

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

2024

№ 69

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
Матрсова 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)
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)
Сущенко 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 Зорян (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

- Масталиев Р.О., Багирлы Н.Д.**
Необходимые условия оптимальности в одной негладкой задаче управления дискретными системами типа Вольтерра 4

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

- Ашурков Н.В., Коробков Д.В., Вавилов О.А., Юркевич В.Д.**
Математическая модель трехфазного AC/DC преобразователя с LC-фильтром 10
- Farkhadov M.P., Vaskovskii S.V., Brokarev I.A.**
Neural network model and software for an information system to intelligently analyze gas quality 22

ОБРАБОТКА ИНФОРМАЦИИ

- Базилевский М.П.**
Интеграция ограничений на коэффициенты интеркорреляций в оптимизационную задачу и условия построения вполне интерпретируемых неэлементарных линейных регрессий 31
- Воробьева Г.Р., Воробьев А.В., Фарваев Э.Ф., Орлов Г.О.**
Элементы системного и вепольного анализа инфраструктуры мониторинга окружающей среды (на примере геомагнитных данных) 42
- Гервас Н.В., Зеркаль С.М.**
Вычислительная кинематическая диагностика скорости зондирующего сигнала с использованием априорной информации об исследуемой среде 54
- Гергет О.М., Лаптев Н.В., Чернявский М.А.**
Нейросетевое моделирование и автоматическая сегментация корня аорты на снимках ангиографии 64
- Nebaba S.G., Markov N.G.**
Detection of flying objects in images using the YOLOv4-CSP convolutional neural network model 72
- Нежелская Л.А., Пономаренко В.Д.**
Обобщенный асинхронный поток событий с продлевающимся мертвым временем 82
- Цициашвили Г.Ш., Боcharников В.Н.**
Интервальное распознавание больших выбросов в экосистеме по наблюдениям с фотоловушек 95

ИНФОРМАТИКА И ПРОГРАММИРОВАНИЕ

- Maltseva M.A.**
Quantum circuit optimization via local qubit reordering by quantum annealing 103
- Обходский А.В., Кульбакин Д.Е., Обходская Е.В., Попов А.С., Родионов Е.О., Сачков В.И., Чернов В.И., Чойнзонов Е.Л.**
Программный комплекс нейросетевой классификации данных газоаналитического обследования дыхательной системы 112
- Ушенина И.В.**
Реализация на современных ПЛИС вычислителя сигмоидной функции активации нейронных сетей табличным методом 124

ПРОЕКТИРОВАНИЕ И ДИАГНОСТИКА ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

- Матросова А.Ю., Савенкова М.М., Провкин В.А., Сухорученко К.А.**
Тестирование многовыходных комбинационных схем в условиях известных реакций на тестовые наборы 134

CONTENTS

CONTROL OF DYNAMICAL SYSTEMS

- Mastaliyev R.O., Bagirli N.D.**
Necessary optimality conditions in a non-smooth control problem for Volterra type discrete systems 4

MATHEMATICAL MODELING

- Ashurkov N.V., Korobkov D.V., Vavilov O.A., Yurkevich V.D.**
Mathematical model for three-phase AC/DC converter with LC-filter 10
- Farkhadov M.P., Vaskovskii S.V., Brokarev I.A.**
Neural network model and software for an information system to intelligently analyze gas quality 22

DATA PROCESSING

- Bazilevskiy M.P.**
Integration of constraints on intercorrelation coefficients into the optimization problem and conditions for constructing quite interpretable non-elementary linear regressions 31
- Vorobeveva G.R., Vorobev A.V., Farvaeve E.F., Orlov G.O.**
Elements of system and Su-field analysis of environmental monitoring tools (using the example of the geomagnetic data) 42
- Gervas N.V., Zerkal S.M.**
Computational kinematic diagnostics of the probing signal speed using a priori information about the medium under study 54
- Gerget O.M., Laptev N.V., Chernyavskiy M.A.**
Segmentation of the aortic root on angiography images 64
- Nebaba S.G., Markov N.G.**
Detection of flying objects in images using the YOLOv4-CSP convolutional neural network model 72
- Nezhel'skaya L.A., Ponomarenko V.D.**
Generalized asynchronous flow of events with extending dead time 82
- Tsitsiashvili G.Sh., Bocharnikov V.N.**
Interval detection of large emissions in the ecosystem based on observations from camera traps 95

INFORMATICS AND PROGRAMMING

- Maltseva M.A.**
Quantum circuit optimization via local qubit reordering by quantum annealing 103
- Obkhodskiy A.V., Kulbakin D.E., Obkhodskaya E.V., Popov A.S., Rodionov E.O., Sachkov V.I., Chernov V.I., Choynzonov E.L.**
Neural network classification software for the gas analytical survey data of respiratory system 112
- Ushenina I.V.**
Realization of the sigmoid activation function for neural networks on current FPGAs by the table-driven method 124

DESIGNING AND DIAGNOSTICS OF COMPUTER SYSTEMS

- Matrosova A.Yu., Savenkova M.M., Provkin V.A., Sukhoruchenko K.A.**
Testing of multi-output combinational circuits when reactions on test patterns are known 134

УПРАВЛЕНИЕ ДИНАМИЧЕСКИМИ СИСТЕМАМИ

CONTROL OF DYNAMICAL SYSTEMS

Научная статья

УДК 517.97

doi: 10.17223/19988605/69/1

Необходимые условия оптимальности в одной негладкой задаче
управления дискретными системами типа ВольтерраРашад Огтай оглы Масталиев¹, Нигар Давид кызы Багирлы²¹ Университет «Азербайджан», Баку, Азербайджан¹ Институт систем управления Министерства науки и образования Азербайджанской Республики, Баку, Азербайджан² Азербайджанский государственный аграрный университет, Гянджа, Азербайджан¹ rashad.mastaliyev@au.edu.az² nigar.bayramova50@gmail.com

Аннотация. Рассматривается задача оптимального управления, описываемая разностным аналогом интегродифференциального уравнения типа Вольтерра с недифференцируемым критерием качества. При этом предполагается, что правая часть исходного уравнения и функционал качества по вектору состояния имеют производные по любому направлению. Сначала, применяя метод приращения и учитывая некоторые понятия негладкого анализа, установлены необходимые условия оптимальности в терминах производных по направлениям, носящие общий характер. Затем, опираясь на них, установлены необходимые условия оптимальности, непосредственно выраженные через параметры рассматриваемой негладкой задачи.

Ключевые слова: разностное уравнение типа Вольтерра; негладкий функционал качества; производная по направлению; оптимальность; задача на минимакс.

