (a) and cathode polarization (b) on current density



Рис. 3. Зависимости равномерности покрытия (a) и катодной поляризации (b) от межэлектродного расстояния
Fig. 3. Dependences of coating uniformity (a) and cathodic polarization (b) on the interelectrode distance

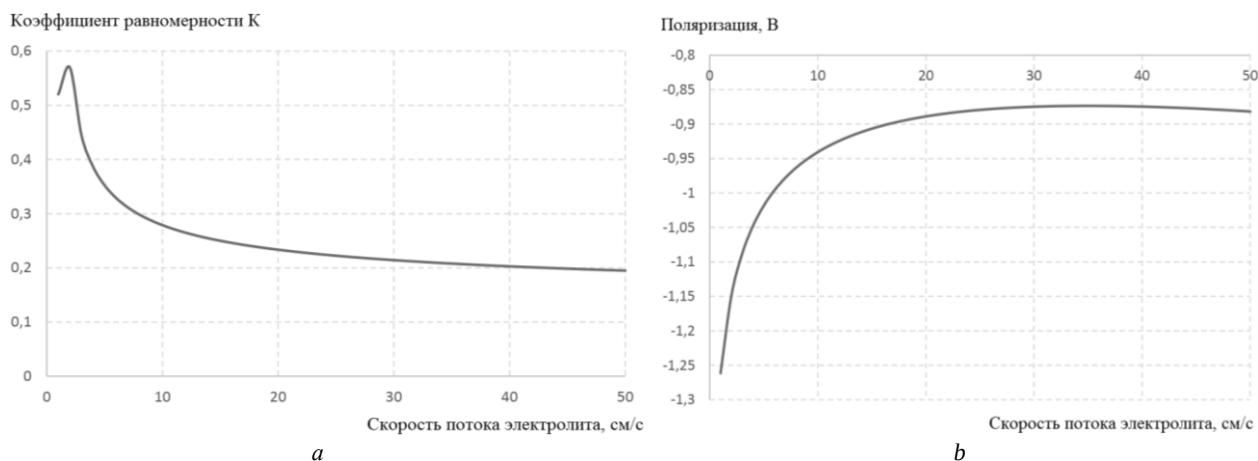


Рис. 4. Зависимости равномерности покрытия (a) и катодной поляризации (b) от скорости потока электролита
Fig. 4. Dependences of coating uniformity (a) and cathodic polarization (b) on the electrolyte flow rate

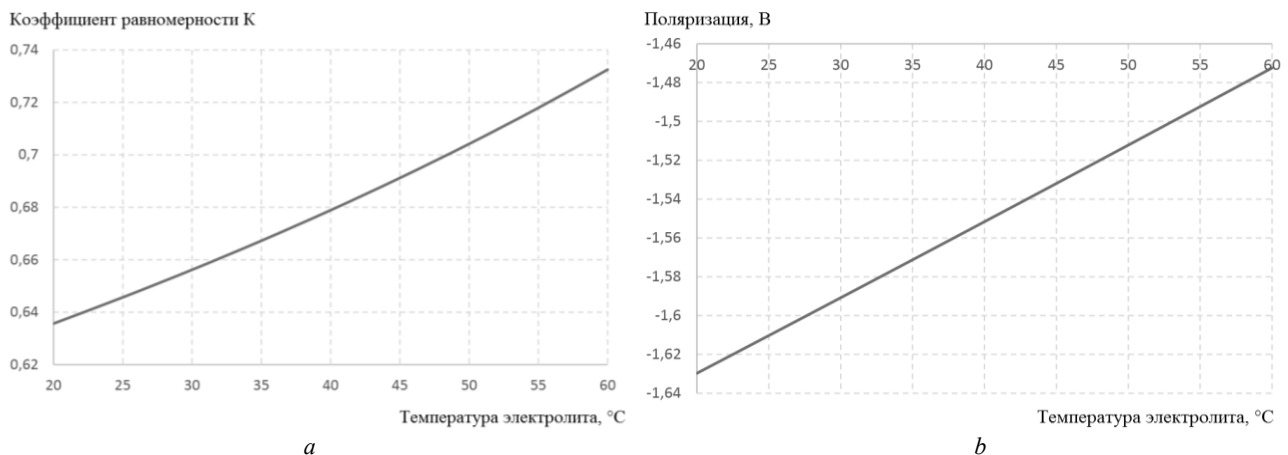


Рис. 5. Зависимости равномерности покрытия (a) и катодной поляризации (b) от температуры электролита
Fig. 5. Dependences of coating uniformity (a) and cathodic polarization (b) on the electrolyte temperature

Из графиков видно, что при более высоких значениях катодной поляризации обеспечивается более равномерное распределение толщины покрытия. Это позволяет предположить, что поддержание более высоких значений катодной поляризации может привести к получению более равномерных покрытий. Однако эти результаты были получены при постоянных значениях скорости потока электролита. При этом зависимости от скорости потока электролита имеют обратную тенденцию по сравнению с зависимостями от других параметров. Это можно объяснить тем, что при низкой интенсивности перемешивания электролита толщина диффузионного слоя больше, особенно на активных участках поверхности катода, где плотность тока выше. В результате межфазовый переход разряжающегося металла затруднен, что приводит к более низким значениям катодной поляризации.

Закключение

В данной работе была выдвинута гипотеза о зависимости равномерности распределения толщины гальванических покрытий на поверхности катода от его поляризации. Для подтверждения данной гипотезы предложена математическая модель, уникальность которой состоит в учете электродной поляризации.

Проведенные исследования с использованием математической модели подтвердили зависимость равномерности покрытия от катодной поляризации.

Исходя из этого, можно сделать вывод, что контроль и регулирование данного параметра в процессе осаждения гальванических покрытий позволит получить покрытия с более высокими показателями качества.

Полученные результаты могут быть использованы для дальнейших исследований процессов осаждения гальванических покрытий и разработки алгоритмов оптимального управления процессом.

Список источников

1. Соловьев Д.С., Потлов А.Ю., Литовка Ю.В. Снижение неравномерности толщины гальванического покрытия, наносимого с циклическим включением анодных секций // Вестник машиностроения. 2016. № 8. С. 78–84.
2. Конкина В.В., Соловьев Д.С., Литовка Ю.В., Мукина И.А. Автоматизированная система управления гальваническими процессами при реверсировании тока с отключаемыми анодными секциями // Проблемы машиностроения и автоматизации. 2016. № 2. С. 67–77.
3. Соловьев Д.С., Мукина И.А., Литовка Ю.В. Особенности оптимального управления гальваническими процессами в многоанодной ванне с различными значениями силы тока // Мехатроника, автоматизация, управление. 2017. Т. 18, № 9. С. 631–636.
4. Литовка Ю.В., Кондрашин Р.С. Моделирование и оптимизация импульсных токовых режимов нанесения гальванических покрытий // Вестник Тамбовского государственного технического университета. 1999. Т. 5, № 1. С. 26–33.
5. Литовка Ю.В., Романенко А.В., Афанасьев А.В. Моделирование и оптимизация процесса нанесения гальванических покрытий в условиях реверсирования тока // Теоретические основы химической технологии. 1998. Т. 32, № 3. С. 301–304.

6. Лютов А.Г., Ишкулова А.Р. Многопараметрическое оптимальное управление процессом нанесения гальванического покрытия с учетом изменения условий электролиза // Вестник Уфимского государственного авиационного технического университета. 2019. Т. 23, № 2 (84). С. 138–144.
7. Лютов А.Г., Ишкулова А.Р. Автоматизированная система экстремального управления процессом нанесения гальванического покрытия в ванне с многосекционными анодами // Мехатроника, автоматизация и управление. 2017. Т. 19, № 3. С. 185–191. doi: 10.17587/mau.18.185-191
8. Pat. 6027631 США. Electroplating system with shields for varying thickness profile of deposited layer / E.K. Broadbent. Оpubл. 22.02.2000.
9. Пуха А.П., Литовка Ю.В., Конкина В.В. Алгоритм управления процессом нанесения гальванического покрытия с подвижными секциями анода // Математические методы в технологиях и технике. 2024. № 3. С. 18–21.
10. Пчелинцева И.Ю., Литовка Ю.В. Система автоматизированного управления процессом нанесения гальванического покрытия в ванне с токопроводящим экраном // Мехатроника, автоматизация, управление. 2022. Т. 23. № 4. С. 188–196. doi: 10.17587/mau.23.188-196
11. Банников А.А., Литовка Ю.В. Оптимальное размещение катодов в автоматизированной системе управления гальванической ванной // Вести высших учебных заведений Черноземья. 2021. № 3 (65). С. 3–9. doi: 10.53015/18159958_2021_3_3
12. Вайнер Я.В., Дасоян М.А. Технология электрохимических покрытий. Л. : Машгиз, 1962. 468 с.
13. Нечаев А.В. Химия : учеб. пособие. Екатеринбург : УрФУ, 2016. Ч. II. 112 с.
14. Каданер Л.И. Справочник по гальваностегии. Киев : Техника, 1976. 254 с.
15. Лукомский Ю.Я., Гамбург Ю.Д. Физико-химические основы электрохимии : учебник. Долгопрудный : Изд. Дом «Интеллект», 2008. 424 с.
16. Левич В.Г. Физико-химическая гидродинамика. М. : Физматгиз, 1959. 700 с.
17. Бахчисарайцян Н.Г., Борисоглебский Ю.В., Буркат Г.К. и др. Практикум по прикладной электрохимии : учеб. пособие для вузов / под ред. В.Н. Варыпаева, В.Н. Кудрявцева. 3-е изд. перераб. Л. : Химия, 1990. 304 с.
18. Мельников П.С. Справочник по гальванопокрытиям в машиностроении. 2-е изд., перераб. и доп. М. : Машиностроение, 1991. 384 с.
19. Кудрявцев Н.Т. Электролитические покрытия металлами. М. : Химия, 1979. 352 с.

References

1. Soloviev, D.S., Potlov, A.Yu. & Litovka, Yu.V. (2016) Snizhenie neravnomernosti tolshchiny gal'vanicheskogo pokrytiya, nanosimogo s tsiklicheskim vklyucheniem anodnykh sektsiy [Reducing the unevenness of the thickness of the galvanic coating applied with cyclic inclusion of anode sections]. *Vestnik mashinostroeniya*. 8. pp. 78–84.
2. Konkina, V.V., Soloviev, D.S., Litovka, Yu.V. & Mukina, I.A. (2016) Avtomatizirovannaya sistema upravleniya gal'vanicheskimi protsessami pri reversirovanii toka s otklyuchaemyimi anodnymi sektsiyami [Automated control system for galvanic processes with current reversal with disconnectable anode sections]. *Problemy mashinostroeniya i avtomatizatsii*. 2. pp. 67–77.
3. Soloviev, D.S., Mukina, I.A. & Litovka, Yu.V. (2017) Osobennosti optimal'nogo upravleniya gal'vanicheskimi protsessami v mnogoanodnoy vanne s razlichnymi znacheniyami sily toka [Optimal control of galvanic processes in a multi-anode bath with different current values]. *Mekhatronika, avtomatizatsiya, upravlenie*. 18(9). pp. 631–636.
4. Litovka, Yu.V. & Kondrashin, R.S. (1999) Modelirovanie i optimizatsiya impul'snykh tokovykh rezhimov naneseniya gal'vanicheskikh pokrytiy [Modeling and optimization of pulsed current modes of applying galvanic coatings]. *Vestnik Tambovskogo gosudarstvennogo tekhnicheskogo universiteta*. 5(1). pp. 26–33.
5. Litovka, Yu.V., Romanenko, A.V. & Afanasyev, A.V. (1998) Modelirovanie i optimizatsiya protsessa naneseniya gal'vanicheskikh pokrytiy v usloviyakh reversirovaniya toka [Modeling and optimization of the process of applying galvanic coatings under current reversal conditions]. *Teoreticheskie osnovy khimicheskoy tekhnologii*. 32(3). pp. 301–304.
6. Lyutov, A.G. & Ishkulova, A.R. (2019) Mnogoparametricheskoe optimal'noe upravlenie protsessom naneseniya gal'vanicheskogo pokrytiya s uchetoм izmeneniya usloviy elektroliza [Multiparameter optimal control of the process of applying galvanic coating taking into account changing electrolysis conditions]. *Vestnik Ufimskogo gosudarstvennogo aviatsionnogo tekhnicheskogo universiteta*. 23(2(84)). pp. 138–144.
7. Lyutov, A.G. & Ishkulova, A.R. (2017) Avtomatizirovannaya sistema ekstremal'nogo upravleniya protsessom naneseniya gal'vanicheskogo pokrytiya v vanne s mnogosektsionnymi anodami [Automated system for extreme control of the process of applying galvanic coating in a bath with multisection anodes]. *Mekhatronika, avtomatizatsiya i upravlenie*. 19(3). pp. 185–191. DOI: 10.17587/mau.18.185-191
8. Patent. 6027631 USA. *Electroplating system with shields for varying thickness profile of deposited layer*. E. K. Broadbent Publ. 22.02.2000.
9. Pukha, A.P., Litovka, Yu.V. & Konkina, V.V. (2024) Algoritм upravleniya protsessom naneseniya gal'vanicheskogo pokrytiya s podvizhnymi sektsiyami anoda [Algorithm for controlling the process of applying a galvanic coating with movable anode sections]. *Matematicheskie metody v tekhnologiyakh i tekhnike*. 3. pp. 18–21.
10. Pchelintseva, I.Yu. & Litovka, Yu.V. (2022). Sistema avtomatizirovannogo upravleniya protsessom naneseniya gal'vanicheskogo pokrytiya v vanne s tokoneprovodyashchim ekranom [Automated control system for the process of applying a galvanic coating in a bath with a non-conductive screen]. *Mekhatronika, avtomatizatsiya, upravlenie*. 23(4). pp. 188–196. DOI: 10.17587/mau.23.188-196

11. Bannikov, A.A. & Litovka, Yu.V. (2021) Optimal'noe razmeshchenie katodov v avtomatizirovannoy sisteme upravleniya gal'vanicheskoy vannoy [Optimal placement of cathodes in an automated control system for a galvanic bath]. *Vesti vysshih uchebnykh zavedeniy Chernozem'ya*. 3(65). pp. 3–9. DOI: 10.53015/18159958_2021_3_3
12. Vayner, Ya.V. & Dasoyan, M.A. (1962) *Tekhnologiya elektrokhimicheskikh pokrytiy* [Technology of Electrochemical Coatings]. Leningrad: Mashgiz.
13. Nechaev, A.V. (2016) *Khimiya* [Chemistry]. Yekaterinburg: UrFU.
14. Kadaner, L.I. (1976) *Spravochnik po gal'vanostegii* [Handbook of Electroplating]. Kyiv: Tekhnika.
15. Lukomsky, Yu.Ya. & Gambur, Yu.D. (2008) *Fiziko-khimicheskie osnovy elektrokhimii* [[Physicochemical Foundations of Electrochemistry]. Dolgoprudnyy: Intellect.
16. Levich, V.G. (1959) *Fiziko-khimicheskaya gidrodinamika* [Physicochemical Hydrodynamics]. Moscow: Fizmatgiz.
17. Bakhchisaraytsyan, N.G., Borisoglebskiy, Yu.V. & Burkat, G.K. (1990) *Praktikum po prikladnoy elektrokhimii* [Practical Training in Applied Electrochemistry]. Leningrad: Khimiya.
18. Melnikov, P.S. (1991) *Spravochnik po gal'vanopokrytiyam v mashinostroenii* [Handbook of Electroplating in Mechanical Engineering]. Moscow: Mashinostroyeniye.
19. Kudryavtsev, N.T. (1979) *Elektroliticheskie pokrytiya metallami* [Electrolytic Coatings of Metals]. Moscow: Khimiya.

Информация об авторах:

Ишкулова Алия Рифовна – кандидат технических наук, доцент кафедры технологии производства летательных аппаратов филиала Уфимского университета науки и технологий в г. Кумертау (Уфа, Россия). E-mail: aliya.ishkulova@yandex.ru

Лютов Алексей Германович – профессор, доктор технических наук, заведующий кафедрой автоматических систем Института искусственного интеллекта МИРЭА – Российского технологического университета (Москва, Россия). E-mail: lutov1@mail.ru

Вклад авторов: все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

Information about the authors:

Ishkulova Aliya R. (Candidate of Technical Sciences, Associate Professor of the Department of Aircraft Production Technology of the Branch of the Ufa University of Science and Technology in Kumertau, Ufa, Russian Federation). E-mail: aliya.ishkulova@yandex.ru

Lyutov Aleksey G. (Professor, Doctor of Technical Sciences, Head of the Department of Automatic Systems of the Institute of Artificial Intelligence MIREA – Russian Technological University, Moscow, Russian Federation). E-mail: lutov1@mail.ru

Contribution of the authors: the authors contributed equally to this article. The authors declare no conflicts of interests.

Поступила в редакцию 25.02.2025; принята к публикации 02.06.2025

Received 25.02.2025; accepted for publication 02.06.2025

Научная статья
УДК 621.928.37
doi: 10.17223/19988605/71/5

Математическая модель и анализ влияния режимных параметров гидроциклона на показатели процесса разделения вязкопластических сред двухстадийной флотацией

Владимир Олегович Яблонский

Волгоградский государственный технический университет, Волгоград, Россия, everest58@mail.ru

Аннотация. Разработана математическая модель разделения вязкопластических сред двухстадийной напорной флотацией в гидроциклоне. Выполнено моделирование поля концентраций частиц твердой фазы, и установлено влияние режимных параметров на показатели работы двухстадийного гидроциклона – флотатора, а также определены оптимальные условия проведения процесса. Показано, что двухстадийная напорная флотация позволяет получить низкую остаточную концентрацию частиц твердой фазы в пленке суспензии. Влияние фактора разделения является более значительным для сред с высоким предельным напряжением сдвига. Установлено, что процесс двухстадийной флотации протекает наиболее эффективно при высоких значениях фактора разделения и выраженных пластических свойствах среды.

Ключевые слова: математическая модель; гидроциклон; разделение суспензий; вязкопластическая жидкость; предел текучести; двухстадийная флотация.

Для цитирования: Яблонский В.О. Математическая модель и анализ влияния режимных параметров гидроциклона на показатели процесса разделения вязкопластических сред двухстадийной флотацией // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2025. № 71. С. 48–57. doi: 10.17223/19988605/71/5

Original article
doi: 10.17223/19988605/71/5

Mathematical model and analysis of the effect of the regime parameters of the hydrocyclone on the indicators of the separation process of viscoplastic media by two-stage flotation

Vladimir O. Yablonskii

Volgograd State Technical University, Volgograd, Russian Federation, everest58@mail.ru

Abstract. A mathematical model of separation of viscoplastic media by two-stage pressure flotation in a hydrocyclone has been developed. The simulation of the field of concentrations of solid phase particles was performed and the influence of operating parameters on the performance of the two - stage hydrocyclone - flotator was established, and the optimal conditions for the process were determined. It is shown that two-stage pressure flotation allows to obtain a low residual concentration of solid phase particles in the suspension film with an increase in the concentration of secondary gas. The influence of the separation factor is more significant for media with high bound shear stress limits. It is established that the flotation process proceeds most effectively at high values of the separation factor and plastic properties of the medium.

Keywords: mathematical model; hydrocyclone; separation of suspensions; viscoplastic fluid; yield strength; two-stage flotation.

For citation: Yablonsky, V.O. (2025) Mathematical model and analysis of the effect of the regime parameters of the hydrocyclone on the indicators of the separation process of viscoplastic media by two-stage flotation. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 71. pp. 48–57. doi: 10.17223/19988605/71/5

Введение

Гидроциклоны, отличающиеся простотой устройства и высокой надежностью, применяются для процессов разделения в химической и других отраслях промышленности. Необходимость интенсификации процессов и достижения высокой степени очистки обрабатываемых сред диктует необходимость применения новых высокоэффективных методов, к которым относится двухстадийная напорная флотация. Процесс напорной флотации широко применяется для разделения продуктов биосинтеза, извлечения тонкодисперсной твердой фазы из суспензий и очистки различных технологических жидкостей.

Многие из разделяемых в химической промышленности продуктов являются вязкопластическими средами, имеющими выраженный предел текучести (предельное напряжение сдвига), характеризующий аномалию неньютоновских свойств среды. К ним относятся суспензии глазури, биомассы белково-витаминных концентратов (паприн, гаприн, меприн), карбид кремния, сточные воды, содержащие нефть, жир и другие вещества, а также буровые растворы и смазочные масла с присадками.

Опубликованные в настоящее время работы по моделированию гидродинамики закрученных пленочных течений вязкопластических сред в поле центробежных сил [1, 2] создают возможность разработки математической модели разделения таких сред двухстадийной напорной флотацией в гидроциклоне.

Авторам работы [3] представляется наиболее целесообразным вместо методов напорной флотации подавать воздух в питающий патрубок гидроциклона, что позволяет начать взаимодействие частиц с пузырьками уже на входе в гидроциклон и значительно упрощает конструкцию аппарата. Однако в данной работе не принимается во внимание увеличение кинетической константы флотации за счет выделения пузырьков газа непосредственно на частицах твердой фазы при напорной флотации. Поэтому для улучшения кинетических характеристик процесса напорной флотации и существенного уменьшения остаточной концентрации частиц твердой фазы целесообразно применение процесса двухстадийной флотации с подачей вторичного газа через пористые стенки гидроциклона, объединяющего преимущества методов центробежного разделения и флотационной очистки.

В работе [4] на основании решения уравнения радиального движения системы частица–пузырек предложена методика расчета показателей разделения гидроциклона–флотатора, а также диаметра цилиндрической части аппарата, при котором обеспечиваются заданные показатели разделения и минимальные капитальные и энергетические затраты. Получено условие устойчивого всплывания комплекса частица–пузырек, определен диаметр частиц твердой фазы, которые могут быть удалены при данных режимных параметрах гидроциклона.

В [5, 6] указывается на зависимость числа вероятных столкновений частиц твердой фазы и пузырьков газа от центробежного ускорения, показано, что центробежное поле является фактором, увеличивающим число столкновений частиц и пузырьков. Установлено, что в условиях напорной флотации преобладающим механизмом образования флотокомплексов является столкновение частиц твердой фазы с пузырьками газа. Эффективность процесса напорной флотации, зависящая от числа столкновений, определяется расстоянием, проходимым частицей твердой фазы в жидкой дисперсионной среде.

Наиболее благоприятные условия для разделения вязкопластических сред напорной флотацией создаются в условиях тонкопленочного течения среды в поле центробежных сил, так как в данном случае имеет место эффект взаимного усиления действия обоих факторов [7].

Ввиду того, что вязкопластические среды обладают высокой эффективной вязкостью, лимитирующей стадией при осуществлении процесса напорной флотации в гидроциклоне является стадия транспортировки комплексов частица–пузырек к поверхности пленки разделяемой среды, что не позволяет достичь хороших показателей разделения. Аппаратом, обеспечивающим наибольшую эффективность проведения такого процесса, является прямоточный цилиндрический гидроциклон [8], в котором процесс разделения осуществляется в пленочном режиме и обеспечивается условие связи частицы с пузырьком газа.

В работе [8] выполнено математическое моделирование поля концентраций частиц твердой фазы при разделении вязкопластических суспензий в гидроциклоне напорной флотацией и проанализировано влияние конструкции корпуса аппарата на эффективность разделения.

Необходимость достижения высокой степени извлечения частиц твердой фазы в ряде химических производств обуславливает применение для этих целей процесса двухстадийной напорной флотации с дополнительным вводом диспергированного вторичного газа через пористую стенку корпуса гидроциклона. Сталкиваясь с флотокомплексами, пузырьки вторичного газа, имеющие значительно больший диаметр, образуют с ними вторичные комплексы, позволяющие быстро извлекать частицы твердой фазы в пенный слой и повысить показатели процесса разделения.

Цель настоящей работы – разработка математической модели разделения вязкопластических сред двухстадийной напорной флотацией в гидроциклоне и исследование влияния фактора разделения при различных реологических свойствах среды на показатели эффективности процесса.

1. Постановка задачи

Процесс двухстадийной напорной флотации протекает следующим образом (рис. 1): предварительно аэрированная суспензия при повышенном давлении (до 0,8 МПа) поступает в цилиндрический прямоточный гидроциклон тангенциально по образующей через входной патрубок 1, установленный в верхней части корпуса гидроциклона. Поступившая в гидроциклон суспензия стекает, образуя вращающуюся пленку 2, по его стенкам вниз, обладая окружной V_ϕ , осевой V_z , и радиальной V_r составляющими скорости.

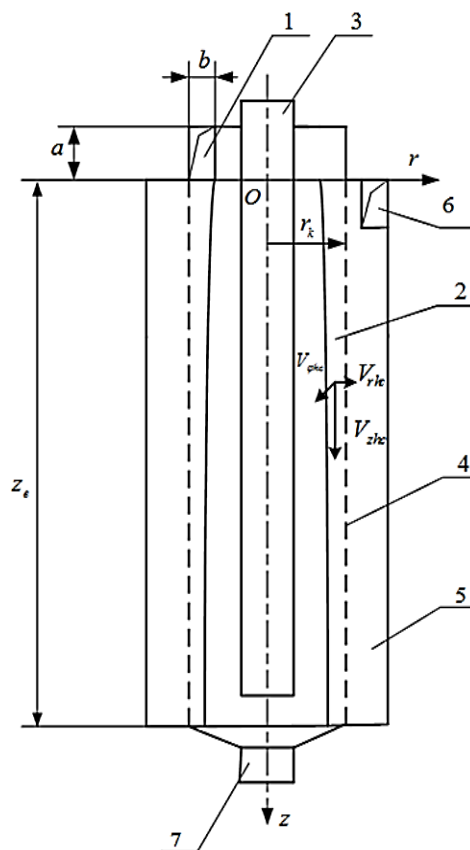


Рис. 1. Схема разделения вязкопластической среды в цилиндрическом гидроциклоне двухстадийной напорной флотацией: 1 – входной патрубок; 2 – вращающаяся пленка среды; 3 – верхний отводящий патрубок;

4 – пористая стенка корпуса; 5 – воздушный коллектор; 6 – патрубок для подачи воздуха; 7 – нижний слив

Fig. 1. Scheme of separation of a viscoplastic medium in a cylindrical hydrocyclone by two-stage pressure flotation:
1 – inlet pipe; 2 – rotating medium film; 3 – upper discharge pipe; 4 – porous body wall; 5 – air collector;
6 – air supply pipe; 7 – bottom drain

При снижении давления до атмосферного создается перенасыщение растворенного газа, и суспензия «вскипает». Частицы твердой фазы под действием центробежной силы движутся к стенке корпуса гидроциклона, а пузырьки газа (воздуха) под действием выталкивающей центростремительной силы Архимеда – навстречу им к поверхности пленки. На первой стадии напорной флотации при столкновении частиц твердой фазы с пузырьками газа происходит образование флотокомплексов, выносящих частицы твердой фазы на поверхность пленки в пенный слой, который удаляется через верхний отводящий патрубок гидроциклона 3. Флотокомплексы имеют окружную $V_{\phi hc}$, осевую V_{zhc} и радиальную V_{rhc} составляющие скорости. Пузырьки вторичного газа, подаваемого через пористую стенку корпуса 4 через воздушный коллектор 5, снабженный патрубком для подачи воздуха 6, имеющие на порядок больший диаметр, чем пузырьки газа, выделяющегося в суспензии при снижении давления, имеют значительно большую скорость всплывания к поверхности пленки, чем образовавшиеся комплексы. На второй стадии флотации пузырьки вторичного газа, сталкиваясь со всплывающими комплексами, образуют с ними вторичные комплексы, в которых частица составляет мостик между пузырьками [9], которые быстро устремляются к поверхности пленки. Осветленная суспензия удаляется из аппарата через нижний слив 7.

Полагаем, что реологические свойства вязкопластической разделяемой среды для случая трехмерного течения могут быть охарактеризованы уравнением Гершеля–Балкли [10], которое при записи в компонентах тензора скоростей деформаций с учетом условия пластичности фон Мизеса имеет вид

$$\tau_{ij} = \left(\frac{\tau_0}{A} + kA^{n-1} \right) \gamma_{ij},$$

$$A = \sqrt{2 \left(\frac{\partial V_r}{\partial r} \right)^2 + 2 \left(\frac{V_r}{r} \right)^2 + 2 \left(\frac{\partial V_z}{\partial z} \right)^2 + \left(\frac{\partial V_\phi}{\partial r} - \frac{V_\phi}{r} \right)^2 + \left(\frac{\partial V_\phi}{\partial z} \right)^2 + \left(\frac{\partial V_z}{\partial r} + \frac{\partial V_r}{\partial z} \right)^2},$$

где τ_{ij} – компоненты тензора напряжений, Па; τ_0 – предельное напряжение сдвига, Па; γ_{ij} – компоненты тензора скоростей деформаций, c^{-1} ; A – интенсивность скоростей деформаций, c^{-1} ; n – показатель нелинейности кривой течения; k – индекс консистентности, $\text{Па} \cdot c^n$; r, z – радиальная и осевая координаты, м; ϕ – окружная координата, рад.

Из-за высокой эффективной вязкости неньютоновских сред течение в цилиндрическом гидроциклоне осуществляется в пленочном режиме [2], что обеспечивает устойчивую связь комплекса частица–пузырек ввиду отсутствия резкого возрастания окружной составляющей скорости вблизи оси гидроциклона.

Математическая модель первой стадии процесса напорной флотации рассмотрена в работе [8]. Математическая постановка задачи для двухстадийной флотации производится при следующих допущениях. Поскольку диаметр пузырьков вторичного газа, подаваемого через пористую стенку корпуса гидроциклона, на порядок превышает диаметр пузырьков газа, выделяющихся на первой стадии напорной флотации, эффективность столкновения свободных частиц твердой фазы с пузырьками вторичного газа согласно формуле Сазерленда [9] мала и ею можно пренебречь. Таким образом, можно считать, что пузырьки вторичного газа, имеющие большую скорость всплывания к поверхности пленки, чем комплексы, образовавшиеся на первой стадии, сталкиваются с ними и, образуя вторичные комплексы, извлекают их в пенный слой.

Для математического моделирования поля концентраций частиц твердой фазы, связанных в комплексы, на второй стадии напорной флотации при разделении вязкопластических суспензий в цилиндрическом гидроциклоне используется дифференциальное уравнение конвективной диффузии в цилиндрических координатах, которое с учетом притока частиц, связанных в комплексы, за счет их образования на первой стадии напорной флотации и стока за счет образования вторичных комплексов на второй стадии может быть записано в векторной форме для частиц твердой фазы и пузырьков вторичного газа следующим образом:

$$\begin{aligned} \operatorname{div}(\bar{V}_c c_{hc}) &= J_{hc1} - J_{hc2}, \\ \operatorname{div}(\bar{V}_{g2} c_{g2}) &= -J_{g2}, \end{aligned} \quad (1)$$

где $\bar{V}_c$ – вектор скорости комплексов; $\bar{V}_{g2}$ – вектор скорости пузырьков вторичного газа; c_{hc} – концентрация частиц твердой фазы, связанных в комплексы, $\text{м}^3/\text{м}^3$; c_{g2} – концентрация пузырьков вторичного газа, $\text{м}^3/\text{м}^3$; J_{hc1} – приток частиц твердой фазы, связанных в комплексы, за счет первой стадии флотации; J_{hc2} – сток частиц твердой фазы, связанных в комплексы, за счет второй стадии флотации; J_{g2} – сток пузырьков вторичного газа за счет второй стадии флотации.

Наиболее распространенным в практике гидроциклонной флотации является случай, когда концентрация пузырьков вторичного газа в пленке суспензии значительно превышает концентрацию частиц твердой фазы, связанных в комплексы, $c_{g2} > c_h$, и можно полагать $c_{g2} = \text{const}$ во всем объеме пленки суспензии в гидроциклоне. Концентрация пузырьков газа, выделяющихся из суспензии на первой стадии флотации, также существенно превышает концентрацию свободных частиц твердой фазы $c_{g1} > c_h$. Для этого случая второе уравнение системы (1) можно исключить, и система (1) сводится к одному уравнению, в котором сток частиц твердой фазы, связанных в комплексы, за счет флотации пузырьками вторичного газа J_{hc2} и их приток за счет первой стадии флотации J_{hc1} согласно [11, 12] имеют вид, аналогичный кинетическому уравнению химической реакции первого порядка:

$$\bar{V}_c \text{grad } c_{hc} + c_{hc} \text{div } \bar{V}_c = A_1 c_h c_{g1} - A_2 c_{hc} c_{g2}, \quad (2)$$

где A_1 – кинетическая константа первой стадии флотации, с^{-1} ; A_2 – кинетическая константа второй стадии флотации, с^{-1} ; c_h – концентрация свободных частиц твердой фазы, $\text{м}^3/\text{м}^3$; c_{g1} – концентрация пузырьков газа, выделяющихся на первой стадии флотации, $\text{м}^3/\text{м}^3$.

Полагаем, что в осевом направлении комплексы движутся без инерции, т.е. осевые составляющие скорости комплексов и дисперсионной среды равны: $V_{zc} = V_{zl}$.

Поскольку радиальные составляющие скорости комплексов и неразрывного потока дисперсионной среды различны, при сделанных допущениях

$$\text{div } \bar{V}_c = \frac{1}{r} \frac{\partial [r(V_{rc} - V_{rl})]}{\partial r},$$

где V_{rc} – радиальная составляющая скорости комплекса, $\text{м}/\text{с}$; V_{rl} – радиальная составляющая скорости дисперсионной среды, $\text{м}/\text{с}$; r – радиальная координата, м , и уравнение (2) принимает вид:

$$V_{rc} \frac{\partial c_{hc}}{\partial r} + V_{zc} \frac{\partial c_{hc}}{\partial z} = A_1 c_h c_{g1} - c_{hc} \left\{ A_2 c_{g2} + \frac{1}{r} \frac{\partial [r(V_{rc} - V_{rl})]}{\partial r} \right\}, \quad (3)$$

где $V_{zc}(r, z)$, $V_{rc}(r, z)$ – осевая и радиальная составляющие скорости комплекса частица–пузырек, $\text{м}/\text{с}$; z – осевая координата, м .

Уравнение (3) решалось совместно с уравнением движения комплекса частица–пузырек.

Будем полагать, что на второй стадии флотации, как и при одностадийной напорной флотации [6, 8], основным механизмом образования вторичных комплексов является столкновение пузырьков вторичного газа с комплексами при их всплывании к поверхности пленки, определяющее воздействие на которое оказывают характеристики дисперсности и физические свойства фаз.

Аналогично допущениям [8] полагаем, что все происходящие столкновения пузырьков вторичного газа с комплексами эффективны (заканчиваются образованием вторичного комплекса), образовавшиеся вторичные комплексы не разрушаются при всплывании в пенный слой и не осыпаются из него. При таких допущениях стадия столкновения пузырьков вторичного газа с комплексами определяет эффективность процесса двухстадийной напорной флотации в условиях пленочного течения разделяемой суспензии в цилиндрическом гидроциклоне ввиду того, что пузырьки вторичного газа, имеющие на порядок больший диаметр, чем пузырьки, образующиеся на первой стадии напорной флотации, обеспечивают значительно большую скорость всплывания вторичных комплексов к поверхности пленки, а стадия транспортировки комплексов в процессе одностадийной напорной флотации является лимитирующей. Кинетические константы первой и второй стадий флотации A_1 и A_2 могут быть определены по формуле, полученной в работах [11, 12].

Радиальная составляющая скорости комплекса V_{rc} определялась из решения дифференциального уравнения движения, которое при условии действия на комплекс центробежной силы инерции, вытал-

кивающей центробежной силы Архимеда и силы сопротивления в радиальном направлении имеет вид:

$$V_{rc} \frac{\partial V_{rc}}{\partial r} + V_{zc} \frac{\partial V_{rc}}{\partial z} = \frac{V_{\phi c}^2}{r} \left[1 - \frac{\rho_l \left(\frac{d_h^3}{d_{g1}^2} + d_{g1} \right)}{\rho_h \frac{d_h^3}{d_{g1}^2} + \rho_g d_{g1}} \right] - \frac{3 K_{cr} \rho_l (V_{rc} - V_{rl}) |V_{rc} - V_{rl}|}{4 \left(\rho_h \frac{d_h^3}{d_{g1}^2} + \rho_g d_{g1} \right) \Phi(c_c)}, \quad (4)$$

где $V_{rc}(r, z)$ – радиальная составляющая скорости комплекса частица–пузырек, м/с; $V_{rl}(r, z)$ – радиальная составляющая скорости жидкости, м/с; K_{cr} – коэффициент сопротивления движению комплекса в радиальном направлении [13]; d_h – диаметр частиц твердой фазы, м; d_{g1} – диаметр пузырьков газа, выделяющихся на первой стадии напорной флотации, м; d_{g2} – диаметр пузырьков вторичного газа, подаваемого через пористую стенку корпуса гидроциклона, м; ρ_h – плотность частиц твердой фазы, кг/м³; ρ_l – плотность дисперсионной среды, кг/м³; ρ_g – плотность газа (воздуха), кг/м³.

С учетом ориентации комплекса частица–пузырек при всплывании к поверхности пленки за характерный размер комплекса d_c был принят диаметр пузырьков газа, выделяющихся из суспензии на первой стадии напорной флотации d_{g1} , так как он существенно превышает диаметр частиц твердой фазы.

Граничное условие для системы обыкновенных дифференциальных уравнений, полученных путем преобразования уравнений (3), (4) имеет вид:

$$z = 0, \quad r_k - b \leq r \leq r_k, \quad c_{hc0} = 0, \quad (5)$$

где c_{hc0} – концентрация частиц твердой фазы, связанных в комплексы, во входном патрубке гидроциклона, м³/м³; r_k – радиус корпуса гидроциклона, м; b – ширина входного патрубка гидроциклона, м.

2. Численное моделирование

Моделирование процесса на второй стадии напорной флотации заключалось в численном решении системы дифференциальных уравнений (3), (4) для различных реологических свойств среды и различных значений фактора разделения при граничном условии (5) с использованием аппроксимации результатов моделирования поля концентраций свободных частиц твердой фазы на первой стадии напорной флотации [8].

С применением метода характеристик [14] система дифференциальных уравнений в частных производных (3), (4) сводится к эквивалентной системе обыкновенных дифференциальных уравнений, одно из которых описывает траекторию движения комплекса частица–пузырек к поверхности пленки суспензии, а другое – изменение остаточной концентрации частиц твердой фазы, связанных в комплексы, вдоль траектории. Полученная система обыкновенных дифференциальных уравнений была решена методом Рунге–Кутты четвертого порядка с фиксированным шагом с использованием программы, составленной на языке *Compaq Visual Fortran*. Расчеты выполнялись по траекториям движения комплекса частица–пузырек до выхода комплекса на поверхность пленки суспензии в пенный слой.

Результаты моделирования поля концентраций свободных частиц твердой фазы в соответствии с математической моделью первой стадии флотации [8] были аппроксимированы уравнением поверхности

$$c(r, z) = \exp(B_0 + (B_1 + B_2/r + B_3/r^2)/z + (B_4 + B_5/r + B_6/r^2)/z^2 + (B_7 + B_8/z + B_9/z^2)r + (B_{10} + B_{11}/z + B_{12}/z^2)r^2 + B_{13} + B_{14}r + B_{15}r^2 + B_{16}r^3 + B_{17}r^4 + B_{18}z + B_{19}z^2 + B_{20}z^3 + B_{21}z^4), \quad (6)$$

где $c(r, z)$ – концентрация свободных частиц твердой фазы, м³/м³; B_0 – B_{21} – коэффициенты регрессии.

Аппроксимация выполнялась в пакете *SPSS* в соответствии с процедурой нелинейной регрессии для использования при моделировании поля концентраций частиц твердой фазы, связанных в комплексы, в соответствии с математической моделью второй стадии напорной флотации. Остаточная сумма квадратов не превышала $7,5 \cdot 10^{-5}$.

Для оценки влияния определяющих чисел подобия и реологических свойств дисперсионной среды на эффективность разделения вязкопластической суспензии двухстадийной напорной флотацией

в гидроциклоне используется среднеинтегральная безразмерная остаточная концентрация S частиц твердой фазы, связанных в комплексы [8], которая определялась с использованием математического пакета *Maple* 2019.

3. Результаты и их обсуждение

Течение неньютоновской жидкости в гидроциклоне характеризуется центробежным числом Фруда (фактором разделения), модифицированным числом Рейнольдса Re_n , безразмерным расходом Q , показателем нелинейности кривой течения n , а также числом пластичности Pl , характеризующим аномалию неньютоновских свойств вязкопластической среды [8]. Основным критерием, характеризующим процессы разделения в поле центробежных сил, является число Фруда (фактор разделения), представляющее собой отношение центробежного ускорения во вращающемся потоке к гравитационному ускорению.

Результаты численного моделирования представлены на рис. 2, 3.

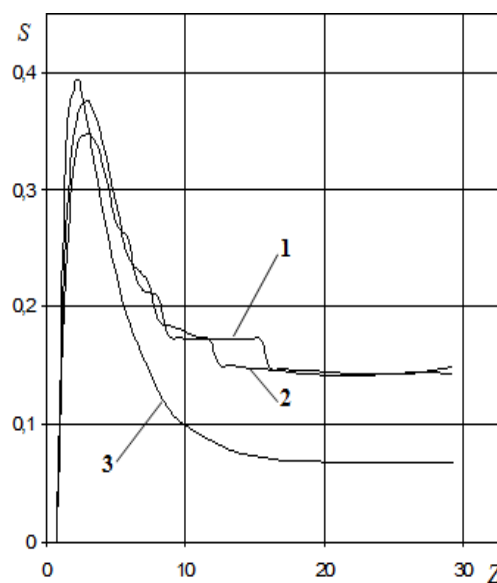


Рис. 2. Распределения среднеинтегральной безразмерной остаточной концентрации частиц твердой фазы S , связанных в комплексы, по оси гидроциклона для $Pl = 8,5 \cdot 10^{-3}$; $Re_n = 4 \cdot 10^3$; $Q = 0,02$; $n = 0,7$; $d_h = 2 \cdot 10^{-5}$ м; $d_{g1} = 3 \cdot 10^{-5}$ м; $d_{g2} = 3 \cdot 10^{-4}$ м; $c_{g1} = 0,1$ м³/м³; $c_{g2} = 0,05$ м³/м³; 1 – $Fr = 30$; 2 – $Fr = 40$; 3 – $Fr = 100$

Fig. 2. Distributions of the average integral dimensionless residual concentration of solid phase particles S bound in complexes along the axis of the hydrocyclone for $Pl = 8,5 \cdot 10^{-3}$; $Re_n = 4 \cdot 10^3$; $Q = 0,02$; $n = 0,7$; $d_h = 2 \cdot 10^{-5}$ м; $d_{g1} = 3 \cdot 10^{-5}$ м; $d_{g2} = 3 \cdot 10^{-4}$ м; $c_{g1} = 0,1$ м³/м³; $c_{g2} = 0,05$ м³/м³; 1 – $Fr = 30$; 2 – $Fr = 40$; 3 – $Fr = 100$

На рис. 2 представлены распределения среднеинтегральной безразмерной остаточной концентрации частиц твердой фазы S , связанных в комплексы, оставшихся в суспензии в данном сечении пленки, по оси гидроциклона (безразмерная осевая координата $Z = z/r_k$) для $Pl = 8,5 \cdot 10^{-3}$, $Re_n = 4 \cdot 10^3$, $Q = 0,02$, $n = 0,7$ при разделении суспензии с выраженной аномалией неньютоновских свойств дисперсионной среды в цилиндрическом прямоточном гидроциклоне двухстадийной напорной флотацией. Из анализа представленных на рис. 2 зависимостей следует, что среднеинтегральная безразмерная остаточная концентрация S частиц твердой фазы, связанных в комплексы, оставшихся в суспензии в данном сечении пленки, сначала возрастает в осевом направлении вследствие образования комплексов на первой стадии флотации, а затем снижается до пренебрежимо малых значений вследствие извлечения комплексов пузырьками вторичного газа. Форма кривых обусловлена взаимодействием указанных процессов образования и извлечения комплексов. Из сравнения кривых 1–3 следует, что среднеинтегральная безразмерная остаточная концентрация S частиц твердой фазы снижается с увеличением числа Fr по всей высоте рабочего пространства гидроциклона.

На рис. 3 представлены распределения среднеинтегральной безразмерной остаточной концентрации частиц твердой фазы S , связанных в комплексы, оставшихся в суспензии в данном сечении пленки, по оси гидроциклона для $Pl = 8,5 \cdot 10^{-2}$, $Re_n = 4 \cdot 10^3$, $Q = 0,02$, $n = 0,7$. Из сравнительного анализа зависимостей, представленных на рис. 3 и рис. 2, следует, что среднеинтегральная безразмерная остаточная концентрация частиц твердой фазы S , связанных в комплексы, оставшихся в суспензии в данном сечении пленки, при $Pl = 8,5 \cdot 10^{-2}$ с увеличением числа Fr снижается более значительно, чем при $Pl = 8,5 \cdot 10^{-3}$, и при $Z = 15 \dots 17$ достигает нулевого значения (см. рис. 3).