Для цитирования: Масталиев Р.О., Багирлы Н.Д. Необходимые условия оптимальности в одной негладкой задаче управления дискретными системами типа Вольтерра // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 69. С. 4–9. doi: 10.17223/19988605/69/1

Original article

doi: 10.17223/19988605/69/1

Necessary optimality conditions in a non-smooth control problem
for Volterra type discrete systemsRashad O. Mastaliyev¹, Nigar D. Bagirli²¹ Azerbaijan University, Baku, Azerbaijan¹ Institute of Control Systems, Baku, Azerbaijan² Azerbaijan State Agricultural University, Ganja, Azerbaijan¹ rashad.mastaliyev@au.edu.az² nigar.bayramova50@gmail.com

Abstract. The optimal control problem described by the difference analog of the Volterra type integro-differential equation with a non-differentiable quality criterion is considered. It is assumed that the right-hand side of the original equation and the quality functional with respect to the state vector have derivatives in any direction. First, using the increment method and taking into account some concepts of non-smooth analysis, the necessary optimality conditions are established in terms of derivatives in directions of a general nature. Then, based on them, the necessary

optimality conditions directly expressed through the parameters of the non-smooth problem under consideration are established.

Keywords: Volterra type difference equation; non-smooth quality functional; directional derivative; optimality; minimax problem.

For citation: Mastaliyev, R.O., Bagirli, N.D. (2024) Necessary optimality conditions in a non-smooth control problem for Volterra type discrete systems. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 69. pp. 4–9. doi: 10.17223/19988605/69/1

Введение

Как известно (см., напр.: [1, 2]), математическое моделирование разнообразных задач экономики и техники приводит к специальным задачам оптимизации с негладкими целевыми функционалами. Такие задачи оптимального управления непрерывными системами изучены в работах [3, 4], где приведена дополнительная библиография.

Наряду с этим дискретные динамические модели управляемых систем являются довольно важным в теоретическом и практическом отношении классом математических моделей, они позволяют охватить очень широкий круг реальных объектов и соответствующих им задач оптимального управления.

Различные аспекты задач оптимального управления, описываемых дискретными системами с негладкими функционалами качества, изучены в работах [5, 6] и др.

Исследованию задач оптимального управления, описываемых разностными уравнениями типа Вольтера с гладким терминальным критерием качества, посвящены работы [7–9].

В предлагаемой же работе подобная задача изучается в случае негладкого функционала качества, при этом с учетом специфики функционала качества доказаны необходимые условия оптимальности в терминах производных по направлениям [10].

1. Постановка задачи

Пусть $T = \{t_0, t_0 + 1, \dots, t_1 - 1\}$ – заданная конечная последовательность, причем разность $t_1 - t_0$ есть натуральное число, а $\Phi(x)$ – заданная скалярная функция, удовлетворяющая условию Липшица и имеющая производные по любому направлению.

Предположим, что управляемый дискретный процесс описывается системой обыкновенных разностных уравнений типа Вольтера

$$x(t+1) = \sum_{\tau=t_0}^t f(t, \tau, x(\tau), u(\tau)), \quad t \in T, \quad x(t_0) = x_0. \quad (1)$$

Здесь $f(t, \tau, x, u)$ – заданная n -мерная вектор-функция, непрерывная по (x, u) с частными производными по x при всех (t, τ) , $u(t)$ – r -мерный вектор управляющих воздействий.

Если управляющая функция $u(t)$ удовлетворяет ограничению

$$u(t) \in U \subset R^r, \quad t \in T, \quad (2)$$

то такую управляющую функцию назовем допустимым управлением. Здесь U – заданное непустое, ограниченное множество.

Задача заключается в минимизации негладкого функционала

$$S(u) = \Phi(x(t_1)) \quad (3)$$

при ограничениях (1), (2).

Допустимое управление $u(t)$, доставляющее минимальное значение функционалу (3) при ограничениях (1), (2), назовем оптимальным управлением, а соответствующий процесс $(u(t), x(t))$ – оптимальным процессом.

Цель настоящей статьи – получение необходимых условий оптимальности первого порядка в рассматриваемой дискретной негладкой задаче (1)–(3).

2. Необходимые условия оптимальности в терминах производных по направлениям

Считая $(u(t), x(t))$ фиксированным допустимым процессом, введем множество

$$f(t, \tau, x(\tau), U) = \{ \alpha : \alpha = f(t, \tau, x(\tau), v), v \in U \}. \quad (4)$$

В дальнейшем будем предполагать, что множество (4) выпукло при всех (t, τ) .

Через $u(t; \varepsilon)$ обозначим произвольное допустимое управление такое, что соответствующее ему состояние $x(t; \varepsilon)$ процесса удовлетворяет соотношению

$$\begin{aligned} x(t+1; \varepsilon) &= \sum_{\tau=t_0}^t f(t, \tau, x(\tau; \varepsilon), u(\tau, \varepsilon)) = \\ &= \sum_{\tau=t_0}^t (1-\varepsilon) f(t, \tau, x(\tau; \varepsilon), u(\tau)) + \sum_{\tau=t_0}^t \varepsilon f(t, \tau, x(\tau; \varepsilon), v(\tau, \varepsilon)), \end{aligned} \quad (5)$$

где $\varepsilon \in [0, 1]$ – произвольное число, а $v(t) \in U$, $t \in T$, – произвольное допустимое управление.

Такое допустимое управление $u(t; \varepsilon)$ существует в силу выпуклости множества (4).

Пусть по определению

$$y(t) = \left. \frac{\partial x(t; \varepsilon)}{\partial \varepsilon} \right|_{\varepsilon=0}. \quad (6)$$

Используя (5) и учитывая гладкость вектор-функции $f(t, x, \tau, u)$, получим, что вектор-функция $y(t)$, определяемая формулой (6), является решением следующего разностного уравнения в вариациях типа Вольтерра:

$$y(t+1) = \sum_{\tau=t_0}^t [f_x(t, \tau, x(\tau), u(\tau)) y(\tau) + (f(t, \tau, x(\tau), v(\tau)) - f(t, \tau, x(\tau), u(\tau)))]. \quad (7)$$

С другой стороны, из (6) ясно, что

$$x(t; \varepsilon) = x(t) + \varepsilon y(t) + o(\varepsilon; t). \quad (8)$$

При этом специальное приращение функционала качества (3), отвечающее допустимым управлениям $u(t; \varepsilon)$ и $u(t)$, записывается в виде:

$$\begin{aligned} S(u(t; \varepsilon)) - S(u(t)) &= \Phi(x(t_1; \varepsilon)) - \Phi(x(t_1)) = \\ &= [\Phi(x(t_1) + \varepsilon y(t_1) + o(t_1; \varepsilon)) - \Phi(x(t_1) + \varepsilon y(t_1))] - \\ &\quad - [\Phi(x(t_1) + \varepsilon y(t_1)) - \Phi(x(t_1))]. \end{aligned} \quad (9)$$