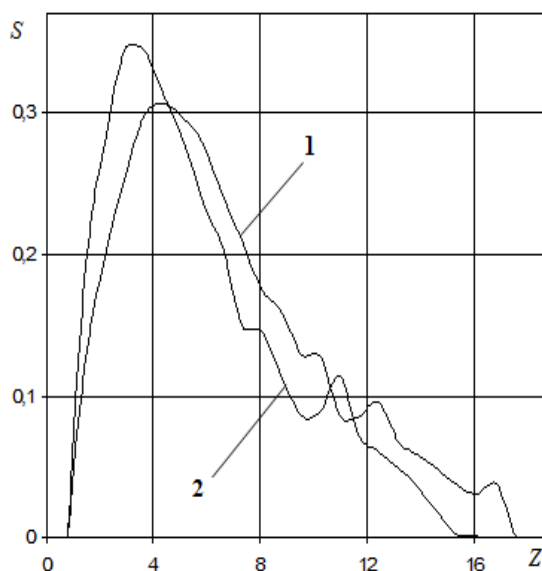


Рис. 3. Распределения среднеинтегральной безразмерной остаточной концентрации частиц твердой фазы S , связанных в комплексы, по оси гидроциклона для $Pl = 8,5 \cdot 10^{-2}$, $Re_n = 4 \cdot 10^3$, $Q = 0,02$, $n = 0,7$; $d_h = 2 \cdot 10^{-5}$ м; $d_{g1} = 3 \cdot 10^{-5}$ м; $d_{g2} = 3 \cdot 10^{-4}$ м; $c_{g1} = 0,1$ м³/м³; $c_{g2} = 0,05$ м³/м³; 1 – $Fr = 30$; 2 – $Fr = 40$

Fig. 3. Distributions of the average integral dimensionless residual concentration of solid phase particles S bound in complexes along the axis of the hydrocyclone for $Pl = 8,5 \cdot 10^{-2}$, $Re_n = 4 \cdot 10^3$, $Q = 0,02$, $n = 0,7$; $d_h = 2 \cdot 10^{-5}$ м; $d_{g1} = 3 \cdot 10^{-5}$ м; $d_{g2} = 3 \cdot 10^{-4}$ м; $c_{g1} = 0,1$ м³/м³; $c_{g2} = 0,05$ м³/м³; 1 – $Fr = 30$; 2 – $Fr = 40$;

Сопоставляя рис. 3 и рис. 2, можно установить, что среднеинтегральная остаточная концентрация частиц твердой фазы S , связанных в комплексы, на выходе из рабочего пространства гидроциклона снижается с увеличением числа пластичности Pl , соответствующем возрастанию предельного напряжения сдвига τ_0 и аномалии неньютоновских свойств среды, что объясняется снижением затухания окружной составляющей скорости среды в осевом направлении [2] и соответствующим возрастанием кинетических констант первой и второй стадий напорной флотации A_1 и A_2 согласно данным [11, 12].

Заключение

Таким образом, с использованием разработанной математической модели разделения вязкопластических сред двухстадийной напорной флотацией в цилиндрическом гидроциклоне выполнено моделирование поля концентраций частиц твердой фазы, связанных в комплексы, установлено влияние фактора разделения и пластических свойств среды на показатели работы гидроциклона-флотатора и определены оптимальные режимы проведения процесса. Показано, что двухстадийная напорная флотация позволяет получить низкую остаточную концентрацию частиц твердой фазы в пленке суспензии. Влияние фактора разделения является более значительным для сред с высоким предельным напряжением сдвига. Установлено, что процесс разделения суспензий с выраженными пластическими свойствами двухстадийной флотацией протекает наиболее эффективно, особенно при высоких значениях фактора разделения, обеспечивая меньшую остаточную концентрацию частиц твердой фазы.

Список источников

1. Матвиенко О.В., Базуев В.П., Асеева А.Е. Математическое моделирование течения закрученного потока псевдопластической жидкости Балкли–Гершеля в цилиндрическом канале // Инженерно-физический журнал. 2019. Т. 92, № 1. С. 215–226. doi: 10.1007/s10780-019-00212-x
2. Яблонский В.О. Гидродинамика нелинейновязкопластической жидкости в цилиндрическом гидроциклоне // Журнал прикладной химии. 2013. Т. 86, вып. 8. С. 1236–1243. doi: 10.1134/S1070427213080107
3. Кутепов А.М., Лагуткин М.Г., Павловский Г.В., Муштаев В.И. Разделение дисперсных систем в гидроциклонах с дополнительным вводом диспергированного газа // Теоретические основы химической технологии. 1999. Т. 33. № 5. С. 571–577.
4. Кутепов А.М., Лагуткин М.Г., Муштаев В.И., Булычев С.Ю. Разделение гетерогенных систем в цилиндрическом прямом точном гидроциклоне // Химическое и нефтегазовое машиностроение. 2002. № 7. С. 14–18. doi: 10.1023/A:1021019131679
5. Орлов С.Л. О столкновении частиц и пузырьков при пневматической флотации в центробежном поле // Развитие методов механической и биологической очистки сточных вод : сб. науч. тр. М. : ВНИИ «Водгео», 1982. С. 132–136.
6. Сотскова Т.З., Баженов Ю.Ф., Голик Г.А. Взаимодействие мелких частиц с газовыми пузырьками при напорной флотации // Химия и технология воды. 1984. Т. 6, № 1. С. 17–22.
7. Щукина А.Г. Математическое моделирование процессов разделения неоднородных систем с неньютоновской дисперсионной средой : автореф. дис. ... канд. техн. наук : 05.13.16. Волгоград, 1996. 33 с.
8. Яблонский В.О. Математическая модель и анализ влияния конструкции корпуса гидроциклона на эффективность разделения вязкопластических сред напорной флотацией // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2023. № 65. С. 47–57. doi: 10.17223/19988605/65/5
9. Дерягин Б.В., Духин С.С., Рулев Н.Н. Микрофлотация: водоочистка, обогащение. М. : Химия, 1986. 112 с.
10. Acary-Robert C., Fernández-Nieto E.D., Narbona-Reina G., Vigneaux P. Viscoplastic free-surface flows // The Herschel-Bulkley case seventh international conference on computational fluid dynamics (ICCFD7), Big Island, Hawaii, July 9–13, 2012. P. 1–15.
11. Тябин Н.В., Дахина Г.Л., Голованчиков А.Б., Мамаков А.А. Расчет аппаратов идеального вытеснения для разделения тонких суспензий электролитическими газами // Теоретические основы химической технологии. 1979. Т. 13, № 6. С. 880–884.
12. Тябин Н.В., Голованчиков А.Б., Дахина Г.Л., Зеленцов В.И. Оптимизация процесса электрофлотационного разделения суспензий // Электронная обработка материалов. 1986. № 5. С. 65–67.
13. Acharya A., Mashelkar R.A., Ulbrecht J. Flow of inelastic and viscoelastic fluids past a sphere // Rheol. Acta. 1976. V. 15 (9). P. 454–463.
14. Пономарёв К.К. Составление и решение дифференциальных уравнений инженерно-технических задач. М. : ЁЁ Медиа, 2024. 184 с.

References

1. Matvienko, O.V., Bazuev, V.P. & Aseeva, A.E. (2019) Mathematical modeling of the flow of a swirling Bulkley-Herschel pseudo-plastic fluid flow in a cylindrical channel. *Engineering Physics Journal*. 92(1). pp. 215–226. DOI: 10.1007/s10780-019-00212-x
2. Yablonsky, V.O. (2013) Hydrodynamics of a nonlinear viscous plastic liquid in a cylindrical hydrocyclone. *Journal of Applied Chemistry*. 86(8). pp. 1236–1243. DOI: 10.1134/S1070427213080107
3. Kutepov, A.M., Lagutkin, M.G., Pavlovsky, G.V. & Mushtaev, V.I. (1999) Separation of dispersed systems in hydrocyclones with additional input of dispersed gas. *Teoreticheskie osnovy khimicheskoy tekhnologii – Theory of Fundamentals of Chemical Technology*. 33(5). pp. 571–577.
4. Kutepov, A.M., Lagutkin, M.G., Mushtaev, V.I. & Bulychiev, S.Yu. (2002) Separation of heterogeneous systems in a cylindrical ramjet hydrocyclone. *Khimicheskoe i neftegazovoe mashinostroenie – Chemical and Oil and Gas Engineering*. 7. pp. 14–18. DOI: 10.1023/A:1021019131679
5. Orlov, S.L. (1982) O stolknovenii chastits i puzyr'kov pri pnevmaticheskoy flotatsii v tsentrobeznom pole [On the collision of particles and bubbles during pneumatic flotation in a centrifugal field]. In: *Razvitie metodov mekhanicheskoy i biologicheskoy ochistki stochnykh vod* [Development of Methods of Mechanical and Biological Wastewater Treatment]. Moscow: Vodgeo. pp. 132–136.
6. Sotskova, T.Z., Bazhenov, Yu.F. & Golik, G.A. (1984) Vzaimodeystvie melkikh chastits s gazovymi puzyr'kami pri napornoy flo-tatsii [Interaction of small particles with gas bubbles during pressure flotation]. *Khimiya i tekhnologiya vody*. 6(1). pp. 17–22.
7. Shchukina, A.G. (1996) *Matematicheskoe modelirovanie protsessov razdeleniya neodnorodnykh sistem s nen'yutonovskoy disper-sionnoy sredoy* [Mathematical modeling of separation processes of inhomogeneous systems with a non-Newtonian dispersion medium]. Abstract of Engineering Cand. Diss. Volgograd: VolGSTU.
8. Yablonsky, V.O. (2013) Mathematical model and analysis of the effect of the hydrocyclone hull design on the efficiency of separation of viscoplastic media by pressure flotation. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 65. pp. 47–57. DOI: 10.17223/19988605/65/5
9. Deryagin, B.V., Dukhin, S.S. & Rulev, N.N. (1986) *Mikroflotatsiya: vodoochistka, obogashchenie* [Microflotation: Water Treatment, Enrichment]. Moscow: Khimiya.

10. Acary-Robert, C., Fernández-Nieto, E.D., Narbona-Reina, G. & Vigneaux, P. (2012) Viscoplastic free-surface flows. *The Herschel-Bulkley Case Seventh International Conference on Computational Fluid Dynamics (ICCFD7)*. Big Island, Hawaii, July 9–13. pp. 1–15.
11. Tyabin, N.V., Dakhina, G.L., Golovanchikov, A.B. & Mamakov, A.A. (1979) Raschet apparatov ideal'nogo vytesneniya dlya razdeleniya tonkikh suspenziy elektroliticheskimi gazami [Calculation of ideal displacement devices for separation of thin suspensions by electrolytic gases]. *Teoreticheskie osnovy khimicheskoy tekhnologii – Theory. Fundamentals of Chemical Technology*. 13(6). pp. 880–884.
12. Tyabin, N.V., Golovanchikov, A.B., Dakhina, G.L. & Zelentsov, V.I. (1986) Optimizatsiya protsessa elektroflotatsionnogo razdeleniya suspenziy [Optimization of the process of electroflotation separation of suspensions]. *Elektronnaya obrabotka materialov – Electronic Processing of Materials*. 5. pp. 65–67.
13. Acharya, A., Mashelkar, R.A. & Ulbrecht, J. (1976) Flow of inelastic and viscoelastic fluids past a sphere. *Rheol. Acta*. 15(9). pp. 454–463.
14. Ponomarev, K.K. (2024) *Sostavlenie i reshenie differentsial'nykh uravneniy inzhenerno-tekhnicheskikh zadach* [Compilation and solution of differential equations of engineering and technical problems]. Moscow: EE Media.

Информация об авторе:

Яблонский Владимир Олегович – кандидат технических наук, доцент кафедры «Процессы и аппараты химических и пищевых производств» Волгоградского государственного технического университета (Волгоград, Россия). E-mail: everest58@mail.ru

Автор заявляет об отсутствии конфликта интересов.

Information about the author:

Yablonskii Vladimir O. (Candidate of Technical Sciences, Associate Professor of the Department “Processes and Devices of Chemical and Food Production” of Volgograd State Technical University (Volgograd, Russian Federation). E-mail: everest58@mail.ru

The author declares no conflicts of interests.

Поступила в редакцию 23.12.2024; принята к публикации 02.06.2025

Received 23.12.2024; accepted for publication 02.06.2025

ОБРАБОТКА ИНФОРМАЦИИ

DATA PROCESSING

Original article

UDC 519.234

doi: 10.17223/19988605/71/6

**On analysis of survival function estimators with its identifiability
with model of right random censorship****Abdurakhim A. Abdushukurov¹, Sukhrob B. Bozorov²**¹ *Moscow State University, Branch in Tashkent, Tashkent, Uzbekistan*² *Gulistan State University, Gulistan, Uzbekistan*¹ *a_abdushukurov@rambler.ru*² *suxrobbek_8912@mail.ru*

Abstract. It is considered random censorship from the right. In this model we investigate basically three types of estimators of exponential, product and power types. We illustrative study all these estimators including its presmoothed modifications comparatively. We demonstrate confidence bands for considered estimators by using natural and artificial censored observation. It is showed that presmoothed power estimators have some peculiarities with respect to exponential and product-type estimators. It is shown that only power estimator has identifiability property with respect to right random censoring model in the case of a finite-size sample.

Keywords: model of right random censorship; survival function; exponential; product and power estimators; presmoothed kernel estimators.

For citation: Abdushukurov, A.A., Bozorov, S.B. (2025) On analysis of survival function estimators with its identifiability with model of right random censorship. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 71. pp. 58–70. doi: 10.17223/19988605/71/6

Научная статья

doi: 10.17223/19988605/71/6

**Анализ оценок функции выживания с учетом их идентифицируемости
с моделью случайного цензурирования справа****Абдурахим Ахмедович Абдушукуров¹, Сухроб Баходирович Бозоров²**¹ *Московский государственный университет, филиал в Ташкенте, Ташкент, Узбекистан*² *Гулистанский государственный университет, Гулистан, Узбекистан*¹ *a_abdushukurov@rambler.ru*² *suxrobbek_8912@mail.ru*

Аннотация. Рассматривается модель случайного цензурирования справа, в которой исследуются сглаженные модификации экспоненциальной, множительной и степенной оценок функции выживания. Демонстрируются доверительные полосы для естественных и искусственных наблюдений для рассматриваемых оценок. Показано, что сглаженные степенные оценки обладают рядом преимуществ по сравнению с экспоненциальной и множественной оценками. Показано, что именно степенная оценка обладает свойством идентифицируемости с моделью случайного цензурирования справа в случае выборки конечного объема.

Ключевые слова: модель случайного цензурирования справа; функция выживания; экспоненциальная; множительная и степенные оценки; сглаженные ядерные оценки.

Для цитирования: Абдушукуров А.А., Бозоров С.Б., Анализ оценок функции выживания с учётом их идентифицируемости с моделью случайного цензурирования справа // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2025. № 71. С. 58–70. doi: 10.17223/19988605/71/6

Introduction

Paper is devoted to consider random censorship model from the right. Suppose that $X_1, X_2, \dots$ and $Y_1, Y_2, \dots$ are two sequences of independent random variables (r.v.'s) with corresponding distribution function (d.f.) $F(t)$ and $G(t)$, $t \in R$, respectively. Both sequences of X_k 's and Y_k 's are mutually independent. Observation is available sample of size n :

$$\mathbb{C}^{(n)} = \{(Z_k, \delta_k), 1 \leq k \leq n\},$$

where $Z_k = \min(X_k, Y_k)$ and indicator $\delta_k = I(X_k \leq Y_k) = I(Z_k = X_k)$.

- If $X_k \leq Y_k$, then $Z_k = \min(X_k, Y_k) = X_k$ is equal to $\delta_k = 1$ and in this case we can observe X .
- Otherwise if $Y_k \leq X_k$, then $Z_k = \min(X_k, Y_k) = Y_k$ is equal to $\delta_k = 0$, and this can be censoring condition.

The main problem has been considered by many authors since 1950s. We comparatively study of survival function estimators taking into account its property of identifiability with model of right random censorship. Observe that in paper [1] authors proposed a new histogram type estimators for the distribution density using the kernel method by using presmoothed estimator for survival function. Remains relevant to considering identifiability of presmoothed power estimator with respect to exponential and product type estimators.

1. Proposed estimators

In sample $\mathbb{C}^{(n)}$, the number of observed X_k 's is equal to summa $v(n) = \sum_{k=1}^n \delta_k$, which is binomial r.v.

$Bi(n; p)$ with $P(\delta_k = 1) = p$. The first, product-limit (PL) estimator has been proposed by E. Kaplan and P. Meier in [1] by formula

$$F_n^{KM}(t) = \begin{cases} 1 - \prod_{\{k: Z_{(k)} \leq t\}} \left(1 - \frac{\delta_{(k)}}{n - k + 1}\right), & t \leq Z_{(n)}, \\ 1, & t > Z_{(n)}, \delta_{(n)} = 1, \\ \text{undefined}, & t > Z_{(n)}, \delta_{(n)} = 0. \end{cases} \quad (1)$$

Here is $Z_{(1)} \leq Z_{(2)} \leq \dots \leq Z_{(n)}$ -ordered sample of $\{Z_k, 1 \leq k \leq n\}$ and $\delta_{(k)}$ accompanies $Z_{(k)}$'s. Estimator (1) has a great influence in development of statistical theory of censored observation. In statistical literature one can meet several modified variants of estimator (1) and they are used by many scientists [2–8]. In order to consider estimator we introduce d.f. of minima as

$$P(Z_k \leq t) = H(t) = 1 - (1 - F(t))(1 - G(t)) = H_{(0)}(t) + H_{(1)}(t), \quad t \in R, \quad (2)$$

where $H_{(m)}(t) = P(Z_k \leq t, \delta_k = m)$, $m = 0, 1$ are subdistribution functions. It is not difficult to see that

$$H_{(0)}(t) = \int_{-\infty}^t (1 - F(u)) dG(u) \quad \text{and} \quad H_{(1)}(t) = \int_{-\infty}^t (1 - G(u)) dF(u).$$

Let's consider integral hazard function corresponding to d.f.'s $F(t)$, $G(t)$ and $H(t)$ by formulas

$$\begin{aligned}\Lambda_{(1)}(t) &= \int_{-\infty}^t \frac{dH_{(1)}(u)}{1-H(u)} = -\log(1-F(t)), \quad \Lambda_{(0)}(t) = \int_{-\infty}^t \frac{dH_{(0)}(u)}{1-H(u)} = -\log(1-G(t)), \\ \Lambda(t) &= \int_{-\infty}^t \frac{dH(u)}{1-H(u)} = \Lambda_{(0)}(t) + \Lambda_{(1)}(t) = -\log(1-H(t))\end{aligned}\quad (3)$$

and its estimators [5, 6, 9, 10]:

$$\begin{aligned}\Lambda_{(1)n}(t) &= \int_{-\infty}^t \frac{dH_{(1)n}^E(u)}{1-H_n^E(u) + \frac{1}{n}}, \quad \Lambda_{(0)n}(t) = \int_{-\infty}^t \frac{dH_{(0)n}^E(u)}{1-H_n^E(u) + \frac{1}{n}}, \\ \Lambda_n(t) &= \int_{-\infty}^t \frac{dH_n^E(u)}{1-H_n^E(u)} = \Lambda_{(0)n}(t) + \Lambda_{(1)n}(t),\end{aligned}\quad (4)$$

where summand $\frac{1}{n}$ in denominators of integrals (4) is adden in order to prevent dividing to zero. Here the empirical estimators of $H(t)$ and $H_{(m)}(t)$, $m=0,1$ are

$$H_n^E(t) = \frac{1}{n} \sum_{k=1}^n I(Z_k \leq t) = H_{(0)n}^E(t) + H_{(1)n}^E(t), \quad t \in R \text{ and } H_{(m)n}^E(t) = \frac{1}{n} \sum_{k=1}^n I(Z_k \leq t, \delta_k = m), \quad m=0,1.$$

By formulas (3) and (4) one can see easily that natural estimator of d.f. $F(t)$ is following B. Altshuler [11] and N. Breslow [3] estimator of exponential form:

$$\begin{aligned}F_n^{AB}(t) &= 1 - \exp(-\Lambda_{(1)n}(t)) = 1 - \exp\left\{-\frac{1}{n} \sum_{k=1}^n \frac{I(Z_k \leq t, \delta_k = 1)}{1-H_n^E(Z_k) + \frac{1}{n}}\right\} = \\ &= 1 - \exp\left\{-\sum_{\{k: Z_{(k)} \leq t\}} \frac{\delta_{(k)}}{n-k+1}\right\} = 1 - \left(\prod_{\{k: Z_{(k)} \leq t\}} \exp\left(-\frac{1}{n-k+1}\right)\right)^{\delta_{(k)}}, \quad t \in R.\end{aligned}\quad (5)$$

In order to compare with estimator (5) we use following modification and very popular form of Kaplan-Meier's estimator (1):

$$F_n^{KM}(t) = \begin{cases} 1 - \prod_{\{k: Z_{(k)} \leq t\}} \left(1 - \frac{\delta_{(k)}}{n-k+1}\right), & t \leq Z_{(n)}, \\ 1, & t > Z_{(n)}, \end{cases} = \begin{cases} \prod_{\{k: Z_{(k)} \leq t\}} \left(\frac{n-k}{n-k+1}\right)^{\delta_{(k)}}, & t \leq Z_{(n)}, \\ 1, & t > Z_{(n)}. \end{cases}\quad (6)$$

Because of approximate equality $e^{-x} \approx 1-x$ under $x \approx 0$ from (5) and (6) we have that

$$\Lambda_{(1)n}(t) \approx -\log(1 - F_n^{KM}(t)), \quad t < T_{H_n^E} = \inf\{t: H_n^E(t) = 1\}.\quad (7)$$

Approximate equality (7) plays a key role in investigating of several asymptotic results for estimator (6) or (1). Author [12–14] proposed a relative risk power estimator $F_n^{RR}(t)$ by formula:

$$F_n^{RR}(t) = 1 - (1 - H_n^E(t))^{R_n(t)} = \begin{cases} 0, & t < Z_{(1)}, \\ 1 - \left(\frac{n-k}{n}\right)^{R_n(t)}, & Z_{(k)} \leq t < Z_{(k+1)}, \quad 1 \leq k \leq n-1, \\ 1, & t \geq Z_{(n)}, \end{cases}\quad (8)$$

where fraction $R_n(t) = \Lambda_{(1)n}(t) \cdot (\Lambda_n(t))^{-1}$ is estimator of relative risk function $R(t) = \Lambda_{(1)}(t) \cdot (\Lambda(t))^{-1}$. It is not difficult to see that $0 \leq R(t), R_n(t) \leq 1$ for all $t \in E$.

Now in order to introduce our next presmoothed estimator we suppose that both d.f. $F(t)$ and $G(t)$ are absolutely continuous with densities $f(t)$ and $g(t)$ respectively. Then d.f. of minima $H(t)$ has density $h(t) = (1 - G(t))f(t) + (1 - F(t))g(t)$. Consider presmoothing problem of estimators (5), (6) and (8).

It is not difficult to see that regression function of indicator δ_k given value of minima $Z_k = t$ is equal to conditional probability

$$p(t) = P(\delta_k = 1 | Z_k = t) = E[\delta_k | Z_k = t], \quad t \in R. \quad (9)$$

Then the integral hazard function $\Lambda_{(1)}(t)$ in (3) can be represented as

$$\Lambda_{(1)}(t) = \int_{-\infty}^t p(u) d\Lambda(u), \quad t \in R, \quad (10)$$

or equivalently for densities

$$h_{(1)}(t) = p(t)h(t), \quad t \in R, \quad (11)$$

$$R(t) = \frac{\Lambda_{(1)}(t)}{\Lambda(t)} = \frac{\int_{-\infty}^t p(u) d\Lambda(u)}{\Lambda(t)}, \quad t \in R. \quad (12)$$

For probability (9) we can use any estimator for regression function. But we prefer Nadaraya-Watson estimator (see, [15])

$$p_n(t) = \frac{\frac{1}{nh(n)} \sum_{k=1}^n \delta_k k \left(\frac{t - Z_k}{h(n)} \right)}{\frac{1}{nh(n)} \sum_{k=1}^n k \left(\frac{t - Z_k}{h(n)} \right)} = \frac{\sum_{k=1}^n \delta_k k \left(\frac{t - Z_k}{h(n)} \right)}{\sum_{k=1}^n k \left(\frac{t - Z_k}{h(n)} \right)}, \quad t \in R, \quad (13)$$

obtaining by formula $p(t) = \frac{h_{(1)}(t)}{h(t)}$ (see, (11)). Here kernel $\{k(t), t \in R\}$ is well known function of density $\{h(n), n \geq 1\}$ is «window width» sequence, tending to zero at $n \rightarrow \infty$. By formulas (9)–(13) function (12) can be estimated as

$$R_n^p(t) = \frac{\Lambda_{(1)n}^p(t)}{\Lambda_n(t)} = \frac{\int_{-\infty}^t p_n(u) d\Lambda_n(u)}{\Lambda_n(t)}, \quad t \in R, \quad (14)$$

where $\Lambda_n(t)$ is estimator in (4), $\Lambda_{(1)n}^p(t)$ is presmoothed estimator of $\Lambda_{(1)}(t)$. Now replacing power $R_n(t)$ by $R_n^p(t)$ in (8) we obtain a new presmoothed relative risk power estimator

$$F_n^{pR}(t) = 1 - (1 - H_n^E(t))^{R_n^p(t)} = \begin{cases} 0, & t < Z_{(1)}, \\ 1 - \left(\frac{n-k}{n} \right)^{R_n^p(t)}, & Z_{(k)} \leq t < Z_{(k+1)}, \quad 1 \leq k \leq n-1, \\ 1, & t \geq Z_{(n)}, \end{cases} \quad (15)$$

Note that authors [15] have been investigated presmoothed variants of estimator (5) and (6).

$$F_n^{pAB}(t) = 1 - \exp(-\Lambda_{(1)n}^p(t)) = 1 - \exp\left(-\sum_{\{k: Z_{(k)} \leq t\}} \frac{p_n(Z_{(k)})}{n-k+1}\right), \quad (16)$$

$$F_n^{pKM}(t) = 1 - \prod_{\{k: Z_{(k)} \leq t\}} \left(1 - \frac{p_n(Z_{(k)})}{n-k+1}\right), \quad t \in R, \quad (17)$$

for estimator (16) and (17) we have asymptotic relation

$$F_n^{pKM}(t) - F(t) = (1 - F(t))(\Lambda_{(1)n}^p(t) - \Lambda(t)) + O_p\left(\frac{1}{n}\right), \quad n \rightarrow \infty.$$

From here we get analogue of (7) for presmoothed estimators (16) and (17)

$$\Lambda_{(1)n}^p(t) \approx -\log(1 - F_n^{pKM}(t)), \quad t < T_{H_n^E}. \quad (18)$$

In particular, in case of independence of indicator δ_k from minima Z_k , from (9) we have $p(t) = p = E\delta_k$ for all $t \in R$, consequently by characterization of simple Proportional Hazards Model (PHM) $1 - F(t) = (1 - H(t))^p$, $p = \frac{1}{1+\beta}$, $\beta > 0$ is some number. Then estimator p by $p_n = \frac{1}{n} \sum_{k=1}^n \delta_k$ we obtain well known power estimator of the from

$$\hat{F}_n(t) = 1 - (1 - H_n^E(t))^{p_n} = \begin{cases} 0, & t < Z_{(1)}, \\ 1 - \left(\frac{n-k}{n}\right)^{p_n}, & Z_{(k)} \leq t < Z_{(k+1)}, 1 \leq k \leq n-1, \\ 1, & t \geq Z_{(n)}, \end{cases} \quad (19)$$

for details, see [16–19]. At the end of section let us remind some theoretical that is asymptotic properties of estimator (15) from [20]. According to the theorem 1 in [20] authors approximate centered and normed estimator (15) by sum of independent and identically distributed random function, in theorem 2 authors get strong uniform consistency result for estimator (15) and in theorem 3 they proved asymptotically normality results for estimator (15).

3. Simulation study

In papers [12–14] authors investigated several asymptotic properties of estimators (1), (5) and (8) at $n \rightarrow \infty$ and showed that all three estimators are asymptotically equivalent in sense of convergence to same centered Gaussian process. Suppose that $G_n^{KM}(t)$, $G_n^{AB}(t)$, $G_n^{RR}(t)$ and $\hat{G}_n(t)$ are estimator of d.f. $G(t)$ defined by analogues of formulas (1), (5), (8) and (19) as follows:

$$G_n^{KM}(t) = \begin{cases} 1 - \prod_{\{k: Z_{(k)} \leq t\}} \left(1 - \frac{1 - \delta_{(k)}}{n - k + 1}\right), & t \leq Z_{(n)}, \\ 1, & t > Z_{(n)}, \delta_{(n)} = 0, \\ \text{undefined}, & t > Z_{(n)}, \delta_{(n)} = 1, \end{cases} \quad (20)$$

$$G_n^{AB}(t) = 1 - \exp(-\Lambda_{(0)n}(t)) = 1 - \exp\left\{-\frac{1}{n} \sum_{k=1}^n \frac{I(Z_k \leq t, \delta_k = 0)}{1 - H_n^E(Z_k) + \frac{1}{n}}\right\}, \quad (21)$$

$$G_n^{RR}(t) = 1 - (1 - H_n^E(t))^{1 - R_n(t)} = \begin{cases} 0, & t < Z_{(1)}, \\ 1 - \left(\frac{n-k}{n}\right)^{1 - R_n(t)}, & Z_{(k)} \leq t < Z_{(k+1)}, 1 \leq k \leq n-1, \\ 1, & t \geq Z_{(n)}, \end{cases} \quad (22)$$

and

$$\hat{G}_n(t) = 1 - (1 - H_n^E(t))^{1 - p_n} = \begin{cases} 0, & t < Z_{(1)}, \\ 1 - \left(\frac{n-k}{n}\right)^{1 - p_n}, & Z_{(k)} \leq t < Z_{(k+1)}, 1 \leq k \leq n-1, \\ 1, & t \geq Z_{(n)}. \end{cases} \quad (23)$$

In considered right random censorship model we have for d.f. of minima Z_i that from (2)

$$1 - H(t) = (1 - F(t))(1 - G(t)), t \in R. \quad (24)$$

Which is easy fulfilled for two power estimator with its analogues: (8) with (22) and (19) with (23):

$$1 - H_n^E(t) = (1 - F_n^{RR}(t))(1 - G_n^{RR}(t)), t \in R,$$

$$1 - H_n^E(t) = (1 - \hat{F}_n(t))(1 - \hat{G}_n(t)), t \in R.$$

But for other two estimators (1) with (20) and (5) with (21) we have relations:

1) $(1 - F_n^{KM}(t))(1 - G_n^{KM}(t)) \neq 1 - H_n^E(t)$ for all $t \in R$, so for $t > Z_{(n)}$ both estimators $F_n^{KM}(t)$ and $G_n^{KM}(t)$ are not defined;

2) $(1 - F_n^{AB}(t))(1 - G_n^{AB}(t)) = \exp(-\Lambda_{(1)n}(t)) \neq 1 - H_n^E(t)$ and $\max\{F_n^{AB}(t), G_n^{AB}(t)\} < 1$.

Hence, only power type estimator (8) and (19) are identifiable with considered right censorship model defined by equality (24).

Now we consider presmoothed estimators (16), (17) and (15) with corresponding variants of estimators of $G(t)$. Here we observe analogue satiation:

3) $(1 - F_n^{pKM}(t))(1 - G_n^{pKM}(t)) \neq 1 - H_n^E(t)$ for all $t \in R$, where

$$G_n^{pKM}(t) = 1 - \prod_{\{k: Z_{(k)} \leq t\}} \left(1 - \frac{1 - \delta_{(k)}}{n - k + 1}\right), t \in R.$$

4) $(1 - F_n^{pAB}(t))(1 - G_n^{pAB}(t)) = \exp(-\Lambda_n^p(t)) \neq 1 - H_n^E(t)$, $t \in R$, where $\Lambda_n^p(t) = \Lambda_{(0)n}^p(t) + \Lambda_{(1)n}^p(t)$.

5) $(1 - F_n^{pR}(t))(1 - G_n^{pR}(t)) = 1 - H_n^E(t)$ for all $t \in \mathbb{R}$.

For demonstrating properties of estimators, at first, we consider following ordered natural censored sample $\{(Z_{(k)}, \delta_{(k)}), k = 1, \dots, 97\}$ of size $n = 97$ from paper [21] (well known Chenning Hause data) (table 1):

Table 1

Ordered natural censored sample $\{(Z_{(k)}, \delta_{(k)}), k = 1, \dots, 97\}$ of size $n = 97$

k	$(Z_{(k)}, \delta_{(k)})$	k	$(Z_{(k)}, \delta_{(k)})$	k	$(Z_{(k)}, \delta_{(k)})$	k	$(Z_{(k)}, \delta_{(k)})$	k	$(Z_{(k)}, \delta_{(k)})$
9	(777,1)	21	(940,0)	41	(977,0)	61	(1018,0)	81	(1059,1)
2	(781,0)	22	(942.5,0)	42	(983,1)	62	(1022,1)	82	(1060,1)
3	(843,0)	23	(943,0)	43	(984,0)	63	(1023,0)	83	(1060,0)
4	(866,0)	24	(945,1)	44	(985,1)	64	(1025,1)	84	(1064,0)
5	(869,1)	25	(945,0)	45	(989,1)	65	(1027,0)	85	(1070,0)
6	(872,1)	26	(948,1)	46	(992.5,1)	66	(1029,1)	86	(1073,0)
7	(876,1)	27	(951,0)	47	(993,1)	67	(1031,1)	87	(1080,1)
8	(893,1)	28	(953,0)	48	(996,1)	68	(1031,0)	88	(1085,1)
9	(894,1)	29	(956,0)	49	(998,1)	69	(1031.5,0)	89	(1093,0)
10	(895,0)	30	(957,1)	50	(1001,0)	70	(1033,1)	90	(1093.5,1)
11	(898,1)	31	(957,0)	51	(1002,0)	71	(1036,1)	91	(1094,1)
12	(906,0)	32	(959,0)	52	(1005,0)	72	(1043,1)	92	(1106,0)
13	(907,1)	33	(960,0)	53	(1006,0)	73	(1043,0)	93	(1107,0)
14	(909,1)	34	(966,1)	54	(1009,1)	74	(1044,1)	94	(1118,0)
15	(911,1)	35	(966,0)	55	(1011.5,1)	75	(1044,0)	95	(1128,1)
16	(911,0)	36	(969,1)	56	(1012,1)	76	(1045,0)	96	(1139,1)
17	(914,0)	37	(970,0)	57	(1012,0)	77	(1047,0)	97	(1153,0)
18	(927,1)	38	(971,1)	58	(1013,0)	78	(1053,1)		
19	(932,1)	39	(972,0)	59	(1015,0)	79	(1055,1)		
20	(936,0)	40	(973,0)	60	(1016,0)	80	(1058,0)		

These data in months, where “1”-corresponds to uncensoring and otherwise “0” to censoring. In the next figures 1–4 we demonstrate estimators $1 - F_n^{AB}(t)$, $1 - F_n^{KM}(t)$, $1 - F_n^{RR}(t)$ and $1 - F_n^{PR}(t)$ separately and in picture 5 all estimators together.

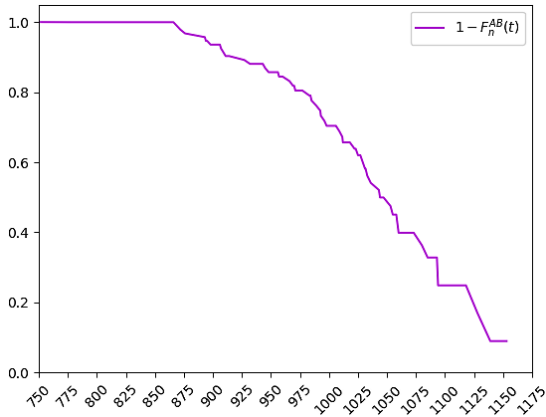


Fig. 1. Estimator $1 - F_n^{AB}(t)$

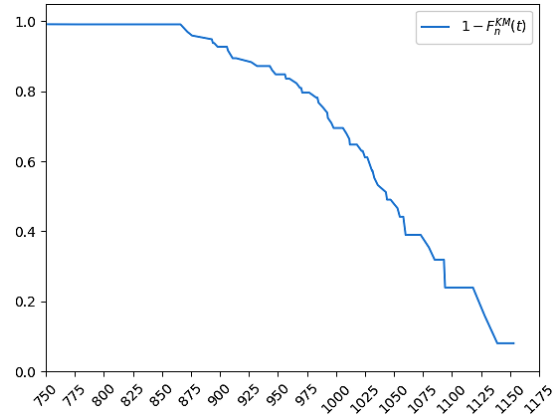


Fig. 2. Estimator $1 - F_n^{KM}(t)$

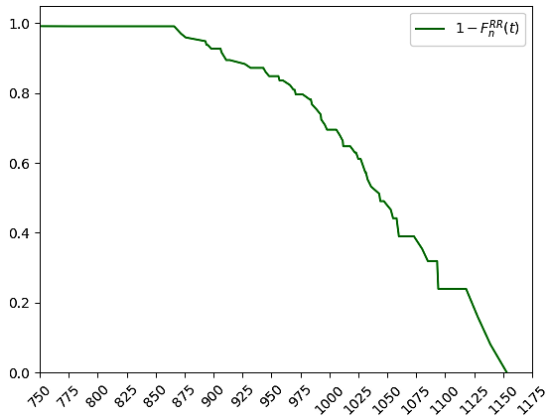


Fig. 3. Estimator $1 - F_n^{RR}(t)$

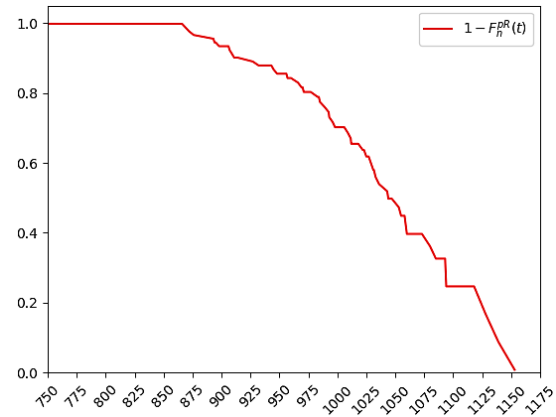


Fig. 4. Estimator $1 - F_n^{PR}(t)$

In figures 3, 4 one can see that only power estimators $F_n^{RR}(t)$ and $F_n^{PR}(t)$ well defined in the whole line.

Now we used Gillespie-Fisher's result about 95%-confidence bands of estimators $1 - F_n^{AB}(t)$, $1 - F_n^{KM}(t)$, $1 - F_n^{RR}(t)$ and $1 - F_n^{PR}(t)$ considering its asymptotically equivalence. Propose band in the form (see, [22]):

$$B_n^*(t, \eta_1, \eta_2) = [B_n^{*-}(t, \eta_1, \eta_2); B_n^{*+}(t, \eta_1, \eta_2)],$$

where

$$B_n^{*-}(t, \eta_1, \eta_2) = F_n^*(t) - n^{-\frac{1}{2}} \cdot (1 - F_n^*(t)) \cdot \left(\eta_1 \cdot d_n^{\frac{1}{2}}(M) + \eta_2 \cdot d_n(t) \cdot d_n^{-\frac{1}{2}}(M) \right),$$

$$B_n^{*+}(t, \eta_1, \eta_2) = \frac{F_n^*(t) + n^{-\frac{1}{2}} \cdot \left(\eta_1 \cdot d_n^{\frac{1}{2}}(M) + \eta_2 \cdot d_n(t) \cdot d_n^{-\frac{1}{2}}(M) \right)}{1 + n^{-\frac{1}{2}} \cdot \left(\eta_1 \cdot d_n^{\frac{1}{2}}(M) + \eta_2 \cdot d_n(t) \cdot d_n^{-\frac{1}{2}}(M) \right)},$$

$$M = 1128; \eta_1 = 1; \eta_2 = 1.37 \text{ and } d_n(t) = \int_{-\infty}^t \frac{dH_{(1)n}^E(u)}{(1 - H_n^E(u) + 1/n)^2}.$$

Here $F_n^*(t)$ one of there estimators $F_n^{AB}(t)$, $F_n^{KM}(t)$, $F_n^{RR}(t)$ and $F_n^{pR}(t)$. These confidence bands we demonstrate in figures 5–8.

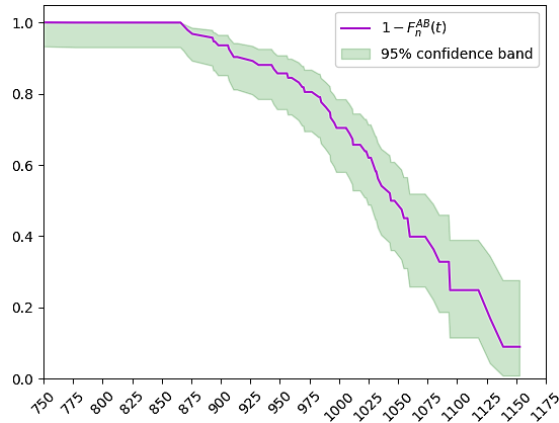


Fig. 5. Estimator $1 - F_n^{AB}(t)$

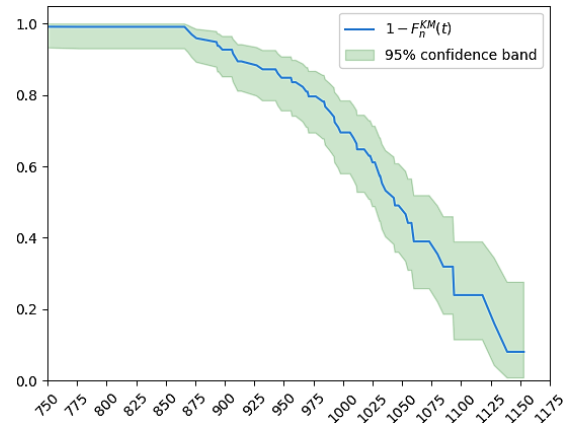


Fig. 6. Estimator $1 - F_n^{KM}(t)$

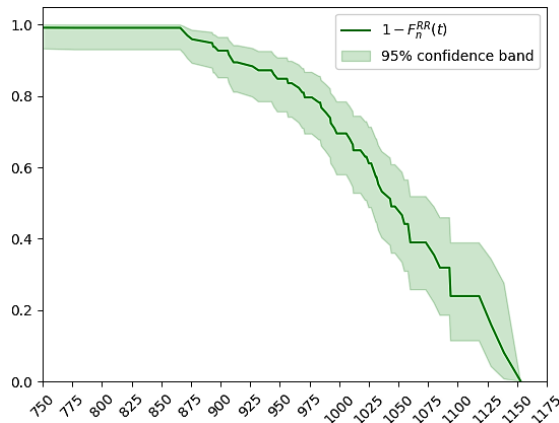


Fig. 7. Estimator $1 - F_n^{RR}(t)$

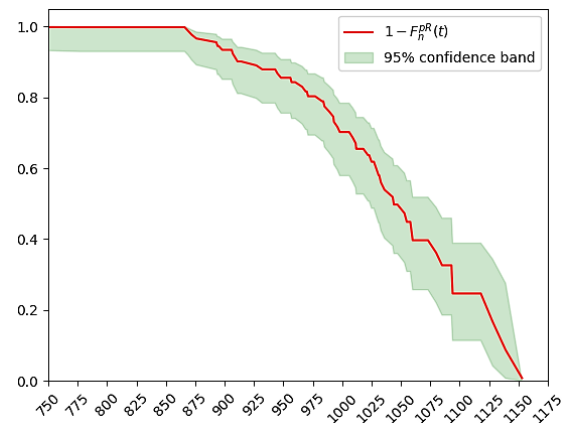


Fig. 8. Estimator $1 - F_n^{pR}(t)$

Now we consider statistics $R_n(t)$, $F_n^{RR}(t)$ and presmoothed estimators $R_n^p(t)$, $F_n^{pKM}(t)$ and $F_n^{pR}(t)$. Artificial data are modelled by exponential distribution $F(t) = 1 - e^{-t}$, $t \geq 0$ for several sample sizes n . In order to investigate estimator $F_n^{KM}(t)$ we choose $n = 500$ with censoring degree 30% (figure 9; table 2). In order to investigate estimator $R_n(t)$ we choose $n = 500$ with censoring degree 30% (figure 10). In the figure 11 we demonstrate estimator $F_n^{RR}(t)$ for artificial data of size $n = 500$ degree of censoring equal 30% (table 3).

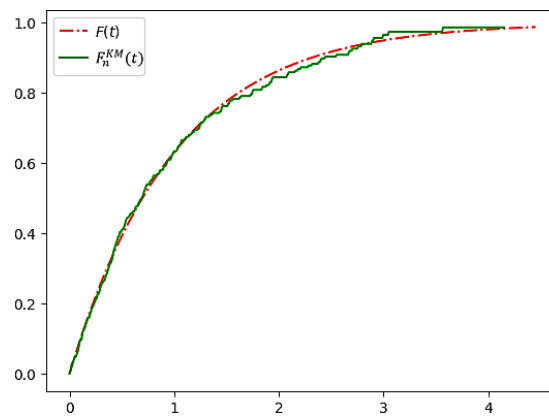


Fig. 9. Estimator $F_n^{KM}(t)$

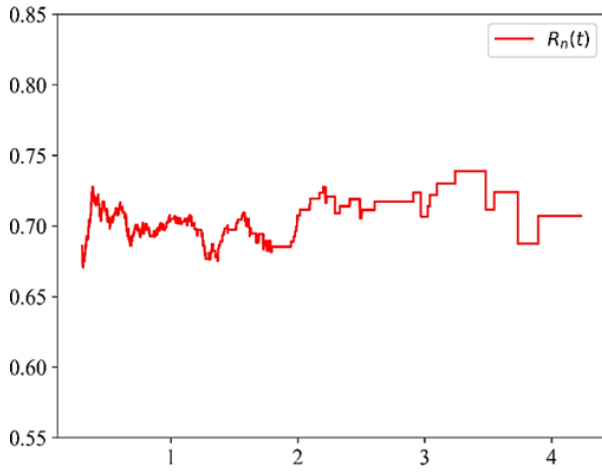


Fig. 10. Estimator $R_n(t)$

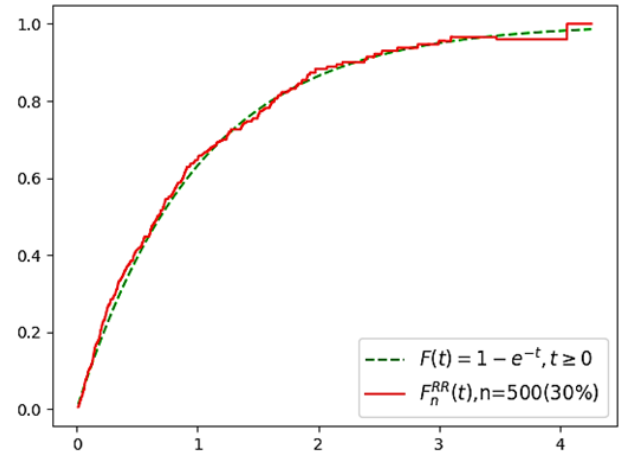


Fig. 11. Estimator $F_n^{RR}(t)$

In figure 12 we demonstrate estimator $R_n^p(t)$ for several smoothing parameter and in figure 13 estimator $F_n^{pKM}(t)$.