Поскольку функция $\Phi(x)$ удовлетворяет условию Липшица, то ясно, что

$$|\Phi(x(t_1) + \varepsilon y(t_1) + o(t_1; \varepsilon)) - \Phi(x(t_1) + \varepsilon y(t_1))| \leq o_1(\varepsilon). \quad (10)$$

Отсюда, используя определение производной по направлениям, получаем

$$\Phi(x(t_1) + \varepsilon y(t_1)) - \Phi(x(t_1)) = \varepsilon \frac{\partial \Phi(x(t_1))}{\partial y(t_1)} + o_2(\varepsilon).$$

Теперь предположим, что допустимое управление $u(t)$ оптимальное, тогда из последнего разложения получаем

$$\frac{\varepsilon \partial \Phi(x(t_1))}{\partial y(t_1)} + o_2(\varepsilon) \geq 0. \quad (11)$$

Таким образом, доказана следующая теорема.

Теорема 1. Если множество (4) выпуклое, то для оптимальности допустимого управления $u(t)$ в рассматриваемой задаче (1)–(3) необходимо, чтобы неравенство

$$\frac{\partial \Phi(x(t_1))}{\partial y(t_1)} \geq 0 \quad (12)$$

выполнялось для всех $v(t) \in U, t \in T$.

Неравенство (12) является неявным критерием оптимальности относительно параметрам задачи (1)–(3). Используя его, удастся получить необходимое условие оптимальности, выраженное непосредственно через параметры задачи (1)–(3).

Уравнение (7) является линейным неоднородным разностным уравнением. Поэтому, применяя результаты, например, работ [7, 11] получаем, что решение $x(t)$ уравнения в вариациях, допускает представление

$$y(t) = \sum_{\tau=t_0}^{t-1} \left[\sum_{s=\tau}^{t-1} R(t, s+1) (f(s, \tau, x(\tau), v(\tau)) - f(s, \tau, x(\tau), u(\tau))) \right]. \quad (13)$$

Здесь $R(t, s)$ является резольвентой уравнения (8), при фиксированном t удовлетворяет по $s (s \leq t)$ матричному разностному уравнению Вольтерра

$$R(t+1, s) = \sum_{\tau=s}^t R(t+1, \tau+1) f_x(\tau, s, x(s)u(s)), \quad s \leq t \quad (14)$$

с начальным условием

$$R(t+1, t+1) = E,$$

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

Учитывая (13), положим

$$L_1(v) = \sum_{\tau=t_0}^{t_1-1} \left[\sum_{s=\tau}^{t_1-1} R(t_1, s+1) (f(s, \tau, x(\tau), v(\tau)) - f(s, \tau, x(\tau), u(\tau))) \right].$$

Следовательно, доказана следующая теорема.

Теорема 2. Если множество (4) выпуклое, то для оптимальности допустимого управления $u(t)$ в задаче (1)–(3) необходимо, чтобы неравенство

$$\frac{\partial \Phi(x(t_1))}{\partial L_1(v)} \geq 0$$

выполнялось для всех $v(t) \in U, t \in T$.

3. Линеаризованное необходимое условие оптимальности

В этом разделе дополнительно предположим, что в задаче (1)–(3) вектор-функция $f(t, \tau, x, u)$ непрерывно дифференцируема по (x, u) , а множество U выпукло. Эти дополнительные условия позволяют нам записать «возмущенную» систему вида:

$$x(t+1; \mu) = \sum_{\tau=t_0}^t f(t, \tau, x(\tau, \mu), u(\tau, \mu)) \equiv \sum_{\tau=t_0}^t (1-\varepsilon) f(t, \tau, x(\tau, \mu), \mu v(\tau) + (1-\mu)u(\tau)), \quad (15)$$

где $\mu \in [0, 1]$ – произвольное число, а $v(t)$ произвольное допустимое управление.

Пусть

$$z(t, v) = \frac{\partial x(t; \mu)}{\partial \mu} \Big|_{\mu=0}. \quad (16)$$

Тогда из (15) следует, что $z(t, v)$, определяемое формулой (16), является решением следующего уравнения в вариациях:

$$z(t+1, v) = \sum_{\tau=t_0}^t [f_x(t, \tau, x(\tau), v(\tau))z(\tau, v) + f_u(t, \tau, x(\tau), u(\tau))(v(\tau) - u(\tau))]. \quad (17)$$

При этом специальное приращение критерия качества (3), отвечающее допустимым управлениям $u(t; \mu)$ и $u(t)$ представляется в виде:

$$S(u(t; \mu)) - S(u(t)) = [\Phi(x(t_1) + \mu z(t_1, v) + o(\mu)) - \Phi(x(t_1) + \mu z(t_1, v))] = \mu \frac{\partial \Phi(x(t_1))}{\partial z(t_1, v)} + o(\mu). \quad (18)$$

Из полученного разложения следует теорема.

Теорема 3. Если множество U выпукло, то для оптимальности допустимого управления $u(t)$ в рассматриваемой задаче (1)–(3) при дополнительном предположении, что функция $f(t, \tau, x, u)$ непрерывно дифференцируема по (x, u) , необходимо, чтобы неравенство

$$\frac{\partial \Phi(x(t_1))}{\partial z(t_1, v)} \geq 0 \quad (19)$$

выполнялось для всех $v(t) \in U, t \in T$.

Как и раньше, конкретизируем полученное необходимое условие оптимальности, другими словами, выразим непосредственно через параметры задачи (1)–(3).

Благодаря результатам работ [7, 11] решение $z(t, v)$ линеаризованного уравнения в вариациях (17) представимо в виде:

$$z(t, v) = \sum_{\tau=t_0}^{t-1} \sum_{s=\tau}^{t-1} [R(t, s) f_u(s, \tau, x(\tau), u(\tau))](v(\tau) - u(\tau)), \quad (20)$$

где $R(\tau, t)$ – решение уравнения (14).

Введя обозначение

$$Q(t, \tau) = \sum_{s=\tau}^t [R(t, s) f_u(s, \tau, x(\tau), u(\tau))],$$

из тождества (20) имеем

$$z(t, v) = \sum_{\tau=t_0}^{t-1} Q(t, \tau)(v(\tau) - u(\tau)).$$

Отсюда следует

$$z(t_1, v) = \sum_{\tau=t_0}^{t_1-1} Q(t_1, \tau)(v(\tau) - u(\tau)).$$

Полагая

$$L_2(v) = \sum_{\tau=t_0}^{t_1-1} Q(t_1, \tau)(v(\tau) - u(\tau)),$$

из неравенства (19) получаем следующий результат.