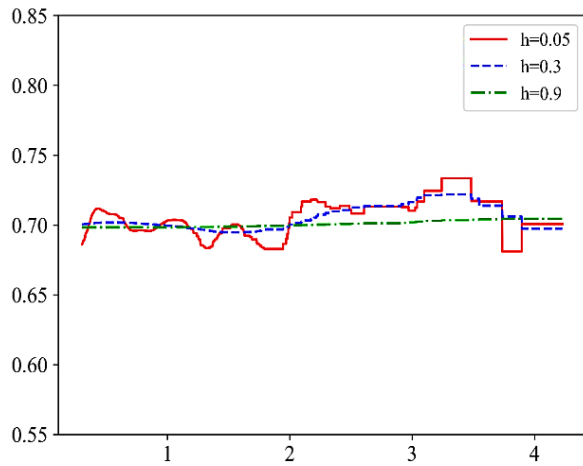


Fig. 12. Estimator $R_n^p(t)$

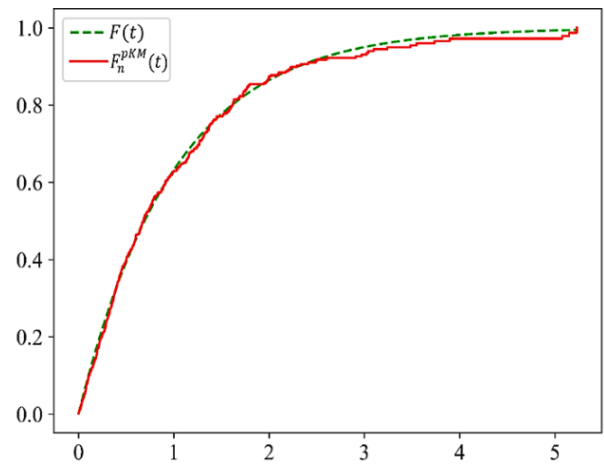


Fig. 13. Estimator $F_n^{pKM}(t)$

In figure 14 we demonstrate presmoothed relative risk power estimator $F_n^{pR}(t)$ for artificial data size $n = 500$ with degree of censoring 30%, (table 4).

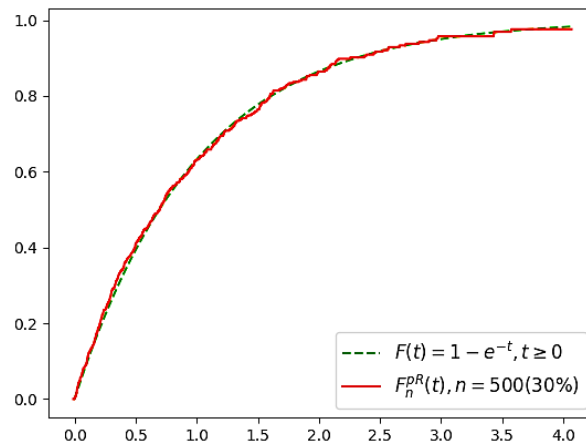


Fig. 14. Estimator $F_n^{pR}(t)$

Table 2

$F_n^{KM}(t)$ deviation of the price from the theoretical distribution

censoring \ n	$n = 100$	$n = 500$	$n = 1\ 000$	$n = 2\ 000$
10 %	0,08703	0,03948	0,02793	0,01977
20 %	0,09014	0,04079	0,02889	0,02041
30 %	0,09482	0,04317	0,03052	0,02166
40 %	0,10512	0,04846	0,03456	0,02453
50 %	0,12387	0,06106	0,04431	0,03233
60 %	0,15258	0,08502	0,06528	0,05020
70 %	0,23185	0,14288	0,11633	0,09412
80 %	0,36522	0,26567	0,23120	0,20212

Table 3

$F_n^{RR}(t)$ deviation of the price from the theoretical distribution

censoring \ n	$n = 100$	$n = 500$	$n = 1\ 000$	$n = 2\ 000$
10 %	0,08576	0,03891	0,02778	0,01946
20 %	0,08620	0,03941	0,02832	0,01977
30 %	0,08721	0,04036	0,03006	0,02124
40 %	0,08845	0,04257	0,03195	0,02207
50 %	0,10239	0,05093	0,03661	0,02596
60 %	0,14121	0,07887	0,05906	0,03708
70 %	0,18716	0,12348	0,10134	0,07319
80 %	0,21807	0,16330	0,14907	0,12187

Table 4

$F_n^{pR}(t)$ deviation of the price from the theoretical distribution

censoring \ n	$n = 100$	$n = 500$	$n = 1\ 000$	$n = 2\ 000$
10 %	0,08561	0,03812	0,02767	0,01932
20 %	0,08617	0,03871	0,02830	0,01956
30 %	0,08712	0,03945	0,02902	0,02017
40 %	0,08785	0,04105	0,03112	0,02197
50 %	0,09034	0,04789	0,03538	0,02487
60 %	0,12056	0,07012	0,05423	0,02718
70 %	0,17120	0,10256	0,09558	0,07018
80 %	0,19823	0,14993	0,13554	0,11956

At the end of section we consider other extension of power estimator (8). In the model of right random censorship we consider other sample $\tilde{C}^{(n)} = \{(Z_k, \delta_k, C_k), k=1, \dots, n\}$ is available for observation, where $Z_k = \min(X_k, Y_k)$, $\delta_k = I(X_k \leq Y_k)$. Here r.v.'s X_k and Y_k are independent given covariate $C_k = x$ with values in the interval $(0, 1]$. For example, this corresponds to experiment, where for tested object (individual or physical system) x denote dose of drug in medicine or temperature, press or voltage in industry. Now the problem consists estimating of conditional d.f.

$$F_x(t) = P(X_k \leq t | C_k = x)$$

under nuisance d.f. $G_x(t) = P(Y_k \leq t | C_k = x)$ and

$$1 - H_x(t) = P(Z_k > t | C_k = x) = (1 - F_x(t))(1 - G_x(t)), t \in R.$$

Consider relative risk power estimator (8) for $F_x(t)$ in presence of covariate in the form

$$F_{xh}(t) = 1 - \left(1 - H_{xh}^E(t)\right)^{R_{xh}(t)} = \begin{cases} 0, & t < Z_{(1)}, \\ 1 - \left(\frac{n-k}{n}\right)^{R_{xh}(t)}, & Z_{(k)} \leq t < Z_{(k+1)}, 1 \leq k \leq n-1, \\ 1, & t \geq Z_{(n)}, \end{cases} \quad (25)$$

where

$$H_{xh}^E(t) = \sum_{k=1}^n \omega_{nk}(x, h) I(Z_k \leq t) = H_{(0)xh}^E(t) + H_{(1)xh}^E(t), \quad H_{(m)xh}^E(t) = \sum_{k=1}^n \omega_{nk}(x, h) I(Z_k \leq t, \delta_k = m), \quad m = 0, 1.$$

are Stoun's type estimators with Gasser-Miiller's wights:

$$\omega_{nk}(x, h) = \frac{1}{C_n(x; h)} \int_{x_{k-1}}^{x_k} \frac{1}{h} k\left(\frac{x-u}{h}\right) du, \quad k = 1, \dots, n,$$

and

$$C_n(x; h) = \int_0^{x_n} \frac{1}{h} k\left(\frac{x-u}{h}\right) du.$$

Usually for large enough n , $C_n(x; h) \approx 1$. For our purposes we put $C_n = 1$. Here

$$R_{xh}(t) = \Lambda_{(1)xh}(t) \cdot (\Lambda_{xh}(t))^{-1}, \quad \Lambda_{xh}(t) = \Lambda_{(0)xh}(t) + \Lambda_{(1)xh}(t), \quad \text{and} \quad \Lambda_{(m)xh}(t) = \int_{-\infty}^t \frac{dH_{(m)xh}^E(u)}{1 - H_{xh}^E(u)}, \quad m = 0, 1.$$

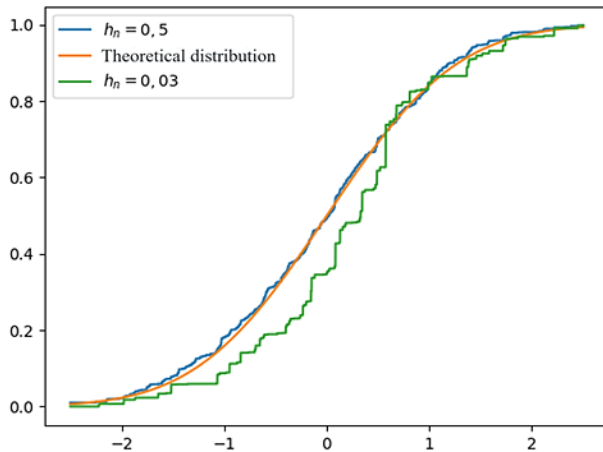


Fig. 15. Estimator

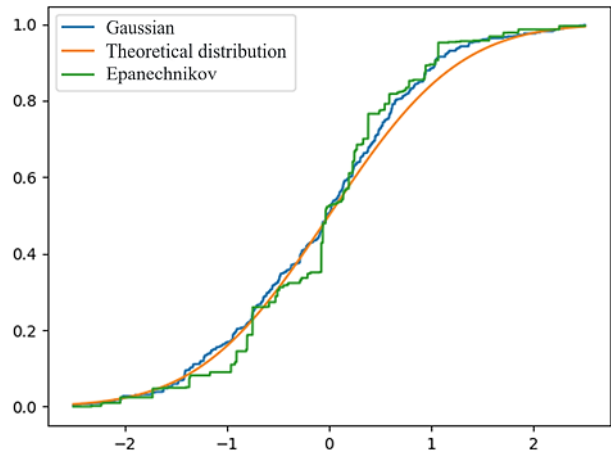


Fig. 16. Estimator

Table 5

Deviation of the estimate from the theoretical distribution

$h(n) \backslash u$	0,1	0,2	0,3	0,5	0,7	0,9	1
Gaussian kernel function							
0,1	0,07232	0,06524	0,06251	0,06360	0,06343	0,07257	0,08896
0,3	0,04902	0,04486	0,04345	0,04075	0,04268	0,04891	0,05133
0,5	0,04095	0,03994	0,03871	0,03813	0,03927	0,04075	0,04200
0,7	0,03933	0,03849	0,03831	0,03781	0,03804	0,03835	0,03913
0,9	0,03826	0,03770	0,03786	0,03704	0,03707	0,03806	0,03832
Epanechnikov kernel function							
0,1	0,09185	0,09226	0,09216	0,09190	0,09182	0,09273	0,12820
0,3	0,06519	0,05647	0,05496	0,05360	0,05267	0,06499	0,07467
0,5	0,05275	0,04857	0,04487	0,04106	0,04551	0,05322	0,05875
0,7	0,04627	0,04323	0,04032	0,03832	0,04076	0,04547	0,04897
0,9	0,04100	0,03923	0,03844	0,03819	0,03865	0,04137	0,04353

In proportional hazards model with covariate we consider conditional d.f. as $F_x(t) = 1 - \exp\{-(1 - 2\log x)t\}$, $t \geq 0$, $0 < x \leq 1$, and let $\varepsilon_x(h) = \sup_{t \geq 0} |F_x(t) - F_{xh}(t)|$.

For generating of $n = 1\,000$ artificial data with degree of censoring 30% for Gaussian and Epanechnikov kernels, we demonstrate estimator (25) for several smoothing parameters h (fig. 15, 16; table 5). The nearest to zero values of distance $\varepsilon_x(h)$, we see for Gaussian kernel.

Conclusion

In this paper authors are comparatively considering in right random censorship model three types of survival function estimators: presmoothed power, exponential (Altshuler-Breslow) and product-types (Kaplan-Meier). In simulation study they used natural sample of size $n=97$ and artificial data for several sample size. It is well known that in right random censoring model survival function of observed minima is the product of survival functions of censored and censoring random variables. Authors are demonstrating that this property is easily fulfilled for two presmoothed power type estimators. But for other two estimators: of exponential and product-limit it is not fully filled, because of their uncertainty in whole line. Hence, only presmoothed power-type estimators are identifiably with considered right censorship model.

References

1. Abdushukurov, A.A., Bozorov, S.B., & Mansurov, D.R. (2024) Optimal kernel density estimation by statistical modeling with random censoring of observations. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravleniye, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 67. pp. 49–59. DOI: 10.17223/19988605/67/6
2. Akritas, M.G. (2000) The central limit theorem under censoring. *Bernoulli*. 6(6). pp. 1109–1120.
3. Breslow, N. (1972). Discussion of professor Cox's paper. *Journal of Royal Statistics Society*. 34(A). pp. 216–217.
4. Breslow, N. & Crowley, J. (1974) A large sample study of the life table and product-limit estimates under random censorship. *Annual Statistics*. 2. pp. 437–453.
5. Burke, M.D., Csörgő, S. & Horváth, L. (1981) Strong approximations some biometric estimates under random censorship. *Z. Wahrscheinlich. Verw. Gebiete*. 56. pp. 87–112.
6. Burke, M.D., Csörgő, S. & Horváth, L. (1988) A correction to and improvement of "Strong approximations of some biometric estimates under random censorship." *Probability Theory and Related Fields*. 79. pp. 51–57.
7. Csörgő, S. (1996). Universal Gaussian approximations under random censorship. *Annual Statistics*. 24(6). pp. 2744–2778.
8. Csörgő, S., & Horváth, L. (1982). On random censorship from the right. *Acta Sci. Math*. 44. pp. 23–34.
9. Kaplan, E.L. & Meier, P. (1958) Nonparametric estimation from incomplete observations. *Journal of American Statistics Association*. 53. pp. 457–481.
10. Abdushukurov, A.A., & Mansurov, D.R. (2023) Estimating the distribution function using parametric methods in informative model of random censorship from both sides. In: Dudin, A., Nazarov, A. & Moiseev, A. (eds) *Communication in Computer and Information Sciences (CCIS)*. Springer. pp. 2025–2037.
11. Altshuler, B. (1970) Theory of measurements for competing risks in animal experiments. *Math. Biosciences*. 6. pp. 1–11.
12. Abdushukurov, A.A. (1998) Nonparametric estimation of the distribution function based on relative risk function. *Communication in Statistics – Theory & Methods*. 27(8). pp. 1991–2012.
13. Abdushukurov, A.A. (1999) On nonparametric estimation of probability indices by censored samples. *Theory of Probability and its Applications*. 43(1). pp. 3–11.
14. Abdushukurov, A.A. (2011) *Estimation of Unknown Distributions by Incomplete Observations*. LAMBERT Academic Publishing (LAP).
15. Jacome, M.A. & Cao, R. (2004) Presmoothed kernel density estimator for censored data. *Journal of Nonparam. Statist.* 16. pp. 289–309.
16. Abdushukurov, A.A. (1987) Nonparametric estimation in Proportional Hazards Model. *Technical Report N3448-B* -87.
17. Abdushukurov, A.A. (1987) Estimating of probability density and intensity functions of the Koziol–Green model of random censoring. *Sankhya Indiran Journal of Statistics*. 48(1). pp. 150–168.
18. Cheng, P.E. & Ling, G.D. (1987) Maximum likelihood estimation of a survival function under Koziol–Green proportional hazards model. *Statist. Probability Letters*. 5. pp. 75–80.
19. Csörgő, S. (1989) Testing for the proportional hazards model of random censorship. *Proceedings IV Prague Symp. Asump. Statist.* Charles Univ. Press. Prague. pp. 41–53.
20. Abdushukurov, A.A., Nurmukhamedova, N.S. & Bozorov, S.B. (2021) Nonparametric estimation of the distribution function under right random censoring based on presmoothed relative risk function. *Lobachevskii Journal of Mathematics*. 42(2). pp. 257–268.
21. Efron, B. (1981) Censored data and the Bootstrap. *The Journal of the Acoustical Society of America*. 76(374). pp. 312–319.
22. Csörgő, S. & Horváth, L. (1986) Confidence bands from censored samples. *Canadian Journal of Statistics*. 14(2). pp. 131–144.

Information about the authors:

Abdushukurov Abdurakhim A. (Doctor of Physical and Mathematical Sciences, Professor, Moscow State University, Branch in Tashkent, Tashkent, Uzbekistan). E-mail: a_abdushukurov@rambler.ru

Bozorov Sukhrob B. (PhD Student, Gulistan State University, Gulistan, Uzbekistan.). E-mail: suxrobbek_8912@mail.ru

Contribution of the authors: the authors contributed equally to this article. The authors declare no conflicts of interests.

Информация об авторах:

Абдушукуров Абдурахим Ахмедович – доктор физико-математических наук, профессор Московского государственного университета, филиала в Ташкенте (Ташкент, Узбекистан). E-mail: a_abdushukurov@rambler.ru

Бозоров Сухроб Баходирович – докторант Гулистанского государственного университета (Гулистан, Узбекистан). E-mail: suxrobbek_8912@mail.ru

Вклад авторов: все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

Received 22.12.2024; accepted for publication 02.06.2025

Поступила в редакцию 22.12.2024; принята к публикации 02.06.2025

Научная статья

УДК 004.95

doi: 10.17223/19988605/71/7

**Мультиосевая модель геофизических данных:
концепция, формализация, реализация****Гульнара Равиловна Воробьева¹, Андрей Владимирович Воробьев², Глеб Олегович Орлов³**^{1,2,3} Уфимский университет науки и технологий, Уфа, Россия¹ Gulnara.vorobeva@gmail.com² Geomagnet@list.ru³ orlovgleb99@mail.ru

Аннотация. Известно, что геофизические поля (геомагнитное, гравитационное и электромагнитное) в контексте регистрируемых или моделируемых данных представляют собой совокупность нескольких соосных векторных составляющих, характеризующих изменение соответствующего параметра в пространстве и во времени. Практикуемые в настоящее время подходы к организации хранения таких данных базируются, как правило, на одной из моделей данных. При этом совместное хранение метаданных источника и непосредственно геофизических данных во многих случаях сопряжено с неэффективным расходованием сопутствующих вычислительных ресурсов. В данной работе рассматривается подход, ориентированный на раздельное хранение, но совместное комплексное применение метаданных и данных средств регистрации геофизических полей с учетом их многовекторной природы и тензорного представления анализируемых параметров. На примере геомагнитного поля рассматривается эффективность предлагаемого подхода в контексте оптимизации хранения данных, с одной стороны, и повышения вычислительной скорости выполнения запросов к ним – с другой.

Ключевые слова: геофизические данные; тензорное поле; векторное поле; хранение данных; обработка и манипулирование прикладными данными.

Благодарности: Исследование выполнено за счет гранта Российского научного фонда № 25-21-00143, <https://rscf.ru/project/25-21-00143/>

Для цитирования: Воробьева Г.Р., Воробьев А.В., Орлов Г.О. Мультиосевая модель геофизических данных: концепция, формализация, реализация // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2025. № 71. С. 71–81. doi: 10.17223/19988605/71/7

Original article

doi: 10.17223/19988605/71/7

Multi-axis model of geophysical data: concept, formalization, implementation**Gulnara R. Vorobeveva¹, Andrey V. Vorobev², Gleb O. Orlov³**^{1,2,3} Ufa University of Science and Technology, Ufa, Russian Federation¹ Gulnara.vorobeva@gmail.com² Geomagnet@list.ru³ orlovgleb99@mail.ru

Abstract. It is known that geophysical fields (geomagnetic, gravitational and electromagnetic) in the context of recorded or modeled data are a set of several coaxial vector components characterizing the change in the corresponding parameter in space and time. Currently practiced approaches to organizing the storage of such data are based, as a rule, on one of the data models. At the same time, joint storage of source metadata and geophysical data directly in many cases is associated with inefficient use of associated computing resources. This paper considers an approach focused on separate storage, but joint complex application of metadata and data of geophysical field recording tools, taking into account their multi-vector nature and tensor representation of the analyzed parameters. Using the example of the

geomagnetic field, the effectiveness of the proposed approach is considered in the context of optimizing data storage, on the one hand, and increasing the computational speed of executing queries to them, on the other.

Keywords: geophysical data; tensor field; vector field; data storage; processing and manipulation of applied data.

Acknowledgments: The research was carried out with the financial support of the Russian Science Foundation grant No. 25-21-00143, <https://rscf.ru/project/25-21-00143/>

For citation: Vorobeva, G.R., Vorobev, A.V., Orlov, G.O. (2025) Multi-axis model of geophysical data: concept, formalization, implementation. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 71. pp. 71–81. doi: 10.17223/19988605/71/7

Введение

На сегодняшний день моделирование и регистрация геофизических полей оказывают большое влияние на исследование процессов, протекающих в недрах и на поверхности Земли. Данная информация используется как в научных исследованиях, так и для поддержки принятия решений в условиях доступа к данным, предоставляемым магнитными обсерваториями и вариационными станциями [1, 2]. При этом формат представления указанной информации также определяет ее владельца. Хотя в соответствующих предметных областях действуют собственные стандарты и спецификации представления геофизических данных (например, спецификация IAGA–2002 для представления данных геомагнитного поля [1]), различные поставщики данных привносят в него собственные «диалекты» (по аналогии с различными СУБД реляционного типа, практикующими различные вариации стандартизованного языка запросов SQL на своих платформах).

В некоторых ситуациях в целях оценки отклонения от нормы требуется моделирование идеального (невозмущенного) геофизического поля. Его параметры могут быть вычислены с использованием таких моделей, как IGRF [1] и WMM [3]. Формат представления результатов вычислений определяется разработчиком программных реализаций соответствующих моделей и в целом может отличаться от общепринятого в соответствующей предметной области, преследуя локальные цели решения поставленной задачи [4].

При этом представляется важным отметить, что формат представления данных должен учитывать большой и непрерывно возрастающий объем геофизической информации. Так, к примеру, магнитные обсерватории регистрируют параметры геомагнитного поля ежеминутно на протяжении многих лет, при этом количество станций только увеличивается, соответственно, и экспоненциально год за годом возрастает объем используемой потребителями геофизической информации. Проблема усложняется тензорным характером рассматриваемых полей: соответствующие векторы не являются атомарными, определяются несколькими (в том числе разнонаправленными) составляющими.

Отличительной особенностью рассматриваемых полей является их тензорный характер, который в упрощенном виде может быть описан некоторой исходной (чаще пространственной) точкой с исходящими из нее векторами, в свою очередь, задающими определенные атрибутивные параметры рассматриваемого процесса. Фактически тензорное поле при этом представляет собой задаваемую функцией поверхность, каждой точке которой ставится в соответствие тензор, привязанный к соответствующей системе координат и имеющий начало в заданной пространственной точке [5, 6]. При этом в качестве независимых элементов тензора могут быть приведены скорости изменения параметров вектора геомагнитного поля по трем направлениям в декартовой системе координат [7–9].

Соответственно, необходимо разработать такую модель хранения данных геофизических полей, которая учитывает их мультиосевую структуру анализируемых параметров с динамически меняющимися значениями в контексте относительно статичных метаданных, в совокупности характеризующих пространственное распределение заданных источников данных. Представляются целесообразными раздельные представление и хранение обозначенных данных для оперативной обработки информации о геофизических полях, в том числе в системах поддержки принятия решений на этой основе. Для оценки эффективности рассматриваемого подхода к хранению данных видится необходимым сравнить

время выполнения простых однопредикатных и сложных многопредикатных запросов к данным с иными известными моделями представления – реляционной и CSV-подобной.

1. Мультиосевая модель хранения данных тензорных полей

Для обеспечения хранения геомагнитных данных в комплексном виде предлагается так называемая мультиосевая структура. Ее центральным компонентом служит некоторый источник геомагнитных данных, для которого осуществляется хранение регистрируемой информации, например геомагнитная станция или ее цифровая модель. Каждому центральному звену многоосевой модели ставятся в соответствие три оси данных, в которых содержатся пары данных в формате «атрибут–значение», при этом количество таких пар неограниченно. Беря во внимание специфику рассматриваемого тензорного геомагнитного поля, предлагается использование следующих осей, логически взаимосвязанных с центральным звеном модели (рис. 1):

- ось метаданных (статичная) содержит основные параметры источника. Если в качестве центрального звена выступает геомагнитная станция или обсерватория, то на данной оси будет такая информация, как наименование станции, ее расположение, принадлежность к ведомству, информация о техническом оснащении, используемые единицы измерения и шаг регистрации;

- ось оперативных данных (динамическая) обеспечивает хранение информации в реальном времени с учетом шага дискретизации, определенного на оси метаданных. Для текущего момента времени результаты регистрации или расчета параметров геофизического поля декомпозируются на отдельные независимые пары вида «атрибут–значение атрибута». Данные этой оси расположены в узлах с временной меткой, каждому из которых ставятся в соответствие дочерние узлы с параметрами геомагнитного поля;

- ось архивных данных (статичная) необходима для инициализации сохранения данных. По истечении времени, заданного шагом дискретизации, ветка динамической оси должна быть перемещена в статичную ось архивных данных. При перемещении она станет новой веткой последней. В результате ось архивных данных будет включать в себя набор узлов в виде временных меток, каждой из которых соответствует дочерняя ветка узлов с атрибутивными параметрами поля.

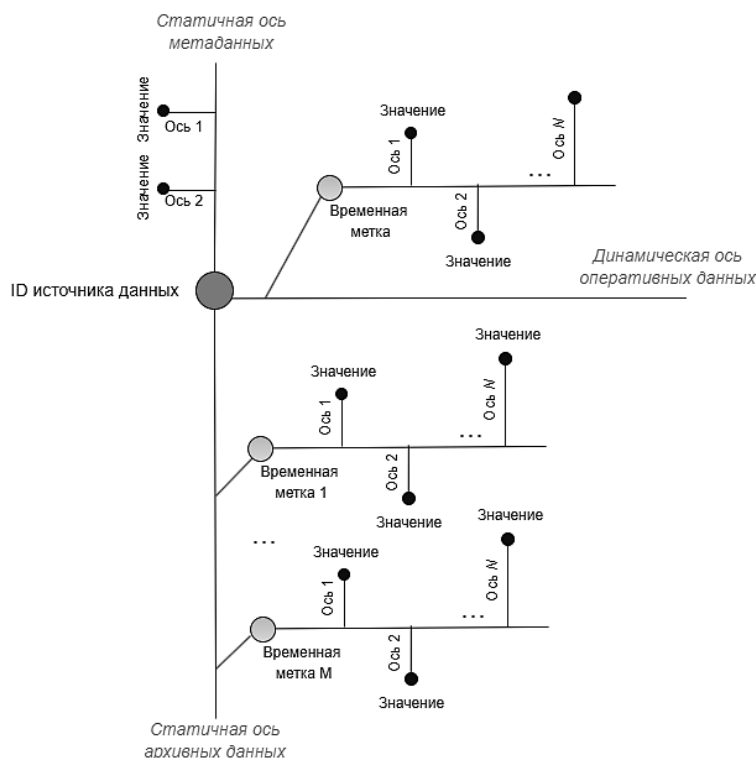


Рис. 1. Обобщенная схема мультиосевой модели хранения данных тензорных геофизических полей
Fig. 1. Generalized scheme of the multi-axis model for storing data of tensor geophysical fields

Обмен данными между статичной и динамичной осями устроен так, что данные динамичной оси перемещаются на нее без изменений (рис. 2), при этом допускается внедрение еще одного узла, который производил бы преобразование данных динамической оси в данные статической оси.

Миграция узла в статичную ось из динамической производится при помощи специального триггера, который настраивается на определенный шаг дискретизации. Этот шаг может быть определен на статической оси. Общее количество узлов на статической оси не ограничено, в то время как на динамической оси одной временной метке может соответствовать только один узел данных.

Узел оси метаданных, соответствующий геопространственной привязке, является основным связующим звеном между экземплярами центрального элемента (компонента) соответствующей модели данных поля. Результатом обозначенной связи нескольких центральных элементов (экземпляров модели) является множество пространственных точек с атрибутивными значениями в некотором пространственно-временном интервале.

Дальнейшее связывание осуществляется по общим узлам соответствующих статичных осей архивных данных таким образом, чтобы сформировать набор геопространственных данных, характеризующих пространственно-временное изменение тех или иных параметров (или их совокупности) геофизического поля.

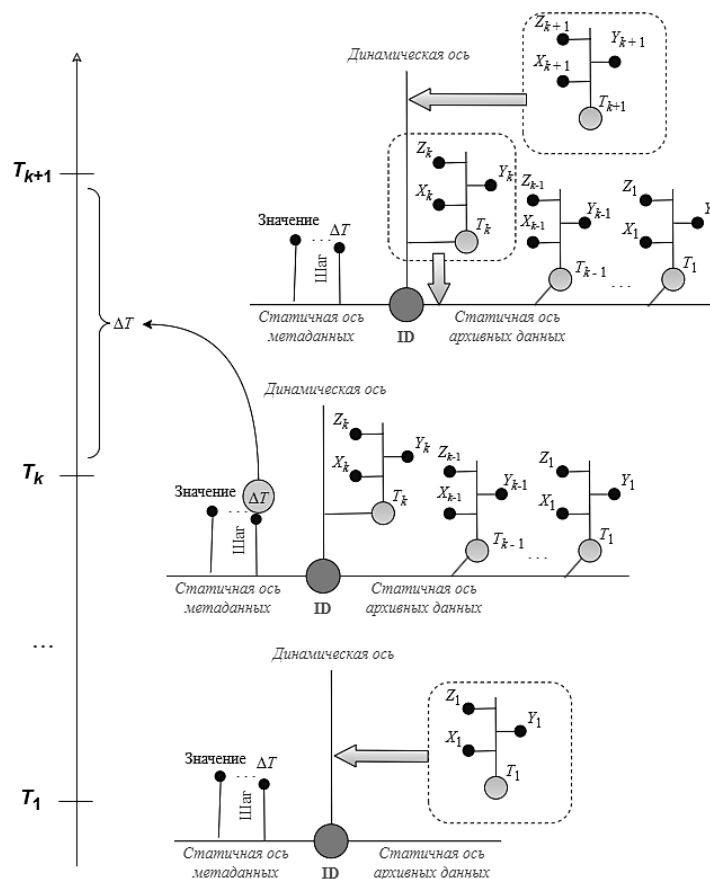


Рис. 2. Схема взаимодействия между статичной и динамической осями модели данных

Fig. 2. Diagram of interaction between static and dynamic axes of the data model

В процессе выборки данных из одного или нескольких экземпляров мультиосевой модели данных геофизических полей выполняется их срез по пространственно-временным параметрам, указываемым в запросе на извлечение данных. При этом представляется необходимым отметить формирование новой мультиосевой структуры, центральный элемент которой определяется спецификой соответствующего запроса на выборку геофизических данных.

В простейшем случае целью запроса является получение данных от единственного источника за определенный временной период. В этом случае выполняется поиск искомого узла в хранилище,

причем в качестве параметров поиска выступают один или несколько атрибутов, представленных в статичной оси метаданных. В результате формируется новая мультиосевая структура, центр которой совпадает с соответствующим экземпляром модели, а оси совпадают только частично: статичная ось метаданных (при необходимости) копируется в новую структуру без изменений (либо не копируется вовсе, тогда результат не содержит статичную ось), динамические оси могут также отсутствовать (если не требуется получения оперативной информации о состоянии поля), статичная ось архивных данных фильтруется таким образом, что в ней остаются только те узлы, которые соответствуют указанным в параметрах запроса временным меткам.

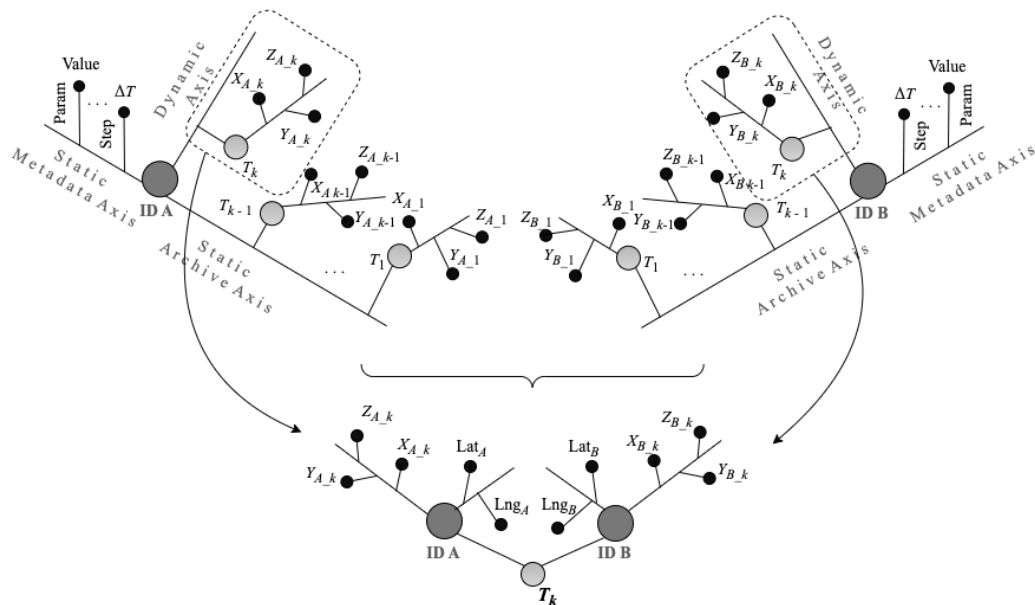


Рис. 3. Схема формирования новой мультиосевой модели при выборке данных геофизических полей
Fig. 3. Scheme of formation of a new multi-axis model when sampling geophysical field data

В более сложном случае необходимо сформировать структуру, базирующуюся на данных регистрации или моделирования геофизического поля несколькими источниками (в разных пространственных точках) за определенный период времени. В этом случае центральным элементом новой структуры является искомая временная метка, для которой необходимо определить комплекс геофизических данных. Каждый отдельный источник, по которому присутствует узел с одноименной временной меткой (в любой из осей данных), формирует новую ось, исходящую из центральной оси с заданной временной меткой. Как и в предыдущем случае, для каждой новой оси (по источнику данных) полностью копируется статичная ось метаданных (при необходимости), либо новая структура обходится полностью без нее. Здесь отметим, что операции формирования новой оси метаданных считаются в данном случае вычислительно нецелесообразными ввиду небольшого числа описываемых атрибутов. Оси данных также формируются в зависимости от искомой временной метки. Фактически в результирующей структуре сохраняется единственная ось данных, содержащая фрагмент с данными, соответствующими искомой временной метке (рис. 3).

2. Формализация мультиосевой модели данных геофизических полей

Мультиосевая модель данных геофизических полей в общем виде представляет собой тройку параметров следующего формата:

$$A = (S_m, D_d, S_d), \quad (1)$$

где S_m – статичная ось метаданных, D_d – динамическая ось оперативных данных, S_d – статическая ось архивных данных, A – центральный компонент модели, характеризующий источник или геопрограммную точку привязки геофизических данных.

Каждая из представленных осей содержит, в свою очередь, элементы, соответствующие описываемым ею данным. Так, статичная ось метаданных S_m представляет собой множество метаданных, заданных совокупность пар вида «атрибут–значение», каждая из которых, в свою очередь, является дочерней ветвью иерархии с родительским элементом в центральном компоненте модели, что может быть описано соотношениями вида:

$$S_m = \bigcup_{i=1}^n \{a_i^s, v_i^s\}; \quad (2)$$

где n – количество атрибутов в оси метаданных модели, a_i^s, v_i^s – атрибут и значение атрибута метаданных соответственно.

Аналогичным образом может быть представлена структура динамической оси оперативных данных и статичной оси архивных данных геофизических полей:

$$D_d = \bigcup_{i=1}^k \{a_i^o, v_i^o\}; \quad (3)$$

$$S_d = \bigcup_{i=1}^l \{a_i^d, v_i^d\};$$

где k, l – количество атрибутов в осях оперативных и архивных данных модели соответственно, a_i^o, v_i^o – атрибут и значение атрибута оперативных данных, a_i^d, v_i^d – атрибут и значение атрибута архивных данных.

Каждому атрибуту оси модели ставится в соответствие определенный домен, характеризующий множество возможных значений рассматриваемого атрибута оси (атрибуты, в свою очередь, могут характеризовать как данные, так и метаданные, однако в последнем случае ввиду статичности и незначительного объема представляемых данных возможно не описывать домен для метаданных). Для динамической и статичной осей данных множества атрибутов, заданных элементами a_i^o и a_i^d (оперативные и архивные соответственно), должны быть эквивалентными, поскольку последующая обработка в соответствии с предлагаемым решением предполагает перенос ветви динамической оси данных в статичную ось архивных данных без каких-либо дополнительных изменений в заданных парах значений и атрибутов. Представляется целесообразным ввести единый домен для одноименных (однотипных) атрибутов обеих осей данных таким образом, чтобы обеспечить перенос данных с одной оси на другую без нормализации соответствующих значений. В общем виде домен данных для атрибута a любой оси данных можно представить следующим образом:

$$\exists D = \{d_1, d_2, \dots, d_p\} : \forall \{a_i^o, v_i^o\}, \{a_j^d, v_j^d\} v_i^o, v_j^d \in D, i=1, \dots, k; j=1, \dots, l. \quad (4)$$

В каждой оси данных (как динамической, так и статичной) формирование отдельных дочерних осей, каждая из которых соответствует определенной временной метке, осуществляется посредством специализированных транзакций. При этом при оперировании осями имеют место транзакции двух видов – чтение и запись соответственно. Для формализованного описания осевых транзакций представляется целесообразным воспользоваться семантикой Эрбранта [10]:

$$S(\Delta T) = r_1(D_d)w_1(S_d)r_2(D_s)w_2(D_d) : D_d \neq \text{null}, D_s \neq \text{null}, \quad (5)$$

где r_i – i -я процедура считывания данных, w_i – i -я процедура записи данных в соответствующие оси.

Переход данных от динамической оси данных к статичной является двухфазным. На первой фазе выполняются две итерации считывания и записи данных: данные считываются в виде дочерней оси динамической оси оперативных данных (операция r_1) и направляются для записи в статичную ось архивных данных (операция w_1). На второй фазе выполняется заполнение динамической оси новой дочерней осью взамен той, что была передана в статичную ось архивных данных. Для этого используется итерация чтения (операция r_2) из внешнего источника данных и записи (операция w_2) непосредственной в качестве новой дочерней оси в искомую ось в динамическую ось оперативных данных. В качестве внешнего источника (D_s) может выступать ресурс соответствующей системы регистрации параметров поля (например, в случае геомагнитного поля в качестве таковых выступают магнитные обсерватории и вариационные станции) либо результат выполнения программного решения, обеспечивающего моделирование и расчет искомых параметров поля. Транзакция считается завершенной тогда, когда завершены обе операции записи в каждую из представленных осей данных. В случае, если на каком-то

из этапов по тем или иным причинам был зафиксирован сбой, выполняется откат соответствующей транзакции.

В каждый момент времени экземпляр мультиосевой модели данных для фиксированного источника (геопространственной привязки) находится в определенном состоянии. В контексте предложенной модели такое состояние представляет собой некоторую модель сигнатуры

$$\Sigma_3 = \{S_m^n, D_d^k, S_d^l\} \quad (6)$$

с основным множеством, представленным соответственно значениями атрибутов в n дочерних узлах-осях статичной оси метаданных, k дочерних узлах-осях динамической оси оперативных данных, l дочерних узлах-осях статичной оси архивных данных.

В общем случае каждая новая итерация фиксирует преобразование экземпляра модели в новое состояние, фактически получаемое переносом узла из одной оси в другую и появлением нового узла в динамической оси по истечении указанного временного интервала ΔT :

$$\Sigma_3 = \{S_m^n, D_d^k, S_d^l\} \xrightarrow{\Delta T} \Sigma'_3 = \{S_m^n, D_d^k, S_d^{l+1}\}. \quad (7)$$

В целом можно рассматривать мультиосевую модель данных как дерево t с конечным числом упорядоченных элементов неопределенной глубины и арности. В общем виде узлы соответствующего дерева t возможно пометить символами некоторого конечного алфавита Σ . В этом контексте домен дерева t , рассматриваемого в качестве некоторой логической структуры, является множеством узлов t , условно обозначаемым $Dom(t)$.

Рассмотрим дерево t как неранжированное дерево таким образом, что имеет место некоторая структура $\sigma(h)$, где $\sigma \in \Sigma$, h – последовательность неранжированных деревьев, определяемая следующим образом:

$$H_\Sigma \ni h ::= \begin{cases} \sigma(h), h', \text{ непустая последовательность деревьев;} \\ \{ \}, \text{ пустая последовательность деревьев.} \end{cases} \quad (8)$$

Множество неранжированных деревьев обозначим как T_Σ^2 . Тогда каждая ось дерева t может в общем виде рассматриваться как бинарное дерево, которое является либо σ -размеченным корнем двух поддеревьев ($\sigma \in \Sigma$), либо пустым деревом:

$$T_\Sigma^2 ::= \begin{cases} \sigma(t, t'), \text{ узел;} \\ \varepsilon, \text{ пустое дерево.} \end{cases} \quad (9)$$

Неранжированные деревья преобразуются в бинарные деревья с помощью функции $\beta(\cdot)$ вида:

$$\begin{aligned} \beta(\cdot) &= H_\Sigma \rightarrow T_\Sigma^2; \\ \beta(\sigma(h), h') &\xrightarrow{\text{def}} \sigma(\beta(h), \beta(h')); \\ \beta(\{ \}) &\xrightarrow{\text{def}} \varepsilon. \end{aligned} \quad (10)$$

Тогда формально $Dom(t)$ может быть представлен как некоторое подмножество $\mathbb{N}^*$, определяемое следующим образом:

$$\begin{aligned} Dom(t) &= \{\varepsilon\} \cup \{iu \mid i \in \{1, \dots, n\}, u \in Dom(t_i)\}; \\ t &= \{t_1, \dots, t_n\}, \sigma \in \Sigma, n \geq 0; t_1, \dots, t_n \in T_\Sigma^2. \end{aligned} \quad (11)$$

Представляется целесообразным рассмотреть отношения между осями и их дочерними осями в мультиосевой модели данных полей в терминах реляционного словаря вида: $(\prec_{ch}, \prec_{sb}, \{O_\sigma \mid \sigma \in \Sigma\})$. Здесь O_σ являются предикатами унарного отношения. Тогда для каждого размеченного узла $\sigma \in \Sigma$ O_σ представляет собой набор узлов, помеченных соответственно как σ . Символы $\prec_{ch}, \prec_{sb}$ являются бинарными предикатами. Первый из них обозначает дочернее отношение (child relation) вида $(v, v \cdot i), v, v \cdot i \in Dom(t)$. Тогда $\prec_{sb}$ является отношением соседства узлов (sibling) вида:

$$(v \cdot i, v \cdot (i+1)), v \cdot i, v \cdot (i+1) \in Dom(t).$$

С учетом сказанного имеют место соотношения вида:

$$\begin{aligned}
 O_{\sigma} &= \{S_m \prec_{sb} D_d \prec_{sb} S_d\}; \\
 S_m &\prec_{ch} \bigcup_{i=1}^n \{a_i^s, v_i^s\}; \{a_i^s, v_i^s\} \prec_{sb} \{a_{i+1}^s, v_{i+1}^s\} \forall i = 1, \dots, n-1; \\
 D_d &\prec_{ch} \bigcup_{i=1}^k \{a_i^o, v_i^o\}; \{a_i^o, v_i^o\} \prec_{sb} \{a_{i+1}^o, v_{i+1}^o\} \forall i = 1, \dots, k-1; \\
 S_d &\prec_{ch} \bigcup_{i=1}^l \{a_i^d, v_i^d\}; \{a_i^d, v_i^d\} \prec_{sb} \{a_{i+1}^d, v_{i+1}^d\} \forall i = 1, \dots, l-1.
 \end{aligned} \tag{12}$$

3. Апробация и оценка эффективности предложенного решения

Как правило, данные геофизических полей содержатся в базах данных с реляционной структурой [11–13], однако наличие метаданных и тензорных данных требует пересмотреть традиционные методы хранения. Поскольку представленная в данной работе многоосевая модель данных имеет древовидную структуру, было принято решение использовать XML (eXtensible Markup Language, расширяемый язык разметки) [14–16]. Необходимо построить рассматриваемую XML-структуру в элементоцентричной манере таким образом, чтобы все данные и отдельные суррогатные узлы были заданы в виде XML-элементов. Сказанное касается в том числе и суррогатного корневого узла всей интегральной структуры данных, поскольку в целом данный компонент вводится искусственно для поддержания иерархической структуры значимых данных и не предполагает никаких иных характеристик (рис. 4).

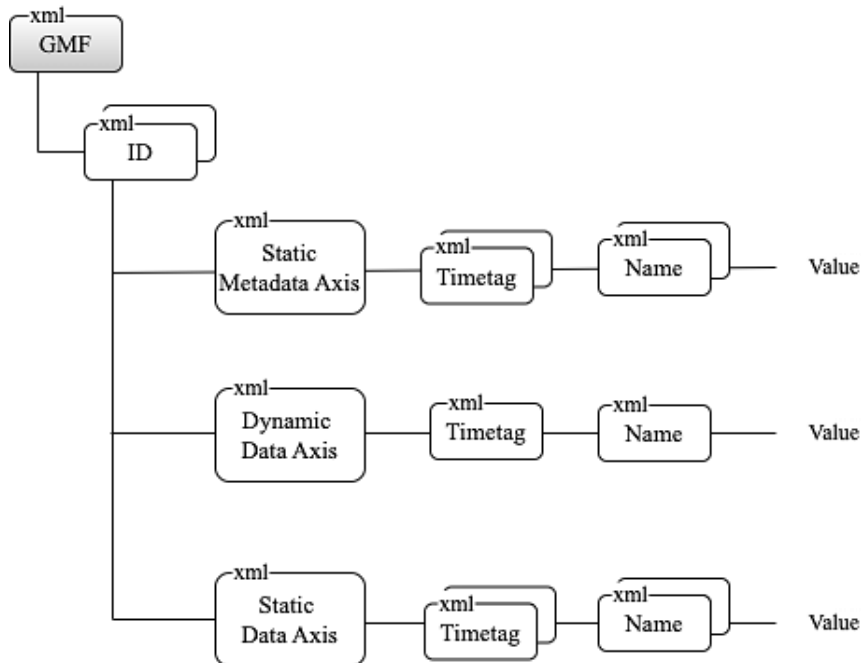


Рис. 4. XML-структура мультиосевой модели данных геофизических полей
Fig. 4. XML structure of the multi-axis model of geophysical field data

В целях создания структуры XML-документа следует описывать ее в виде XML-схемы [17], что позволит производить валидацию итоговой схемы после внесения изменений и добавления новых источников данных, при этом сохранив элементоцентричную структуру. С учетом древовидной структуры целесообразно производить запросы, основываясь на спецификациях XPath [18, 19].