Теорема 4. Если множество U выпукло, то для оптимальности допустимого управления $u(t)$ в рассматриваемой задаче (1)–(3) необходимо, чтобы неравенство

$$\frac{\partial \Phi(x(t_1))}{\partial L_2(v)} \geq 0$$

выполнялось для всех $v(t) \in U, t \in T$.

Заключение

В предлагаемой работе рассматривается негладкая задача оптимального управления дискретными системами Вольтерра и при помощи метода явной линеаризации установлено необходимое условие оптимальности первого порядка в терминах производных по направлениям.

Список источников

1. Конюховский П.В. Математические методы исследования операций в экономике. СПб. : Питер, 2020. 208 с.
2. Черноморов Г.А. Теория принятия решений. Новочеркасск, 2002. 276 с.
3. Отакулов С., Хайдаров Т.Т. Условия оптимальности в негладкой задаче управления для динамической системы с параметрами // J. Physics and Math. 2020. V. 13 (65). С. 7–11. doi: 10.24411.2520-6990-2020-11847

4. Отакулов С., Жуманов К.С. Негладкая задача оптимального управления для линейных модели динамических систем // J. Science and Innovation. 2022. № 3. С. 252–259. doi: 10.5281/zenodo.6809643
5. Гараева Э.А., Мансимов К.Б. Необходимое условие оптимальности в задаче управления с дискретным временем при недифференцируемом критерии качества // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2017. № 38. С. 4–10. doi: 10.17223/19988605/38/1
6. Shvartsman I. Necessary optimality conditions in discrete non-smooth optimal control // J. Optimization Theory and Applications. 2011. V. 153. P. 578–586.
7. Мансимов К.Б., Масталиев Р.О. Оптимизация процессов, описываемых разностными уравнениями Вольтерра. Saarbrücken : Lambert Academic Publishing, 2017. 263 с.
8. Масталиев Р.О. Особые управления в дискретных системах Вольтерра // Вестник Бакинского университета. Сер. физ.-мат. наук. 2008. № 1, С. 44–49.
9. Mansimov K.B., Mastaliyev R.O. Necessary first and second order optimality conditions in problems of control described by a system of Volterra difference equations // J. Automatic Control and Computer Sciences, 2008. V. 42. P. 71–76.
10. Демьянов В.Ф., Рубинов А.М. Основы негладкого анализа и квазидифференциальное исчисление. М. : Наука, 1990. 432 с.
11. Колмановский В.Б. Об асимптотических свойствах решений некоторых нелинейных систем Вольтерра // Автоматика и телемеханика. 2000. № 4. С. 4–12.

References

1. Konyukhovskiy, P.V. (2020) *Matematicheskie metody issledovaniya operatsiy v ekonomike* [Mathematical methods of operations research in economics]. St. Petersburg: Piter.
2. Chernomorov, G.A. (2002) *Teoriya prinyatiya resheniy* [Theory of decision making]. Novocherkassk: [s.n.].
3. Otakulov, S. & Khaidarov, T.T. (2020) Optimality conditions in a nonsmooth control problem for a dynamic system with parameters. *Journal of Physics and Mathematics*. 13(65). pp. 7–11. DOI: 10.24411.2520-6990-2020-11847
4. Otakulov, S. & Zhumanov, K.S. (2022) Nonsmooth optimal control problem for linear models of dynamic systems *Journal of Science and Innovation*. 3. pp. 252–259. (In Russian). DOI: 10.5281/zenodo.6809643
5. Garaeva, E.A. & Mansimov, K.B. (2017) Necessary condition for optimality in a discrete-time control problem with a non-differentiable quality criterion. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 38. pp. 4–10. DOI: 10.17223/19988605/38/1
6. Shvartsman, I. (2011) Necessary optimality conditions in discrete nonsmooth optimal control. *Journal of Optimization Theory and Applications*. 153. pp. 578–586.
7. Mansimov, K.B. & Mastaliyev, R.O. (2017) *Optimizatsiya protsessov, opisivaemykh raznostnymi uravneniyami Vol'terra* [Optimization of processes described by Volterra difference equations]. Saarbrücken: Lambert Academic Publishing.
8. Mastaliyev, R.O. (2008) Osobyie upravleniya v diskretnykh sistemakh Vol'terra [Singular controls in discrete Volterra systems]. *Vestnik Bakinskogo universiteta. Ser. fiz.-mat. Nauk*. 1. pp. 44–49.
9. Mansimov, K.B. & Mastaliyev, R.O. (2008) Necessary first and second order optimality conditions in problems of control described by a system of Volterra difference equations. *Journal of Automatic Control and Computer Sciences*. 42. pp. 71–76.
10. Demyanov, V.F. & Rubinov, A.M. (1990) *Osnovy negladkogo analiza i kvazidifferentsial'noe ischislenie* [Fundamentals of Nonsmooth Analysis and Quasidifferential Calculus]. Moscow: Nauka.
11. Kolmanovskiy, V.B. (2000) Ob asimptoticheskikh svoystvakh resheniy nekotorykh nelineynykh sistem Vol'terra [On asymptotic properties of solutions of some Volterra nonlinear systems]. *Avtomatika i telemekhanika–Automation and Remote Control*. 4(4). pp. 25–33.

Информация об авторах:

Масталиев Рашад Огтай оглы – доктор философии по математике (кандидат физико-математических наук), доцент, заведующий кафедрой математики и информатики университета «Азербайджан»; ведущий научный сотрудник Института систем управления Министерства науки и образования Азербайджанской Республики (Баку, Азербайджан). E-mail: rashad.mastaliyev@au.edu.az

Багирлы Нигар Давид кызы – ассистент кафедры математики и физики Азербайджанского государственного аграрного университета (Гянджа, Азербайджан). E-mail: nigar.bayramova50@gmail.com

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

Information about the authors:

Mastaliyev Rashad O. (Doctor of Philosophy in Mathematics (Candidate of Physical and Mathematical Sciences), Associate Professor, Head of the Department of Mathematics and Informatics of the Azerbaijan University, Leading Researcher, Institute of Control Systems Ministry of Science and Education of the Republic of Azerbaijan, Baku, Azerbaijan). E-mail: rashad.mastaliyev@au.edu.az

Bagirly Nigar D. (Assistant of the Department of Mathematics and Physics, Azerbaijan State Agricultural University Ganja, Azerbaijan). E-mail: nigar.bayramova50@gmail.com