Эффективность использования мультиосевой модели для хранения геофизических данных оценивалась путем сравнения с другими известными подходами (а именно с реляционным представлением и текстовыми CSV-файлами) с использованием набора геомагнитных данных из проекта SuperMAG, а также метрик скорости получения и внесения данных. Мультиосевая модель при этом была построена в формате XML с использованием СУБД Sedna [20, 21]. Результирующие данные представлены в формате XML/KML [22, 23].

Было установлено, что использование мультиосевой структуры данных для хранения геофизической информации увеличивает скорость выполнения однопредикативных запросов на ~ 25,7%, мультипредикативных запросов на ~ 20,1%, а скорость вставки новых записей на ~ 21,3% в среднем, что подтверждает целесообразность реализации предложенного решения.

Заключение

Информация о геофизических полях на текущий момент играет большую роль в понимании процессов, протекающих на поверхности и за пределами Земли, а также широко применяется в поддержке принятия решений, в том числе для снижения негативного влияния геофизических явлений на технологическую безопасность. Поэтому совершенствование известных методов хранения и организации геофизических данных в целях повышения производительности запросов к БД и снижения времени отклика представляется актуальной задачей.

В данной работе была представлена мультиосевая модель хранения и организации геофизических данных, которая основывается на элементоцентричной разметке и иерархической структуре хранения информации. Модель отличается использованием статичной и динамической осей, которые вводятся в соответствии с тензорной спецификой геофизического поля. Центральный узел осей соответствует источнику данных, в то время как сами оси используются для представления метаданных, а также архивных и оперативных данных, взаимодействие между которыми реализовано с использованием триггеров и темпоральных предикатов на уровне фоновых процессов.

Список источников

1. Geomagnetic Observations and Models / M. Manda, M. Korte (eds). Dordrecht : Springer, 2011. (IAGA Special Sopron Book Series; 5).
2. INTERMAGNET technical reference manual. Version 4.6 / ed. by S.-L. Benoît. Edinburgh : INTERMAGNET. BGS, 2012.
3. Califf S., Alken P., Chulliat A., Anderson B., Rock K., Vines S., Barnes R., Liou K. Investigation of geomagnetic reference models based on the Iridium constellation // *Earth, Planets and Space*. 2022. V. 1 (74). P. 1–21. doi: 10.1186/s40623-022-01574-w
4. Lazzeri C., Samsonov A., Forsyth C., Branduardi-Raymont G., Bogdanova Yu. A Statistical Study of the Properties of, and Geomagnetic Responses to, Large, Rapid Southward Turnings of the Interplanetary Magnetic Field // *J. of Geophysical Research*:
5. Peng Z., Laramée S. Higher Dimensional Vector Field Visualization. A Survey // *Theory and Practice of Computer Graphics (TPCG '09)*. 2009. P. 149–163.
6. Hergl C., Blecha C., Kretschmar V., Raith F., Gunther F., Stommel M., Jankowai J., Hotz I., Nagel T., Scheuermann G. Visualization of Tensor Fields in Mechanics // *Computer Graphics Forum*. 2021. V. 40 (6). P. 135–161. doi: 10.1111/cgf.14209
7. Jilesh V., Pournami A. On a generalization of Laplace distribution with applications. *International // J. of Data Science and Analytics*. 2025. V. 19 (1). P. 1–10. doi: 10.1007/s41060-024-00706-7
8. He Z., Hu X., Teng Yu., Zhang X., Shen X. Data agreement analysis and correction of comparative geomagnetic vector observations // *Earth, Planets and Space*. 2022. V. 74 (1). Art. 29. doi: 10.1186/s40623-022-01583-9
9. Huang Y., Wu L., Li D. Theoretical Research on Full Attitude Determination Using Geomagnetic Gradient Tensor // *The J. of Navigation*. 2015. V. 68 (5). P. 951–961. doi: 10.1017/S0373463315000259
10. Paulo O., Chuangjie X. On the Herbrand functional interpretation // *Mathematical Logic Quarterly*. 2020. V. 66 (1). P. 91–98. doi: 10.1002/malq.201900067
11. Воробьев А.В., Воробьева Г.Р. Индуктивный метод восстановления временных рядов геомагнитных данных // *Труды СПИИРАН*. 2018. № 2 (57). С. 104–133. doi: 10.15622/sp.57.5
12. Воробьев А.В., Воробьева Г.Р. Оценка влияния геомагнитной активности на метрологические характеристики инклинометрических информационно-измерительных систем // *Измерительная техника*. 2017. № 6. С. 21–24.
13. Vorobev A.V., Vorobeva G.R., Yusupova N.I. Conception of geomagnetic data integrated space // *SPIIRAS Proceedings*. 2019. V. 18 (2). P. 390–415. doi: 10.15622/sp.18.2.390-415
14. Ali M., Khan M. Enhancing XML Data Parsing and Querying Performance on Multi-Core Architectures // *Statistics Computing and Interdisciplinary Research*. 2024. V. 6 (1). P. 75–89. doi: 10.52700/scir.v6i1.158
15. Воробьев А.В., Воробьева Г.Р. Подход к обнаружению и устранению артефактов пространственных изолиний в приложениях Веб-ГИС // *Компьютерная оптика*. 2023. Т. 47, № 1. С. 126–136. doi: 10.18287/2412-6179-CO-1127
16. Воробьев А.В., Воробьева Г.Р. Подход к динамической визуализации разнородных геопространственных векторных изображений // *Компьютерная оптика*. 2024. Т. 48, № 1. С. 123–138. doi: 10.18287/2412-6179-CO-1279
17. Jamal S., Rahman C., Abdulkarim M. XML Schema Validation Using Java API for XML Processing // *UKH Journal of Science and Engineering*. 2022. № 6. P. 33–41. doi: 10.25079/ukhjse.v6n1y2022

18. Chen R., Wang Z., Su H., Xie S., Wang Z. Parallel XPath query based on cost optimization // *The Journal of Supercomputing*. 2022. V. 78 (2). P. 1–30. doi: 10.1007/s11227-021-04074-y
19. Areces C., Fervari R. Axiomatizing Hybrid XPath with Data // *Logical Methods in Computer Science*. 2021. V. 17 (3). P. 1–37. doi: 10.46298/lmcs-17(3:5)2021
20. Chernenkiy V.M., Gapanyuk Y.E., Kaganov Y.T., Dunin I.V., Lyaskovsky M.A., Larionov V. Storing Metagraph Model in Relational, Document-Oriented, and Graph Databases // *Selected Papers of the XX International Conference on Data Analytics and Management in Data Intensive Domains*. 2018. V. 2277. P. 82–89.
21. Terekhov V., Gapanyuk Y., Kanev A. Metagraph representation for overcoming limitations of existing knowledge bases // 28th Conference of Open Innovations Association (FRUCT). IEEE, 2021. P. 458–464.
22. Zichar M. Geovisualization based upon KML // *Journal of Agricultural Informatics*. 2021. V. 3 (1). P. 19–26. doi: 10.17700/jai.2012.3.1.49
23. Ortinez A., Ruiz S., Luis G., Ortega E., Garcia-Reynoso A., Peralta O., López-Gaona A., Castro T., Martinez Arroyo A. Emission inventory point source visualization on Google Earth and integrated with HYSPLIT model // *Atmósfera*. 2020. V. 34 (2). P. 1–36. doi: 10.20937/ATM.52834

References

1. Mandea, M. & Korte, M. (eds) (2011) *Geomagnetic Observations and Models*. Dordrecht: Springer. pp. 149–181.
2. Benoît, S.-L. (ed.) (2012) *Intermagnet Technical Reference Manual. Version 4.6 (2012)*. Edinburgh: INTERMAGNET.
3. Califf, S., Alken, P., Chulliat, A., Anderson, B., Rock, K., Vines, S., Barnes, R. & Liou, K. (2022) Investigation of geomagnetic reference models based on the Iridium constellation. *Earth, Planets and Space*. 1(74). pp. 1–21. DOI: 10.1186/s40623-022-01574-w
4. Lazzeri, C., Samsonov, A., Forsyth, C., Branduardi-Raymont, G. & Bogdanova, Yu. (2024) A Statistical Study of the Properties of, and Geomagnetic Responses to, Large, Rapid Southward Turnings of the Interplanetary Magnetic Field. *Journal of Geophysical Research: Space Physics*. 129(9). pp. 1–15. DOI: 10.1029/2023JA032160
5. Peng, Z. & Laramée, S. (2009) Higher Dimensional Vector Field Visualization. A Survey. *Theory and Practice of Computer Graphics (TPCG '09)*. pp. 149–163.
6. Hergl, C., Blecha, C., Kretzschmar, V., Raith, F., Gunther, F., Stommel, M., Jankowai, J., Hotz, I., Nagel, T. & Scheuermann, G. (2021) Visualization of Tensor Fields in Mechanics. *Computer Graphics Forum*. 40(6). pp. 135–161. DOI: 10.1111/cgf.14209
7. Jilesh, V. & Pournami, A. (2025). On a generalization of Laplace distribution with applications. *International Journal of Data Science and Analytics*. 19(1). pp. 1–10. DOI: 10.1007/s41060-024-00706-7
8. He, Z., Hu, X., Teng, Yu., Zhang, X. & Shen, X. (2022) Data agreement analysis and correction of comparative geomagnetic vector observations. *Earth, Planets and Space*. 74. DOI: 10.1186/s40623-022-01583-9
9. Huang, Y., Wu, L. & Li, D. (2015) Theoretical Research on Full Attitude Determination Using Geomagnetic Gradient Tensor. *The Journal of Navigation*. 68(5). pp. 951–961. DOI: 10.1017/S0373463315000259
10. Paulo, O. & Chuangjie, X. (2020). On the Herbrand functional interpretation. *Mathematical Logic Quarterly*. 66(1). pp. 91–98. DOI: 10.1002/malq.201900067
11. Vorobiev, A.V. & Vorobieva, G.R. (2018) Induktivnyy metod vosstanovleniya vremennykh ryadov geomagnitnykh dannykh [Inductive method for reconstructing time series of geomagnetic data]. *Trudy SPIIRAN*. 2(57). pp. 104–133. DOI: 10.15622/sp.57.5
12. Vorobiev, A.V. & Vorobieva, G.R. (2017) Otsenka vliyaniya geomagnitnoy aktivnosti na metrologicheskie kharakteristiki inklinometricheskikh informatsionno-izmeritel'nykh sistem [Assessment of the influence of geomagnetic activity on the metrological characteristics of inclinometric information-measuring systems]. *Izmeritel'naya tekhnika*. 6. pp. 21–24.
13. Vorobiev, A.V., Vorobieva, G.R. & Yusupova, N.I. (2019) Conception of geomagnetic data integrated space. *Trudy SPIIRAN*. 18(2). pp. 390–415. DOI: 10.15622/sp.18.2.390-415
14. Ali, M. & Khan, M. (2024). Enhancing XML Data Parsing and Querying Performance on Multi-Core Architectures. *Statistics Computing and Interdisciplinary Research*. 6(1). pp. 75–89. DOI: 10.52700/scir.v6i1.158
15. Vorobiev, A.V. & Vorobieva, G.R. (2023) Podkhod k obnaruzheniyu i ustraneniyu artefaktov prostranstvennykh izolinii v pri-lozheniyakh Veb-GIS [Approach to detection and elimination of spatial isoline artifacts in Web GIS applications]. *Komp'yuternaya optika*. 47(1). pp. 126–136. DOI: 10.18287/2412-6179-CO-1127
16. Vorobiev, A.V. & Vorobieva, G.R. (2024) Podkhod k dinamicheskoy vizualizatsii raznorodnykh geoprostranstvennykh vektornykh izobrazheniy [Approach to dynamic visualization of heterogeneous geospatial vector images]. *Komp'yuternaya optika*. 48(1). pp. 123–138. DOI: 10.18287/2412-6179-CO-1279
17. Jamal, S., Rahman, C. & Abdulkarim, M. (2022) XML Schema Validation Using Java API for XML Processing. *UKH Journal of Science and Engineering*. 6. pp. 33–41. DOI: 10.25079/ukhjse.v6n1y2022
18. Chen, R., Wang, Z., Su, H., Xie, S. & Wang, Z. (2022) Parallel XPath query based on cost optimization. *The Journal of Supercomputing*. 78(2). pp. 1–30. DOI: 10.1007/s11227-021-04074-y
19. Areces, C. & Fervari, R. (2021). Axiomatizing Hybrid XPath with Data. *Logical Methods in Computer Science*. 17(3). pp. 1–37. DOI: 10.46298/lmcs-17(3:5)2021
20. Chernenkiy, V.M., Gapanyuk, Y.E., Kaganov, Y.T., Dunin, I.V., Lyaskovsky, M.A. & Larionov, V. (2018) Storing Metagraph Model in Relational, Document-Oriented, and Graph Databases. *Selected Papers of the XX International Conference on Data Analytics and Management in Data Intensive Domains*. 2277. pp. 82–89.

21. Terekhov, V., Gapanyuk, Y. & Kanev, A. (2021) Metagraph representation for overcoming limitations of existing knowledge bases. *28th Conference of Open Innovations Association (FRUCT)*. IEEE. pp. 458–464.
22. Zichar, M. (2012). Geovisualization based upon KML. *Journal of Agricultural Informatics*. 3(1). pp. 19–26. DOI: 10.17700/jai.2012.3.1.49
23. Ortinez, A., Ruiz, S., Luis, G., Ortega, E., Garcia-Reynoso, A., Peralta, O., López-Gaona, A., Castro, T. & Martinez Arroyo, A. (2020) Emission inventory point source visualization on Google Earth and integrated with HYSPLIT model. *Atmósfera*. 34(2). pp. 1–36. DOI: 10.20937/ATM.52834

Информация об авторах:

Воробьева Гульнара Равиелевна – доцент, доктор технических наук, профессор кафедры вычислительной математики и кибернетики Уфимского университета науки и технологий (Уфа, Россия). E-mail: Gulnara.vorobeva@gmail.com

Воробьев Андрей Владимирович – доцент, доктор технических наук, заведующий кафедрой информатики Уфимского университета науки и технологий (Уфа, Россия). E-mail: Geomagnet@list.ru

Орлов Глеб Олегович – аспирант кафедры вычислительной математики и кибернетики Уфимского университета науки и технологий (Уфа, Россия). E-mail: orlovgleb99@mail.ru

Вклад авторов: все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

Information about the authors:

Vorobeva Gulnara R. (Associate Professor, Doctor of Technical Sciences, Professor of the Department of Computational Mathematics and Cybernetics, Ufa University of Science and Technology, Ufa, Russian Federation). E-mail: Gulnara.vorobeva@gmail.com

Vorobev Andrey V. (Associate Professor, Doctor of Technical Sciences, Head of the Department of Computer Science, Ufa University of Science and Technology, Ufa, Russian Federation). E-mail: Geomagnet@list.ru

Orlov Gleb O. (Post-graduate Student of the Department of Computational Mathematics and Cybernetics, Ufa University of Science and Technology, Ufa, Russian Federation). E-mail: orlovgleb99@mail.ru

Contribution of the authors: the authors contributed equally to this article. The authors declare no conflicts of interests.

Поступила в редакцию 18.01.2025; принята к публикации 02.06.2025

Received 18.01.2025; accepted for publication 02.06.2025

Original article

UDC 681.516

doi: 10.17223/19988605/71/8

**Shortages of perishables control for a stochastic inventory system
in retail through dynamic pricing****Anna V.¹ Kitaeva, Yu Cao²**^{1,2} National Research Tomsk State University, Tomsk, Russian Federation¹ kit1157@yandex.ru² ch.cy@stud.tsu.ru

Abstract. We address the management of a single product over a finite time horizon, with customer arrivals modeled as a heterogeneous compound Poisson process. The arrival rate depends on the inventory level and is adjusted by a time-dependent, unknown weight function, incorporating scenario of shortages. Retail price serves as the control variable for regulating this rate. Within the framework of diffusion approximation for the inventory process and a linear dependence of arrival intensity on price, we formulate and solve the expected revenue maximization problem concerning the weight function. Furthermore, two types of near-optimal weight functions are proposed, and the results of numerical optimization and simulation are presented.

Keywords: dynamic pricing; price sensitive demand; compound Poisson demand; diffusion approximation; shortages.

For citation: Kitaeva, A.V., Yu, Cao (2025) Shortages of perishables control for a stochastic inventory system in retail through dynamic pricing. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 71. pp. 82–92. doi: 10.17223/19988605/71/8

Научная статья

doi: 10.17223/19988605/71/8

**Контроль дефицита скоропортящихся товаров в стохастической системе
управления запасами с использованием динамического ценообразования****Анна Владимировна Китаева¹, Юй Цао²**^{1,2} Национальный исследовательский Томский государственный университет, Томск, Россия¹ kit1157@yandex.ru² ch.cy@stud.tsu.ru

Аннотация. Рассматривается задача управления запасами единичного продукта на конечном временном горизонте с гетерогенным пуассоновским процессом спроса. Интенсивность спроса зависит от уровня запаса и весовой функции от времени, создающей возможность дефицита товара. Управляющим параметром выступает розничная цена. В рамках диффузионной аппроксимации уровня запасов и линейной зависимости интенсивности спроса от розничной цены решается задача максимизации средней прибыли относительно весовой функции и партии товара. Рассмотрено два вида весовых функций, приведены численные результаты оптимизации и имитационного моделирования.

Ключевые слова: динамическое ценообразование; чувствительный к цене спрос; составной пуассоновский спрос; диффузионная аппроксимация; дефицит.

Для цитирования: Китаева А.В., Юй Цао. Контроль дефицита скоропортящихся товаров в стохастической системе управления запасами с использованием динамического ценообразования // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2025. № 71. С. 82–92. doi: 10.17223/19988605/71/8

1. Introduction and problem statement

Research on inventory models for perishables dates back to the 1960s. A comprehensive review of fresh product supply chain management was provided in [1, 2], highlighting key developments in the field. Recent advancements in inventory theory have increasingly incorporated dynamic pricing mechanisms, with particular emphasis on demand-driven market environments [3].

In [4], we proposed a basic dynamic pricing control model, which allow us, triggering purchases, to sell all the perishable product at hand during its lifetime almost surely. In [5], we introduced a multiplicative coefficient into the basic model. In [6], we considered more complicated weight function introducing a power-law coefficient.

The presence of the adjustable coefficients in the weight functions allows us to use a linear approximation of the intensity-of-price dependence and optimize the zero-ending inventory retailing process. Moreover, the coefficients can help us also to solve the problem of fitting the demand rate in a real-life situation. This problem is the most challenging one in dynamic pricing, see, for example, [7].

In this paper, we do not specify the type of weight function, and we get the equation for the optimal weight function, which maximizes the expected revenue, and consider its approximate solution as one near-optimal weight function. This near-optimal weight function almost surely implies shortages depending on the range of the unknown constant it contains. Firstly, we consider the task of the expected revenue maximization with respect to the constant considering lost sales. Secondly, we modify the optimal weight function approximation so that the price at the beginning of the sales period becomes close to the base price and obtain the expected revenue in case of possible shortages.

Inventory shortages systems are generally classified into two categories: backlog systems and lost-sales systems [8]. In lost-sales systems, actual sales often underestimate true demand, resulting in inaccurate demand forecasts [9]. Additionally, lost sales can give rise to rationing and gaming behaviors, both of which contribute to the so-called bullwhip effect in supply chains. This effect causes instability, undermining operational efficiency across industries [10]. From both academic and practical perspectives, the assumption of lost sales is often considered more appropriate than backlogging [11]. In this study, we adapt a framework where unmet demand is treated as lost.

We begin by outlining the model's assumptions and notations. The supply chain under consideration consists of a single vendor and customers. Acting as a monopolist, the vendor aims to maximize revenue by procuring a fixed lot size Q_0 at a unit cost d and selling it over a predetermined period T . Inventory replenishment is not permitted within this fixed period.

The demand is assumed to be a compound Poisson process with intensity $\lambda(c(t))$, where $c = c(t)$ is a dynamic retail price per unit, the orders are independent identically distributed continuous random variables with the first and second moments a_1 and a_2 respectively.

The stochastic properties of the sales process and the expected revenue are determined using a diffusion approximation of the stock level process. This approach enables us to derive analytically manageable expressions. Thus, we assume that the stock level process satisfies the following equation:

$$dQ(t) = -a_1\lambda(c(t))dt + \sqrt{a_2\lambda(c(t))}dw(t),$$

where $w(\cdot)$ is the Wiener process.

2. Near-Optimal weight function for a large lot size

We consider the following model for controlling the intensity of customers' flow through dynamic pricing:

$$a_1\lambda(c(t)) = \frac{Q(t)}{T\varphi(t/T)}, \quad (1)$$

where $\varphi(\cdot)$ is an unknown weight function, $t \in [0, T]$. As stated in [12], "formulating problems in the framework of intensity control is considered a promising approach." We will call control model (1) as a general one.

The idea of such a model of intensity control as well as a lot of results in inventory control modelling belongs to Alexander Fedorovich Terpugov (1939-2009), former Head of the Department of Probability Theory and Mathematical Statistics at Tomsk State University, outstanding scientist and teacher.

Thus, the stock level process satisfies the following equation:

$$dQ(t) = -\frac{Q(t)}{T\varphi(t/T)}dt + \sqrt{\frac{a_2}{a_1} \frac{Q(t)}{T\varphi(t/T)}}dw(t).$$

2.1. The expected revenue and near-optimal weight function

It is easy to show that the expectation

$$E\{Q(t)\} = \bar{Q}(t) = Q_0 \exp\{-\psi(t/T)\}, \quad (2)$$

where $\psi(z) = \int_0^z \varphi^{-1}(x)dx$.

Denote $\bar{Q}^2(t) = E\{Q^2(t)\}$. Applying Ito's formula and averaging, we get

$$\frac{d\bar{Q}^2(t)}{dt} = -2\frac{\bar{Q}^2(t)}{T\varphi(t/T)} + \frac{a_2 Q_0 \exp\{-\psi(t/T)\}}{a_1 T\varphi(t/T)}$$

subject to $\bar{Q}^2(0) = Q_0^2$.

It follows that the variance

$$\text{Var}\{Q(t)\} = \frac{a_2 Q_0}{a_1} \exp\{-\psi(t/T)\} (1 - \exp\{-\psi(t/T)\}). \quad (3)$$

Let us consider linear approximation of the intensity-of-price dependence

$$\lambda(c) = \lambda_0 - \lambda_1 \frac{c(t) - c_0}{c_0}, \quad (4)$$

where c_0 is a stationary price corresponding stationary intensity λ_0 and parameter $\lambda_1 > 0$ characterizes the sensitivity of $\lambda(\cdot)$ to relative price's deviations from stationary price c_0 .

Linear dependence of the customers' flow intensity on the price is common in literature, for example, in [13] the demand rate is supposed to be a linear function of the price.

From (1) and (4) we get

$$c(t) = c_0 \left(1 + \frac{\lambda_0}{\lambda_1} - \frac{Q(t)}{a_1 \lambda_1 T \varphi(t/T)} \right).$$

The average revenue at time unit

$$E\{c(t)a_1\lambda(c)\} = c_0 E\left\{\left(1 + \frac{\lambda_0}{\lambda_1} - \frac{1}{a_1 \lambda_1} \frac{Q(t)}{T\varphi(t/T)}\right) \frac{Q(t)}{T\varphi(t/T)}\right\} = c_0 \left(1 + \frac{\lambda_0}{\lambda_1}\right) \frac{\bar{Q}(t)}{T\varphi(t/T)} - \frac{c_0}{a_1 \lambda_1} \frac{\bar{Q}^2(t)}{T^2 \varphi(t/T)^2}.$$

The expected revenue over the cycle

$$\begin{aligned} \bar{S} &= \int_0^T E\{a_1 c(t) \lambda(t)\} dt = \\ &= \frac{c_0 Q_0}{\lambda_1} \left[(\lambda_0 + \lambda_1) \int_0^1 e^{-\psi(z)} \psi'(z) dz - \frac{1}{a_1 T} \left(Q_0 - \frac{a_2}{a_1} \right) \int_0^1 e^{-2\psi(z)} \psi'^2(z) dz - \frac{a_2}{a_1^2 T} \int_0^1 e^{-\psi(z)} \psi'^2(z) dz \right]. \end{aligned}$$

Taking into account $\int_0^1 e^{-\psi(z)} \psi'(z) dz = 1 - e^{-\psi(1)}$, we get

$$\bar{S} = \frac{c_0 Q_0}{\lambda_1} \left[(\lambda_0 + \lambda_1) (1 - e^{-\psi(1)}) - \frac{1}{a_1 T} \left(Q_0 - \frac{a_2}{a_1} \right) \int_0^1 e^{-2\psi(z)} \psi'^2(z) dz - \frac{a_2}{a_1^2 T} \int_0^1 e^{-\psi(z)} \psi'^2(z) dz \right].$$

Let us solve the task $-\int_0^1 \left[\left(\frac{a_1 Q_0}{a_2} - 1 \right) e^{-2\psi(z)} + e^{-\psi(z)} \right] \psi'^2(z) dz \Rightarrow \max_{\psi(\cdot)}$ subject to $\psi(0) = 0$.

Following Euler's equation, optimal function $\psi(\cdot)$ satisfies

$$Ae^{-\psi}(\psi'' - \psi'^2) + \left(\psi'' - \frac{1}{2}\psi'^2 \right) = 0, \quad (5)$$

where coefficient $A = a_1 Q_0 / a_2 - 1$.

Coefficient $A \gg 1$, because Q_0 is usually large, so neglecting the last term in (5) we get $\psi'' - \psi'^2 = 0$. It follows that approximate solution of (5) has form $\varphi(z) = C - z$, where C is a constant.

If $C = 1$, then $E\{Q(T)\} = \text{Var}\{Q(T)\} = 0$, so the lot will be sold during time T almost surely. This case was considered in [4]. This is, so called, the basic zero-ending inventory dynamic price control model.

If $C > 1$, then the leftovers are possible almost surely. The expected stock level at the end of the period $\bar{Q}(T) = Q_0((C - 1)/C)$.

If $C < 1$, then almost surely the vender will be out of the resources before the end of the period, that is, shortages are possible. Here, we consider this case.

2.2. The selling period's duration

Let us consider the Laplace transform of the probability density function of the stock level $\Phi(p, t) = E\{\exp(-pQ(t))\}$. By applying Ito's formula, we obtain

$$d \exp(-pQ(t)) = p \frac{Q(t)}{CT - t} \exp(-pQ(t)) \left(1 + \frac{a_2}{2a_1} p \right) dt - p \exp(-pQ(t)) \sqrt{\frac{a_2}{a_1} \frac{Q(t)}{CT - t}} dw(t). \quad (6)$$

After averaging (6), we get

$$(CT - t) \frac{\partial \Phi}{\partial t} + p \left(1 + \frac{a_2}{2a_1} p \right) \frac{\partial \Phi}{\partial p} = 0. \quad (7)$$

Solution of (7) has the form $\Phi(p, t) = \varphi\left(\frac{p(CT - t)}{p + \beta}\right)$, where $\varphi(\cdot)$ is an unknown function and parameter $\beta = 2a_1 / a_2$. The density function $f(q, 0) = \delta(q - Q_0)$, it follows $\Phi(p, 0) = \varphi\left(\frac{pCT}{p + \beta}\right) = \exp(-pQ_0)$ or

$\varphi(z) = \exp\left(-\frac{\beta z Q_0}{CT - z}\right)$. Finally, we get $\Phi(p, t) = \varphi\left(\frac{p(CT - t)}{p + \beta}\right) = \exp\left(-\frac{\beta p(CT - t)}{pt + \beta CT} Q_0\right)$.

By employing the inverse Laplace transform, we derive

$$f(q, t) = \exp\left(-\beta \frac{(CT - t)Q_0}{t}\right) \times \left[\delta(q) + \beta \exp\left(-\beta \frac{CTQ_0}{t}\right) \sqrt{\frac{CT(CT - t)Q_0}{t^2 q}} I_1\left(2\sqrt{\frac{CT(CT - t)}{t^2}} \beta^2 Q_0 q\right) \right]. \quad (8)$$

From (8) it follows that the cumulative distribution function of the length of time τ it takes to sell lot Q_0

$$F_\tau(t) = P(\tau \leq t) = \exp\left(-\beta \frac{CT - t}{t} Q_0\right).$$

Thus, $P(\tau \leq CT) = 1$, that is, the lot will be sold by the time CT almost surely.

The average of the selling period's duration $E\{\tau\} = \int_0^{CT} (1 - F_\tau(t)) dt = CT \left(1 - \exp(\beta Q_0) \int_0^1 \exp\left(\frac{-\beta Q_0}{x}\right) dx \right)$.

Let us consider the case $\beta Q_0 \gg 1$. In this case, we obtain

$$E\{\tau\} = CT \left(1 - \frac{1}{\beta Q_0} + o\left(\frac{1}{\beta Q_0}\right) \right). \quad (9)$$

2.3. The expected revenue taking shortages into account

Let us consider time point T_1 , $T_1 = Tx_0$, where $x_0 = C(1 - 1/\beta Q_0)$, that is, taking the main part of (9) as the approximation of the average of the selling period's duration. In simulation, we take the moment when residual inventory fails to meet the requested purchase quantity of the terminal consumer.

In Figure 1 the results of illustrative simulation of the selling period's duration are presented. The thinning algorithm is used to generate a non-homogeneous Poisson process, and simulations conducted with 1000 iterations, and the average outcomes reported; $0.9 < C < 1$, $T = 10$, $Q_0 = 500$, $a_1 = 5$, and $\lambda_0 / \lambda_1 = 4$, $\lambda_1 T = 100$; purchases are uniformly distributed over $(0, 10)$ or exponentially distributed. The black curve is the theoretical result, while the red curve represents the results of simulation in case of uniformly distributed purchases and the blue curve represents the results in case of exponentially distributed purchases.

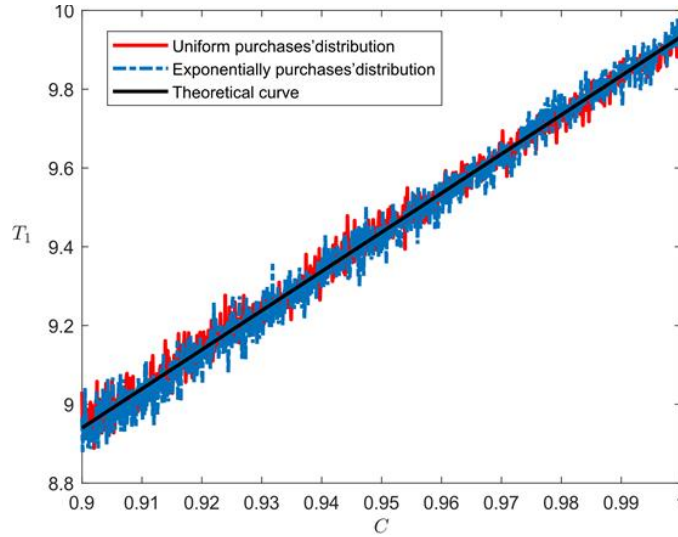


Fig. 1. T_1 simulation, $T = 10$, $Q_0 = 500$, purchases are Uniform $(0, 10)$ or Exp(5)

The parameter C significantly affects the duration of shortages. When $C < 0.9$ caution is advised, as the duration of shortages exceeds 10% of the entire period.

We consider the setting where unsatisfied demand is lost and assume that the demand intensity during shortages' period is static; $\lambda(c(T_1)) \approx \frac{a_2}{2a_1^2} \frac{1}{T(C - x_0)}$ and $c(T_1) \approx c_0 \left(1 + \frac{\lambda_0}{\lambda_1} - \frac{a_2}{2a_1^2} \frac{1}{\lambda_1 T(C - x_0)} \right)$.

We define the average shortage penalty as follow

$$\bar{S}_s = -a_1 c(T_1) \lambda(c(T_1)) (T - T_1) = -a_1 c_0 \left(1 + \frac{\lambda_0}{\lambda_1} - \frac{a_2}{2a_1^2} \frac{1}{\lambda_1 T(C - x_0)} \right) \frac{1 - x_0}{C - x_0}.$$

Thus, we get

$$\begin{aligned} \bar{S}_{C < 1} = \int_0^{T_1} E\{c(t)\lambda(t)\} dt + \bar{S}_s = c_0 Q_0 \left[\left(1 + \frac{\lambda_0}{\lambda_1} \right) \frac{x_0}{C} - \frac{1}{C \lambda_1 T} \left(\left(\frac{Q_0 a_1 - a_2}{a_1^2} \right) \frac{x_0}{C} - \frac{a_2}{a_1^2} \ln \left(\frac{C - x_0}{C} \right) \right) \right] - \\ - a_1 c_0 \left(1 + \frac{\lambda_0}{\lambda_1} - \frac{a_2}{2a_1^2} \frac{1}{\lambda_1 T(C - x_0)} \right) \frac{1 - x_0}{C - x_0}. \end{aligned} \quad (10)$$

The expected revenue is monotonically increasing with respect to C , and weighted expected revenue $\frac{\bar{S}_{C<1}}{a_1 c_0}$ depends on four dimensionless system's parameters Q_0 / a_1 , λ_0 / λ_1 , $\lambda_1 T$, a_2 / a_1^2 , except C . In Figure 2 the results of illustrative simulation of weighted revenue's dependence on C are presented for different sets of the system's parameters. We apply the thinning algorithm to generate a non-homogeneous Poisson process, execute 1000 simulation iterations and take the average values. The black curves represent the theoretical results, and the red curves are the simulation results.

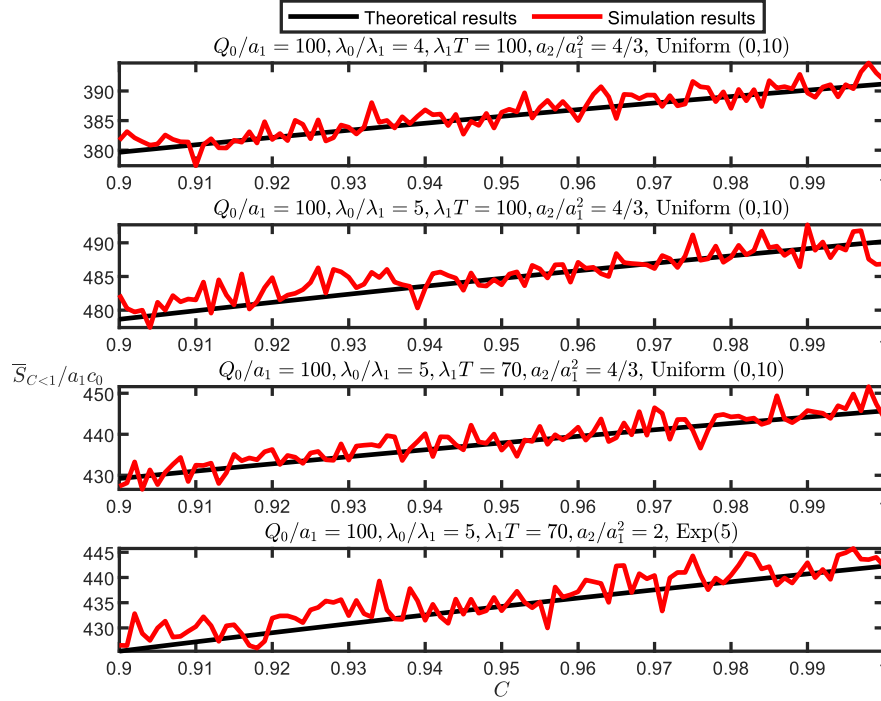


Fig. 2. $\bar{S}_{C<1} / a_1 c_0$ dependence on C for different sets of the system's parameters

Comparing the first three subplots in Figure 2 we see that the increasing ratio λ_0 / λ_1 and $\lambda_1 T$ leads to an increase in price, thereby increasing the revenue significantly. Ratio a_2 / a_1^2 determines the coefficient of variation of purchases and has no significant impact on revenue, compare the last two subplots in Figure 2.

There exists a monotonically increasing relationship between the weighted revenue $\bar{S}_{C<1} / a_1 c_0$ and parameter C , that is, as C tends to 1 from below, the duration of shortages decreases proportionally, thus enabling retailers to achieve higher revenue margins.

In Table numerical results of relative revenues calculation for the basic model $\bar{S}_{basic} / c_0 Q_0$ and the general one $\bar{S}_{C=1} / c_0 Q_0$ in case of $\lambda_0 T a_1 = Q_0$ are presented; $\bar{S}_{basic} / c_0 Q_0 \approx 1 - \frac{\lambda_0}{\lambda_1 \beta Q_0}$ and

$$\bar{S}_{C=1} / c_0 Q_0 \approx 1 + \frac{2\lambda_0(1 - 1/\beta Q_0 - \ln(\beta Q_0))}{\lambda_1 \beta Q_0} - \frac{1}{\lambda_1 T} \left(\frac{\lambda_1}{\lambda_0} + 1 - \frac{1}{(\beta Q_0)^2} \right).$$

Revenues for the basic and general models

λ_0 / λ_1	1	1	2	4	4
$\lambda_1 T$	400	400	200	100	100
βQ_0	2000/3	600	600	600	400
$\bar{S}_{basic} / c_0 Q_0$	0,9985	0,9983	0,9967	0,9933	0,9900
$\bar{S}_{C=1} / c_0 Q_0$	0,9790	0,9775	0,9600	0,9251	0,9034

The relative revenues for these two models are close to each other, the difference arises because of taking into account the lost sales.

The weighed revenue is a concave function with respect to a lot size. Figure 3 depicts weighted revenue $\bar{S}_{C<1} / a_1 c_0$ dependence on Q_0/a_1 for $C = 0,9, 0,95$ and $0,99$; $\lambda_0 / \lambda_1 = 4$, $\lambda_1 T = 100$, $a_2 / a_1^2 = 4/3$.

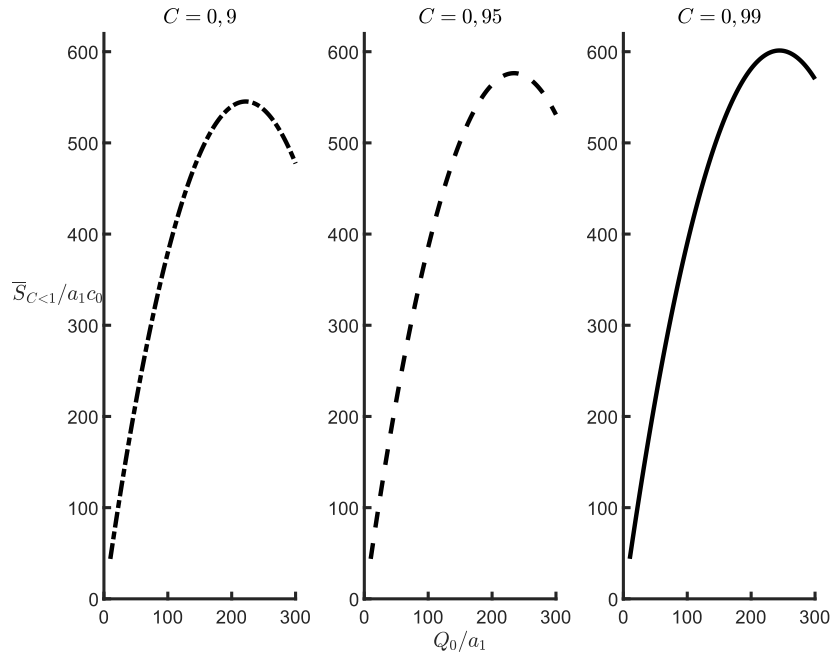


Fig. 3. $\bar{S}_{C<1} / a_1 c_0$ dependence on Q_0/a_1

As C tends to 1 from below, the optimal inventory level exhibits an upward trend, accompanied by a corresponding enhancement in the revenue.

3. Lot-weighted control model

Here, motivated by the form of the optimal weight function approximation, we are going to consider the following intensity of price dependence:

$$a_1 \lambda(c(t)) = \frac{CQ(t)}{CT-t}, \quad (11)$$

where $C > 0$ is a constant, $t < CT$. We will refer to this function as a lot-weighted one.

Stock level process is described by the following stochastic differential equation:

$$dQ(t) = -\frac{CQ(t)}{CT-t} dt + \sqrt{\frac{a_2}{a_1} \frac{CQ(t)}{CT-t}} dw(t). \quad (12)$$

Below we consider probabilistic characteristics of the stock level process and give the expression for the expected revenues for $C < 1$.

3.1. Probabilistic characteristics of the stock level process

Expectation and variance of $Q(t)$

$$E\{Q(t)\} = Q_0 g(t), \quad \text{Var}\{Q(t)\} = \frac{a_2}{a_1} Q_0 g(t)(1-g(t)),$$

where $g(t) = \left(1 - \frac{t}{CT}\right)^C$, $t \leq CT$.

Consistent with the preceding context, we obtain the density function

$$f(q, t) = \exp\left(\frac{-\beta Q_0 g(t)}{1-g(t)}\right) \times \left(\delta(q) + \beta \exp\left(\frac{-\beta Q_0}{1-g(t)}\right) \frac{\sqrt{g(t)Q_0/q}}{1-g(t)} I_1\left(\frac{2\beta\sqrt{g(t)Q_0q}}{1-g(t)}\right) \right).$$

It follows that the cumulative distribution function of τ

$$F_\tau(t) = P(\tau \leq t) = \exp\left(-\beta Q_0 \frac{g(t)}{1-g(t)}\right). \quad (13)$$

Note, that from (13) we get $P(\tau \leq CT) = 1$, that is, the shortages occur almost surely for $C < 1$. The average of the selling period's duration for $C < 1$

$$E\{\tau\} = \int_0^{CT} (1 - F_\tau(t)) dt = CT \left(1 - \int_0^1 \exp\left(-\beta Q_0 \frac{z^C}{1-z^C}\right) dz \right).$$

For $\beta Q_0 \gg 1$

$$E\{\tau\} \approx CT \left(1 - \int_0^1 \exp(-\beta Q_0 z^C) dz \right) = CT \left(1 - \frac{\Gamma(1+1/C)}{(\beta Q_0)^{1/C}} \right), \quad (14)$$

where $\Gamma(\cdot)$ is the gamma function.

3.2. The expected revenue taking shortages into account

Let us find the expected revenue for large lot size. Define the stationary (basic) price c_0 as a price satisfying equation $a_1 \lambda(c_0) = Q_0 / T$. Denote deviations from the stationary price $\Delta c(t) = c(t) - c_0$.

Using the Taylor expansion, we get $a_1 \lambda(c(t)) = CQ(t) / (CT - t) = a_1 \lambda(c_0) + a_1 \lambda'(c_0) \Delta c(t) + \dots$ and

$$\Delta c(t) \approx \frac{1}{a_1 \lambda'(c_0)} \left(\frac{CQ(t)}{CT - t} - \frac{Q_0}{T} \right).$$

Let us find conditional expectation $E\{Q(t) | Q(t) > 0\}$, $E\{Q(t)\} = (1 - \pi_0(t)) E\{Q | Q(t) > 0\} = Q_0 g(t)$,

$$\text{where } \pi_0(t) = P(\tau \leq t) = \exp\left(-\beta Q_0 \frac{g(t)}{1-g(t)}\right).$$

$$\text{Consequently, } E\{Q(t) | Q(t) > 0\} = \frac{Q_0 g(t)}{(1 - \pi_0(t))}, \quad E\left\{ \frac{CQ}{CT - t} - \frac{Q_0}{T} \middle| Q(t) > 0 \right\} = \frac{CQ_0 g(t)}{(CT - t)(1 - \pi_0(t))} - \frac{Q_0}{T},$$

$$E\{\Delta c(t) | Q(t) > 0\} = \frac{CQ_0 g(t)}{a_1 \lambda'(c_0)(CT - t)(1 - \pi_0(t))} - \frac{Q_0}{a_1 \lambda'(c_0)T}.$$

We take (14) as the approximation of the moment of shortages occurrence, $T_1 = Tx_0$, where

$$x_0 = C \left(1 - \frac{\Gamma(1+1/C)}{(\beta Q_0)^{1/C}} \right), \text{ and } Q(T_1) = \frac{Q_0 \Gamma(1+1/C)}{(\beta Q_0)^{1/C}}. \text{ We assume the setting where unsatisfied demand is lost}$$

and the demand intensity during shortages' period $T - T_1$ is static and defined by the intensity

$$\lambda(c(T_1)) = \frac{CQ(T_1)}{a_1(CT - T_1)} = \frac{Q_0}{a_1 T} \frac{C\Gamma(1+1/C)}{(C - x_0)(\beta Q_0)^{1/C}} = \frac{Q_0}{a_1 T} = \lambda_0; \text{ the retail price during the shortage period}$$

$$\begin{aligned} c(T_1) &= c_0 + \Delta c(T_1) = c_0 + \frac{1}{a_1 \lambda'(c_0)} \left(\frac{CQ(T_1)}{CT - T_1} - \frac{Q_0}{T} \right) = \\ &= c_0 + \frac{1}{a_1 \lambda'(c_0)} \frac{Q_0}{T} \left(\frac{C\Gamma(1+1/C)}{(C - x_0)(\beta Q_0)^{1/C}} - 1 \right) = c_0. \end{aligned}$$

The expected revenue considering the lost sales at $[CT, T]$ for $\beta Q_0 \gg 1$

$$\begin{aligned} \bar{S}_{C<1} &= a_1 c_0 \lambda(c_0) \int_0^{CT} (1 - \pi_0(t)) dt + a_1 (\lambda(c_0) + c_0 \lambda'(c_0)) \int_0^{CT} (1 - \pi_0(t)) \times \\ &\times \left(\frac{C Q_0 g(t)}{a_1 \lambda'(c_0) (CT - t) (1 - \pi_0(t))} - \frac{Q_0}{a_1 \lambda'(c_0) T} \right) dt - a_1 c_0 \lambda(c_0) T (1 - x_0) = \\ &= a_1 c_0 \lambda(c_0) T \left(1 + \frac{\lambda(c_0)}{c_0 \lambda'(c_0)} (1 - x_0) \right) - a_1 c_0 \lambda(c_0) T (1 - x_0). \end{aligned} \quad (15)$$

Using linear approximation (4) and substituting $Q_0 / T = a_1 \lambda_0$, we can rewrite (15) as follows

$$\bar{S}_{C<1} = a_1 c_0 \lambda_0 T \left(1 - \frac{\lambda_0}{\lambda_1} (1 - x_0) \right) - a_1 c_0 \lambda_0 T (1 - x_0), \text{ or relative revenue}$$

$$\bar{S}_{C<1} / c_0 Q_0 = C \left(1 - \frac{\Gamma(1 + 1/C)}{(\beta Q_0)^{1/C}} \right) \left(\frac{\lambda_0}{\lambda_1} + 1 \right) - \frac{\lambda_0}{\lambda_1}.$$

Relative revenue $\bar{S}_{C<1} / c_0 Q_0$ monotonically increases with respect to C . Tending C to one from below we get $\bar{S}_{C<1} / c_0 Q_0 \rightarrow 1 - \frac{\lambda_0}{\lambda_1 \beta Q_0} - \frac{1}{\beta Q_0}$. The result is the same as $\bar{S}_{basic} / c_0 Q_0$ taking into account that the last term appears due to the lost sales consideration.

Figure 4 depicts the relative revenues dependence on C for general (black lines) and lot-weighted (red lines) weight functions.

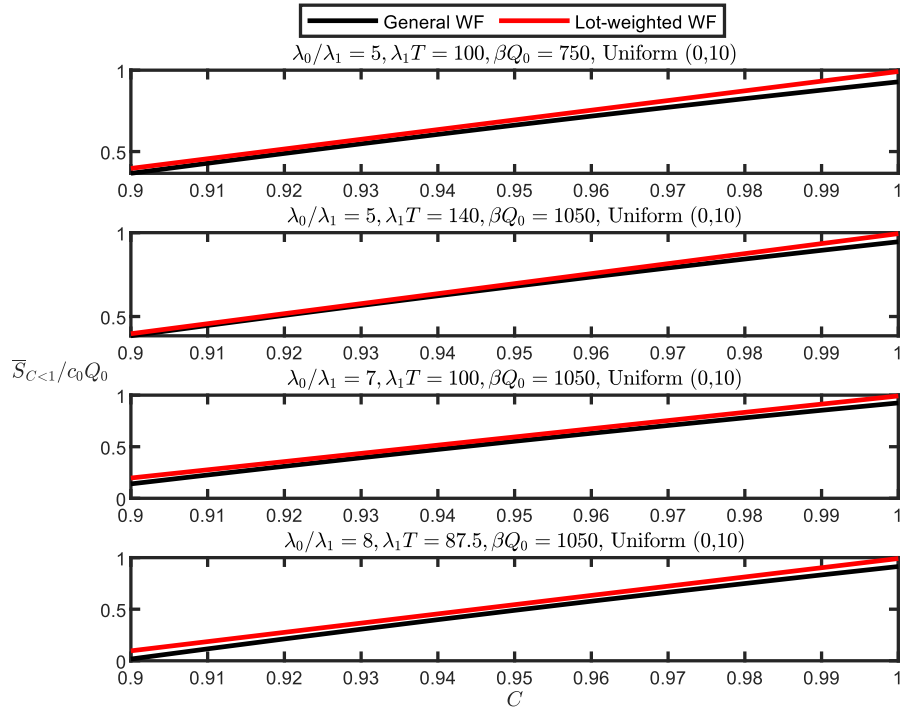


Fig. 4. Relative revenue $\bar{S}_{C<1} / c_0 Q_0$ dependence on C

Our analysis reveals a negligible revenue gap between the two models. The first two subplots in Figure 4 demonstrate that larger lot sizes mitigate this disparity. An increasing ratio λ_0 / λ_1 and decreasing $\lambda_1 T$ result in reduced revenue significantly, accompanied by a widening revenue gap between two models, compare the last three subplots in Figure 4. However, as C tends to 1 from below, the relative revenues of both models converge to 1, with the lot-weighted model consistently generating higher revenues.

Conclusion

This study investigates dynamic pricing for perishable products under a heterogeneous compound Poisson demand, employing a diffusion approximation framework of the stock level process. Such approximation works well for large lot sizes. Two near-optimal weight function models – general and lot-weighted – are proposed to optimize revenues addressing shortages by adjusting coefficient $C < 1$ in the weight functions.

The general model, derived analytically using the linear dependence of the intensity on price, provides closed-form expressions for the expected revenue and shortage duration, validated through simulation. It exhibits heightened sensitivity to C and depends on three dimensionless inventory system's parameters: βQ_0 , λ_0 / λ_1 , and $\lambda_1 T$. The modified, lot-weighted model, ensures more stable operation of the inventory system, depends only two parameters βQ_0 , λ_0 / λ_1 , and achieves a little higher revenue compared to the general model for C close to 1.

Numerical analyses demonstrate that both models enable the revenue optimization while balancing the shortages occurrence. Small values of the adjustable coefficient can lead to a significant reduction in revenue. These two control models can be used in case of the supply chain disruption during the sales period, which may lead to the need to shorten the period. Using these models allows us to sell all inventory by the new deadline and estimate the losses. The lot-weighted model has some advantages comparing with the general one: its relative revenue does not depend on $\lambda_1 T$, and at the beginning of the period we set basic price c_0 corresponding to basic intensity λ_0 .

References

1. Krishna Manasvi, J. & Matai, R. (2022) Agri-fresh supply chain management: a systematic literature review. *Recent Advances in Industrial Production: Select Proceedings of ICEM*. pp. 449–457.
2. Shaabani, H. (2022) A literature review of the perishable inventory routing problem. *The Asian Journal of Shipping and Logistics*. 38(3). pp. 143–161. DOI: 10.1016/j.ajsl.2022.05.002
3. Yang, N. & Zhang, R. (2014) Dynamic pricing and inventory management under inventory-dependent demand. *Operations Research*. 62(5). pp. 1077–1094.
4. Kitaeva, A.V., Stepanova, N.V. & Zhukovskaya, A.O. (2019) Zero ending inventory dynamic pricing model under stochastic demand, fixed lifetime product, and fixed order quantity. *IFAC-PapersOnLine*. 52(3). pp. 2482–2487.
5. Kitaeva, A.V., Stepanova, N.V. & Zhukovskaya, A.O. (2020) Profit optimization for zero ending inventories dynamic pricing model under stochastic demand and fixed lifetime product. *IFAC-PapersOnLine*. 53(2). pp. 10505–10510.
6. Kitaeva, A.V., Stepanova, N.V. & Zhukovskiy, O.I. (2022) Profit Optimization with a Power-Law Adjustable Coefficient for Zero Ending Inventories Dynamic Pricing Model, Stochastic Demand, and Fixed Lifetime Product. *IFAC-PapersOnLine*. 55(10). pp. 1793–1797. DOI:10.1016/j.ifacol.2020.12.2796
7. Dolgui, A. & Proth, J.M. (2010) *Supply chain engineering: useful methods and techniques*. London: Springer.
8. Gayon, J.P., Massonnet, G., Rapine, C. & Stauffer, G. (2016) Constant approximation algorithms for the one warehouse multiple retailers problem with backlog or lost-sales. *European Journal of Operational Research*. 250(1). pp. 155–163.
9. Gruen, T.W., Corsten, D.S. & Bharadwaj, S. (2002) *Retail out-of-stocks: A worldwide examination of extent, causes and consumer responses*. Washington, DC: Grocery Manufacturers of America.
10. Hu, Q. (2019) Bullwhip effect in a supply chain model with multiple delivery delays. *Operations Research Letters*. 47(1). pp. 36–40.
11. Lim, A.E., Wei, Z.X. & Zhang, H. (2024) *Partial Backorder Inventory System: Asymptotic Optimality and Demand Learning*. arXiv preprint arXiv:2404.00046.
12. Gallego, G. & Van Ryzin, G. (1994) Optimal dynamic pricing of inventories with stochastic demand over finite horizons. *Management Science*. 40(8). pp. 999–1020.
13. Li, S., Zhang, J. & Tang, W. (2015) Joint dynamic pricing and inventory control policy for a stochastic inventory system with perishable products. *International Journal of Production Research*. 53(13). pp. 2937–2950.

Information about the authors:

Kitaeva Anna V. (Doctor of Physical and Mathematical Sciences, Professor of Institute of Applied Mathematics and Computer Science, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: kit1157@yandex.ru

Yu Cao (Post-graduate Student, Institute of Applied Mathematics and Computer Science, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: ch.cy@stud.tsu.ru

Contribution of the authors: the authors contributed equally to this article. The authors declare no conflicts of interests.

Информация об авторах:

Китаева Анна Владимировна – доктор физико-математических наук, профессор Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: kit1157@yandex.ru

Юй Цао – аспирант Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: ch.cy@stud.tsu.ru

Вклад авторов: все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

Received 17.02.2025; accepted for publication 02.06.2025

Поступила в редакцию 17.02.2025; принята к публикации 02.06.2025

Научная статья
УДК 515.142.33
doi: 10.17223/19988605/71/9

Быстрая реализация алгоритма невыпуклого полосового слияния триангуляции Делоне

Марина Игоревна Литовченко¹, Юрий Леонидович Костюк²

^{1,2} *Национальный исследовательский Томский государственный университет, Томск, Россия*

¹ *litovchenkom21@yandex.ru*

² *kostyuk.yu48@mail.ru*

Аннотация. Рассматривается быстрая реализация полосового алгоритма триангуляции Делоне набора точек плоскости. Вначале область с точками делится на полосы некоторой ширины, и производится сортировка точек в каждой полосе. Далее в каждой полосе строится отдельная невыпуклая триангуляция методом заметающей прямой. Затем соседние полосы сшиваются между собой, а после получения триангуляции всей области выполняется перестроение треугольников по критерию Делоне. Приведены результаты работы алгоритма для равномерного и неравномерного (фрактального) распределения точек внутри области. Время работы предложенного алгоритма по сравнению с наиболее быстрым алгоритмом динамического кэширования оказалось меньше в 1,7 раза. Это объясняется существенно меньшим количеством проверок и перестроений треугольников по критерию Делоне в предложенном алгоритме.

Ключевые слова: триангуляция Делоне; полосовой алгоритм; алгоритм заметающей прямой; сравнение алгоритмов.

Для цитирования: Литовченко М.И., Костюк Ю.Л. Быстрая реализация алгоритма невыпуклого полосового слияния триангуляции Делоне // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2025. № 71. С. 93–102. doi: 10.17223/19988605/71/9

Original article
doi: 10.17223/19988605/71/9

Fast implementation of a non-convex stripe merging algorithm for Delaunay triangulation

Marina I. Litovchenko¹, Yuriy L. Kostyuk²

^{1,2} *National Research Tomsk State University, Tomsk, Russian Federation*

¹ *litovchenkom21@yandex.ru*

² *kostyuk.yu48@mail.ru*

Abstract. The paper presents a fast implementation of a Stripe Merging algorithm for Delaunay triangulation of a set of points in the plane. Initially, the area containing the points is divided into stripes of a specified width, and the points within each stripe are sorted. Subsequently, a separate non-convex triangulation is constructed in each stripe using the Sweep Line algorithm. Neighboring stripes are then merged together, and upon obtaining a triangulation of the entire area, the triangles are rebuilt according to the Delaunay criterion. The performance of the algorithm is demonstrated on both uniform and non-uniform (fractal) point distributions within the domain. The execution time of the proposed algorithm turned out to be 1,7 times faster than that of the most efficient iterative algorithm with dynamic hashing for triangle searches. This improvement can be attributed to the significantly reduced number of Delaunay criterion checks and triangles rebuilds performed by the proposed method.

Keywords: Delaunay triangulation; stripe merging algorithm; sweep line algorithm; algorithm comparison.

For citation: Litovchenko, M.I., Kostyuk, Yu.L. (2025) Fast implementation of a non-convex Stripe Merging algorithm for Delaunay triangulation. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 71. pp. 93–102. doi: 10.17223/19988605/71/9

Введение

Задача построения триангуляции по набору заданных точек является одной из самых трудоемких задач, которые повсеместно используются в различных областях. Например, для распознавания объектов и построения их поверхностей при обработке данных лазерного сканирования [1, 2], для идентификации пространственной сцены в системе технического зрения робота [3], при моделировании местности и распознавании образов [4].

Классический алгоритм триангуляции Делоне для точек на плоскости по методу «разделяй и властвуй» имеет порядок трудоемкости в наихудшем случае $O(n \cdot \log n)$ [5]. На данный момент создано достаточно много эффективных реализаций алгоритмов для триангуляции [6], один из самых эффективных – алгоритм динамического кэширования поиска треугольников со средней трудоемкостью $O(n)$. Тем не менее исследования по разработке новых подходов или модификации уже существующих для определенных задач остаются актуальными. Обзор основных современных алгоритмических подходов можно найти в [7].

Модификация алгоритма заметающей прямой изложена в [8], основная идея заключается в учете свойств описанной вокруг треугольника окружности, что позволяет уменьшить количество вытянутых треугольников.

Например, в работе [9] описывается альтернативный подход для построения триангуляции Делоне – генетический алгоритм, который вычисляет приближенную оптимальную триангуляцию, но работает дольше традиционных алгоритмов. Этот алгоритм позволяет динамически добавлять новые точки в триангуляцию, что может потребоваться в ряде задач. Кроме того, он может быть распараллелен.

Не каждый алгоритм триангуляции можно реализовать параллельно, поэтому подобные исследования также являются актуальными. Однако из-за вычислительной сложности некоторых этапов алгоритмов их последовательная реализация иногда может отрабатывать быстрее параллельной. Так, в работе [10] последовательная реализация триангуляции произвольных областей методом «разделяй и властвуй» [5] оказалась эффективнее параллельной.

Одним из наиболее подходящих алгоритмов для исследования возможности распараллеливания является алгоритм полосового слияния [11].

В данной работе представлена такая модификация алгоритма полосового слияния, которая позволила его существенно ускорить по сравнению с алгоритмом динамического кэширования даже при последовательной реализации.

1. Общая схема алгоритма полосового слияния

Алгоритм полосового слияния содержит следующие этапы:

- 1) разбиение рабочей области с точками на полосы, например по координате Y ;
- 2) сортировка точек по координате X в каждой полосе;
- 3) триангуляция по точкам внутри каждой полосы методом заметающей прямой;
- 4) сшивание полученных триангуляций в одну;
- 5) проверка и перестроение всех треугольников по критерию Делоне [6].

При реализации алгоритма в работе [11] были обнаружены следующие проблемы:

- 1) на этапе триангуляции внутри полосы могли создаваться вырожденные треугольники, когда все три точки лежат на одной прямой;
- 2) при сшивании двух полос из-за невыпуклых границ триангуляций в полосах затруднительно создавать корректные треугольники.

Первая проблема в работе [11] решалась добавлением в алгоритм дополнительных проверок и перестроений, а вторая – достраиванием границ полос в одном варианте до полной выпуклости, а в другом варианте – до частичной выпуклости с дополнением многочисленных проверок. В результате появлялось большое количество вытянутых треугольников, которые позже снова перестраивались

по критерию Делоне. Из-за этого оба варианта алгоритма существенно проигрывали по быстродействию алгоритму динамического кэширования [11].

В данной работе предлагается ряд изменений на 1-м, 3-м и 4-м этапах работы алгоритма, позволивших значительно ускорить его выполнение.

2. Обработка точек и структуры данных

Рабочая область для построения триангуляции делится на некоторое количество полос одинаковой ширины. Ширина полос подбирается таким образом, чтобы обеспечить минимум перестроений треугольников: стороны треугольников должны быть близки к средней величине. Оптимальное количество полос, приведенное в работе [11], можно рассчитать по упрощенной формуле (с округлением до целого)

$$m = \left\lceil 0,5 \sqrt{\frac{b}{a} n} + 0,5 \right\rceil, \quad (1)$$

где m – количество полос, a – размер области по координате X , b – размер области по координате Y , n – количество точек в области.

Для каждой полосы формируется список входящих точек, а также дополнительные четыре граничные точки. Стоит отметить, что координаты всех точек преобразованы к целым числам, а минимальные и максимальные границы рабочей области – $x_{\min}$, $x_{\max}$ по координате X , $y_{\min}$, $y_{\max}$ по координате Y – расширяются на 1 по всем направлениям. Таким образом, для всех граничных точек следует вычислить значения только по координате Y , так как по координате X значения будут равны соответственно $x_{\min} - 1$, $x_{\max} + 1$. Значения по координате Y для граничных точек вычисляются по следующему рекуррентному соотношению:

$$\begin{cases} P_1 = y_{\min} - 1; S_1 = P_1; \\ P_i = P_{i-1} + p; S_i = \lfloor P_i \rfloor, i = \overline{2, m}; \\ S_{m+1} = y_{\max} + 1; \end{cases} \quad (2)$$

где $p = (y_{\max} - y_{\min} + 2)/m$ – средняя ширина полосы, P_i – вещественная переменная, S_i – нижняя граница по оси Y для i -й полосы. Тогда в i -ю полосу попадают точки с координатами по оси Y : $S_i \leq y < S_{i+1}$.

Важным этапом является сортировка точек по оси X для каждой полосы, а при одинаковых координатах X – по оси Y , что гарантирует отсутствие вырожденных треугольников при построении невыпуклой триангуляции методом заматающей прямой в каждой полосе. При использовании метода цифровой сортировки можно добиться линейного порядка трудоемкости $O(n)$ для первых двух этапов.

На трудоемкость алгоритмов часто влияет сама структура представления триангуляции. В [6] подробно описаны различные структуры такого представления и их влияние на трудоемкость. Одна из наиболее распространенных структур триангуляции хранит в себе массив из трех номеров вершин треугольника и массив из трех номеров соседних треугольников. Обход сторон треугольников в такой структуре совершается всегда по часовой стрелке. Такая структура используется и в предлагаемой реализации.

3. Частичная триангуляция в полосах

Когда полосы готовы, в каждой из них методом заматающей прямой строится отдельная невыпуклая триангуляция. Первый треугольник строится из первых двух граничных точек полосы и первой «физической» точки полосы. Далее, продвигаясь по оси X заматающей прямой, берется следующая отсортированная точка. Для нее определяется ближайшее ребро последнего построенного треугольника, строится новый треугольник с вершинами ближайшего ребра и т.д. Кроме того, на этом этапе для каждой полосы создаются списки номеров треугольников, образующих верхнюю и нижнюю границы полосы. Они необходимы для последующего сшивания триангуляций. Трудоемкость данного этапа также имеет порядок $O(n)$.

4. Сшивание триангуляций

Сшивание двух соседних полос производится по граничным ребрам невыпуклых триангуляций. При объединении приходится учитывать множество случаев вырожденности треугольников вследствие невыпуклости границ самих полос. Особенность новой реализации состоит в том, что при объединении полос на их некоторых участках границы полос достраиваются до кусочно-выпуклой оболочки только в особых, нечасто встречающихся случаях.

Идея сшивания сводится к рассмотрению конструкции из трех ребер, как на рис. 1. Здесь S_0, S_1, S_2, S_3 – граничные точки нижней полосы, S_2, S_3, S_4, S_5 – граничные точки верхней полосы. Ведущее ребро t_1b_1 – ребро последнего построенного треугольника при сшивании, t_1t_2 – просматриваемая нижняя граница верхней полосы, а b_1b_2 – соответственно верхняя граница нижней полосы. Во избежание появления вырожденных треугольников необходимо найти такую точку t_n или b_n , чтобы выполнялось условие $t_n > b_2$ или $b_n > t_2$. То есть такую точку верхней полосы, которая по координате X будет лежать дальше, чем текущая точка b_2 нижней полосы. И наоборот, для нижней полосы точка должна лежать по координате X дальше, чем текущая точка t_2 верхней полосы.

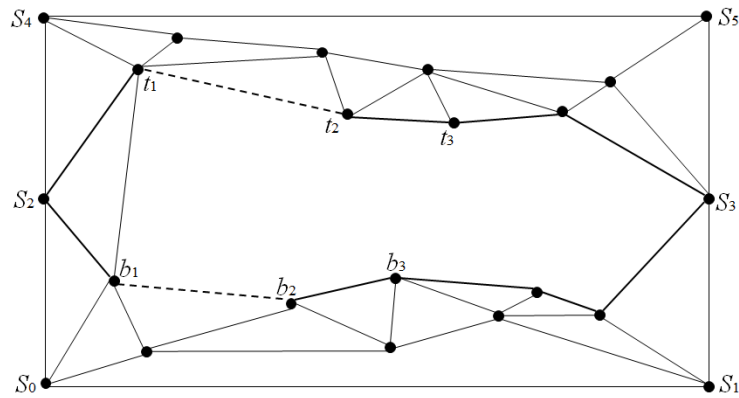


Рис. 1. Пример выбора построения треугольника при сшивании полос. Ситуация $t_3 > b_2, b_3 > t_2$

Fig. 1. An example of choosing the construction of a triangle when stripes are merging. Situation $t_3 > b_2, b_3 > t_2$

Для простоты опишем алгоритм для модификации нижней границы (для верхней границы – аналогично). Находим точку b_n , лежащую дальше точки t_2 по координате X . Если $n = 3$, то можно построить хотя бы один невырожденный треугольник $t_2b_1t_1$ или $t_1b_2b_1$. Выбор, какой треугольник построить, определяется минимальной длиной нового ведущего ребра t_2b_1 или t_1b_2 .

Если же $n > 3$, строится кусочно-выпуклая оболочка на участке границ, содержащих точки $[b_2, b_n]$ (для верхней полосы – аналогично на участке $[t_2, t_n]$). При этом изменяется список номеров треугольников, образующих границу полосы. Данная ситуация проиллюстрирована на рис. 2.

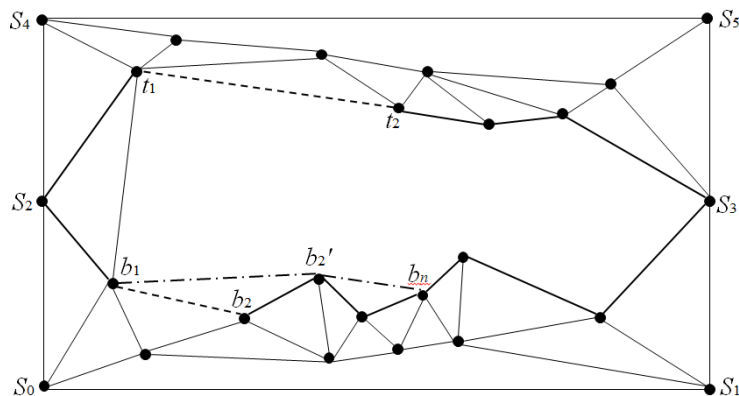


Рис. 2. Достраивание кусочно-выпуклой оболочки при $n > 3$ на участке границ $[b_2, b_n]$.

Новое рассматриваемое ребро при построении треугольников – граница b_1b_2'

Fig. 2. Completion of the piecewise convex hull at $n > 3$ on the boundary segment $[b_2, b_n]$.

Boundary b_1b_2' is a new considered edge at building triangles

Стоит отметить, что первоначально в качестве ведущего ребра t_1b_1 выбирается первое ребро верхней или нижней границы, правая точка которого имеет меньшую координату по оси X . Это вносит единообразие в обработку множества проверок на вырожденность треугольников, в том числе самого первого, при сшивании полос. Тогда следующее ребро выбранной границы становится рассматриваемым, тем самым образуется та же конструкция из ребер t_1t_2 , t_1b_1 , b_1b_2 .

5. Проверка и перестроение треугольников

После сшивания всех полос в единую триангуляцию запускается этап проверки и перестроения треугольников по критерию Делоне. Это позволяет уменьшить количество проверок и перестроений по сравнению с алгоритмом динамического кэширования, в котором перестроение проводится после образования каждого нового треугольника.

На рис. 3 представлены триангуляции до и после этапов сшивания полос и перестроения треугольников на наборе из 5 тыс. точек, полученных при наземном лазерном сканировании участка местности.

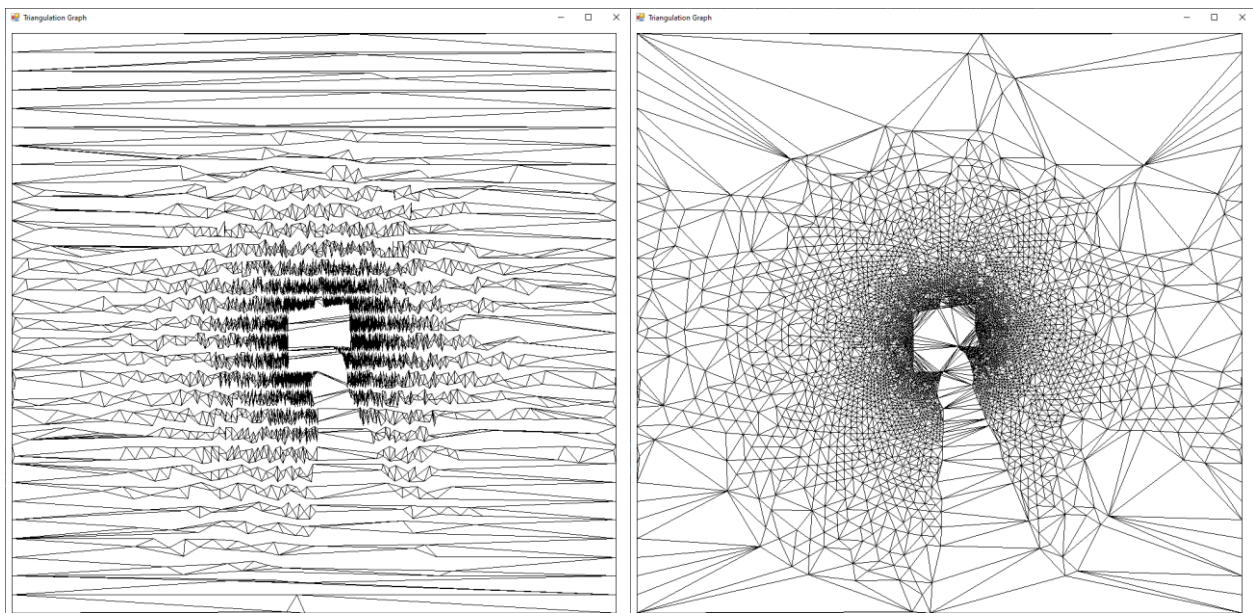


Рис. 3. Полученные триангуляции: до сшивания полос (слева), после сшивания полос и перестроения треугольников (справа)

Fig. 3. The computed triangulations before stripe merge (on the left), the final merged triangulation (on the right)

6. Сравнение алгоритмов

Опишем вычислительный эксперимент, в ходе которого было проведено качественное сравнение двух алгоритмов: динамического кэширования и невыпуклого полосового слияния. Реализация алгоритмов выполнена на языке C++, для замеров времени используется библиотека <chrono>. Отображения итоговых триангуляций получены программой, написанной на языке C#.

В табл. 1 приведены сравнительные характеристики вычислительного эксперимента для двух алгоритмов триангуляции: DH (Dynamic Hash) – алгоритма с динамическим кэшированием поиска треугольников, SM (Stripe Merging) – алгоритма невыпуклого полосового слияния. Количество точек n , сгенерированных по равномерному распределению, в эксперименте варьирует от 500 тыс. до 4 млн. Качественные характеристики для оценки алгоритмов слева направо: количество точек n , время в секундах t , сумма ребер всех треугольников в триангуляции ($\sum$ ребер), среднее значение минимального угла треугольников (min угол), среднее количество проверок для перестроения треугольников на точку, среднее количество перестроений треугольников на точку. При этом не учитывались треугольники на границах триангуляций, у которых хотя бы одна вершина является граничной точкой. Большинство

таких треугольников очень вытянутые, поэтому они заметно влияют на качественные характеристики итоговой триангуляции. Поэтому в практических случаях эти треугольники, содержащие «искусственные» граничные точки, обычно удаляют.

В алгоритмах использовано два критерия перестроения треугольников – по условию Делоне и по минимальному ребру [5], когда для пары соседних по ребру треугольников, т.е. в четырехугольнике, строится диагональ меньшей длины. При таких проверке и перестроении требуется меньше вычислений, чем по условию Делоне, но полученная триангуляция будет отличаться от триангуляции Делоне. В ряде практических применений может использоваться любая из таких триангуляций, поэтому в табл. 1 приведены характеристики обоих вариантов.

Таблица 1

Качественные характеристики двух алгоритмов построения триангуляции

Алгоритм, критерий	n , млн	t , с	Σ ребер	min угол, °	Кол-во проверок	Кол-во перестроений
ДН, Делоне	0,5	3,99	14 447 526	31,54	20,76	2,79
ДН, мин. ребро		3,62	14 286 233	31,93	19,34	2,45
SM, Делоне		2,34	14 649 084	31,56	11,05	1,01
SM, мин. ребро		2,22	14 465 991	32,03	10,75	0,96
ДН, Делоне	1	8,11	20 271 128	32,35	20,01	2,61
ДН, мин. ребро		7,46	20 097 236	32,73	18,69	2,3
SM, Делоне		4,67	20 671 102	32,37	10,35	0,91
SM, мин. ребро		4,33	20 465 094	32,82	10,06	0,85
ДН, Делоне	2	16,41	28 096 636	33,78	18,59	2,31
ДН, мин. ребро		14,97	27 936 010	34,17	17,51	2,06
SM, Делоне		8,81	28 970 376	33,82	9,33	0,77
SM, мин. ребро		8,49	28 773 724	34,24	9,16	0,73
ДН, Делоне	4	29,21	38 158 768	36,35	16,18	1,87
ДН, мин. ребро		26,91	37 996 544	36,75	15,45	1,7
SM, Делоне		15,51	39 799 132	36,41	7,75	0,56
SM, мин. ребро		15,14	39 617 344	36,82	7,7	0,55

Проанализируем работу алгоритмов на примере набора из 500 тыс. точек. Как видно из табл. 1, предлагаемый алгоритм SM обрабатывает быстрее алгоритма ДН в 1,7 раза. При этом отличия триангуляций, полученные этими алгоритмами, незначительны и объясняются тем, что в алгоритме ДН добавляется всего 4 дополнительных граничных точки, а в алгоритме SM – по 4 точки в каждой полосе. Среднее количество проверок по критерию Делоне на одну точку в алгоритме SM почти в 2 раза меньше, чем в алгоритме ДН, а количество перестроений почти в 3 раза меньше.

Если провести анализ по различиям полученных результатов по критериям перестроения Делоне и минимального ребра, то можно увидеть, что алгоритм с перестроением по условию минимального ребра работает чуть быстрее и требует чуть меньше проверок. Интересно, что среднее значение минимальных углов треугольников у этого алгоритма больше, т.е. в совокупности все треугольники оказались ближе к равнобедренным.

Таблица 2

Распределение времени на этапах полосового алгоритма

Распределение времени	0,5 млн	% от t	1 млн	% от t
Разбиение пространства и точек на полосы	0,16	7,33	0,31	7,35
Общее время сортировки точек во всех полосах	0,31	14,23	0,69	16,51
Общее время частичной триангуляции по полосам	0,37	16,83	0,68	16,26
Общее время объединения всех полос	0,34	15,55	0,67	15,91
Проверка и перестроение всех треугольников	1,01	46,06	1,84	43,96
Общее время выполнения алгоритма, t	2,19	–	4,19	–

Так как предлагаемый алгоритм состоит из нескольких этапов, то интересно отследить вклад каждого этапа в общее время работы. Рассмотрим время выполнения основных этапов на двух примерах: из 500 тыс. и 1 млн точек. Как видно из табл. 2, наибольшее влияние на общее время алгоритма имеет этап проверки и перестроения треугольников, его время составляет примерно 45% от общего времени. Этапы сортировки точек в полосах, построения триангуляции в полосах, объединения полос имеют примерно одинаковый вклад в процентном соотношении.

7. Поведение алгоритма при неравномерном распределении точек

Также было исследовано, как поведет себя алгоритм при построении триангуляции на точках, сгенерированных по неравномерному распределению. Моделью такого распределения внутри области, например квадрата, может служить фрактальное распределение с точкой сгущения в центре квадрата.

Один из вариантов одномерного фрактального распределения в диапазоне $[0, 1]$ задается следующей интегральной функцией распределения [12]:

$$\begin{cases} y = x^{1/r}, 0 \leq x \leq 1, \\ y = 0, x \leq 0, \\ y = 1, x \geq 1, \end{cases} \quad (3)$$

где $r \geq 1$ – ранг распределения.

Если $r = 1$, то это распределение превращается в равномерное. Чем больше r , тем больше распределение отличается от равномерного. Это распределение самоподобно, так как если взять часть кривой в диапазоне $[0, x_0]$, умножить все $x \leq x_0$ на $1/x_0$, то соответствующие значения функции совпадут с исходными на всем диапазоне $[0, 1]$.

При генерации координат случайных точек по этому распределению нужно датчиком случайных чисел сгенерировать две независимые случайные величины d_0 и d_1 с равномерным распределением в диапазоне $[0, 1)$. Затем вычислить случайное расстояние от центра квадрата $D \times D$ через функцию, обратную к функции (3):

$$R = d_0^r D / \sqrt{2}. \quad (4)$$

Тогда координаты точки относительно центра квадрата вычисляются по формулам

$$x = R \cos(2\pi d_1), \quad y = R \sin(2\pi d_1), \quad (5)$$

причем если точка выходит за пределы квадрата, то она отбрасывается.

В табл. 3 приведены качественные характеристики работы двух алгоритмов для фрактального распределения при $r = 1,5$ на наборах из 500 тыс. и 1 млн точек.

Таблица 3

Характеристики алгоритмов при обработке точек фрактального распределения, $r = 1,5$

Алгоритм, критерий	n , млн	t , с	Σ ребер	min угол, °	Кол-во проверок	Кол-во перестроений
DH, Делоне	0,5	4,27	12 237 120	32,85	17,78	2,32
DH, мин. ребро		3,91	12 102 218	33,23	16,65	2,04
SM, Делоне		2,42	12 281 369	32,87	10,83	1,13
SM, мин. ребро		2,44	12 124 610	33,31	10,68	1,10
DH, Делоне	1	8,24	17 010 980	33,73	16,41	2,08
DH, мин. ребро		7,56	16 853 486	34,09	15,41	1,84
SM, Делоне		5,1	17 109 162	33,75	9,51	0,93
SM, мин. ребро		4,72	16 925 260	34,16	9,39	0,91

Как и ожидалось, время работы обоих алгоритмов увеличилось, но всего на 5%. Другие характеристики построенных триангуляций изменились незначительно. Таким образом, предложенный алго-

ритм работает достаточно эффективно как для равномерных, так и для неравномерных (фрактальных) распределений.

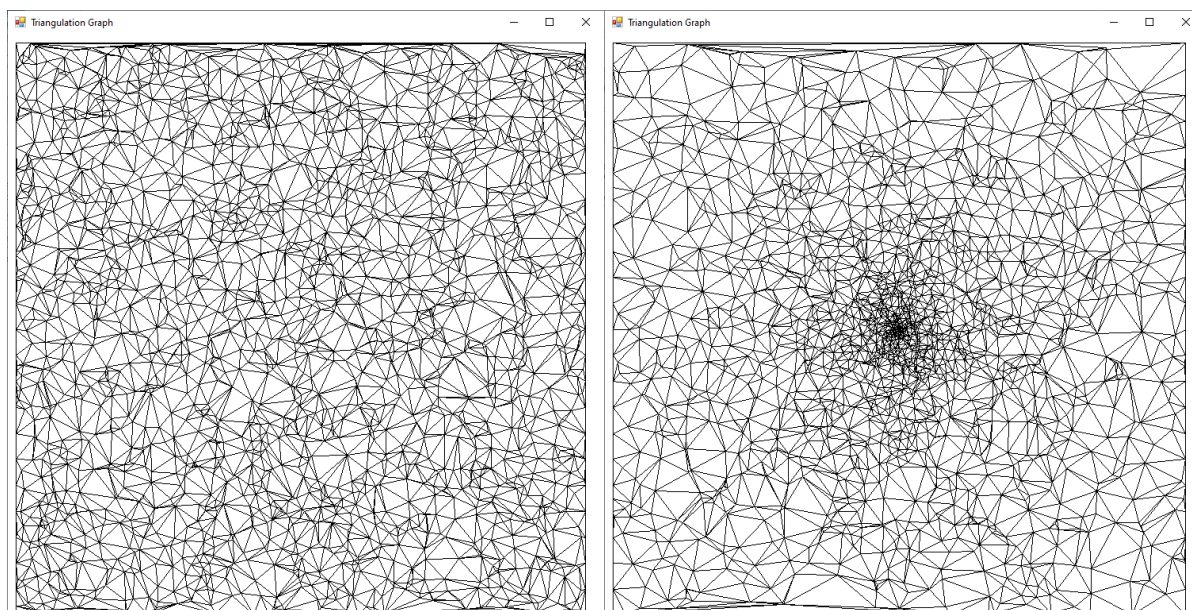


Рис. 4. Триангуляция по набору из 2 000 точек с равномерным распределением (слева) и с неравномерным (фрактальным, $r = 1,5$) распределением (справа)

Fig. 4. Delaunay triangulation of a set of 2000 points with uniform distribution (on the left) and with non-uniform (fractal, $r = 1.5$) distribution (on the right)

На рис. 4 приведен пример построенной триангуляции предложенным полосовым алгоритмом по набору из 2 000 точек для равномерного и неравномерного (фрактального) распределения.

Заключение

В данной статье описана быстрая реализация полосового алгоритма триангуляции Делоне, время работы которого в 1,7 раза меньше по сравнению с известным наиболее быстрым алгоритмом динамического кэширования поиска треугольников. Этого удалось добиться за счет следующих модификаций:

1) к каждой полосе точек добавлено 4 граничные точки; кроме того, отсортированные внутри полосы точки по оси X при одинаковых координатах X отсортированы по Y , что гарантирует отсутствие вырожденных треугольников при триангуляции внутри полосы;

2) в процессе сшивания полос при выборе очередного сшивающего ребра вначале выполняются самые простые проверки расположения точек на границах полос, гарантирующие построение невырожденных треугольников, и только если эти проверки безуспешны, выполняются более сложные проверки и перестроение участка одной из границ до частичной выпуклости.

Благодаря этим модификациям количество проверок треугольников уменьшилось почти в 2 раза, а количество перестроений треугольников – почти в 3 раза по сравнению с алгоритмом динамического кэширования. Как показал вычислительный эксперимент, эти выводы справедливы для области точек как с равномерным, так и неравномерным (фрактальным) распределением.

Кроме того, в вычислительном эксперименте проверялась работа алгоритмов с перестроением треугольников не только по критерию Делоне, но и по минимальному ребру, когда получается триангуляция, вполне пригодная для практического применения. При этом время работы алгоритмов сокращается на 3–5%, а такие общие характеристики триангуляции, как сумма длин всех ребер всех треугольников и среднее значение минимальных углов во всех треугольниках, изменяется незначительно.

Преимуществом предложенного алгоритма является не только его более высокая скорость работы, но и возможность распараллеливания на десятки или даже сотни потоков на основных этапах выполнения.

Список источников

1. Костюк Ю.Л., Гульбин К.Г., Пешехонов С.В. Построение поверхностной триангуляции и выделение пространственных фигур по данным лазерного сканирования // Вестник Томского государственного университета. Сер. «Информатика, кибернетика, математика». 2006. № 293. С. 151–155.
2. Костюк Ю.Л., Литовченко М.И. Распознавание граней трехмерных объектов по данным лазерного сканирования // Информационные технологии и математическое моделирование (ИТММ–2017) : материалы XVI Междунар. конф. им. А.Ф. Терпугова. Томск : Изд-во НТЛ, 2017. С. 55–61.
3. Храмов В.В. Способ описания объектов сцены системы технического зрения робота средствами нечеткой триангуляции Делоне с использованием адаптивной сетки // Современные информационные технологии и ИТ-образование. 2020. Т. 16, № 4. С. 833–840. doi: 10.25559/SITITO.16.202004.833-840
4. Крамаров С.О., Митясова О.Ю., Темкин И.О., Храмов В.В. Методология интеллектуальной навигации для управления автономными подвижными объектами на основе триангуляции Делоне // Горный информационно-аналитический бюллетень. 2021. № 2. С. 87–98. doi: 10.25018/0236-1493-2021-2-0-87-98
5. Препарата Ф., Шеймос М. Вычислительная геометрия: введение : пер. с англ. М. : Мир, 1989.
6. Скворцов А.В. Триангуляция Делоне и ее применение. Томск : Изд-во Том. ун-та, 2002.
7. Elshakhs Yu.S. et al. A comprehensive survey on Delaunay triangulation: applications, algorithms, and implementations over CPUs, GPUs, and FPGAs // IEEE Access. 2024. V. 12. P. 12562–12585. doi: 10.1109/ACCESS.2024.3354709
8. Hadi N.A., Farhani A., Dahalan W.M. An Improved Simple Sweep Line Algorithm for Delaunay Refinement Triangulation // Advanced Engineering for Processes and Technologies II / A. Ismail, W.M. Dahalan, A. Öchsner (eds.). Springer, 2021. С. 263–270. doi: 10.1007/978-3-030-67307-9_23
9. Dimitriou P., Karyotis V. On the computation of Delaunay triangulations via genetic algorithms // Evolutionary Intelligence. 2024. V. 17 (4). P. 2413–2432. doi:10.1007/s12065-023-00893-5
10. Бикбулатов Т.Х., Тумаков Д.Н. Использование частичной параллелизации для триангуляции двумерных областей // Программные продукты и системы. 2022. Т. 35, № 3. С. 293–304. doi: 10.15827/0236-235X.139.293-304
11. Скворцов А.В., Костюк Ю.Л. Эффективные алгоритмы построения триангуляции Делоне // Геоинформатика. Теория и практика. 1998. Вып. 1. С. 22–47.
12. Костюк Ю.Л. Эффективные алгоритмы обработки и отображения графических данных и их реализация в программных комплексах : автореф. ... дис. д-ра техн. наук. Томск, 2002. URL: <http://vital.lib.tsu.ru/vital/access/manager/Repository/vtls:000139431> (дата обращения: 26.12.2024).

References

1. Kostyuk, Yu.L., Gulbin, K.G. & Peshekhonov, S.V. (2006) Postroenie poverkhnostnoy triangulyatsii i vydelenie prostranstvennykh figur po dannym lazernogo skanirovaniya [Construction of surface triangulation and selection of spatial figures based on laser scanning data]. *Vestnik Tomskogo gosudarstvennogo universiteta. Ser. "Informatika, kibernetika, matematika."* 293. pp. 151–155.
2. Kostyuk, Yu.L. & Litovchenko, M.I. (2017) Raspoznavanie graney trekhmernykh ob"ektov po dannym lazernogo skanirovaniya [Recognition of edges of three-dimensional objects based on laser scanning data]. *Informatsionnye tekhnologii i matematicheskoe modelirovanie (ITMM–2017)* [Information Technologies and Mathematical Modeling (ITMM–2017)]. Proc. of the 16th International Conference named after A.F. Terpugov. Tomsk: NTL. pp. 55–61.
3. Khramov, V.V. (2020) Sposob opisaniya ob"ektov stseny sistemy tekhnicheskogo zreniya robota sredstvami nechetkoy triangulyatsii Delone s ispol'zovaniem adaptivnoy setki [A Method of Describing Objects in a Scene of a Robot Vision System by Means of Fuzzy Delaunay Triangulation Using an Adaptive Mesh]. *Sovremennye informatsionnye tekhnologii i IT-obrazovanie*. 16(4). pp. 833–840. DOI: 10.25559/SITITO.16.202004.833-840
4. Kramarov, S.O., Mityasova, O.Yu., Temkin, I.O. & Khramov, V.V. (2021) Metodologiya intellektual'noy navigatsii dlya upravleniya avtonomnymi podvizhnymi ob"ektami na osnove triangulyatsii Delone [Delaunay triangulation-based methodology of intelligent navigation and control of mobile objects]. *Gornyy informatsionno-analiticheskiy byulleten'*. 2. pp. 87–98. DOI: 10.25018/0236-1493-2021-2-0-87-98
5. Preparata, F.P. & Shamos, M.I. (1989) *Vychislitel'naya geometriya: vvedenie* [Computational Geometry: An Introduction]. Translated from English). Moscow: Mir.
6. Skvortsov, A.V. (2002) *Triangulyatsiya Delone i ee primenenie* [Delaunay Triangulation and its Application]. Tomsk: Tomsk State University.
7. Elshakhs, Yu.S. et al. (2024) A comprehensive survey on Delaunay triangulation: applications, algorithms, and implementations over CPUs, GPUs, and FPGAs. *IEEE Access*. 12. DOI: 10.1109/ACCESS.2024.3354709
8. Hadi, N.A., Farhani, A. & Dahalan, W.M. (2021) An improved simple sweep line algorithm for delaunay refinement triangulation. In: Ismail, A., Dahalan, W.M. & Öchsner, A. (eds) *Advanced Engineering for Processes and Technologies II*. Springer. pp. 263–270. DOI: 10.1007/978-3-030-67307-9_23
9. Dimitriou, P. & Karyotis, V. (2024) On the computation of Delaunay triangulations via genetic algorithms. *Evolutionary Intelligence*. 17(4). pp. 2413–2432. DOI: 10.1007/s12065-023-00893-5

10. Bikbulatov, T.Kh. & Tumakov, D.N. (2022) Ispol'zovanie chastichnoy parallelizatsii dlya triangulyatsii dvumernykh oblastey [Using partial parallelization to triangulate 2D domains]. *Programmnye produkty i sistemy – Software & Systems*. 35(3). pp. 293–304. DOI: 10.15827/0236-235X.139.293-304
11. Skvortsov, A.V. & Kostyuk, Yu.L. (1998) Effektivnye algoritmy postroeniya triangulyatsii Delone [Efficient algorithms for constructing Delaunay triangulation]. *Geoinformatika. Teoriya i praktika – Geoinformatics. Theory and Practice*. 1. pp. 22–47.
12. Kostyuk, Yu.L. (2002) *Effektivnye algoritmy obrabotki i otobrazheniya graficheskikh dannykh i ikh realizatsiya v programmnykh kompleksakh* [Effective algorithms for processing and displaying graphic data and their implementation in software systems]. Abstract of Engineering Dr. Diss. Tomsk: Tomsk State University.