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

Поступила в редакцию 28.07.2024; принята к публикации 02.12.2024

Received 28.07.2024; accepted for publication 02.12.2024

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

MATHEMATICAL MODELING

Научная статья
УДК 681.5 (681.5.013)
doi: 10.17223/19988605/69/2

Математическая модель трехфазного AC/DC преобразователя с LC-фильтром

Никита Владимирович Ашурков¹, Дмитрий Владиславович Коробков²,
Олег Андреевич Вавилов³, Валерий Дмитриевич Юркевич⁴

^{1, 2, 3, 4} Новосибирский государственный технический университет, Новосибирск, Россия

¹ 4445@list.ru

² korobkov@corp.nstu.ru

³ vavilov.oleg.rp462@gmail.com

⁴ yurkev@mail.ru

Аннотация. Рассматриваются математические модели трехфазного двухуровневого инвертора напряжения с расщепленной емкостью, работающего в обратном режиме (активного выпрямителя), в составе бортовой авиационной системы электроснабжения с первичным четырехпроводным источником переменного тока переменной частоты. Предлагается математическая модель преобразователя в стационарной системе $a-b-c$ -координат с использованием модели силовых полупроводниковых ключей в виде переключающих функций, а также математическая модель в пространстве состояний ключей схемы преобразователя. Проведен анализ электромагнитных процессов в преобразователе, получены диаграммы мгновенных значений токов и напряжений. Проведен сравнительный анализ полученных диаграмм с результатами имитационного моделирования, подтверждающий адекватность полученных математических моделей.

Ключевые слова: математическая модель; электроснабжение; инвертор напряжения; активный выпрямитель; широтно-импульсная модуляция; имитационное моделирование; верификация модели.

Благодарности: Исследование выполнено в рамках программы «Приоритет-2030» НГТУ по стратегическому направлению «Силовая электроника и интеллектуальная энергетика», продуктовая линейка «Гибридные микросборки энергопреобразующей аппаратуры летательных аппаратов».

Для цитирования: Ашурков Н.В., Коробков Д.В., Вавилов О.А., Юркевич В.Д. Математическая модель трехфазного AC/DC преобразователя с LC-фильтром // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 69. С. 10–21. doi: 10.17223/19988605/69/2

Original article
doi: 10.17223/19988605/69/2

Mathematical model for three-phase AC/DC converter with LC-filter

Nikita V. Ashurkov¹, Dmitry V. Korobkov²,
Oleg A. Vavilov³, Valery D. Yurkevich⁴

^{1, 2, 3, 4} Novosibirsk State Technical University, Novosibirsk, Russian Federation

¹ 4445@list.ru

² korobkov@corp.nstu.ru

³ vavilov.oleg.rp462@gmail.com

⁴ yurkev@mail.ru

Abstract. The article is devoted to the problem of developing mathematical models of a three-phase two-level voltage inverter with a split capacitance. An inverter operating in reverse mode (active rectifier) is considered, working

as part of on-board power supply system with a primary four-wire source of alternating current with variable frequency. Mathematical models have been developed in a stationary a-b-c coordinate system using a model of power semiconductor keys in the form of switching functions, and in the state space of the converter circuit power switches. An analysis of electromagnetic processes in the circuit has been carried out, and inverter currents and voltages diagrams have been obtained. A comparative analysis of the obtained diagrams with the results of simulation modeling was carried out, and the adequacy of the obtained mathematical models was confirmed.

Keywords: mathematical model; power supply; voltage inverter; active rectifier; pulse width modulation; simulation modeling; model verification.

Acknowledgments: The research was carried out as a part of the “Priority-2030” program of NSTU in the strategic direction “Power electronics and smart energy”, the product line “Hybrid microassemblies of energy-converting aircraft equipment.”

For citation: Ashurkov, N.V., Korobkov, D.V., Vavilov, O.A., Yurkevich, V.D. (2024) Mathematical model for three-phase AC/DC converter with LC-filter. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 69. pp. 10–21. doi: 10.17223/19988605/69/2

Введение

Современные системы электроснабжения мобильных объектов представляют собой много-агентные сложно интегрированные комплексы, включающие несколько видов первичных источников электрической энергии [1–9]. Согласование характеристик первичных источников электрической энергии и нагрузки осуществляется статическими преобразователями переменного и постоянного напряжений. Для улучшения энергетических и удельных характеристик элементов систем преобразования из переменного напряжения (Alternating Current – AC) в постоянное (Direct Current – DC) применяются инверторы напряжения (ИН), работающие как активные выпрямители (АВ), т.е. AC/DC-преобразователи [10].

Функционирование активного выпрямителя направлено на поддержание заданного уровня напряжения в выходном звене постоянного напряжения при одновременном обеспечении высокого качества энергии, потребляемой от источника переменного напряжения, что достигается с помощью системы автоматического управления преобразователя.

Одним из наиболее широко распространенных вариантов регулирования выходных параметров ИН в настоящее время является использование метода широтно-импульсной модуляции (ШИМ). В подавляющем большинстве случаев применения инверторов с ШИМ используются системы автоматического управления с обратными связями по току и напряжению с различного вида регуляторами. Решение задачи синтеза закона управления и систем автоматического регулирования (САР) требует разработки математических моделей преобразователей [11–23].

Так, например, в работах [11–14] используется математическая модель АВ в синхронной вращающейся системе d - q -координат. В работе [14] геометрическая модель для больших сигналов АВ в синхронной вращающейся системе d - q -координат применена для получения естественных траекторий средних на такте ШИМ значений переменных состояния преобразователя с целью прогнозирования движения рабочей точки в переходных процессах по заранее определенным траекториям.

В работе [15] уравнения мгновенных значений токов первичного источника переменного напряжения в трехфазной системе координат используются в технологии прямого управления мощностью и моментом (Direct Power and Torque Control – DPTC) с использованием контроллера с адаптивным нейро-нечетким интерфейсом управления системой (Adaptive Neuro-Fuzzy Inference System – ANFIS) с активным выпрямителем, инвертором напряжения и индукторным двигателем.

В работе [16] выражения токов ШИМ-выпрямителя в стационарной системе α - β -координат используются для выбора наилучшего момента переключения полупроводниковых элементов АВ в алгоритме, основанном на сетях Петри с одновременным ограничением ошибки управления активной и реактивной мощностями. В работе [17] для реализации бездатчикового метода виртуального потока с применением обобщенного интегратора второго порядка также применено преобразование трехфазной математической модели АВ в стационарную систему α - β -координат.