Информация об авторах:

Литовченко Марина Игоревна – ассистент, аспирант кафедры теоретических основ информатики Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: litovchenkom21@yandex.ru

Костюк Юрий Леонидович – профессор, доктор технических наук, профессор кафедры теоретических основ информатики Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: kostyuk.yu48@mail.ru

Вклад авторов: все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

Information about the authors:

Litovchenko Marina I. (Assistant, Post-graduate Student, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: litovchenkom21@yandex.ru

Kostyuk Yuriy L. (Professor, Doctor of Technical Sciences, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: kostyuk.yu48@mail.ru

Contribution of the authors: the authors contributed equally to this article. The authors declare no conflicts of interests.

Поступила в редакцию 16.01.2025; принята к публикации 02.06.2025

Received 16.01.2025; accepted for publication 02.06.2025

Научная статья
УДК 621.396.96
doi: 10.17223/19988605/71/10

Применение УПС-фильтра при решении задачи обнаружения в режиме активной локации

Игорь Михайлович Рудько

*Институт проблем управления им. В.А. Трапезникова Российской академии наук,
Москва, Россия, migor-rudko@mail.ru*

Аннотация. Для режима активной локации проводится сравнение двух алгоритмов обнаружения при использовании ими энергетического критерия обнаружения – «классической» задачи проверки двух простых гипотез и применения фильтра на основе порядковой статистики (УПС-фильтра), которым заменяется интегратор в «классической» задаче. Рассматривается два режима – обработка единичного импульса и обработка пачки импульсов. Показано, что при фиксированной ложной тревоге применение УПС-фильтра обеспечивает существенный выигрыш в вероятности обнаружения.

Ключевые слова: активная локация; порядковая статистика; проверка статистических гипотез; системы обнаружения.

Для цитирования: Рудько И.М. Применение УПС-фильтра при решении задачи обнаружения в режиме активной локации // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2025. № 71. С. 103–111. doi: 10.17223/19988605/71/10

Original article
doi: 10.17223/19988605/71/10

Using the TOS-filter in solving the detection problem in active location mode

Igor M. Rudko

*V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences,
Moscow, Russian Federation, igor-rudko@mail.ru*

Abstract. For the active location mode, a comparison is made of two detection algorithms using the energy detection criterion - the "classical" task of checking two simple hypotheses and the use of a filter based on order statistics (TOS- filter), which replaces the integrator in the "classical" task. Two modes are considered - processing a single pulse and processing a pulse train. It is shown that with a fixed false alarm, the use of the TOS-filter provides a significant gain in the detection probability.

Keywords: active location; order statistics; testing statistical hypotheses; detection systems.

For citation: Rudko, I.M. (2025) Using the TOS-filter in solving the detection problem in active location mode. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 71. pp. 103–111. doi: 10.17223/19988605/71/10

Введение

Задача активной локации (АЛ) – обнаружение лоцируемого объекта-цели, определение расстояния до нее, угловых координат цели, направления ее движения и ее скорости. Задача решается при использовании сигналов, зондирующих цель неконтактным методом, – облучением ее, получением отраженного от цели (рассеянного целью) сигнала и его обработкой, которая на *этапе первичной обработки* включает операции обнаружения и измерения параметров сигналов. Совокупность оценок параметров сигналов образует локационную отметку цели.

Известно два основных способа обработки зондирующих сигналов (ЗС) в режиме АЛ. Первый – непосредственное измерение временного интервала между моментами излучения и возвращения отраженного от цели сигнала. Второй способ состоит в вычислении функции взаимной корреляции между посланным и возвращенным сигналами. Преимуществом этого способа является возможность работы с длительными ЗС [1]. Для решения задач АЛ применяются различные виды ЗС: импульсные, непрерывные, с внутриимпульсной модуляцией и без таковой, когерентные и некогерентные, одиночные и пачечные [2. С. 33], узкополосные и сверхширокополосные (широкополосные хаотические) ЗС [1].

На этапе первичной обработки для достижения удовлетворительной вероятности обнаружения ($P_{обн}$) для всех вариантов ЗС бывает необходимо производить энергетическое накопление (интегрирование, суммирование) возвращенного сигнала. Во многих системах АЛ для достижения удовлетворительной достоверности обнаружения бывает необходимо облучать цель серией импульсов. [3. С. 293].

В настоящей работе исследуется возможность замены интегратора (сумматора) в энергетическом приемнике (ЭП) системы АЛ на фильтр на основе усеченных порядковых статистик (УПС-фильтр) [4, 5], который позволяет существенно увеличить $P_{обн}$ по сравнению с ЭП.

Цифровая обработка сигнала, к которой относится УПС-фильтрация, в настоящее время возможна только для относительно низкочастотных сигналов, ограниченных сотнями килогерц (гидроакустическая АЛ) [1].

До настоящего времени нет работ, посвященных использованию порядковых статистик в задачах обнаружения в режиме АЛ.

1. Обнаружение сигнала в задачах активной локации

Решение задачи обнаружения (*этап первичной обработки*) сводится к принятию решения о наличии или отсутствии цели. В реальных условиях на входе приемного устройства локатора наблюдается смесь полезного сигнала и шума. Это приводит к случайному характеру результатов обнаружения и указывает на статистический характер задачи обнаружения. Полной статистической характеристикой колебаний шума является плотность вероятности. Колебания шума описывают также с помощью корреляционной функции и спектральной плотности мощности. При синтезе оптимальных алгоритмов обработки АЛ сигнала, кроме корреляционных и спектральных характеристик шума, требуется знание плотности вероятности его распределения.

Задача обнаружения сводится к задаче проверки двух гипотез: H_0 – принимается только шум, H_1 – принимается смесь сигнала и шума. Решение о наличии сигнала принимается по критерию Неймана–Пирсона, согласно которому оптимальная система обнаружения должна максимизировать вероятность правильного обнаружения $P_{обн}$ при фиксированной вероятности ложной тревоги $P_{лт}$, т.е. $P_{обн} = \max$ при $P_{лт} = \text{const}$ [2. С. 77].

Разрешающая способность АЛ определяется величиной полосы частот ΔF , занимаемой сигналом. Произведение $B = \Delta F \tau$, где τ – длительность посылки, определяющее энергетическую эффективность зондирующего сигнала, называют базой сигнала. База простого импульсного сигнала равна единице. Характерная база сложных сигналов, используемых в АЛ, составляет величину порядка 100 [6]. Критерием возможности обнаружения сигнала на фоне шумов различного происхождения является соотношение сигнал/шум q^2 , определяемое как

$$q^2 = 2 \int_{\tau} P_{in} dt / P_N, \quad (1)$$

где P_{in} – мощность полезного сигнала на входе приемного устройства локатора, а P_N – спектральная плотность мощности шумов. Интегрирование проводится по времени посылки сигнала τ . Если излучаемая и, соответственно, принимаемая мощность зондирующего сигнала в течение времени посылки τ постоянна, а помеха реализуется в полосе частот Δf , занимаемой сигналом, то формулу (1) можно записать следующим образом [6]:

$$q^2 = \frac{P_{in}}{P_N} \Delta f \tau = \frac{P_{in}}{P_N} B. \quad (2)$$

Таким образом, возможность обнаружить сигнал при оптимальном приеме с заданными значениями $P_{\text{обн}}$ и $P_{\text{лт}}$ определяется лишь отношением энергии сигнала к спектральной плотности шума [2. С. 95].

2. Фильтр на основе усеченных порядковых статистик (УПС-фильтр)

Пусть имеется выборка из m элементов случайного процесса $X = \{X_1, \dots, X_i, \dots, X_m\}$. Порядковая статистика $X_{(i)}$, где $1 \leq i \leq m$, – это упорядоченные величины статистики X_i такие, что $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(i)} \leq \dots \leq X_{(m)}$. Если случайные величины X_i статистически независимы и одинаково распределены, то случайные величины $X_{(i)}$ зависимы из-за неравенств между ними.

Известно [7. С. 79], что при больших значениях m достаточно точным приближением к математическому ожиданию элемента порядковой статистики с номером μ_j является значение, удовлетворяющее уравнению $P(x) = j/(m+1)$, где $P(x)$ – функции распределения случайной величины x . Если ввести в рассмотрение функцию Q такую, что $Q[P(x)] = x$, то получим асимптотическую формулу $\mu_j \sim Q(j/(m+1))$, где $1 \leq j \leq m$, т.е. математические ожидания μ_j приближенно равны квантилям функции распределения $P(x)$, взятым с равномерным шагом.

Как этот факт можно использовать покажем на **примере**.

Рассмотрим случайные процессы X_0 и X_1 , имеющие χ^2 -распределение с n степенями свободы, но разные дисперсии $\sigma_0^2 = 1$ и $\sigma_1^2 = 1,2$. Сравним выборки этих процессов при $m = 8$ и $n = 16$ и определим математические ожидания их порядковых статистик.

Математические ожидания μ_j имеют следующие значения

$\mu_{0j} = \{ 9,553 \ 11,498 \ 13,083 \ 14,580 \ \mathbf{16,122} \ \mathbf{17,840} \ \mathbf{19,956} \ \mathbf{23,099} \};$

$\mu_{1j} = \{ 11,464 \ 13,797 \ \mathbf{15,699} \ \mathbf{17,496} \ \mathbf{19,347} \ \mathbf{21,408} \ \mathbf{23,947} \ \mathbf{27,720} \}.$

Сравним статистические свойства случайных величин Z и W , сформированных из случайной выборки $X_{(i)}$, $1 \leq i \leq m$, двумя различными способами:

1) $Z = \sum_{i=1}^m X_{(i)} = \sum_{i=1}^m X_i$ – имеет χ^2 -распределение с $2N = 2nm$ степенями свободы, математическим ожиданием $\mu_Z = 2nm\sigma^2$ и дисперсией $\sigma_Z^2 = 4nm\sigma^2$, и в силу центральной предельной теоремы при достаточно больших значениях m ее функция плотности распределения нормализуется: $Z \sim N(\mu_Z, \sigma_Z^2)$.

2) $W = \sum_{i=k}^m X_{(i)}$, где k – **плавающий** порог, определяемый из условия

$$k = \arg \min_i |h - X_{(i)}|, \quad (3)$$

h – **фиксированный** порог. При достаточно больших значениях m функция плотности распределения W также нормализуется: $W \sim N(\mu_W, \sigma_W^2)$.

Например, если задать фиксированный порог h равным медиане $X_0 - h = 15,338$, то, рассматривая приведенные выше значения математических ожиданий μ_j как выборочные значения, из (3) получим для $X_0 - k_0 = 5$, а для $X_1 - k_1 = 3$.

Если датчиком случайных чисел сгенерировать выборки рассматриваемых процессов и построить их гистограммы, то получим гистограммы, вид которых приведен на рис. 1, *a*. На рис. 1, *b* – эти же гистограммы, нормированные с учетом «веса» отброшенных членов вариационного ряда. (Вес отброшенных членов можно учесть, введя δ -функцию в нуле с соответствующим весом $P_0 = \delta(0) \int_0^h f(x) dx$.)

Из приведенных рисунков видно, что гистограммы процессов W_0 и W_1 «раздвинулись» относительно процессов Z_0 и Z_1 . Следовательно, если случайный процесс X_0 описывает помеху, а случайный процесс X_1 – смесь помехи и сигнала, то (при фиксированной $P_{\text{лт}}$) $P_{\text{обн}}$ сигнала для процесса W больше, чем для процесса Z .

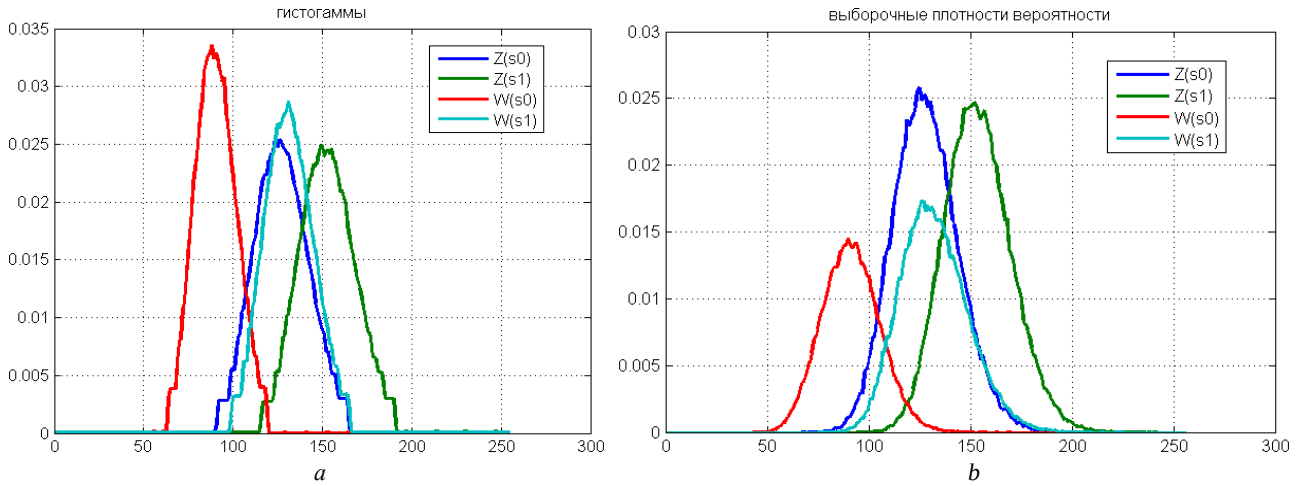


Рис. 1. Гистограммы процессов (Z_0 и Z_1) и (W_0 и W_1): a – исходные нормированные гистограммы, b – гистограммы, нормированные с учетом отброшенных членов вариационного ряда
 Fig. 1. Histograms of processes (Z_0 and Z_1) and (W_0 and W_1): a) initial normalized histograms, b) histograms normalized taking into account the discarded terms of the variation series

Алгоритм работы УПС-фильтра основан на этом свойстве порядковых статистик. Теоретические основы алгоритма и результаты его моделирования приведены в работах [4, 5], примеры использования – локация в пассивном режиме в [8], обнаружение дискретных составляющих спектра в частотной области в [9].

Алгоритм работы УПС-фильтра:

1. Время интегрирования (равное длительности посылки) разбивается на m одинаковых интервалов, в каждом из которых проводится интегрирование (накопление энергии), т.е. вычисляется последовательность выборок $\mathbf{X} \sim \{X_1, \dots, X_i, \dots, X_m\}$. Накапливается c оцениваемых выборок $\mathbf{X}$.

2. По накопленным выборкам строится матрица $\mathbf{X}_{ik}$ размерностью m строк на c столбцов (c – «глубина» матрицы памяти) – $\{X_1, \dots, X_i, \dots, X_m\}^T_k$, где $1 \leq k \leq c$.

3. В каждом столбце матрицы $\mathbf{X}_{ik}$ строится порядковая статистика $X_{(i)k}$, где $1 \leq i \leq m$, – упорядоченные величины статистики X_i такие, что $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(i)} \leq \dots \leq X_{(m)}$.

4. В каждой строке полученной матрицы $\mathbf{X}_{(i)k}$ определяются оценки математических ожиданий (вектор $\hat{m}$) $\hat{m}_i = \frac{1}{c} \sum_{k=1}^c X_{(i)k}$, где $1 \leq i \leq m$.

5. Порог отсеечения k (первый порог) определяется из условия

$$k = \arg \min_i |h_0 - \hat{m}_i|,$$

где $1 \leq i \leq m$, а порог h_0 определяется по формуле $\alpha = \int_0^{h_0} f_{\text{ш}}(x, \sigma_0^2) dx$, где α – заданный квантиль, $f_{\text{ш}}$ – плотность вероятности шума, σ_0^2 – дисперсия шума (известный параметр).

6. Вычисляется оценка

$$W = \sum_{j=k}^m X_{(j)} . \quad (4)$$

На выходе УПС-фильтра получаем последовательность отфильтрованных оценок W , задача обнаружения по которым решается «классическим» алгоритмом проверки двух гипотез.

УПС-фильтр работает по принципу скользящего окна, т.е. каждый новый вектор $\mathbf{X}_k$ с индексом $c + 1$ вытесняет из матрицы $\mathbf{X}_{(i)k}$ вектор $\mathbf{X}_k$ с индексом 1.

Если результаты применения интегратора зависят от двух параметров (времени интегрирования τ и отношения сигнал/помеха q^2), то для УПС-фильтра результаты зависят, кроме того, еще от трех параметров (см.: [5]):

- соотношения величин n и m при условии $nm = N = \text{const}$,
- величины порога отсечения h_0 ,
- глубины с матрицы памяти $\mathbf{X}_{(i)k}$.

Применительно к задаче обнаружения УПС-фильтр заменяет в алгоритме обнаружения интегратор (сумматор).

3. Моделирование применения УПС-фильтра в задаче активной локации

Рассмотрим применение УПС-фильтра на примере обработки широкополосного хаотического ЗС для двух вариантов посылки:

- обработка единичного импульса;
- обработка пачки импульсов.

Модель принимаемого сигнала: смесь белого шума и широкополосного сигнала, сгенерированного из белого шума путем двойной спектральной обработки (аналогично [1]) в заданном частотном диапазоне и имеющего заданную длительность τ с последующим квадратичным детектированием этой смеси. Генерируется M импульсных посылок со скважностью $T_0 \gg \tau$.

Обработка единичного импульса на интервале времени T_0 (например, с линии задержки) производится скользящим окном (интегратором) размерностью $\tau = N\Delta t$, где Δt – квантование по времени. То есть рассчитывается изменение энергии принимаемого сигнала за время T_0 и строится матрица $\mathbf{M}|N, L|$, где $L = T_0/\Delta t$. Матрица $\mathbf{M}$ обрабатывается согласно проведенному выше алгоритму УПС-фильтра. «Глубина» (число столбцов) матрицы УПС-фильтра определяется временем интегрирования $c\Delta t = \tau_0 \leq \tau$, т.е. за время, меньшее или равное времени интегрирования.

Порядковые статистики расчета $\mathbf{X}_{(i)}$ сохраняются в памяти и используются при обработке пачки.

Обработка пачки импульсов производится независимо для каждого временного канала t_j , где $1 \leq j \leq L$. Вариационный ряд $\{X_1, \dots, X_i, \dots, X_m\}_j$, использовавшийся для расчета W во временном канале j , является вектор-столбцом матрицы УПС-фильтра, число векторов равно числу импульсов в пачке – M . Сформированная УПС-матрица размерностью $|m, M|$ обрабатывается согласно алгоритму УПС-фильтра. Полученные по формуле (4) оценки W_i для всех импульсов пачки ($1 \leq i \leq M$) накапливаются для получения оценки УПС всей пачки $W_j = \sum_{i=k}^M W_{ij}$, где $1 \leq j \leq L$. Таким образом, параллельно рассчитываются L УПС-фильтров, на выходе которых получаем оценки энергии в t_j временном канале.

Статистическая обработка результатов (для каждого импульса или каждого временного канала) производится согласно блок-схеме, приведенной на рис. 2 (Д – детектор).

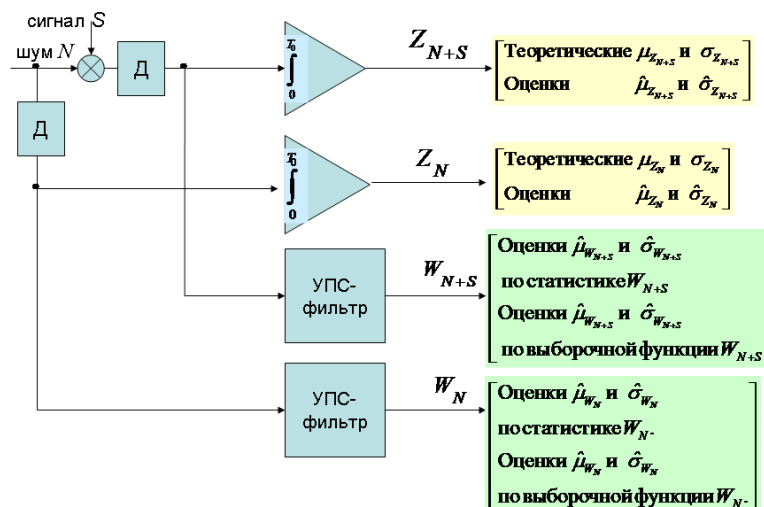


Рис. 2. Блок-схема сравнения статистик Z и W
Fig. 2. Block diagram for comparison of Z and W statistics

Моделирование проводилось в среде MATLAB.

Параметры модели:

- квантование по времени $\Delta t = 1$;
- длина реализации (скважность следования импульсов) $L\Delta t = T_0 = 256$;
- разрешение по частоте $\Delta f = 1/T_0 = 1/256$;
- максимальная частота $f_{\max} = N\Delta f/2 = L/(2T_0) = 1/(2\Delta t) = 1/2$;
- длительность посылки (импульса) $\tau\Delta t = 32$;
- начало посылки $\tau_n = L/2 - \tau/2 + 1$;
- разбиение посылки (согласно п. 1 алгоритма) $m = 8$, длительностью $n = 4$;
- база $B = \tau\Delta t(f_2 - f_1) = \tau\Delta t\Delta F = 32(64/256 - 32/256) = 4$;
- число импульсов в пачке $M = 10$;
- статистическая обработка по 1 000 замерам.

Результаты статистического моделирования *обработки единичного импульса* представлены на следующих рисунках.

На рис. 3 приведены результаты моделирования обработки единичного импульса:

а) УПС-фильтром со скользящим окном $\mathbf{M}[m, \tau']$ ($\tau' \leq \tau$) (статистика W) и интегратором (скользящим) за время интегрирования τ (статистика Z) – энергетическое накопление на выходе интегратора и УПС-фильтра в зависимости от времени;

б) УПС-фильтром со скользящим окном (матрицей) $\mathbf{M}[m, \tau']$ (статистика W) и интегратором за время τ (статистика Z) – вероятность обнаружения на выходе интегратора и УПС-фильтра.

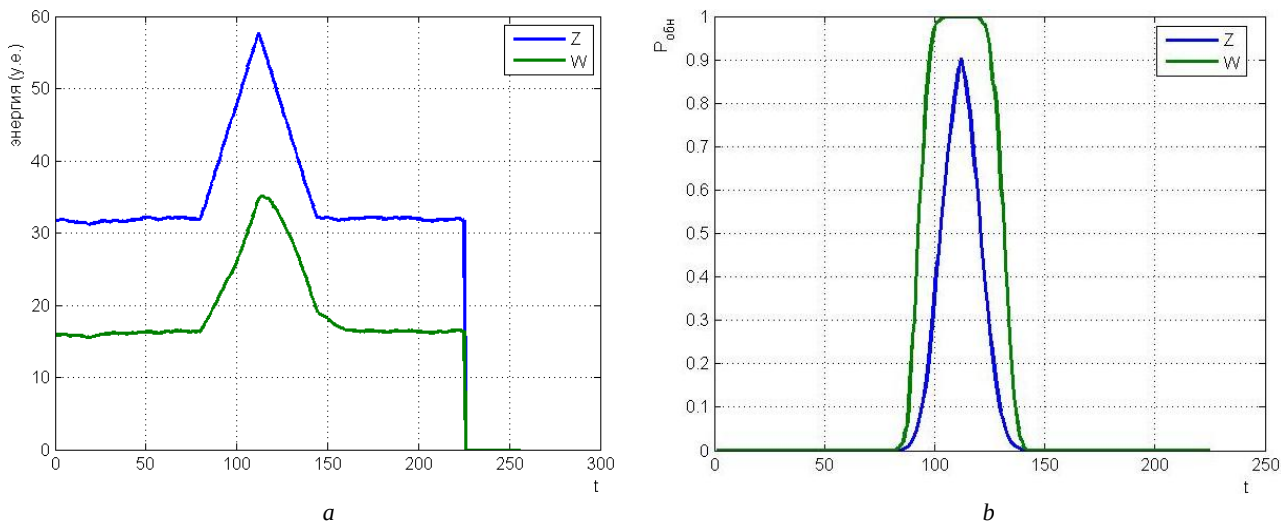


Рис. 3. Энергия на выходе интегратора (Z) и УПС-фильтра (W) – а;
вероятность обнаружения на выходе интегратора (Z) и УПС-фильтра (W) – б
Fig. 3. a) Energy at the output of the integrator (Z) and the TOS-filter (W).
b) Probability of detection at the output of the integrator (Z) and TOS-filter (W)

На рис. 4 приведены кривые обнаружения $P_{\text{обн}}(q^2)$ для интегратора и УПС-фильтра (q^2 определяется по формуле (2)). Рассматриваются зависимости $P_{\text{обн}}$ от «глубины» матрицы памяти c ($\tau' = c\Delta t$) и от порога отсеечения h_0 .

В работах [4, 5] показано, что когда время наблюдения сигнала достаточно велико, $P_{\text{обн}}$ сигнала увеличивается с увеличением «глубины» матрицы памяти c , но в случае импульсной посылки величина c ограничена длительностью посылки τ ($c\Delta t \leq \tau$). При этом необходимо учитывать, что при $c\Delta t = \tau$ скользящее окно сгладит максимум энергии импульса (см. рис. 3, а) и $P_{\text{обн}}$ по максимуму энергии уменьшится, как видно на рис. 4, а для $c = 32$. Уменьшение c ($c = 4$) также приводит к уменьшению $P_{\text{обн}}$. Оптимальной величиной c является $\tau/4 \leq c\Delta t \leq \tau/2$.

Оптимальной величиной порога отсеечения h_0 является величина, близкая к медиане плотности вероятности шума ($h_0 = 8$), что видно из рис. 4, б и хорошо согласуется с теоретическим результатом из [4].

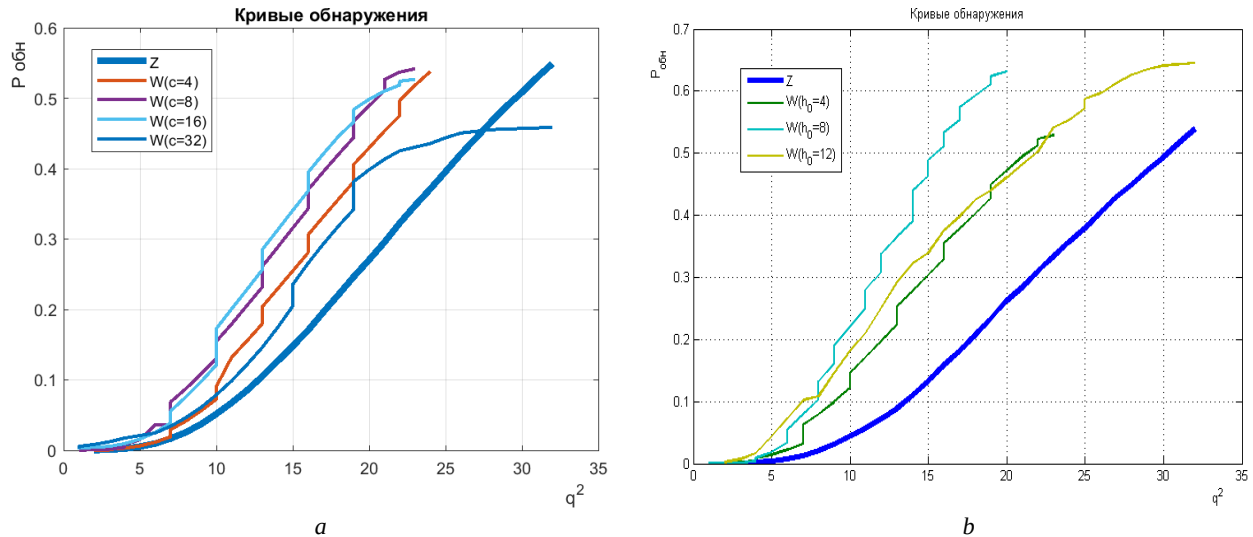


Рис. 4. Кривые обнаружения $P_{обн}(q^2)$ для интегратора (Z) и УПС-фильтра (W):
 а – зависимость от «глубины» матрицы памяти c ; б – зависимость от порога отсека h_0
 Fig. 4. Detection curves $P_{обн}(q^2)$ for the integrator (Z) and TOS-filter (W):
 а) dependence on the “depth” of the memory matrix c ; б) dependence on the cutoff threshold h_0

При обработке пачки импульсов формируется трехмерная матрица $\mathbf{M}[N, L, M]$, причем описанная выше матрица $\mathbf{M}[N, L]$ является ее частью как результат обработки k -го импульса пачки. Обработка трехмерной матрицы $\mathbf{M}[N, L, M]$ может производиться двумя способами:

- M двумерных матриц $\mathbf{M}[N, L]_k = \mathbf{M}[N, L, k]$, $1 \leq k \leq M$, обрабатываются как единичные импульсы, а результаты оценок энергии накапливаются по M импульсам;
- для фиксированного временного отсчета Δt_i , $1 \leq i \leq L$, выбирается двумерная матрица $\mathbf{M}[N, M]_i = \mathbf{M}[N, i, M]$. При обработке пачки импульсов используется не скользящее окно (матрица) $\mathbf{M}[m, \tau]$, а L матриц фиксированного размера, число векторов в которых определяется числом импульсов в пачке $\mathbf{M}[N, M]$.

На рис. 5 приведены кривые обнаружения $P_{обн}(q^2)$ для интегратора и УПС-фильтра пачки импульсов, когда УПС-фильтр обрабатывает пачку по первому способу – $P_{обн1}$ (график W_1), и когда УПС-фильтр обрабатывает пачку по второму способу – $P_{обн2}$ (график W_2).

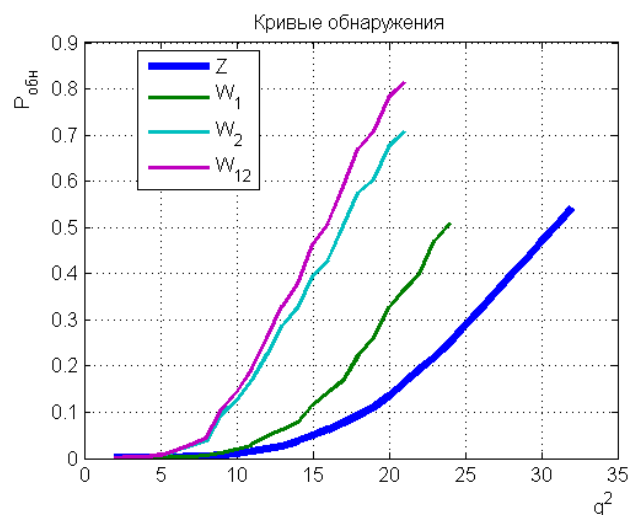


Рис. 5. Кривые обнаружения $P_{обн}(q^2)$ для интегратора и УПС-фильтра для пачки импульсов
 Fig. 5. Detection curves $P_{обн}(q^2)$ for the integrator and TOS-filter for a pulse train

Так как оценки для W определяются по разным массивам данных: по первому способу по «строкам» матрицы $\mathbf{M}[:, L, M]$, а по второму – по «столбцам» матрицы $\mathbf{M}[N, :, M]$, можно считать

полученные оценки независимыми и для расчета вероятности правильного обнаружения пачки использовать формулу интегральной (кумулятивной) вероятности $P_{\text{обн}} = 1 - (1 - P_{\text{обн1}})(1 - P_{\text{обн2}})$ (график W_{12}).

Из приведенных кривых обнаружения видно, что второй способ (обработка по «столбцам») обеспечивает лучшую $P_{\text{обн}}$, чем первый способ (обработка по «строкам»). Это, видимо, объясняется тем, что если скользящее окно «сглаживает острые углы» исходного процесса Z (см. рис. 4), то при обработке пачки по «столбцам» сглаживания нет.

Заключение

Для систем АЛ разработан алгоритм, заменяющий интегратор (сумматор) в энергетическом приемнике системы АЛ на фильтр на основе усеченных порядковых статистик (УПС-фильтр), который позволяет увеличить $P_{\text{обн}}$ при фиксированной $P_{\text{лт}}$ по сравнению с ЭП.

Статистическое моделирование при сравнении ЭП и УПС-фильтра показывает, что при малых отношениях сигнал/помеха применение УПС-фильтра обеспечивает существенный выигрыш в $P_{\text{обн}}$ по сравнению с ЭП.

Список источников

1. Залогин Н.Н., Скнаря А.В. Зондирующие сигналы для активной локации высокого пространственного разрешения // РЕНСИТ. 2016. Т. 8, № 1. С. 3–8; doi: 10.17725/rensit.2016.08.003
2. Бердышев В.П., Гарин Е.Н., Фомин А.Н. и др. Радиолокационные системы / общ. ред. В.П. Бердышева. Красноярск : Сиб. федер. ун-т, 2011.
3. Ван Трис Г. Теория обнаружения, оценок и модуляции : в 4 т. М. : Сов. радио, 1977. Т. 3.
4. Рудько И.М. Применение порядковых статистик в задачах обнаружения // Управление большими системами. 2012. Вып. 37. С. 63–83.
5. Рудько И.М. Исследование влияния параметров фильтра на основе усеченных порядковых статистик (УПС-фильтра) на его выходные характеристики // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2022. № 60, С. 42–51. doi: 10.17223/19988605/60/5
6. Залогин Н.Н., Калинин В.И., Скнаря А.В. Активная локация с использованием широкополосных хаотических сигналов // РЕНСИТ. 2011. Т. 3, № 1. С. 3–17.
7. Дэйвид Г. Порядковые статистики. М. : Наука, 1979.
8. Рудько И.М. Алгоритм обнаружения звукоядов в спектре шумоизлучения морских судов // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2023. № 62. С. 76–84. doi: 10.17223/19988605/62/8
9. Рудько И.М. Обнаружение подвижного объекта с использованием порядковых статистик // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2019. № 48. С. 42–50. doi: 10.17223/19988605/48/5

References

1. Zalogin, N. & Sknarya, A. (2016) Zondiruyushchie signaly dlya aktivnoy lokatsii vysokogo prostranstvennogo razresheniya [Sounding signals for the active location of high spatial resolution]. *RENSIT*. 8(1). pp. 3–8.
2. Berdyshev, V.P., Garin, E.N., Fomin, A.N. et al. (2011) *Radiolokatsionnye sistemy* [Radar systems]. Krasnoyarsk: Siberian Federal University.
3. Van Trees, H. (1977) *Teoriya obnaruzheniya, otsenok i modulyatsii* [Theory of Detection, Estimation and Modulation]. Vol. 3. Moscow: Sovetskoe radio.
4. Rudko, I.M. (2012) *Primenenie poryadkovykh statistik v zadachakh obnaruzheniya* [Application of ordinary statistics in detection problems]. *Upravlenie bol'shimi sistemami*. 37. pp. 63–83.
5. Rudko, I.M. (2022) Investigating the effect of filter parameters based on truncated order statistics (TOS-filter) on its output characteristics. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 60. pp. 42–51. DOI: 10.17223/19988605/60/5
6. Zalogin, N., Kalinin, V. & Sknarya, A. (2011) Aktivnaya lokatsiya s ispol'zovaniem shirokopolosnykh khaoticheskikh signalov [The active location with the use of ultrawide-band chaotic signals]. *RENSIT*. 3(1). pp. 3–17.
7. David, H.A. (1979) *Poryadkovye statistiki* [Order Statistics]. Translated from English. Moscow: Nauka.
8. Rudko, I.M. (2023) Algorithm for detecting harmonic series in the noise emission spectrum of sea vessels. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 62. pp. 76–84. DOI: 10.17223/19988605/62/8

9. Rudko, I.M. (2019) Detection of moving object using order statistics. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 48. pp. 42–50. DOI: 10.17223/19988605/48/5

Информация об авторе:

Рудько Игорь Михайлович – кандидат технических наук, старший научный сотрудник Института проблем управления им. В.А. Трапезникова Российской академии наук (Москва, Россия). E-mail: igor-rudko@mail.ru

Автор заявляет об отсутствии конфликта интересов.

Information about the author:

Rudko Igor M. (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.

Contribution of the authors: the authors contributed equally to this article. The authors declare no conflicts of interests.

Поступила в редакцию 19.02.2025; принята к публикации 02.06.2025

Received 19.02.2025; accepted for publication 02.06.2025

Научная статья

УДК 681.5

doi: 10.17223/19988605/71/11

**Идентификация в дискретных системах со случайными
и интервальными параметрами с использованием ЕКФ****Валерий Иванович Смагин¹, Константин Станиславович Ким²**^{1,2} *Национальный исследовательский Томский государственный университет, Томск, Россия*¹ *vsm@mail.tsu.ru*² *kks93@rambler.ru*

Аннотация. Рассматривается задача идентификации для дискретного объекта со случайными и интервальными параметрами. Случайные параметры в модели задаются как гауссовские случайные последовательности. Для представления интервальных параметров используется вероятностный подход, в основе которого лежит замена неопределенных параметров интервального типа независимыми случайными величинами с равномерным распределением. Для решения задачи идентификации используется расширенный фильтр Калмана (ЕКФ). Представлены результаты моделирования.

Ключевые слова: дискретные системы; идентификация; случайные параметры; интервальные параметры; расширенный фильтр Калмана.

Для цитирования: Смагин В.И., Ким К.С. Идентификация в дискретных системах со случайными и интервальными параметрами с использованием ЕКФ // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2025. № 71. С. 112–119. doi: 10.17223/19988605/71/11

Original article

doi: 10.17223/19988605/71/11

Identification in discrete systems with random and interval parameters using EKF**Valery I. Smagin¹, Konstantin S. Kim²**^{1,2} *National Research Tomsk State University, Tomsk, Russian Federation*¹ *vsm@mail.tsu.ru*² *kks93@rambler.ru*

Abstract. The identification problem for a discrete object with random and interval parameters is considered. The random parameters in the model are set as Gaussian random sequences. To represent the interval parameters, a probabilistic approach is used, which is based on the replacement of indeterminate interval-type parameters with independent random variables with a uniform distribution. The extended Kalman filter (EKF) is used to solve the identification problem. The simulation results are presented.

Keywords: discrete systems; identification; random parameters; interval parameters; extended Kalman filter.

For citation: Smagin, V.I., Kim, K.S. (2025) Identification in discrete systems with random and interval parameters using EKF. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnik i informatika* – Tomsk State University Journal of Control and Computer Science. 71. pp. 112–119. doi: 10.17223/19988605/71/11

Введение

Проблема оценки параметров стохастических динамических систем исследуется в значительной степени из-за ее важности для построения моделей [1–6]. Задачи идентификации стохастических систем, описываемых дифференциальными и разностными уравнениями, приводят к задачам нелинейного

оценивания. Было предложено несколько методов решения задач идентификации, большинство из которых основаны на алгоритмах нелинейной фильтрации ЕКФ, сигма-точечных алгоритмах нелинейной фильтрации Unscented Kalman filter (UKF) [3, 4] и кубатурных алгоритмах Cubature Kalman filter (CKF) [5, 6].

Области применения алгоритма ЕКФ варьируют от систем управления до обработки сигналов, мониторинга работоспособности систем, навигации и эконометрики. Однако точность оценки алгоритма ЕКФ может быть серьезно снижена из-за неопределенностей в описании модели, отклонений в измерениях, вызванных аномалиями датчиков и ошибками передачи данных. Попытки решить эту проблему предприняты в работах [7–14].

В работе [7] рассмотрены методы робастной идентификации в статических моделях.

В [8, 9] рассмотрены задачи синтеза алгоритмов идентификации с использованием ЕКФ при аномальных ошибках и выбросах в канале измерений.

В [10] была поставлена вспомогательная задача фильтрации по критерию H_∞ , в этой работе вводятся дополнительные данные для учета членов более высокого порядка, которыми пренебрегали при решении классической задачи ЕКФ.

В [11, 12] представлен робастный фильтр для нелинейных систем с дискретным временем, подверженных неопределенностям. В работе [11] предполагается, что нелинейные функции являются неопределенными, но принадлежат конической области. Это условие характеризуется как условие Липшица для состояния системы и остаточных значений. Предложенный алгоритм также допускает, что динамические помехи и шумы измерений могут иметь неизвестные статистические характеристики. В работе [12] исследуется задача робастной фильтрации для нелинейных систем с дискретным временем. Представлено явное решение задачи робастной фильтрации Калмана, основанное на уравнении Риккати.

В статье [13] предлагается робастный ЕКФ для нелинейных систем с дискретным временем с параметрическими неопределенностями, неизвестным входом и коррелированными шумами в объекте и измерениях. Предлагается расширенная модель для одновременной оценки неизвестных входных данных и состояний системы. Разработанный фильтр гарантирует верхнюю границу ковариации ошибок оценки. В [14] предложен подход построения нелинейной робастной оценки при наличии неизвестных входных данных и параметрических неопределенностей, основанный на минимизации верхней границы оценки.

В настоящей работе рассматривается задача идентификации в дискретных системах с аддитивными возмущениями со случайными и интервальными параметрами, причем в математическом описании случайных параметров постоянная составляющая может быть задана с интервальной неопределенностью. Для интерпретации интервальной неопределенности в работе используется метод, в основе которого лежит вероятностный подход, заключающийся в том, что интервальный параметр заменяется равномерно распределенной случайной величиной [15].

1. Постановка задачи

Рассмотрим линейную дискретную систему с неизвестными постоянными параметрами, модель которой дополнительно зависит от некоторых случайных и интервальных параметров. В этом случае систему можно представить в виде следующего разностного уравнения:

$$x(k+1) = (\tilde{A}(\theta) + \sum_{i=1}^m B_i \xi_i(k))x(k) + f(k) + q(k), \quad x(0) = x_0, \quad (1)$$

где $x(k) \in R^n$ – вектор состояния в момент времени k , $f(k)$ – известный входной вектор; x_0 – случайный вектор с известным математическим ожиданием и ковариационной матрицей $N_0 = M\{(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T\}$; $\theta \in R^p$ – вектор неизвестных постоянных параметров, подлежащих определению; $\tilde{A}(\theta)$ – матрица переходов состояний с интервальной неопределенностью (с заданными нижними и верхними границами $\underline{A}(\theta)$ и $\bar{A}(\theta)$ соответственно), $\xi(k) \in R^m$ – случайные возмущения со

следующими характеристиками: $M\{\xi(k)\} = 0$, $M\{\xi(k)\xi^T(j)\} = I_m \delta_{kj}$, $q(k) \in R^{n_1}$ – случайные возмущения со следующими характеристиками: $M\{q(k)\} = 0$, $M\{q(k)q^T(j)\} = Q\delta_{kj}$, Q – неотрицательно определенная матрица, I_m – единичная матрица размерности $m \times m$, δ_{kj} – символ Кронекера.

Модель наблюдений имеет вид:

$$y(k) = Sx(k) + v(k), \quad (2)$$

где $y(k) \in R^l$ – вектор наблюдений, S – матрица канала наблюдений, $v(k) \in R^l$ – шум наблюдений ($M\{v(k)\} = 0$, $M\{v(k)v^T(j)\} = V\delta_{kj}$).

Предполагается, что случайные последовательности $\xi(k)$, $q(k)$, $v(k)$ и вектор x_0 взаимно независимы, система (1) наблюдаема при параметрических возмущениях матрицы динамики $\tilde{A}$. Используя информацию, доступную к моменту времени $k \in [0; T]$, необходимо построить оценку параметров $\hat{\theta}(k)$ на основе минимизации следующего критерия:

$$J(0; T) = M\left\{\sum_{k=0}^T (\theta - \hat{\theta}(k))^T (\theta - \hat{\theta}(k))\right\}. \quad (3)$$

2. Алгоритм идентификации

Для решения задачи воспользуемся рекуррентным фильтром Калмана, при этом будем использовать вероятностный подход для нахождения его матрицы перехода. Суть метода заключается в том, что интервальные параметры заменяются независимыми случайными величинами, равномерно распределенными по интервалам неопределенности.

Затем, используя вероятностный подход, заменим неопределенные интервальные матрицы $\tilde{A}$ матрицами, элементы которых зависят от случайных величин

$$A(\theta, \psi) = (A(\theta) + \sum_{j=1}^{m_1} A_j \psi_j), \quad (4)$$

где ψ_j – независимые равномерно распределенные случайные величины на интервале $[-1, +1]$ ($-1 \leq \psi_j \leq 1$ ($j = \overline{1, m_1}$)). Будем считать, что случайные величины ψ_j независимы от x_0 , $\xi(k)$, $q(k)$ и $v(k)$. В (4) матрица $A(\theta) = \frac{1}{2}(\underline{A}(\theta) + \bar{A}(\theta))$, является медианой интервальной матрицы $\tilde{A}(\theta)$. Матрицы A_j , B_i ($i = \overline{1, m}$; $j = \overline{1, m_1}$) можно задать так, чтобы один элемент, соответствующий интервальному элементу матрицы $\tilde{A}(\theta)$, оставался ненулевым (если один и тот же интервальный элемент стоит на нескольких позициях матрицы $\tilde{A}(\theta)$, тогда матрицы A_j , B_i будут иметь несколько соответствующих ненулевых элементов). Значение матриц A_j , B_i можно определить по ширине интервала неопределенности элементов матрицы $\tilde{A}(\theta)$.

Тогда модель системы (1) примет вид:

$$x(k+1) = (A(\theta, \psi) + \sum_{i=1}^m B_i \xi_i) x(k) + f(k) + q(k), \quad x(0) = x_0. \quad (5)$$

Для построения оценки $\hat{\theta}(k)$ воспользуемся рекуррентным расширенным фильтром Калмана (ЕКФ). Построим оценки $\hat{x}(k)$ и $\hat{\theta}(k)$ при условии, что θ является неизвестным постоянным вектором.

В качестве динамической модели вектора θ можно использовать разностное уравнение

$$\theta(k+1) = \theta(k), \quad \theta(0) = \theta_0, \quad (6)$$

где θ_0 – случайная величина с математическим ожиданием $\bar{\theta}_0$ и дисперсионной матрицей P_{θ_0} .

Введем в рассмотрение расширенный вектор

$$X(k) = \begin{pmatrix} x(k) \\ \theta(k) \end{pmatrix}. \quad (7)$$

Уравнения динамики модели (1) с учетом (5) можно записать в виде:

$$X(k+1) = F(X(k), k) + \bar{q}(k), \quad (8)$$

где векторная функция с учетом (4) примет вид:

$$F(X(k), k) = \begin{pmatrix} (A(\theta(k))x(k) + f(k)) \\ \theta(k) \end{pmatrix}. \quad (9)$$

В (8) $\bar{q}(k)$ – последовательность случайных векторов, определяемая формулой

$$\bar{q}(k) = \begin{pmatrix} (\sum_{j=1}^{m_1} A_j \psi_j + \sum_{i=1}^m B_i \xi_i) x(k) + q(k) \\ 0 \end{pmatrix}. \quad (10)$$

Канал наблюдений примет вид:

$$y(k) = \bar{S}X(k) + v(k), \quad (11)$$

где $\bar{S} = [S \ 0]$.

Оценку вектора найдем с помощью рекуррентного алгоритма:

$$\begin{aligned} \hat{X}(k+1) &= \begin{pmatrix} \hat{x}(k+1) \\ \hat{\theta}(k+1) \end{pmatrix} = \begin{pmatrix} A(\hat{\theta}(k))\hat{x}(k) + f(k) \\ \hat{\theta}(k) \end{pmatrix} - K(k)[y(k+1) - \\ &- S(A(\hat{\theta}(k))\hat{x}(k) + f(k))], \quad \hat{X}(0) = X_0 = \begin{pmatrix} \bar{x}_0 \\ \bar{\theta}_0 \end{pmatrix}. \end{aligned} \quad (12)$$

Коэффициенты матрицы передачи фильтра равны

$$K(k) = P_f(k) \bar{S}^T [\bar{S} P_f(k) \bar{S}^T + V(k)]^{-1}, \quad (13)$$

где

$$\begin{aligned} P_f(k) &= \bar{A}(k)P(k)\bar{A}^T(k) + \tilde{Q}(k), \\ \bar{A}(k) &= \frac{\partial F}{\partial X} \Big|_{X=\hat{X}(k)} = \begin{pmatrix} A(\hat{\theta}(k)) & \frac{\partial A(\theta)}{\partial \theta} \Big|_{\theta=\hat{\theta}(k)} \hat{x}(k) \\ 0 & I_p \end{pmatrix}. \end{aligned} \quad (14)$$

Дисперсионная матрица $P(k)$ определится из уравнения

$$\begin{aligned} P(k+1) &= (I_p - K(k)\bar{S})\bar{A}(k)P(k)\bar{A}^T(k)(I_p - K(k)\bar{S})^T + \\ &+ (I_p - K(k)\bar{S})\tilde{Q}(k)(I_p - K(k)\bar{S})^T + \\ &+ K(k)V(k)K(k)^T, \quad P(0) = P_0, \end{aligned} \quad (15)$$

где

$$\begin{aligned} \tilde{Q}(k) &= \sum_{i=1}^m (B_i P(k) B_i^T + B_i \hat{X}(k) \hat{X}^T(k) B_i^T) + \frac{1}{3} \sum_{j=1}^{m_1} (A_j P(k) A_j^T + A_j \hat{X}(k) \hat{X}^T(k) A_j^T) + \bar{Q}(k), \\ \bar{Q} &= \begin{pmatrix} Q & 0 \\ 0 & 0 \end{pmatrix}, \quad P_0 = \begin{pmatrix} N_0 & 0 \\ 0 & P_{\theta_0} \end{pmatrix}. \end{aligned}$$

Оценки параметров θ для дискретной системы с интервальными и случайными параметрами строятся с помощью алгоритма робастного расширенного фильтра Калмана (REKF) и определяются по рекуррентной схеме (12). Матричные коэффициенты перехода $K(k)$ рассчитываются по формуле (13), а матрица $P(k)$ определится из уравнения (15).

3. Результаты моделирования

Моделирование алгоритма идентификации выполнено для модели объекта, в которой интервальный параметр заменен на равномерно распределенную случайную величину ψ_1 на интервале $[-\gamma_1, +\gamma_1]$:

$$\begin{aligned}x_1(k+1) &= (0,8 - \theta_1)x_1(k) + (0,3 + \psi_1 + \xi_1(k))x_2(k) + f_1(k), \\x_2(k+1) &= \theta_1 x_1(k) + (1 - \theta_2)x_2(k) + f_2(k).\end{aligned}\quad (16)$$

В (16) θ_1, θ_2 – неизвестные параметры, подлежащие определению, $\xi_1(k)$ – гауссовская случайная последовательность с характеристиками

$$M\{\xi_1(k)\} = 0 \text{ и } M\{\xi_1(k)\xi_1(j)\} = \gamma_z^2 \delta_{kj}.$$

Модель (16) может быть представлена в векторно-матричном виде (5) с матрицами

$$\begin{aligned}A &= \begin{pmatrix} 0,65 & 0 \\ 0,35 & 0,9 \end{pmatrix}, S = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, A_1 = \begin{pmatrix} 0 & \gamma_2 \\ 0 & 0 \end{pmatrix}, B_1 = \begin{pmatrix} 0 & \gamma_1 \\ 0 & 0 \end{pmatrix}, \\Q &= \begin{pmatrix} 14 & 0 \\ 0 & 15 \end{pmatrix}, V = \begin{pmatrix} 2,5 & 0 \\ 0 & 1,5 \end{pmatrix}, f(k) = \begin{pmatrix} 0,2 \\ 0 \end{pmatrix}.\end{aligned}\quad (17)$$

Начальные условия:

$$x(0) = \begin{pmatrix} 100 \\ 250 \end{pmatrix}, \hat{x}(0) = \begin{pmatrix} 96 \\ 245 \end{pmatrix}, N_0 = \begin{pmatrix} 10 & 0 \\ 0 & 10 \end{pmatrix}, P_{\theta_0} = \begin{pmatrix} 10 & 0 \\ 0 & 10 \end{pmatrix}.$$

Матрица S канала наблюдений (2) принималась равной I_2 . При моделировании точные значения параметров следующие:

$$\theta_1 = 0,15; \quad \theta_2 = 0,95.$$

Качество идентификации оценивалось посредством сравнения стандартных ошибок отклонений

$$\sigma_i = \sqrt{\frac{\sum_{k=1}^T (\theta_i(k) - \hat{\theta}_i(k))^2}{T-1}}, \quad (i=1,2) \quad (18)$$

для следующих алгоритмов:

- (EKF) используется расширенный фильтр Калмана для системы, учитывающей номинальные значения параметров матрицы A (интервальная неопределенность и случайные параметры не учитываются);
- (REKF) реализуется предложенный фильтр (6) с коэффициентами перехода $K(k)$, которые рассчитываются по формуле (12).

Стандартные ошибки отклонений оценок параметров для первой и второй компонент σ_1 и σ_2 приведены в таблице для методов EKF и REKF. Моделирование выполнено при $T = 100$ с усреднением по 1 000 реализациям.

Результаты сравнения стандартных ошибок отклонений оценок σ_i вектора параметров

№ п/п	γ_1	γ_2	i номер компоненты	EKF	REKF
				σ_i	σ_i
1	0,1	0,1	1	0,073	0,050
			2	0,141	0,138
2	0,25	0,25	1	0,266	0,102
			2	0,189	0,145
3	0,5	0,1	1	0,362	0,140
			2	0,209	0,149
4	0,1	0,5	1	0,257	0,111
			2	0,183	0,145
5	0,5	0,5	1	0,517	0,154
			2	0,265	0,153

На рис. 1 приведены реализации оценок параметров для алгоритмов ЕКФ и РЕКФ при $\gamma_1 = 0,5$ и $\gamma_2 = 0,5$.

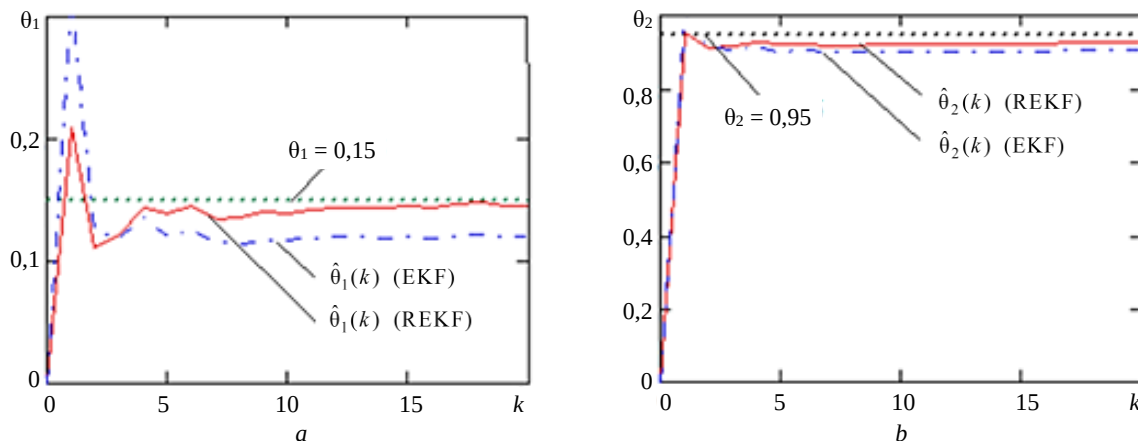


Рис. 1. Оценки параметров для алгоритмов РЕКФ (сплошная линия), ЕКФ (штрих-пунктирная линия), точные значения (пунктирная линия); а – оценки θ_1 , б – оценки θ_2

Fig. 1. Parameter estimates for REKF algorithm (solid line), EKF (dashed line), exact values (dotted line), a) estimates θ_1 , b) estimates θ_2

Из таблицы видно, что алгоритм РЕКФ, учитывающий случайные параметры и интервальную неопределенность параметров в матрице A , имеет преимущество в точности по сравнению расширенным фильтром Калмана, построенным по номинальным значениям параметров (ЕКФ), при этом с ростом параметров γ_1 и γ_2 , определяющих интенсивность случайного процесса $\xi_1(k)$ и величину интервала неопределенности параметра ψ_1 , преимущество алгоритма РЕКФ над алгоритмом ЕКФ увеличивается.

Заключение

В статье предложен алгоритм идентификации для дискретного объекта со случайными и интервальными параметрами. Случайные параметры в модели задаются как гауссовские случайные последовательности. Для представления интервальных параметров используется вероятностный подход, в основе которого лежит замена параметров интервального типа независимыми случайными величинами с равномерным распределением. Оценки строятся на основе рекуррентных схем калмановской фильтрации.

Результаты моделирования показали, что совместный учет в модели случайных и интервальных параметров (алгоритм РЕКФ) позволяет повысить точность идентификации по сравнению с классическим алгоритмом идентификации, построенным с использованием расширенного фильтра Калмана (алгоритм ЕКФ), рассчитанного по номинальным значениям параметров модели.

Список источников

1. Astrom K.J. Introduction to stochastic control theory. New York : Academic Press, 1970.
2. Sage A.P., Melsa J.L. Estimation Theory with Applications to Communications and Control. McGraw-Hill, 1971.
3. Julier S.J., Uhlmann J.K., Durrant-Whyte H. A new approach for filtering nonlinear systems // Proc. of the American Control Conference. Seattle, WA. 1995. P. 1628–1632.
4. Särkkä S. On unscented Kalman filter for state estimation of continuous-time nonlinear systems // IEEE Trans. Automat. Control. 2007. V. 52 (9). P. 1631–1641. doi: 10.1109/TAC.2007.904453
5. Arasaratnam I., Haykin S. Cubature Kalman filters // IEEE Trans. Automat. Control. 2009. V. 54 (6). P. 1254–1269. doi: 10.1109/TAC.2009.2019800
6. Цыганова Ю.В., Куликова М.В. Об эффективных методах параметрической идентификации линейных дискретных стохастических систем // Автоматика и телемеханика. 2012. № 6. С. 34–51.
7. Poljak B.T., Tsytkin Ja.Z. Robust Identification // Automatica. 1980. V. 16. P. 53–63.
8. Fang H., Haile M.A., Wang Y. Robust Extended Kalman Filtering for Systems with Measurement Outliers // Transactions on Control Systems Technology. 2021. V. 99. P. 1–8.

9. Смагин В.И., Ким К.С. Идентификация с использованием расширенного экстраполятора Калмана при аномальных ошибках в наблюдениях // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 68. С. 92–99. doi: 10.17223/19988605/68/10
10. Einicke G.A., Langford B., White L.B. Robust Extended Kalman Filtering // IEEE Transactions on Signal Processing, 1999. V. 47 (9). P. 2596–2599.
11. Souto R.F., Ishihara J.Y., Borges G.A. A robust extended Kalman filter for discrete-time systems with uncertain dynamics, measurements and correlated noise // 2009 American Control Conference, 10–12 June 2009, St. Louis, MO, USA. P. 1–6. doi: 10.1109/ACC.2009.5160604
12. Xiong K., Wei C.L., Liu L.D. Robust Kalman filtering for discrete-time nonlinear systems with parameter uncertainties // Aerospace Science and Technology. 2012. V. 18, is. 1. P. 15–24.
13. Asgari M., Khaloozadeh H. Robust extended Kalman filtering for non-linear systems with unknown input: a UBB model approach // IET Radar Sonar Navig. 2020. V. 14, is. 11. P. 1837–1844. doi: 10.1049/iet-rsn.2020.0258www.ietdl.org
14. Jahanian M., Ramezani A., Moarefianpour A., Shoorehdeli M.A. Robust extended Kalman filtering for nonlinear systems in the presence of unknown inputs and correlated noises // Optimal Control of Networked and Transportation Systems. 2022. V. 43, is. 1. P. 243–256. doi: 10.1002/oca.2786
15. Barmish B.R., Polyak B.T. A new approach to open robustness problems based on probabilistic predication formulae // Proc. 13th World IFAC Congr. 30 June – 5 July. San Francisco, USA. 1996. V. H. P. 1–6.

References

1. Astrom, K.J. (1970) *Introduction to Stochastic Control Theory*. New York: Academic Press.
2. Sage, A.P. & Melsa, J.L. (1971) *Estimation Theory with Applications to Communications and Control*. McGraw-Hill.
3. Julier, S.J., Uhlmann, J.K. & Durrant-Whyte, H. (1995) A new approach for filtering nonlinear systems. In: *Proceedings of the American Control Conference*. Seattle. WA. pp. 1628–1632.
4. Särkkä, S. (2007) On unscented Kalman filter for state estimation of continuous-time nonlinear systems. *IEEE Trans. Automat. Control*. 52(9). pp. 1631–1641. DOI: 10.1109/TAC.2007.904453
5. Arasaratnam, I. & Haykin, S. (2009) Cubature Kalman filters. *IEEE Trans. Automat. Control*. 54(6). pp. 1254–1269. DOI: 10.1109/TAC.2009.2019800
6. Tsyganova, Yu.V. & Kulikova, M.V. (2012) Ob effektivnykh metodakh parametriceskoy identifikatsii lineynykh diskretnykh stokhasticheskikh sistem [On effective methods of parametric identification of linear discrete stochastic systems]. *Avtomatika i Telemekhanika*. 6. pp. 34–51.
7. Poljak, B.T. & Tsytkin, Ja.Z. (1980) Robust Identification. *Automatica*. 16. pp. 53–63.
8. Fang, H., Haile, M.A. & Wang, Y. (2021) Robust Extended Kalman Filtering for Systems with Measurement Outliers. *Transactions on Control Systems Technology*. 99. pp. 1–8.
9. Smagin, V.I. & Kim, K.S. (2024) Identification using the extended Kalman extrapolator in case of anomalous errors in observations. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 68. pp. 92–99. DOI: 10.17223/19988605/68/10
10. Einicke, G.A., Langford, B. & White, L.B. (1999) Robust Extended Kalman Filtering. *IEEE Transactions On Signal Processing*. 47(9). pp. 2596–2599.
11. Souto, R.F., Ishihara, J.Y. & Borges, G.A. (2009) A robust extended Kalman filter for discrete-time systems with uncertain dynamics, measurements and correlated noise. *Proceedings 2009 American Control Conference*. 10–12 June 2009. St. Louis, MO, USA. pp. 1–6. DOI: 10.1109/ACC.2009.5160604
12. Xiong, K., Wei, C.L. & Liu, L.D. (2012) Robust Kalman filtering for discrete-time nonlinear systems with parameter uncertainties. *Aerospace Science and Technology*. 18(1). pp. 15–24.
13. Asgari, M. & Khaloozadeh, H. (2020) Robust extended Kalman filtering for non-linear systems with unknown input: a UBB model approach. *IET Radar Sonar Navig*. 14(11). pp. 1837–1844. DOI: 10.1049/iet-rsn.2020.0258www.ietdl.org
14. Jahanian, M., Ramezani, A., Moarefianpour, A. & Shoorehdeli, M.A. (2022) Robust extended Kalman filtering for nonlinear systems in the presence of unknown inputs and correlated noises. *Optimal Control of Networked and Transportation Systems*. 43(1). pp. 243–256. DOI:10.1002/oca.2786
15. Barmish, B.R. & Polyak, B.T. (1996) A new approach to open robustness problems based on probabilistic predication formulae. *Proceedings 13th World IFAC Congr*. 30 June – 5 July. vol. H. San Francisco. USA. pp. 1–6.

Информация об авторах:

Смагин Валерий Иванович – профессор, доктор технических наук, профессор кафедры прикладной математики Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: vsm@mail.tsu.ru

Ким Константин Станиславович – кандидат физико-математических наук, доцент Высшей IT школы Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: kks93@rambler.ru

Вклад авторов: все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

Information about the authors:

Smagin Valery I. (Doctor of Technical Science, Professor, National Research Tomsk State University, Tomsk, Russian Federation).
E-mail: vsm@mail.tsu.ru

Kim Konstantin S. (Candidate of Physical and Mathematical Sciences, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: kks93@rambler.ru

Contribution of the authors: the authors contributed equally to this article. The authors declare no conflicts of interests.

Поступила в редакцию 12.01.2025; принята к публикации 02.06.2025

Received 12.01.2025; accepted for publication 02.06.2025

ИНФОРМАТИКА И ПРОГРАММИРОВАНИЕ

INFORMATICS AND PROGRAMMING

Научная статья

УДК 621.382

doi: 10.17223/19988605/71/12

Реализация на ПЛИС и сравнительный анализ вычислителей сигмоида, работающих с полным диапазоном аргумента с учетом симметрии**Инна Владимировна Ушенина***Пензенский государственный технологический университет, Пенза, Россия, ivl23@yandex.ru*

Аннотация. Реализованы на ПЛИС и проанализированы две схемы вычислителей функции сигмоида, одна из которых использует симметрию функции, а вторая работает с полным диапазоном аргумента. Каждая из схем реализована при разрядности от 7 до 10 бит. Показано, что по сравнению со второй схемой первая расходует меньшее количество ресурсов. Однако в первой схеме используются вспомогательные блоки с длинными цепями переноса, которые существенно удлиняют критический путь распространения сигнала и таким образом влияют на время вычислений. По сравнению с автоматическим размещением на кристалле элементов вспомогательных блоков, принадлежащих критическому пути, их ручное размещение строго по порядку сокращает время работы блоков на величину до 25%, а время работы первой схемы в целом – до 17%. Но в любом случае вторая схема будет работать как минимум вдвое быстрее первой.

Ключевые слова: сигмоид; нейронная сеть; ПЛИС; метод поразрядного отображения; симметрия; критический путь.

Благодарности: Исследование выполнено за счет гранта Российского научного фонда и Пензенской области № 24-21-20100, <https://rscf.ru/project/24-21-20100/>

Для цитирования: Ушенина И.В. Реализация на ПЛИС и сравнительный анализ вычислителей сигмоида, работающих с полным диапазоном аргумента с учетом симметрии // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2025. № 71. С. 120–129. doi: 10.17223/19988605/71/12

Original article

doi: 10.17223/19988605/71/12

FPGA implementation and comparative analysis of sigmoid calculators processing the full argument range in view of symmetry**Inna V. Ushenina***Penza State Technological University, Penza, Russian Federation, ivl23@yandex.ru*

Abstract. This paper implements on FPGA and analyses two sigmoid calculation circuits: the first circuit uses the function's symmetry, and the second one processes the full argument range. Each circuit is implemented with a bit width of 7 to 10 bits. It is shown that, compared to the second circuit, the first one consumes a smaller amount of resources. But the first circuit uses auxiliary blocks containing long carry chains. sigmoid function; neural network; FPGA; bit-level mapping method; symmetry; critical path.

Keywords: sigmoid function; neural network; FPGA; bit-level mapping method; symmetry; critical path.

Acknowledgments: The research was carried out with the financial support of the Russian Science Foundation and Penza region under Grant № 24-21-20100 <https://rscf.ru/project/24-21-20100/>

For citation: Ushenina, I.V. (2025) FPGA implementation and comparative analysis of sigmoid calculators processing the full argument range in view of symmetry. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 71. pp. 120–129. doi: 10.17223/19988605/71/12

Введение

Аппаратная реализация нейронных сетей необходима в тех случаях, когда они должны работать в реальном времени, а также когда требуется ускорение их работы или обучения. Среди аппаратных платформ, подходящих для реализации нейронных сетей – графических процессоров, заказных микросхем, ПЛИС – последние могут проигрывать другим платформам в производительности и энергоэффективности, но являются наиболее доступным вариантом и имеют возможность перепрограммирования [1–3].

Нейронные сети, обучающиеся по алгоритму обратного распространения ошибки, используют в качестве функций активации сигмоид

$$s(x) = \frac{1}{1 + e^{-x}}. \quad (1)$$

или гиперболический тангенс, которые имеют подходящую форму, непрерывно дифференцируемы и требуют небольших дополнительных аппаратных затрат на вычисление производных в ходе обучения.

Современные ПЛИС не позволяют напрямую вычислять частное и экспоненту. Для вычисления на ПЛИС сигмоида и гиперболического тангенса к настоящему моменту предложен ряд методов, среди которых – метод поразрядного отображения [4]. Он заключается в том, что каждый разряд двоичного представления значения функции отделяется от других и вычисляется как булева функция от разрядов двоичного представления аргумента. Это позволяет выполнять вычисления относительно быстро, без дополнительной погрешности и использовать для вычислений блоки программируемой логики, имеющиеся на ПЛИС в большом количестве [5].

Дополнительным преимуществом функций сигмоида и гиперболического тангенса, актуальным при их аппаратной реализации, является симметрия. Так, $s(x)$ при отрицательных аргументах благодаря симметрии можно вычислить по формуле

$$s(x) = 1 - s(|x|). \quad (2)$$

В [4], где предложен метод поразрядного отображения, и в [5–9], использующих этот метод для вычисления функций сигмоида или гиперболического тангенса, описаны вычисления только при неотрицательных аргументах. В [4, 6], где вычисляется $s(x)$, отрицательные аргументы предлагается обрабатывать согласно (2). При этом в [4] схема вычислений по формуле (2) не описывается, а в [6] представлена только функциональная схема, без описания ее реализации.

Дополнение вычислителей отдельных разрядов значений $s(x)$, работающих только с неотрицательными аргументами, вычислениями по (2) может привести к существенному снижению скорости вычислений. Это связано с тем, что аппаратная реализация сложения и вычитания требует переноса битов переполнения. Перенос выполняют цепочки комбинационной логики, длина которых соответствует разрядности обрабатываемых чисел. Вместе с вычислителем младшего разряда значения $s(|x|)$ цепь переноса будет образовывать критический путь, время прохождения сигнала по которому будет влиять на скорость вычислителя $s(x)$ в целом. Также, поскольку отрицательные аргументы представлены в дополнительном коде, их представления необходимо перевести в прямой код, что требует установки еще одного сумматора и приводит к удлинению критического пути еще одной цепью переноса.

Существует возможность выполнения вычислений по формуле (2) на сумматорах блоков цифровой обработки сигналов (ЦОС) ПЛИС, которые работают значительно быстрее блоков программируемой логики. Однако блоки ЦОС обычно выполняют линейную часть вычислений в нейронах, а именно суммирование произведений входных сигналов нейронов и масштабирующих коэффициентов. То есть

придется вводить временное разделение использования блоков ЦОС и усложнять схему управления нейронной сетью. Кроме того, в слоях нейронных сетей, реализуемых аппаратно, часто есть параллелизм на уровне нейронов, но его нет на уровне синапсов. Это означает, что каждый нейрон в слое представлен собственным блоком ЦОС, на который поступают по очереди сигналы от нейронов предыдущего слоя [10]. В таких случаях вычислитель функции активации устанавливается общим для всех нейронов слоя, и он отделен от блоков ЦОС регистром сдвига [11] или мультиплексором [12].

Способом избежать удлинения критического пути может стать отказ от использования симметрии функции активации и вычислений по (2) и работа методом поразрядного отображения с полным диапазоном аргумента.

Настоящая работа является продолжением [5], где метод поразрядного отображения использован для реализации на блоках программируемой логики ПЛИС схем вычислителей отдельных разрядов значений $s(x)$. Эти схемы работают с аргументами в диапазоне $[0; 8)$ и предполагают использование симметрии сигмоида при обработке отрицательных аргументов. В настоящей работе представлены и проанализированы две схемы вычислителей значений $s(x)$, использующие схемы вычислителей их отдельных разрядов, полученные в [5]. Первая схема использует симметрию функции и имеет длинный критический путь, образованный цепями переноса. Вторая схема не использует симметрию и содержит удвоенный набор вычислителей отдельных разрядов значений $s(x)$: для работы с неотрицательными и отрицательными аргументами.

Представленные схемы работают с числами в формате с фиксированной запятой с разрядностью от 7 до 10 бит. Разрядность 7–8 бит достаточна при развертывании обученной нейронной сети, а разрядность 9–10 бит – при обучении нейронной сети [13]. Разрядности представлений аргументов и значений функции сигмоида во всех случаях равны. При этом аргумент находится в диапазоне $[-8; 8)$, и из всех представляющих его разрядов один является знаковым, три отводятся на представление целой части, а все остальные разряды – на представление дробной части. При указании разрядности аргумента знаковый разряд не учитывается. Значение $s(x)$ находится в диапазоне $(0; 1)$, и все представляющие его разряды относятся к дробной части. Обе схемы используют только блоки программируемой логики ПЛИС, оставляя ресурсы блоков ЦОС для линейной части вычислений в нейронах. Схемы описаны на языке VHDL и реализованы на ПЛИС GW2AR от Gowin Semiconductor [14]. Разработка, реализация и анализ схем выполнены в среде Gowin EDA.

1. Схемы вычислителей функции сигмоида

Метод поразрядного отображения предполагает представление отдельных разрядов значений $s(x)$ в виде булевых функций разрядов ее аргумента. Для их получения сначала формируется полный список аргументов и соответствующих им значений $s(x)$, представленных в двоичном коде с заданной разрядностью. На основе этого списка для каждого разряда $s(x)$ формируются таблицы истинности. Вычислители разрядов $s(x)$ можно реализовать непосредственно по данным таблицам. В таком случае вычислители должны отличать каждую из комбинаций разрядов x , расходуя на это максимальное количество ресурсов.

По таблицам истинности для расчета разрядов $s(x)$ можно получить булевы функции, представленные в совершенной дизъюнктивной нормальной форме, а затем минимизировать их. В [5] и в настоящей работе минимальные дизъюнктивные нормальные формы (МДНФ) булевых функций для расчета разрядов $s(x)$ получены с использованием алгоритма Куайна–МакКласки.

Значения старших разрядов $s(x)$ часто совпадают для целых групп соседних значений x ; особенно это выражено для $s(x)$, по модулю близких к нулю и единице. За счет этого свойства представление булевых функций для расчета старших разрядов $s(x)$ в МДНФ сводится к сумме нескольких простых импликант, вычисление которой требует значительно меньшего количества ресурсов, чем работа с таблицей истинности.

По мере продвижения к младшим разрядам $s(x)$ булевы функции в МДНФ содержат все большее количество импликант и литералов в каждой из них, что в конечном итоге сравнивает сложность их вычисления с вычислениями по таблицам истинности [5]. В схемах, представленных в данной работе,

от двух до четырех младших разрядов значений $s(x)$ вычисляются по таблицам истинности, остальные разряды – по булевым функциям, представленным в МДНФ.

1.1. Схема вычислителя функции сигмоида, использующая ее симметрию (схема 1)

Структурная схема вычислителя функции сигмоида, использующего ее симметрию, представлена на рис. 1. Схема содержит, помимо набора описанных в [5] вычислителей отдельных разрядов значений функции сигмоида $s_i(|x|)$, два вспомогательных блока, на которые поступает знаковый разряд аргумента $sign$. Первый блок принимает от линейной части нейрона аргумент сигмоида, представленный в дополнительном коде $x_{доп}$ и формирует представление $|x|$: если аргумент неотрицательный, он остается без изменений; если отрицательный – он переводится в прямой код, т.е. инвертируется и складывается с единицей в младшем разряде. Второй блок при $x \geq 0$ приравнивает $s(x)$ к $s(|x|)$, а при $x < 0$ вычисляет $s(x)$ согласно (2): переводит $-s(|x|)$ в дополнительный код и складывает результат с единицей.

Сумматоры, необходимые во вспомогательных блоках, можно реализовать на базе программируемого примитива арифметико-логического устройства ALU, предлагаемого в [15]. Однако ALU содержит многочисленные настройки, которые не будут использоваться, но займут ресурсы и удлинит критический путь.

С учетом сказанного схемы вспомогательных блоков были построены на базе табличных преобразователей (look-up tables, LUTs) в строгом соответствии с решаемой задачей. LUT работают по таблицам истинности, составленным с учетом особенностей выполняемых операций (диапазоны аргумента и значения $s(x)$, обработка только одного неизвестного слагаемого при сложении, необходимость операции инверсии). Таким образом, функции, требуемые от вспомогательных блоков, экономно реализованы на небольшом количестве LUT.

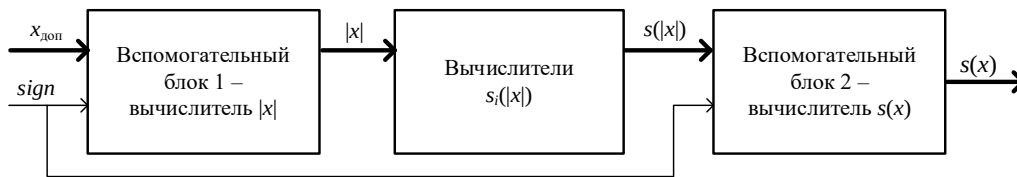


Рис. 1. Структурная схема вычислителя функции сигмоида, использующего ее симметрию

Fig. 1. Block diagram of the sigmoid function calculator that uses the function's symmetry

Вспомогательные блоки представляют собой последовательно соединенные звенья с одним или двумя LUT, у которых может быть два или три входа (LUT2, LUT3). В звеньях с двумя LUT каждый из них принимает один и тот же набор сигналов. Схема 8-разрядного вычислителя $s(x)$ с вспомогательными блоками на базе LUT представлена на рис. 2.

Во вспомогательном блоке 1 (на рис. 2 слева) каждое звено выполняет при $sign = 1$ инверсию одного из разрядов $x_{доп}$, формируя из него один из разрядов представления модуля x в прямом коде $-x_{i пр}$. Нижнее звено, обрабатывающее младший разряд аргумента, $x_{0 доп}$, при $sign = 1$ складывает его с единицей и выдает $x_{0 пр}$ и бит переполнения $carry_1$ на следующее звено. Остальные звенья при $sign = 1$ обрабатывают бит переполнения $carry_i$, полученный от предыдущего звена, и $x_{i доп}$, формируя бит переполнения для следующего звена $carry_{i+1}$ и $x_{i пр}$. Если $sign = 0$, $x_{i доп} = x_{i пр}$ и $carry_{i+1} = 0$. Нижнее и промежуточные звенья содержат по два LUT, один из которых формирует $x_{i пр}$, а другой – $carry_{i+1}$. Верхнее звено содержит один LUT3 и формирует $x_{7 пр}$.

Полученные разряды $x_{7 пр} \dots x_{0 пр}$ поступают на вычислители $s_i(|x|)$, описанные в [5]. Как и в [5], $s_8(|x|)$ соответствует младшему разряду $s(|x|)$. Также $s_8(x)$ соответствует младшему разряду $s(x)$.

Знаковый разряд аргумента $sign$ и $s_i(|x|)$ поступают на звенья вспомогательного блока 2 – вычислителя $s(x)$ по формуле (2), изображенного на рис. 2 справа. Звено, формирующее старший разряд $s_1(x)$, работает как инвертор знакового разряда $sign$, так как при отрицательном аргументе $0 < s(x) < 0,5$, а при неотрицательном – $0,5 \leq s(x) < 1$. При $sign = 0$ остальные разряды $s(x)$ приравниваются соответствующим разрядам $s(|x|)$, а все биты переноса – нулю. При $sign = 1 - s(|x|)$ переводится в дополнительный код

и складывается с единицей, согласно (2). Звено, формирующее разряд $s_2(x)$, не выдает бита переноса и состоит из одного LUT3. Остальные звенья содержат пару LUT, один из которых формирует бит переноса $carry_{i-1}$, а другой – разряд $s_i(x)$. Все звенья, кроме формирующего $s_1(x)$, связаны между собой цепями переноса $carry_2$ – $carry_7$.

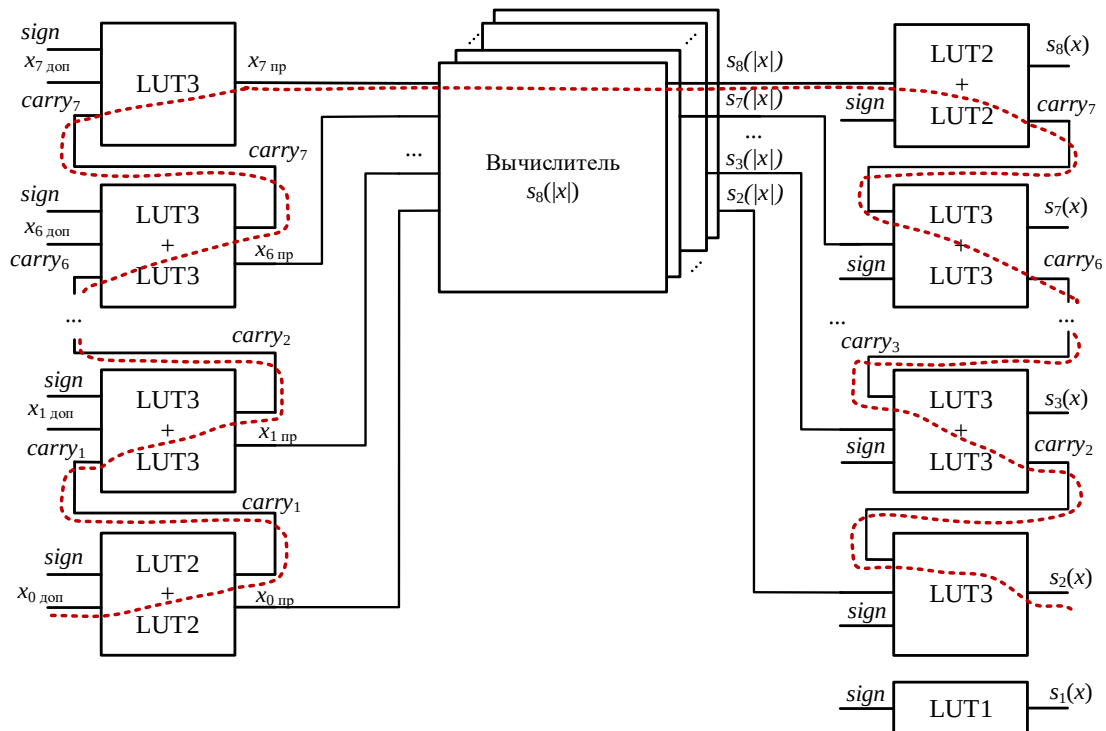


Рис. 2. Схема 1. 8-разрядный вычислитель функции сигмоида, использующий ее симметрию.

Вспомогательные блоки выполнены на табличных преобразователях. Критический путь показан штриховой линией

Fig. 2. Circuit 1. 8-bit calculator of the sigmoid function that uses the function's symmetry.

The auxiliary blocks are implemented on the LUTs. The critical path is shown by the dashed line

В обоих вспомогательных блоках LUT, формирующие x_i пр и $s_i(x)$, реализуют одинаковые таблицы истинности. То же касается LUT, формирующих биты переноса в обоих вспомогательных блоках. Таблицы истинности, реализуемые LUT во вспомогательных блоках 1 и 2, представлены в табл. 1, 2.

На рис. 2 показано, что цепи переноса вспомогательных блоков вместе с вычислителем разряда $s_8(|x|)$ образуют критический путь. При других разрядностях x и $s(x)$ схема 1 строится аналогично, отличия будут только в количествах вычислителей $s_i(|x|)$ и звеньев вспомогательных блоков. Таким образом, время вычисления $s(x)$ в схеме 1 определяется критическим путем, содержащим две цепи переноса, количество звеньев в которых соответствует разрядности x и $s(x)$, и вычислитель младшего разряда значения $s(|x|)$.

Таблица 1

Таблицы истинности LUT2 и LUT3 вспомогательного блока 1

Входы			Выходы $carry_{i+1}$ (LUT3)	Выходы x_i пр (LUT3)	Выход $carry_1$ (LUT2)	Выход x_0 пр (LUT2)
$carry_i$ (только LUT3)	$sign$	x_i доп				
0	0	0	0	0	0	0
0	0	1	0	1	0	1
0	1	0	0	1	1	0
0	1	1	0	0	0	1
1	0	0	0	0	—	—
1	0	1	0	1	—	—
1	1	0	1	0	—	—
1	1	1	0	1	—	—

Таблицы истинности LUT2 и LUT3 вспомогательного блока 2

Входы			Выходы $carry_{i-1}$ (LUT3)	Выходы $s_i(x)$ (LUT3)	Выход $carry_7$ (LUT2)	Выход $s_8(x)$ (LUT2)
$carry_i$ (только LUT3)	$sign$	$s_i(x)$				
0	0	0	0	0	0	0
0	0	1	0	1	0	1
0	1	0	0	1	1	0
0	1	1	0	0	0	1
1	0	0	0	0	–	–
1	0	1	0	1	–	–
1	1	0	1	0	–	–
1	1	1	0	1	–	–

1.2. Схема вычислителя функции сигмоида с двумя наборами вычислителей $s_i(x)$ (схема 2)

В схеме, представленной на рис. 3 (схема 2), аргумент обрабатывается одновременно двумя наборами вычислителей отдельных разрядов значений функции сигмоида. Первый набор, тот же что и в схеме 1, работает, предполагая $x \geq 0$; второй – предполагая $x < 0$. При этом вычислители первого набора работают с x , представленным в прямом коде, а вычислители второго набора – с x , представленным в дополнительном коде. То есть преобразователь представления x из дополнительного кода в прямой в данном случае не нужен. Также упраздняются и вычисления по формуле (2), вместо которых в схеме 2 устанавливается многоразрядный мультиплексор 2-1. На его входы подаются значения функции сигмоида $sp(x)$ и $sm(x)$, рассчитанные наборами вычислителей, предполагающими соответственно $x \geq 0$ и $x < 0$. Нужное значение выбирается знаковым разрядом аргумента $sign$, поступающим на управляющий вход мультиплексора.

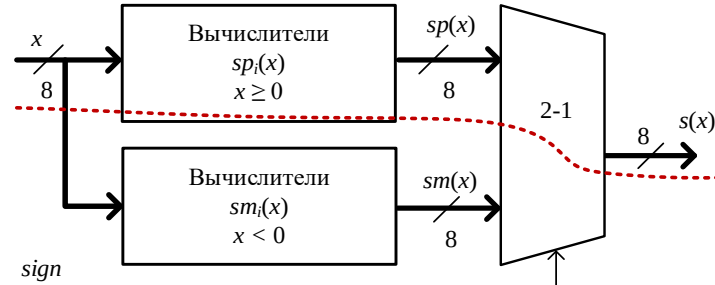


Рис. 3. Схема 2. 8-разрядный вычислитель функции сигмоида с двумя наборами вычислителей отдельных разрядов ее значений. Критический путь показан штриховой линией

Fig. 3. Circuit 2. 8-bit calculator of the sigmoid function with two sets of output bit calculators.

The critical path is shown by the dashed line

Таким образом, в схеме 2 критический путь составлен вычислителем одного из разрядов значения $s(x)$ и мультиплексором, – например, как показано на рис. 3.

2. Результаты реализации схем

Основные результаты реализации схем 1 и 2 при их разрядности 7–10 бит показаны в табл. 3–5. В табл. 3 представлена ресурсоемкость схем, выраженная в требуемом количестве основных ресурсов ПЛИС – 4-входовых табличных преобразователей (LUT4) и содержащих их блоков программируемой логики (configurable function unit, CFU). Видно, что схема 2 требует на 27–86% большего количества LUT4 по сравнению со схемой 1. Возрастание разницы в ресурсоемкости двух схем с повышением их разрядности связано с тем, что большая часть задействованных в схемах ресурсов используется вычислителями $s_i(|x|)$, $sm_i(x)$ и $sp_i(x)$. Булевы функции, реализуемые ими, с повышением разрядности усложняются и требуют все большего количества ресурсов.

Таблица 3

Ресурсоемкость схем 1 и 2

Разрядность, бит	Схема 1 (рис. 2)		Схема 2 (рис. 3)	
	CFU	LUT4	CFU	LUT4
7	8	59	10	75
8	13	98	19	148
9	24	189	41	326
10	52	414	97	771

Таблица 4

Временные характеристики схем 1 и 2

Разрядность, бит	Схема 1 (рис. 2)		Схема 2 (рис. 3)	
	$F_{\max}$, МГц	$t_{\text{кр}}$, нс	$F_{\max}$, МГц	$t_{\text{кр}}$, нс
7	80	10,005	155	2,893
8	69	11,03	130	3,723
9	66	12,948	108	4,748
10	58	15,171	80	6,924

Таблица 5

Время, затрачиваемое вспомогательным блоком 1 схемы 1 на вычисление $|x|$

Разрядность, бит	Автоматическое размещение			Ручное размещение (LUT цепей переноса организованы в столбцы и расположены строго по порядку)		
	Элементы, нс	Соед., нс	Сумма, нс	Элементы, нс	Соед., нс	Сумма, нс
7	4,444	1,168	5,612	4,152	0,033	4,186
8	4,999	0,959	5,958	4,893	0,036	4,929
9	5,564	0,454	6,018	5,368	0,261	5,629
10	5,842	2,091	7,933	5,584	0,267	5,851

В табл. 4 для схем 1 и 2 показаны максимальная тактовая частота $F_{\max}$ и время $t_{\text{кр}}$, затрачиваемое на вычисление $s(x)$. Время $t_{\text{кр}}$ включает в себя время работы всех находящихся на критическом пути логических элементов и задержки распространения сигналов по соединениям между ними. Из табл. 4 видно, что при разрядностях 7 и 8 бит переход от схемы 1 к схеме 2 повышает $F_{\max}$ почти в два раза, но при большей разрядности этот выигрыш уже не так велик. Это объясняется тем, что при разрядностях 7 и 8 бит вычислители младших разрядов значений $s(|x|)$ в схеме 1 относительно просты, и их вклад в $t_{\text{кр}}$ невелик по сравнению с вкладом вспомогательных блоков. Например, у 7-разрядной схемы 1 время вычисления $s_7(|x|)$ составляет 1,775 нс, а время работы вспомогательных блоков 1 и 2 – как минимум 4,186 и 3,42 нс соответственно. С повышением разрядности схем 1 и 2 вычислители $s_i(|x|)$, $sm_i(x)$ и $sp_i(x)$ усложняются, и время вычислений в них приближается к времени работы вспомогательных блоков. Тем не менее даже при разрядности 10 бит $t_{\text{кр}}$ в большей степени определяется вспомогательными блоками, и переход от схемы 1 к схеме 2 дает выигрыш по $F_{\max}$ 37%.

Из табл. 4 следует, что переход от схемы 1 к схеме 2 дает больший выигрыш по $t_{\text{кр}}$, чем по $F_{\max}$, и с повышением разрядности схем это все более заметно. Это объясняется тем, что для расчета $F_{\max}$ входные и выходные сигналы схем 1 и 2 синхронизируются. Пути распространения сигналов, в том числе используемые в расчетах критические пути, при этом дополняются соединениями схем с входными и выходными синхронизирующими регистрами. В случае перехода от схемы 1 к схеме 2 площадь, занимаемая схемой, увеличивается до 86%, поэтому соединения регистров с элементами схемы 2 являются более длинными.

Проектируемое устройство может быть размещено на ПЛИС автоматически, т.е. средой проектирования, или вручную. При ручном размещении проектируемого устройства место каждого его элемента выбирает разработчик. Ручное размещение может быть частичным, например выполняться только для критических путей. В табл. 4 для схемы 1 приведены параметры, полученные при ручном

размещении элементов (LUT) ее вспомогательных блоков (см. рис. 2). При автоматическом размещении, выполняемом средой проектирования, LUT располагаются хоть и недалеко друг от друга, но в произвольном порядке, что удлиняет критический путь. При ручном размещении LUT цепей переноса, образующие критический путь, были организованы в столбцы и расположены строго по порядку, что позволило минимизировать длину соединений между LUT и сократить критический путь.