В работе [18] дискретная математическая модель АВ, полученная с помощью преобразования Кларка, представленная в системе отсчета (α - β) и дополненная таблицей ограниченного по количеству состояний полупроводниковых элементов АВ, использована для цифровой реализации модельного прогнозирующего управления (Model Predictive Control – MPC) в предиктивном контроллере с пространственно-векторным модулятором (Space-Vector Modulator – SVM) генерации оптимального вектора напряжения.

Необходимо отметить, что в приведенных выше работах [11–18] рассматривается первичный источник трехфазного переменного напряжения без нулевого провода с постоянной частотой и симметричными напряжениями.

Несбалансированность фаз первичного источника переменного напряжения без нулевого провода с упомянутыми способами управления АВ ведет к появлению гармонических искажений во входных токах, что требует модификации законов управления [19]. В работе [19] для синтеза закона управления при несбалансированности фаз первичного источника переменного напряжения используются математические модели АВ в синхронной вращающейся системе d - q -координат для прямой и обратной последовательностей входных токов АВ. В работе [20] компоненты входного напряжения и тока АВ в стационарной ортогональной системе α - β -координат используются для вычисления корректирующих коэффициентов дополнительного внешнего контура с целью балансировки входных фазных токов. В работе [21] дискретная модель АВ в стационарной системе α - β -координат используется для вычисления желаемых мгновенных значений активной и реактивной мощностей при управлении амплитудой и фазой первой гармоники входного переменного тока АВ с целью стабилизации выходного постоянного напряжения.

В четырехпроводных системах с АВ [22–24] при небалансе напряжений первичного источника и / или нагрузки в нулевом проводе возникают токи нулевой последовательности. Поэтому математические модели АВ дополняются уравнением тока нулевого провода. В [22] используется нелинейная динамическая математическая модель преобразователя пятого порядка в синхронной системе d - q -координат с тремя контролируемыми параметрами с последующим понижением порядка до трех и процедурой линеаризации для расчета параметров ПИ-регуляторов. В [23] математическая модель расчета фазных координат и постоянных напряжений расщепленной емкости АВ используется для синтеза параметров трех скалярных пропорциональных регуляторов входных фазных токов, ПИ-регуляторов постоянного напряжения и разницы напряжений плеч емкости постоянного напряжения.

В упомянутых выше работах исследуются системы с входным индуктивным фильтром, который снижает пульсационные составляющие входного тока АВ и, соответственно, первичного источника переменного напряжения. В системах преобразования электрической энергии авиационных комплексов первичный источник переменного тока переменной частоты характеризуется не только наличием несимметрии уровней фазных напряжений и наличием нейтрали первичного источника, но и предъявлением требований к спектральному составу напряжений [25]. Наличие несимметрии напряжений и нагрузки делает рациональным применение скалярных методов независимого управления фазами АВ. Требования к спектральному составу напряжений делают необходимым повышение порядка фильтра.

Цель данной работы – разработка математической модели для описания динамических процессов по токам и напряжениям в перспективной схеме трехфазного активного выпрямителя авиационной системы электропитания. Особенность представленной математической модели состоит в учете наличия нейтралей нагрузки и первичного источника переменного тока переменной частоты с введением LC -фильтра входного переменного напряжения AC/DC преобразователя (активного выпрямителя). Разработанная математическая модель в дальнейшем может быть использована для синтеза эффективной системы управления AC/DC преобразователем в рамках решения актуальной научно-технической задачи, связанной с повышением качественных показателей преобразования электрической энергии в авиационных системах электропитания при наличии нейтрали первичного источника и нагрузки с введением LC -фильтра входного переменного напряжения преобразователя [26–27].

В настоящей работе представлена математическая модель модели трехфазного преобразователя с учетом наличия нейтралей нагрузки и первичного источника переменного тока переменной частоты

с LC-фильтром входного переменного напряжения AC/DC преобразователя (активного выпрямителя). Осуществлена верификация модели путем проведения математического моделирования и сравнения с результатами имитационного моделирования, полученных с помощью внутренних инструментов пакета ПО Matlab/Simulink и SimPowerSystem [28].

1. Математические модели активного выпрямителя

Рассматривается трехфазный двухуровневый инвертор напряжения с расцепленной емкостью, работающий в обратном режиме (активный выпрямитель), в составе бортовой авиационной системы электроснабжения с первичным четырехпроводным источником переменного тока переменной частоты, схема замещения которого представлена на рис. 1. Здесь L , R_L и C_f – сопротивление, индуктивность и емкость входного Г-образного LC-фильтра; R_f и L_f – сопротивление и индуктивность фидера сети; $e_{(A)}...e_{(C)}$ – ЭДС сети переменного тока переменной частоты; $e_{и(A)}$, $e_{и(B)}$, $e_{и(C)}$ – внутренние фазные ЭДС инвертора напряжения; $i_{C(A)}$ – ток емкости C_f фазы A ; $u_{C(A)}$ – напряжение емкости C_f фазы A ; $i_{L(A)}$ – входной переменный ток фазы A ; C_d – расцепленная емкость звена постоянного напряжения; i_{d1} и i_{d2} – выходные токи звена постоянного напряжения преобразователя; U_1 и U_2 – напряжения на плечах расцепленной емкости C_d ; U_H – постоянное напряжение нагрузки; R_H – сопротивление нагрузки; $VT_1...VT_6$ – полупроводниковые ключи в составе трехфазной мостовой схемы преобразователя. Средние точки конденсаторов C_d выходного звена постоянного напряжения и конденсаторов C_f входного Г-образного LC-фильтра соединены. Такое соединение при $C_d \gg C_f$ обеспечивает независимую работу фаз АВ.

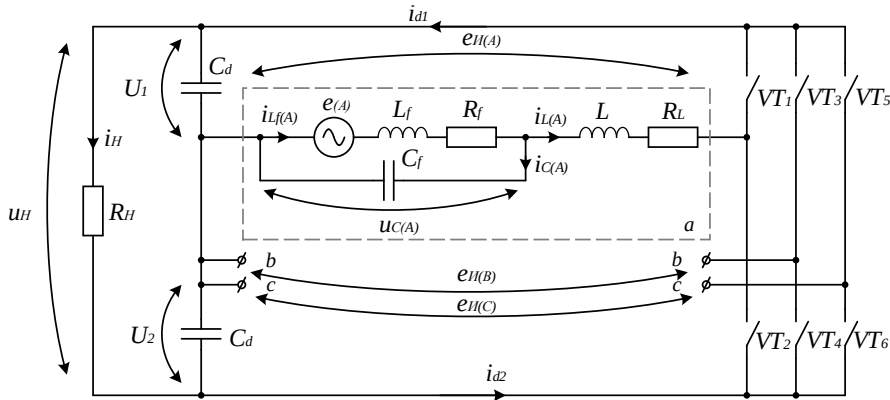


Рис. 1. Электрическая схема трехфазного активного выпрямителя
Fig. 1. Electrical schematic diagram of a three-phase active rectifier

При решении задач анализа электромагнитных процессов и синтеза САПР при скалярном пофазном управлении с ШИМ и оптимизации режимов переключения $VT_{1...6}$ в SVM необходимо получить математическую модель для описания процессов в схеме на рис. 1. При этом используются общепринятые допущения [1–24]: время переключения полупроводниковых ключей нулевое, соответственно, мертвое время (временная пауза между включенными состояниями ключей одной стойки) можно установить также нулевым; потери мощности в ключах принимаются нулевыми.