В табл. 5 для примера приведены составляющие времени вычисления $|x|$ вспомогательным блоком 1 схемы 1 при автоматическом и ручном размещении элементов (LUT) критического пути, а именно: время работы элементов (Элементы, нс), задержка распространения сигналов по соединениям между элементами (Соед., нс) и сумма этих составляющих (Сумма, нс). Видно, что за счет ручного размещения сокращение времени вычисления $|x|$ может составить до 25%. Использование ручного размещения в обоих вспомогательных блоках схемы 1 позволяет сократить ее $t_{кр}$ на 5–17% по сравнению с результатами автоматического размещения. При этом ручное размещение вычислителей $s_i(|x|)$ не выполнялось: они не имеют единственного пути распространения сигналов значительно длиннее остальных, который следовало бы оптимизировать. Несколько путей могут иметь небольшие отличия во времени распространения, и укорочение одного пути при оптимизации может дать удлинение другого на примерно ту же величину.

Заключение

ПЛИС используются для реализации нейронных сетей при необходимости их ускорения или работы в реальном времени, и сохранение высокой скорости вычислений является в этих случаях актуальной задачей. Поэтому, если нет дефицита ресурсов, для вычисления сигмоидной функции активации следует выбирать схему 2, работающую методом поразрядного отображения с полным диапазоном аргумента функции. Схема 1, использующая симметрию функции сигмоида, проигрывает в скорости схеме 2 из-за двух цепей переноса в составе вспомогательных блоков, образующих совместно с вычислителем младшего разряда двоичного представления значения функции сигмоида длинный критический путь. Ручное размещение на кристалле элементов вспомогательных блоков, принадлежащих критическому пути, строго по порядку позволяет сократить время работы вспомогательных блоков максимум на 25% по сравнению результатом их автоматического размещения; при этом общее время вычисления сигмоида схемой 1 снижается на 5–17%. Однако даже при использовании ручного размещения элементов вспомогательных блоков схемы 1 она расходует на вычисления как минимум в два раза больше времени по сравнению со схемой 2.

Список источников

1. Nurvitadhi E., Venkatesh G., Sim J., Marr D., Huang R., Ong Gee Hock J., Boudoukh G. Can FPGAs beat GPUs in accelerating next-generation deep neural networks? // Proc. of the 2017 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays (FPGA'17), 22–24 February, Monterey, CA, USA. 2017. P. 5–14.
2. Шашев Д.В., Шатравин В.В. Реализация сигмоидной функции активации с помощью концепции перестраиваемых вычислительных сред // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2022. № 61. С. 117–127. doi: 10.17223/19988605/61/12
3. Шипицин С.П., Ямаев М.И. Развитие аппаратно-ориентированных нейронных сетей на FPGA и ASIC // Вестник Пермского национального исследовательского политехнического университета. Электротехника, информационные технологии, системы управления. 2019. № 31. С. 177–192.
4. Tommiska M.T. Efficient digital implementation of the sigmoid function for reprogrammable logic // IEE Proceedings-Computers and Digital Techniques. 2003. V. 150 (6). P. 403–411. doi: 10.1049/ip-cdt:20030965
5. Ушенина И.В. Реализация на современных ПЛИС вычислителя сигмоидной функции активации нейронных сетей табличным методом // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 69. С. 124–133. doi: 10.17223/19988605/69/13
6. Li X.J., Li L. IP core based hardware implementation of multi-layer perceptrons on FPGAs: a parallel approach // Advanced Materials Research. 2012. V. 433. P. 5647–5653. doi: 10.4028/www.scientific.net/AMR.433-440.56471

7. Yang T., Wei Y., Tu Z., Zeng H., Kinsy M.A., Zheng N., Ren P. Design Space Exploration of Neural Network Activation Function Circuits // *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*. 2019. V. 38 (10). P. 1974–1978. doi: 10.1109/TCAD.2018.2871198
8. Rajput G., Raut G., Chandra M., Vishvakarma S.K. VLSI implementation of transcendental function hyperbolic tangent for deep neural network accelerators // *Microprocessors and Microsystems*. 2021. V. 84. Art. 104270. doi: 10.1016/j.micpro.2021.104270
9. Saranya S., Elango B. Implementation of PWL and LUT based approximation for hyperbolic tangent activation function in VLSI // *Proc. of 2014 International Conference on Communication and Signal Processing*, 03–05 April, Melmaruvathur, India. 2014. P. 1778–1782.
10. Omondi A.R., Rajapakse J.C. *FPGA implementations of neural networks*. New York : Springer, 2006.
11. Thasnimol V.S., George M.A. Hardware Accelerator Implementation of Multilayer Perceptron // *Proc. of Congress on Intelligent Systems (CIS 2020)*. 2021. V. 1. P. 107–119.
12. Vu H.M., Thang H.V. A Customized Hardware Architecture for Multi-layer Artificial Neural Networks on FPGA // *Proc. of Fourth International Conference on Information Systems Design and Intelligent Applications*, India. 2018. P. 637–644.
13. Holt J.L., Hwang J.N. Finite precision error analysis of neural network hardware implementations // *IEEE Transactions on Computers*. 1993. V. 42 (3). P. 281–290. doi: 10.1109/12.210171
14. GW2AR series of FPGA Products : Data Sheet. URL: <https://cdn.gowinsemi.com.cn/DS226E.pdf> (accessed: 10.09.2024).
15. Configurable Function Unit (CFU) : User Guide. URL: https://www.gowinsemi.com/upload/database_doc/1777/document/62e351486e153.pdf (accessed: 10.09.2024).

References

1. Nurvitadhi, E., Venkatesh, G., Sim, J., Marr, D., Huang, R., Ong Gee Hock, J. & Boudoukh, G. (2017) Can FPGAs beat GPUs in accelerating next-generation deep neural networks? *Proceedings of the 2017 ACM/SIGDA international symposium on field-programmable gate arrays (FPGA '17)*. 22–24 February. Monterey California. USA. pp. 5–14.
2. Shashev, D.V. & Shatravin, V.V. (2022) Implementation of the sigmoid activation function using the reconfigurable computing environments. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 61. pp. 117–127. DOI: 10.17223/19988605/61/12
3. Shipitsin, S.P. & Iamaev, M.I. (2019) Hardware neural networks progress on FPGA and ASIC. *Vestnik Permskogo natsional'nogo issledovatel'skogo politekhnicheskogo universiteta. Elektrotehnika, informatsionnye tekhnologii, sistemy upravleniya*. 31. pp. 177–192.
4. Tommiska, M.T. (2003) Efficient digital implementation of the sigmoid function for reprogrammable logic. *IEEE Proceedings-Computers and Digital Techniques*. 150(6). pp. 403–411. DOI: 10.1049/ip-cdt:20030965
5. Ushenina, I.V. (2024) Implementation of the sigmoid activation function for neural networks on current FPGAs by the table-driven method. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 69. pp. 124–133. DOI: 10.17223/19988605/69/13
6. Li, X.J. & Li, L. (2012) IP core-based hardware implementation of multi-layer perceptrons on FPGAs: a parallel approach. *Advanced Materials Research*. 433. pp. 5647–5653. DOI: 10.4028/www.scientific.net/AMR.433-440.5647
7. Yang, T., Wei, Y., Tu, Z., Zeng, H., Kinsy, M.A., Zheng, N. & Ren, P. (2019) Design Space Exploration of Neural Network Activation Function Circuits. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*. 38(10). pp. 1974–1978. DOI: 10.1109/TCAD.2018.2871198.
8. Rajput, G., Raut, G., Chandra, M., & Vishvakarma, S.K. (2021) VLSI implementation of transcendental function hyperbolic tangent for deep neural network accelerators. *Microprocessors and Microsystems*. 84. pp. 104270. DOI: 10.1016/j.micpro.2021.104270
9. Saranya S. & Elango B. (2014) Implementation of PWL and LUT based approximation for hyperbolic tangent activation function in VLSI. *Proceedings of 2014 International Conference on Communication and Signal Processing*. 03-05 April. Melmaruvathur. India. pp. 1778–1782.
10. Omondi, A.R. & Rajapakse, J.C. (2006) *FPGA implementations of neural networks*. New York: Springer.
11. Thasnimol, V.S. & George, M. (2021) A Hardware Accelerator Implementation of Multilayer Perceptron. *Proceedings of Congress on Intelligent Systems (CIS 2020)*. 1. pp. 107–119.
12. Vu, H.M. & Thang, H.V. (2018). A Customized Hardware Architecture for Multi-layer Artificial Neural Networks on FPGA. *Proceedings of Fourth International Conference on Information Systems Design and Intelligent Applications*. India. pp. 637–644.
13. Holt, J.L. & Hwang, J.N. (1993) Finite precision error analysis of neural network hardware implementations. *IEEE Transactions on Computers*. 42(3). pp. 281–290. DOI: 10.1109/12.210171
14. GW2AR series of FPGA Products. [Online] Available from: <https://cdn.gowinsemi.com.cn/DS226E.pdf>. (Accessed: 10th September 2024).
15. Configurable Function Unit (CFU) User Guide. [Online] Available from: https://www.gowinsemi.com/upload/database_doc/1777/document/62e351486e153.pdf. (Accessed: 10th September 2024).

Информация об авторе:

Ушенина Инна Владимировна – доцент, кандидат технических наук, доцент кафедры «Программирование» Пензенского государственного технологического университета (Пенза, Россия). E-mail: ivl23@yandex.ru

Автор заявляет об отсутствии конфликта интересов.

Information about the author:

Ushenina Inna V. (Candidate of Technical Sciences, Associate Professor, Penza State Technological University, Penza, Russian Federation). E-mail: ivl23@yandex.ru.

The author declares no conflicts of interests.

Поступила в редакцию 07.10.2024; принята к публикации 02.06.2025

Received 07.10.2024; accepted for publication 02.06.2025

**ПРОЕКТИРОВАНИЕ И ДИАГНОСТИКА
ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ****DESIGNING AND DIAGNOSTICS OF COMPUTER SYSTEMS**

Научная статья

УДК 519.7

doi: 10.17223/19988605/71/13

Проверка эквивалентности частично построенных комбинационных схем**Анжела Юрьевна Матросова¹, Марина Максимовна Савенкова²**^{1,2} Национальный исследовательский Томский государственный университет, Томск, Россия¹ mau11@yandex.ru² marina1412_11@mail.ru

Аннотация. Предложен алгоритм проверки эквивалентности частично построенной схемы и ее спецификации. Алгоритм основан на использовании операций пересечения однокоренных ROBDD-графов многовыходных подсхем частично построенной схемы, зависящих как от входных и выходных переменных подсхем, так и только от их выходных переменных. Для обнаружения неисправности подсхемы разработан алгоритм поиска тестовых наборов от входных переменных схемы, содержащей подсхему. Рассматриваемый в данной работе подход позволяет отказаться от использования Quantified Boolean Function (QBF) решателей или их модификаций, требующих, как правило, более сложных вычислений, чем операции над ROBDD-графами.

Ключевые слова: программное обеспечение; компьютерная томография; промышленная дефектоскопия; эвристический алгоритм; эталонный образец; многопоточная обработка.

Для цитирования: Матросова А.Ю., Савенкова М.М. Проверка эквивалентности частично построенных комбинационных схем // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2025. № 71. С. 130–139. doi: 10.17223/19988605/71/13

Original article

doi: 10.17223/19988605/71/13

Partial equivalence checking for combinational circuits**Anzhela Yu. Matrosova¹, Marina M. Savenkova²**^{1,2} National Research Tomsk State University, Tomsk, Russian Federation¹ mau11@yandex.ru² marina1412_11@mail.ru

Abstract. This paper suggests an algorithm for partial equivalence checking for combinational circuit with its specification. The algorithm is based on the intersection operations of single-root ROBDDs of multi-output sub-circuits, which depend on both the input and output variables of sub-circuits, as well as only the output variables of these sub-circuits. To provide fault detection within the sub-circuits that are parts of more complicate combinational circuit, an algorithm for generating test patterns depending on the input variables of the combinational circuit has been developed. The approach proposed eliminates the need for the necessity of employing Quantified Boolean Function (QBF) solvers or their modifications, which obviously entails more complex calculations than those associated with operations on ROBDDs.

Keywords: combinational circuits; test patterns; ROBDDs.

For citation: Matrosova, A.Yu., Savenkova, M.M. (2025) Partial equivalence checking for combinational circuits. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 71. pp. 130–139. doi: 10.17223/19988605/71/13

Введение

При создании интегральных схем возможна передача производства фрагментов схем различным разработчикам с последующим включением фрагментов в основную схему. Это позволяет одновременно создавать фрагменты и основную схему. Спецификация и основная схема (частично построенная схема) могут быть построены в различных элементных базисах. Возникает необходимость проверки частично построенной схемы (the Partial Equivalence Checking Problem, PEC) в условиях, когда отдельные фрагменты схемы еще не созданы, известны лишь их спецификации. Спецификация комбинационной схемы в целом также известна.

Традиционные подходы к тестированию неисправностей логических схем позволяют находить тестовые наборы для заданной неисправности многовыходной схемы в виде множества входных наборов схемы, причем различия реакций исправной и неисправной схем на каждом из наборов неизвестны и могут быть определены двоичным моделированием. В случае когда необходимо знать все реакции схемы и их связь с порождающими входными наборами схемы, выполнить двоичное моделирование для каждого из входных наборов исправной и неисправной схем не всегда возможно.

В работе [1] предложены алгоритм поиска всех тестовых наборов с соответствующими реакциями и их компактное представление на основе задания исправной и неисправной схем однокоренными ROBDD-графами [2]. Будем называть эти графы в дальнейшем однокоренными графами многовыходных схем. В работе [1] показана возможность применения к однокоренным графам многовыходных схем операций конъюнкции и инверсии подобно обычным ROBDD-графам.

В данной работе предлагается применить алгоритм построения тестовых наборов из [1] к задаче проверки эквивалентности частично построенной схемы и известной спецификации. Подстановка функций, сопоставляемых выходам предшествующей подсхемы вместо подмножества входных переменных последующей подсхемы, сводится к перемножению соответствующих однокоренных графов этих многовыходных подсхем. Рассмотрены цепи из подсхем, в которых выходы предыдущей подсхемы цепи являются входами ее последующей подсхемы. Предложены алгоритм поиска тестовых наборов от входных переменных подсхемы и алгоритм их трансляции на входы комбинационной схемы C , содержащей эту подсхему. Эти алгоритмы также основаны на выполнении операций над однокоренными графами многовыходных подсхем. Предлагаемый в данной работе подход позволяет отказаться от использования Quantified Boolean Function (QBF) решателей или их модификаций [3–5], требующих, как правило, более сложных вычислений, чем операции над ROBDD-графами, характеризующиеся полиномиальной сложностью.

Рассмотрен пример применения такого подхода к комбинационной схеме, из которой исключена подсхема, отданная стороннему разработчику.

1. Постановка задачи

Задана комбинационная схема C (рис. 1), с множеством $y_1, y_2, \dots, y_m$ выходов и множеством $x_1, x_2, \dots, x_n$ входов. Схема C содержит подсхему, отданную стороннему разработчику, на рис. 1 она окрашена в черный цвет. В дальнейшем договоримся называть ее черным ящиком в отличие от других подсхем. Выделенная подсхема имеет множество выходов $v_1, v_2, \dots, v_s$ и множество входов $u_1, u_2, \dots, u_r$. Множества $u_1, u_2, \dots, u_r$ и $v_1, v_2, \dots, v_s$ в общем случае являются множествами внутренних полюсов схемы C . Структура черного ящика считается известной и рассматривается в качестве спецификации. Частично построенная схема может быть реализована в ином элементном базисе, чем спецификации подсхем U и V схемы C .

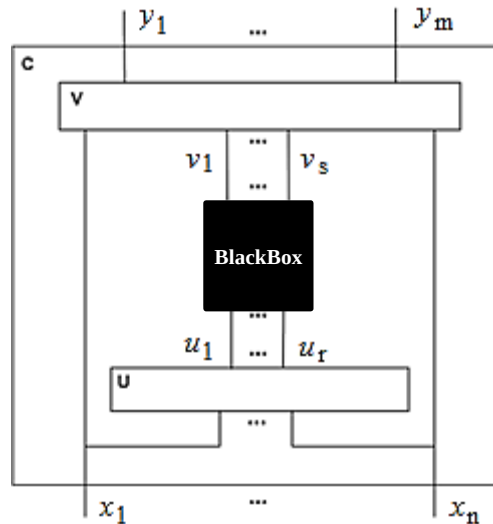


Рис. 1. Комбинационная схема C
Fig. 1. Combinational circuit C

В схеме C выходами подсхемы U являются входы $u_1, u_2, \dots, u_r$ черного ящика, а ее входами – входы $x_1, x_2, \dots, x_n$ схемы C . Выходами подсхемы V являются выходы схемы C , входами подсхемы V являются подмножество входов схемы C и выходы черного ящика.

В случае когда однокоренной граф подсхемы, содержащейся в схеме C , не совпадает с однокоренным графом спецификации этой подсхемы, заключаем, что построенная подсхема неисправна. Для нее определяются тестовые наборы от входных переменных схемы C . Если все рассмотренные подсхемы частично построенной схемы оказались исправными, делается заключение об исправности частично построенной схемы в целом.

2. Построение ROBDD-графа входо-выходных переменных многовыходной схемы

Представим векторную булеву функцию многовыходной схемы C^* виде системы из m булевых функций:

$$y_1 = f_1(x_1, x_2, \dots, x_n), \quad y_2 = f_2(x_1, x_2, \dots, x_n), \quad \dots, \quad y_m = f_m(x_1, x_2, \dots, x_n).$$

Построим функцию f , которая зависит от входных и выходных переменных этой схемы:

$$f(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m) = \bigwedge_{i=1}^m [y_i \sim f_i(x_1, x_2, \dots, x_n)],$$

где $\sim$ – логическая операция «эквивалентность».

В дальнейшем будем называть эту функцию входо-выходной функцией многовыходной схемы C^* .

Для получения всех достижимых реакций схемы и условий их обеспечения можно построить ROBDD-граф входо-выходной функции схемы C^* , предложенный в работе [2].

Если известна структура схемы C^* , то ROBDD-граф для входо-выходной функции может быть получен непосредственно по ее структуре. При построении графа сначала выполняются разложения по выходным переменным. Приведем пример построения такого графа. Назовем этот граф в дальнейшем однокоренным графом многовыходной схемы.

Пусть поведение схемы C^* описывается системой из трех булевых функций. Достижимые выходные наборы (реакции) вместе с условиями их достижимости (входными наборами) представлены таблицей истинности (табл. 1) входо-выходной функции $f(x_1, x_2, x_3, x_4, y_1, y_2, y_3)$.

Достижимым состояниям соответствуют единичные значения входо-выходной функции схемы C^* . Нулевые значения этой функции в таблице не представлены. На рис. 2 задан ROBDD-граф

входо-выходной функции схемы C , зависящий от входных и выходных переменных системы булевых функций, описывающий поведение схемы C .

Таблица 1

Таблица достижимых состояний схемы C

x_1	x_2	x_3	x_4	y_1	y_2	y_3
0	0	0	0	0	1	1
0	0	0	1	1	1	0
0	0	1	0	1	0	1
0	0	1	1	0	1	1
0	1	0	0	1	1	0
0	1	0	1	1	0	1
0	1	1	0	0	1	1
0	1	1	1	1	1	0
1	0	0	0	1	0	1
1	0	0	1	0	1	1
1	0	1	0	1	1	0
1	0	1	1	1	0	1
1	1	0	0	0	1	1
1	1	0	1	1	1	0
1	1	1	0	1	0	1
1	1	1	1	0	1	1

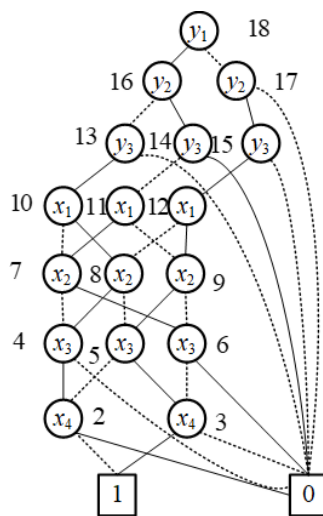


Рис. 2. ROBDD-граф функции $f(x_1, x_2, x_3, x_4, y_1, y_2, y_3)$

Fig. 2. ROBDD of function $f(x_1, x_2, x_3, x_4, y_1, y_2, y_3)$

Простая цепь $\overline{y_1} \overline{y_2} \overline{y_3} \overline{x_1} x_2 \overline{x_3} x_4$ в ROBDD-графе, ведущая в 1-концевую вершину, представляет входной набор 0101, обеспечивающий реакцию 101 выходных переменных схемы.

Простая цепь $\overline{y_1} \overline{y_2} \overline{y_3} \overline{x_1} x_2 x_3$, ведущая в 0-концевую вершину, представляет входные наборы 0111, 0110, не обеспечивающие реакцию 101 выходных переменных схемы C .

Простая цепь $\overline{y_1} \overline{y_2} \overline{y_3}$, ведущая в 0-концевую вершину, представляет недостижимую реакцию выходных переменных схемы C на выходном наборе, обращающем порождаемую цепью конъюнкцию в единицу.

Итак, простые цепи однокоренного ROBDD-графа схемы C^* , ведущие из корня в 1-концевую вершину, задают достижимые входо-выходные состояния этой схемы, а простые цепи, ведущие в 0-концевую вершину, задают ее недостижимые входо-выходные состояния.

В работе [1] предложен алгоритм построения множества всех тестовых наборов с соответствующими реакциями исправной и неисправной схем. Каждой паре реакций сопоставляется обычный ROBDD-граф от входных переменных схемы.

3. Цепи связанных подсхем

3.1. Подключение к подмножеству выходных переменных предшествующей подсхемы входов последующей подсхемы

Обратимся к рис. 1. Пусть выходы подсхемы U (предшествующей подсхемы) подключены к входам черного ящика (последующей подсхемы). Имеем соединение подсхем, при котором одна подсхема (в нашем случае U) предшествует подсхеме (в нашем случае подсхеме, переданной стороннему разработчику). В общем случае это могут быть любые две подсхемы, одна из которых непосредственно предшествует другой, причем выходы предшествующей подсхемы подключены к входам последующей подсхемы в общей схеме S . Понятно, что в результате подключения некоторые реакции черного ящика (последующей подсхемы) могут оказаться недостижимыми. Найдем достижимые реакции последующей подсхемы в рассматриваемой паре подсхем, используя следующую теорему.

Теорема 1. Достижимые реакции последующей подсхемы и условия их обеспечения представляются однокоренным графом, полученным от перемножения однокоренного графа многовыходной последующей подсхемы и однокоренного графа достижимых реакций предшествующей подсхемы.

Доказательство. Сопоставим однокоренному графу последующей подсхемы соответствующую ортогональную ДНФ, задаваемую путями из корня графа в его 1-концевую вершину. Таким же образом строим ортогональную ДНФ по однокоренному графу достижимых реакций предшествующей подсхемы. Каждая конъюнкция первого графа умножается на каждую конъюнкцию второго графа. Пустые пересечения возможны только за счет общих переменных, т.е. по выходным переменным предшествующей подсхемы, являющихся входными переменными последующей подсхемы. Наличие пустых пересечений приводит к сокращению области обеспечения реакций последующей подсхемы. Если область обеспечения некоторой реакции последующей подсхемы оказывается пустой, то реакция последующей подсхемы становится недостижимой. Перемножение графов эквивалентно перемножению порожденных ими ортогональных ДНФ. Утверждение доказано.

3.2. Поиск входного набора схемы, обеспечивающего реакцию подсхемы в цепи из подсхем

Имеем цепь из s подсхем следующего вида (рис. 3). Выходы последней подсхемы цепи, в частности, могут быть выходами схемы S . Однако это условие необязательно, поскольку для наблюдения доступны выходы всех подсхем цепи. Что касается входов подсхем, не совпадающих с входами схемы S , то на них сигналы могут поступать от входов схемы S через предшествующие подсхемы цепи. Для получения достижимых реакций каждой из последующих подсхем цепи достаточно выполнять перемножение однокоренного графа многовыходной последующей подсхемы и однокоренного графа реакций предшествующей подсхемы. Будем сохранять однокоренные графы, полученные в результате такого перемножения, зависящие от входных и выходных переменных последующей подсхемы. Это необходимо для получения тестовых наборов от входных переменных схемы S при доставке соответствующих тестовых наборов на подсхему цепи.

Пусть мы построили граф, представляющий результат пересечения однокоренного графа реакций предпоследней $(s - 1)$ -й подсхемы цепи и однокоренного графа последней s -й подсхемы цепи. Выбираем из построенного графа достижимое состояние, т.е. реакцию последней s -й подсхемы цепи, и соответствующий набор входных переменных этой подсхемы, обеспечивающий выбранную реакцию. Набор входных переменных, обеспечивающий выбранную реакцию, представляется вектором $\alpha_x : \alpha_v$. Здесь α_x – составляющая по входным переменным схемы S , α_v – составляющая по выходным пере-

менным $(s - 1)$ -й подсхемы. Найдем входной набор схемы C , обеспечивающий достижение вектора α_v в рассматриваемой цепи, который в то же время порождает составляющую α_x .

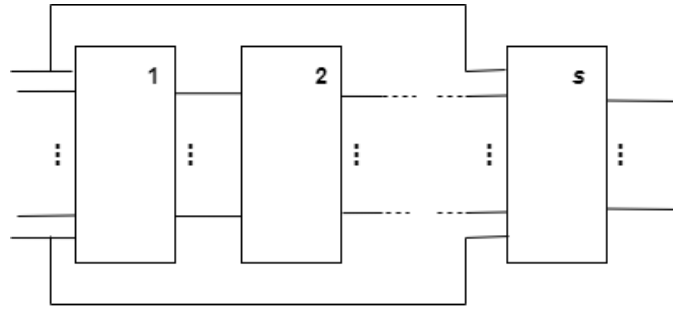


Рис. 3. Цепь подсхем длины s
Fig. 3. Chain of length s subcircuit

Алгоритм поиска входного вектора схемы C , обеспечивающего составляющую α_v входного вектора последней подсхемы:

1. В предпоследней подсхеме выделяем обычный ROBDD-граф, обеспечивающий достижение вектора α_v последней подсхемы цепи. Обозначим его R_{s-1}^* .

2. Выбираем в нем произвольный вектор $\alpha_{(s-1)}$, представляемый путем из корня графа R_{s-1}^* в его 1-концевую вершину. В однокоренном графе R_{s-2} подсхемы $(s - 2)$, который мы запомнили, при движении от первой подсхемы рассматриваемой цепи к ее s -й подсхеме выделяем обычный ROBDD-граф R_{s-2}^* , обеспечивающий достижение вектора $\alpha_{(s-1)}$ и т.д.

3. Двигаясь таким образом по графам подсхем цепи, выполняем пункты 1, 2 алгоритма. Приходим к первой подсхеме цепи. В однокоренном графе R_1 первой подсхемы выделяем обычный ROBDD-граф R_1^* , обеспечивающий достижение вектора α_2 , который является реакцией первой подсхемы рассматриваемой цепи. Произвольный вектор α_1 , представляемый путем из корня графа R_1^* первой подсхемы в его 1-концевую вершину, обеспечивает достижение вектора α_v s -й подсхемы. Граф R_1^* представляет множество входных наборов схемы C , обеспечивающих вектор α_v последней подсхемы цепи. Этот граф порождает уникальный путь достижимых реакций цепи, обеспечивающий составляющую α_v . Соответствующие компоненты вектора α_1 представляют вектор α_x s -й подсхемы.

Заметим, что $\alpha_x: \alpha_v$ может быть тестовым набором, обнаруживающим неисправность последней подсхемы рассматриваемой цепи. Аналогичным образом могут быть найдены значения входных переменных схемы C , обеспечивающие тестовые наборы для любой из подсхем рассматриваемой цепи.

Будем иметь в виду, что цепь на рис. 1 состоит из подсхемы U , подсхемы черного ящика и подсхемы V . Подсхема V связана не только с выходами черного ящика, но и с входами схемы C , которые пересекаются с входами подсхемы U .

При вычислении входного набора схемы C , обеспечивающего одну из реакций схемы V за счет рассматриваемой цепи подсхем, выполняем представленный выше алгоритм.

Теорема 2. Составляющая по входным переменным схемы C для s -й (последней) подсхемы рассматриваемой цепи обеспечивает генерацию второй составляющей по входным переменным этой же подсхемы.

Доказательство. Имеем цепь из s подсхем схемы C . Входы первой подсхемы совпадают с входами схемы C . Часть входов схемы C (возможно, все) поступают в то же время на входы s -й подсхемы. Выходы первой подсхемы совпадают с входами 2-й подсхемы. На вторую подсхему поступают только достижимые реакции первой подсхемы, т.е. часть всевозможных входных наборов 2-й подсхемы. Множество всех входных наборов схемы C разбивается на подмножества входных наборов, обеспечивающих достижимые реакции первой подсхемы рассматриваемой цепи. Реакции первой подсхемы далее

разбиваются на подмножества, обеспечивающие достижимые реакции 2-й подсхемы. Наборы входных переменных схемы C разбиваются на подмножества, обеспечивающие реакции 2-й подсхемы и т.д., вплоть до входных наборов s -й подсхемы, являющихся реакциями $(s - 1)$ -й подсхемы

Эти реакции в общем случае являются частичными входными компонентами векторов, поступающих на s -ю подсхему. Входные переменные схемы C , поступающие на входы s -й подсхемы, обеспечивают генерацию выходного вектора $(s - 1)$ -й подсхемы, т.е. реакцию этой подсхемы. При перемножении графа реакций $(s - 1)$ -й подсхемы и однокоренного графа входных и выходных переменных s -й подсхемы получаем достижимые реакции s -й подсхемы, т.е. достижимые выходы схемы C . Выходы рассматриваемой цепи схемы C обеспечиваются разбиением входов схемы C на подмножества, сопоставляемые каждой из реакций s -й подсхемы. Будем иметь в виду, что каждый входной вектор подмножества этого разбиения обеспечивает совпадение составляющей входных переменных схемы C , поступающих на s -ю подсхему с вектором, обеспечивающим генерацию второй составляющей входного вектора s -й подсхемы, являющейся реакцией $(s - 1)$ -й подсхемы рассматриваемой цепи. Утверждение доказано.

Итак, пусть для каждой подсхемы рассматриваемой цепи построен однокоренной граф, зависящий от ее входных и выходных переменных, полученный при движении от 1-й подсхемы к s -й подсхеме рассматриваемой цепи

Теорема 3. Входной набор схемы C , обеспечивающий состояние (реакцию и обеспечивающий входной набор) s -й подсхемы, находится последовательным рассмотрением однокоренных графов подсхем, начиная от $(s - 1)$ -го и вплоть до графа первой подсхемы цепи, причем для каждого графа уже выбранная реакция обеспечивается произвольным выбором входного набора среди множества входных наборов этой же подсхемы, обеспечивающих выбранную реакцию.

Доказательство. При рассмотрении соседних подсхем цепи выбираем для реакции последующей подсхемы произвольный входной набор этой подсхемы, который гарантированно обеспечивается соответствующей реакцией предшествующей подсхемы: выбранный входной набор обязательно содержится среди реакций предшествующей подсхемы. Теорема доказана.

Будем иметь в виду, что, имея цепь подсхем, начинаем анализировать первую подсхему цепи, потом вторую и т.д. Если очередная подсхема оказалась неисправной, то при переходе к следующей подсхеме мы используем однокоренные графы подсхем, построенные на основе их спецификаций.

4. Алгоритм проверки эквивалентности подсхем схемы C

Речь идет о подсхемах, которые входят в схему C и не отданы сторонним разработчикам. Перечислим шаги алгоритма на примере схемы, представленной на рис. 1.

1. Строим $R_{U_{\text{спец.}}}^{\Pi}(U, X)$ – однокоренной граф многовыходной схемы, полученный по спецификации схемы U . Граф зависит от множеств $U = \{u_1, u_2, \dots, u_r\}$ выходных и $X = \{x_1, x_2, \dots, x_n\}$ входных переменных подсхемы U .

2. Строим $R_{U_{\text{реализ.}}}^{\Pi}(U, X)$ – однокоренной граф многовыходной схемы, полученный по реализации схемы U .

3. Находим различия функций, реализуемых схемой U согласно спецификации и реализации, построенной разработчиком, возможно, в новом элементном базисе. Для этого строим два графа – $R_{U_{\text{спец.}}}^{d\Pi}(U, X)$ и $R_{U_{\text{реализ.}}}^{d\Pi}(U, X)$, содержащих тестовые наборы и соответствующие реакции исправной и неисправной схем соответственно [1]. Полученные графы зависят от входных переменных схемы C и выходных переменных подсхемы U . Если графы совпадают, схема U исправна. Иначе получаем множество тестовых наборов

$$R_{U_{\text{спец.}}}^{d\Pi}(U, X) = R_{U_{\text{спец.}}}^{\Pi}(U, X) \wedge \bar{R}_{U_{\text{реализ.}}}^{\Pi}(U, X),$$

$$R_{U_{\text{реализ.}}}^{d\Pi}(U, X) = R_{U_{\text{реализ.}}}^{\Pi}(U, X) \wedge \bar{R}_{U_{\text{спец.}}}^{\Pi}(U, X).$$

4. Строим $R_{\text{ч.я.спец}}^{\Pi}(V, U)$ – ROBDD-граф для многовыходной подсхемы, реализуемой черным ящиком согласно спецификации. Граф зависит от множеств $V = \{v_1, v_2, \dots, v_s\}$ выходных и $U = \{u_1, u_2, \dots, u_r\}$ входных переменных черного ящика.

5. Умножаем граф $R_{\text{ч.я.спец}}^{\Pi}(V, U)$ на граф реакций для подсхемы U , затем получаем граф реакций для полученного произведения. Обозначим полученный граф $R_{\text{ч.я.,}U\text{спец}}^*(V)$.

6. Строим $R_{V\text{спец}}^{\Pi}(Y, V, X)$ – ROBDD-граф для многовыходной подсхемы V , реализуемой согласно спецификации. Граф зависит от множества $Y = \{y_1, y_2, \dots, y_m\}$ выходных переменных схемы S и множеств входных переменных подсхемы V . Умножаем его на граф $R_{\text{ч.я.,}U\text{спец}}^*(V)$. Получаем граф $R_{V\text{спец}}^{\Pi*}(Y, V, X)$.

7. Аналогичным образом получаем ROBDD-граф $R_{V\text{реализ}}^{\Pi}(Y, V, X)$ для многовыходной подсхемы V , построенной разработчиком. Граф зависит от множества выходных переменных и множеств входных переменных подсхемы V . Умножаем его на граф $R_{\text{ч.я.,}U\text{спец}}^*(V)$. Получаем граф $R_{V\text{реализ}}^{\Pi*}(Y, V, X)$.

8. Находим различия функций, реализуемых схемой V согласно спецификации и реализации, построенной разработчиком, возможно, в новом элементном базисе. Для этого строим два графа $R_{V\text{спец}}^{d\Pi}(Y, V, X)$ и $R_{V\text{реализ}}^{d\Pi}(Y, V, X)$, содержащих тестовые наборы и соответствующие реакции исправной и неисправной схем соответственно:

$$R_{V\text{спец}}^{d\Pi}(Y, V, X) = R_{V\text{спец}}^{\Pi*}(Y, V, X) \wedge \bar{R}_{V\text{реализ}}^{\Pi*}(Y, V, X),$$

$$R_{V\text{реализ}}^{d\Pi}(Y, V, X) = R_{V\text{реализ}}^{\Pi*}(Y, V, X) \wedge \bar{R}_{V\text{спец}}^{\Pi*}(Y, V, X).$$

Таким образом, в результате применения приведенного выше алгоритма извлекаются множества тестовых наборов для подсхем U и V от входных переменных этих подсхем [1].

Заметим, что обычно нет необходимости использовать все тестовые наборы и их реакции для определения места неисправности. Достаточно иметь несколько тестовых наборов, реакции которых как можно больше отличаются расстоянием по Хеммингу [1].

5. Пример проверки исправности частично построенной схемы

Пусть дана схема S (рис. 4, а). Реализованная разработчиком схема имеет неисправность на выходе f_1 подсхемы V , характеризующуюся тем, что вместо дизъюнкции на этом выходе реализуется функция, принимающая единичные значения на входных наборах: $x_1 = 0, XOR_2 = 0$; $x_1 = 1, XOR_2 = 0$; $x_1 = 1, XOR_2 = 1$. Заштрихованная подсхема отдана стороннему разработчику.

Используя теорему 1 раздела 3, найдем множество достижимых наборов для подсхемы V . Таких наборов на входах подсхемы V 12, недостижимыми являются наборы 0011, 0111, 1011 и 1111 по переменным x_1 и выходам вентилей XOR_2, AND_2, AND_3 . Далее находим один из графов различий (рис. 4, б), применяя алгоритм, предложенный в работе [1]. Граф различий содержит 6 путей, ведущих в 1-терминальную вершину (табл. 2). Эти наборы входных переменных подсхемы V являются тестовыми наборами для заданной неисправности, т.е. из 12 достижимых наборов 6 являются тестовыми. На рис. 4, б представлен граф различий подсхемы V с реакциями неисправной схемы в виде ROBDD с инверсными дугами, полученный с помощью пакета dd [6].

Инверсная метка на ребре ROBDD с инверсными ребрами обозначает инверсию функции, соответствующей подграфу с вершиной, в которую входит помеченная дуга. Использование ROBDD с инверсными дугами связано с улучшением производительности в процессе выполнения операций над графами и является особенностью данной программной библиотеки, основанной на пакете CUDD (Colorado University Decision Diagram) [7] предыдущей разработки.

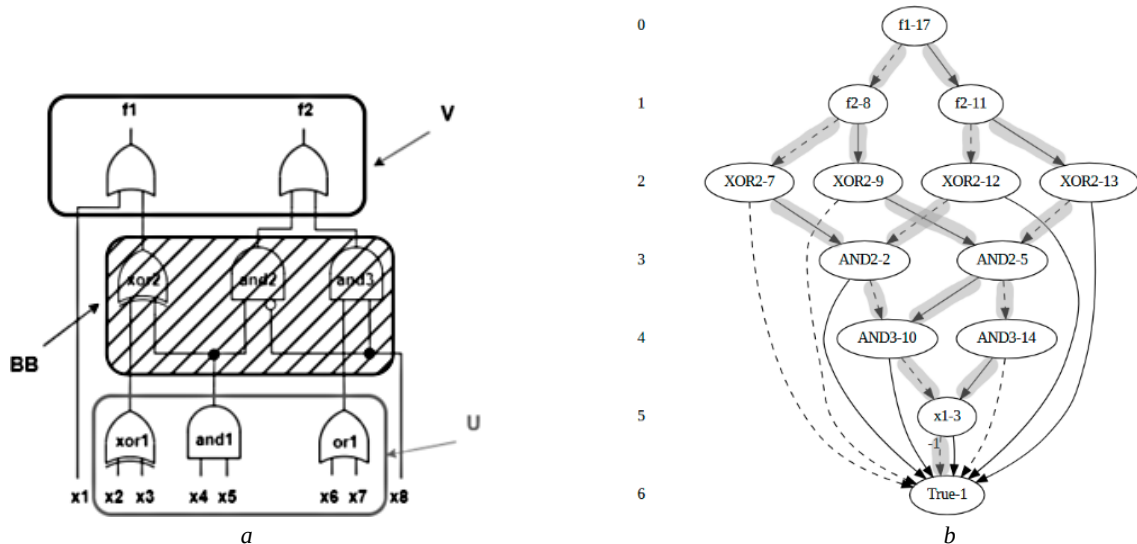


Рис. 4. Подсхема V: структура комбинационной схемы C (a); граф различий с реакциями неисправной подсхемы (b)
 Fig. 4. Subcircuit V: the structure of the combination circuit C (a); graph of differences with reactions of the faulty subcircuit (b)

Таблица 2

Таблица тестовых наборов для подсхемы V

x_1	XOR_2	AND_2	AND_3	f_1	f_2
0	1	0	0	0	0
0	1	1	0	0	1
0	1	0	1	0	1
0	0	0	0	1	0
0	0	1	0	1	1
0	0	0	1	1	1

Возьмем какой-либо тестовый набор, например 0000, и найдем значения входных переменных, которые обеспечивают данные значения на элементах XOR_2 , AND_2 , AND_3 , используя алгоритм раздела 3.2. Заметим, что для обнаружения неисправности достаточно одного набора входных переменных схемы C , доставляющего тестовый набор на входы схемы V . Поскольку неисправность задана только на одном выходе подсхемы, то расстояние по Хеммингу между реакциями исправной и неисправных схем для всех найденных тестовых наборов одинаково, поэтому выбираем любой набор из найденного множества. Выполнив шаги алгоритма поиска входного вектора схемы C , обеспечивающего составляющую α_v входного вектора подсхемы V в условиях использования однокоренных графов подсхем BB и U , найдем входной вектор 00000000 схемы C для входных переменных $x_1, \dots, x_8$.

Подставив набор 00000000, имеем: значение выхода f_1 равно 0 в исправной схеме и 1 в неисправной схеме, следовательно, данный набор обнаруживает заданную неисправность.

Заключение

Предложен алгоритм проверки эквивалентности частично построенной комбинационной схемы и соответствующей спецификации. Спецификации подсхем схемы C и подсхем, отданных стороннему разработчику, также известны. Схема C содержит цепь из подсхем, среди которых некоторые подсхемы отданы стороннему разработчику. На основе операций пересечения однокоренных ROBDD-графов многовыходных последующих подсхем от их входных и выходных переменных и однокоренных графов реакций предшествующих подсхем от их выходных переменных цепи (речь идет о паре соседних подсхем цепи) разработан алгоритм проверки эквивалентности подсхем, созданных разработчиком, и спецификаций этих подсхем. Для обнаружения неисправности подсхемы предложен алгоритм обеспечения доставки входных тестовых наборов подсхемы цепи на входы схемы C .

Список источников

1. Матросова А.Ю., Савенкова М.М., Провкин В.А., Сухорученко К.А. Тестирование многовыходных схем // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 69. С. 134–143. doi: 10.17223/19988605/69/14
2. Провкин В.А., Матросова А.Ю. Графовые представления множеств всех достижимых реакций комбинационной схемы // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2022. № 61. С. 128–138. doi: 10.17223/19988605/61/13
3. Scholl C., Becker B. Checking equivalence for circuits containing incompletely specified boxes // Int'l Conf. on Computer Design (ICCD). IEEE, 2002. P. 56–63.
4. Herbstritt M., Becker B., Scholl C. Advanced SAT-techniques for bounded model checking of blackbox designs // Int'l Workshop on Microprocessor Test and Verification (MTV). IEEE, 2006. P. 37–44.
5. Bryant R.E. Graph-Based Algorithms for Boolean Function Manipulation // IEEE Transactions on Computers. 1986. V. C-35, is. 8. P. 677–691.
6. dd package // Github.com. URL: <https://github.com/tulip-control/dd> (accessed: 25.04.2025).
7. The CUDD package, BDD, ADD Tutorial and examples // David Kebo Hounounou. URL: <https://www.davidkebo.com/cudd#cudd1> (accessed: 25.04.2025).

References

1. Matrosova, A.Yu., Savenkova, M.M., Provkin, V.A. & Sukhoruchenko, K. A. (2024) Testing of multi-output combinational circuits when reactions on test patterns are known. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 69. pp. 134–143. DOI: 10.17223/19988605/69/14
2. Provkin, V.A. & Matrosova, A.Yu. (2022) Graph representations of the sets of all reachable reactions of the combinational circuit. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 61. pp. 128–138. DOI: 10.17223/19988605/61/13
3. Scholl, C. & Becker B. (2002) Checking equivalence for circuits containing incompletely specified boxes. *IEEE Int'l Conf. on Computer Design (ICCD)*. pp. 56–63.
4. Herbstritt, M., Becker, B. & Scholl C. (2006) Advanced SAT-techniques for bounded model checking of blackbox designs. *IEEE Int'l Workshop on Microprocessor Test and Verification (MTV)*. pp. 37–44.
5. Bryant, R.E. (1986) Graph-Based Algorithms for Boolean Function Manipulation. *IEEE Transactions on Computers*. C-35(8). pp. 677–691.
6. dd package. [Online] Available from: <https://github.com/tulip-control/dd> (Accessed: 25th April 2025).
7. The CUDD package, BDD, ADD Tutorial and examples. [Online] Available from: <https://www.davidkebo.com/cudd#cudd1>. (Accessed: 25th April 2025).

Информация об авторах:

Матросова Анжела Юрьевна – профессор, доктор технических наук, профессор кафедры компьютерной безопасности Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: mau11@yandex.ru

Савенкова Марина Максимовна – магистрант кафедры компьютерной безопасности Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: marina1412_11@mail.ru

Вклад авторов: все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

Information about the authors:

Matrosova Anzhela Yu. (Doctor of Technical Sciences, Professor, Institute of Applied Mathematics and Computer Science, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: mau11@yandex.ru

Savenkova Marina M. (Master's Student, Institute of Applied Mathematics and Computer Science, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: marina1412_11@mail.ru

Contribution of the authors: the authors contributed equally to this article. The authors declare no conflicts of interests.

Поступила в редакцию 15.02.2025; принята к публикации 02.06.2025

Received 15.02.2025; accepted for publication 02.06.2025

Научный журнал

**ВЕСТНИК
ТОМСКОГО
ГОСУДАРСТВЕННОГО
УНИВЕРСИТЕТА**

**УПРАВЛЕНИЕ,
ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА
И ИНФОРМАТИКА**

**TOMSK STATE UNIVERSITY
JOURNAL OF CONTROL AND COMPUTER SCIENCE**

2025. № 71

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 30.06.2025 г. Формат 60x84^{1/8}.
Гарнитура Times. Усл. печ. л. 16,3.
Тираж 250 экз. Заказ № 6359. Цена свободная.

Дата выхода в свет 10.07.2025 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