Система дифференциальных уравнений, описывающая электромагнитные процессы в трех входных фазах преобразователя, имеет следующий вид:

$$\begin{cases} e_{(j)} - L_f \frac{di_{L_f(j)}}{dt} - i_{L_f(j)} R_f - L \frac{di_{L(j)}}{dt} - i_{L(j)} R_L = e_{и(j)}, \\ e_{(j)} - L_f \frac{di_{L_f(j)}}{dt} - i_{L_f(j)} R_f = u_{C(j)}, \\ C \frac{du_{C(j)}}{dt} = i_{L_f(j)} - i_{L(j)}, \end{cases} \quad (1)$$

где $j = A, B, C$.

Для звена постоянного напряжения справедлива следующая система уравнений:

$$\begin{cases} C_d \frac{dU_1}{dt} = i_{d1} - \frac{U_1 + U_2}{R_n}, \\ C_d \frac{dU_2}{dt} = i_{d2} - \frac{U_1 + U_2}{R_n}. \end{cases} \quad (2)$$

Введем переключающие функции Ψ_{VTk} ключей VT_k , $k = 1 \dots 6$

$$\Psi_{VTk} = \begin{cases} 1, & \text{если } VT_k \text{ замкнут и проводит ток,} \\ 0, & \text{если } VT_k \text{ разомкнут и не проводит ток.} \end{cases} \quad (3)$$

Процесс формирования Ψ_{VTk} при алгоритме скалярной синусоидальной ШИМ пояснен диаграммами рис. 2 на примере верхних ключей VT_1 , VT_3 , VT_5 . Для фазы A формируется двухполярный синусоидальный сигнал, называемый модулирующим сигналом $U_{mA}(t) = M \sin(\omega t)$, частота которого ω задается равной частоте первой гармоники входного переменного напряжения АВ. Амплитуда данного сигнала, обозначенная как M , должна удовлетворять условию $0 \leq M \leq 1$, в качестве примера на рис. 2 она равна единице. Для фаз B и C независимо формируются аналогичные синусоидальные сигналы $U_{mB}(t)$ и $U_{mC}(t)$ со сдвигом фаз на 120° и 240° соответственно. Сформированные модулирующие сигналы сравниваются с опорным пилообразным двухполярным сигналом $U_{on}(t)$. Переключающие функции Ψ_{VTk} указанных ключей равны 1 при $U_{mj}(t) \geq U_{on}(t)$. Видно, что на периоде модулирующего сигнала изменяется длительность состояния, в котором $\Psi_{VTk} = 1$.

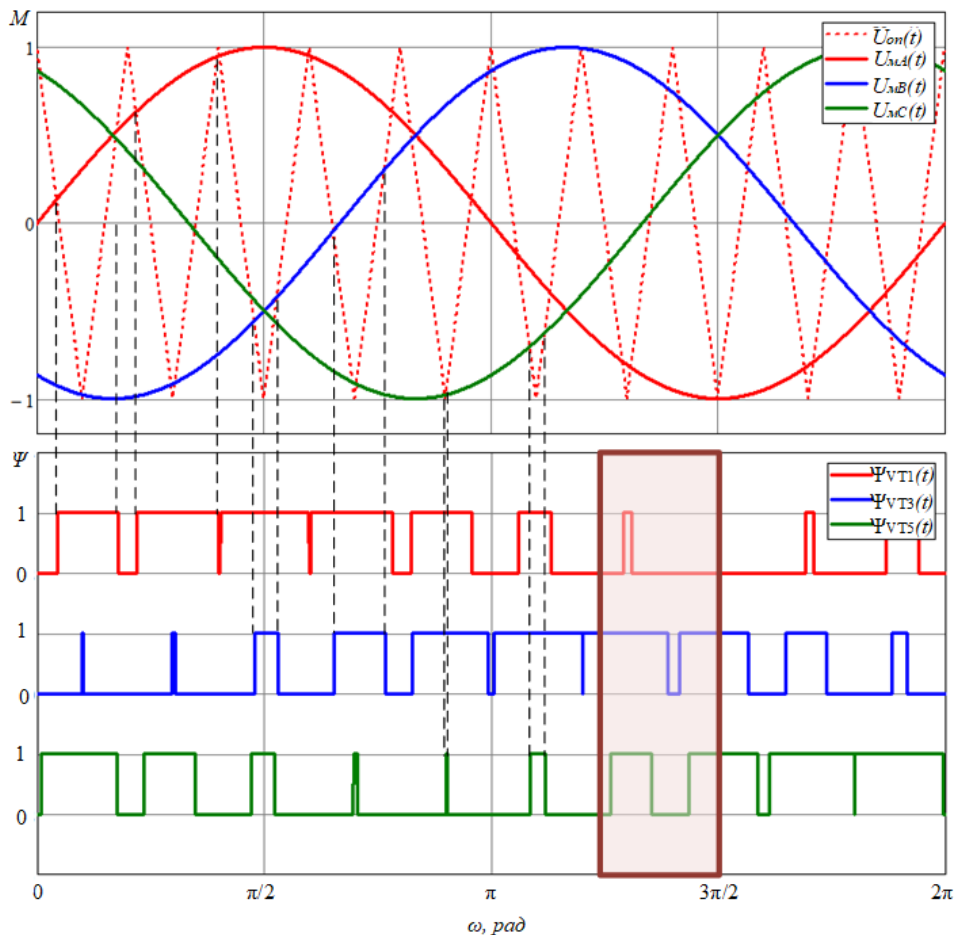


Рис. 2. Формирование переключающих функций Ψ_{VTk}

Fig. 2. Ψ_{VTk} switching functions generation process

В любой момент времени в каждой из фазных стоек последовательно соединенных ключей может быть замкнут только один ключ, и, соответственно, можно записать, что

$$\begin{cases} \Psi_{VT_1} = 1 - \Psi_{VT_2}, \\ \Psi_{VT_3} = 1 - \Psi_{VT_4}, \\ \Psi_{VT_5} = 1 - \Psi_{VT_6}. \end{cases} \quad (4)$$

Анализируя процесс работы схемы на рис. 1, системы уравнений (1) и (2) с использованием переключающих функций Ψ_{VTk} (3), (4) можно переписать следующим образом:

$$\begin{cases} L_f \frac{di_{f(A)}}{dt} = e_{(A)} - u_{C(A)} - i_{f(A)} R_f, \\ L \frac{di_{L(A)}}{dt} = -U_1 \Psi_{VT1} + U_2 \Psi_{VT2} + u_{C(A)} - i_{L(A)} R_L, \\ C \frac{du_{C(A)}}{dt} = i_{f(A)} - i_{L(A)}. \end{cases} \quad (5)$$

$$\begin{cases} L_f \frac{di_{f(B)}}{dt} = e_{(B)} - u_{C(B)} - i_{f(B)} R_f, \\ L \frac{di_{L(B)}}{dt} = -U_1 \Psi_{VT3} + U_2 \Psi_{VT4} + u_{C(B)} - i_{L(B)} R_L, \\ C \frac{du_{C(B)}}{dt} = i_{f(B)} - i_{L(B)}. \end{cases} \quad (6)$$

$$\begin{cases} L_f \frac{di_{f(C)}}{dt} = e_{(C)} - u_{C(C)} - i_{f(C)} R_f, \\ L \frac{di_{L(C)}}{dt} = -U_1 \Psi_{VT5} + U_2 \Psi_{VT6} + u_{C(C)} - i_{L(C)} R_L, \\ C \frac{du_{C(C)}}{dt} = i_{f(C)} - i_{L(C)}. \end{cases} \quad (7)$$

$$\begin{cases} C_d \frac{dU_1}{dt} = i_{L(A)} \Psi_{VT1} + i_{L(B)} \Psi_{VT3} + i_{L(C)} \Psi_{VT5} - \frac{U_1 + U_2}{R_n}, \\ C_d \frac{dU_2}{dt} = -i_{L(A)} \Psi_{VT2} - i_{L(B)} \Psi_{VT4} - i_{L(C)} \Psi_{VT6} - \frac{U_1 + U_2}{R_n}. \end{cases} \quad (8)$$

Полученные системы уравнений (5)–(8) являются математической моделью в стационарной трехфазной системе a – b – c -координат, которая позволяет провести анализ электромагнитных процессов с учетом характерных особенностей работы трехфазного управляемого моста.

В процессе работы преобразователя по схеме рис. 1 можно выделить 8 состояний полупроводниковых ключей [18]. Введем переключающие функции $\Psi_1 \dots \Psi_6$, Ψ_+ , Ψ_- по числу данных состояний. Принцип формирования этих функций пояснен на рис. 3, который представляет собой выделенный на рис. 2 участок периода модулирующего сигнала ШИМ. В верхней части рис. 3 приведена двоичная кодировка номера состояния схемы.

Так, например, при $\Psi_+ = 1$ замкнуты ключи VT_1 , VT_3 , VT_5 (схема замещения инвертора рис. 4, a , двоичный код 111), при $\Psi_- = 1$ замкнуты ключи VT_2 , VT_4 , VT_6 (схема замещения инвертора рис. 4, b , двоичный код 010).

Переключающие функции состояний схемы на периоде входного переменного напряжения удовлетворяют следующим условиям:

$$\begin{cases} \Psi_1 + \Psi_2 + \Psi_3 + \Psi_4 + \Psi_5 + \Psi_6 + \Psi_+ + \Psi_- = 1, \\ \Psi_1 \cdot \Psi_2 \cdot \Psi_3 \cdot \Psi_4 \cdot \Psi_5 \cdot \Psi_6 \cdot \Psi_+ \cdot \Psi_- = 0. \end{cases} \quad (9)$$

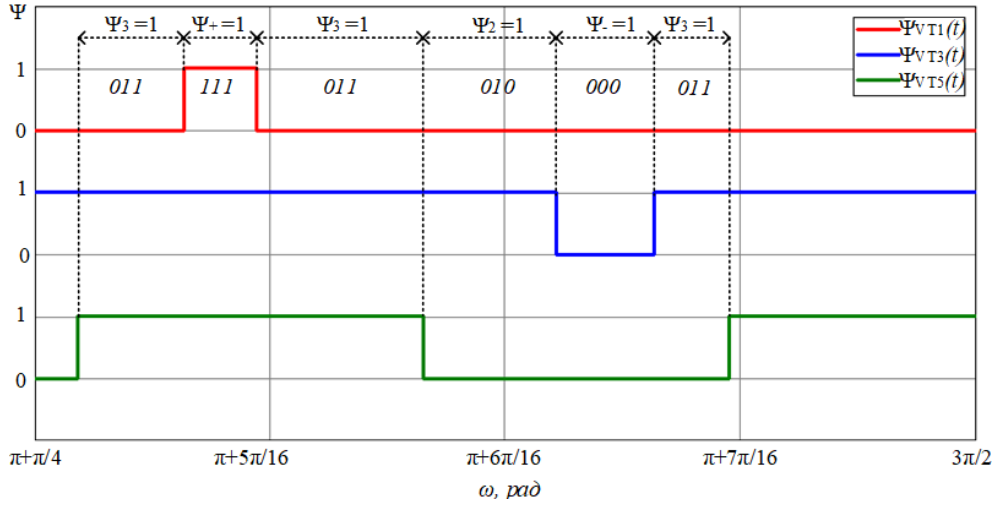

 Рис. 3. Диаграммы переключающих функций $\Psi_1, \Psi_3, \Psi_+, \Psi_-$ состояний ключей

 Fig. 3. $\Psi_1, \Psi_3, \Psi_+, \Psi_-$ switching functions of the keys status

 Рис. 4. Схема замещения инвертора (a – при $\Psi_+=1$, b – при $\Psi_2=1$)

 Fig. 4. Equivalent circuit for inverter bridge (a – for $\Psi_2=1$, b – for $\Psi_+=1$)

В свою очередь, связь переключающих функций ключей и состояний схемы можно отразить следующим образом:

$$\begin{cases} \Psi_{VT1} = \Psi_4 + \Psi_5 + \Psi_6 + \Psi_+, \\ \Psi_{VT2} = \Psi_1 + \Psi_2 + \Psi_3 + \Psi_-, \\ \Psi_{VT3} = \Psi_2 + \Psi_3 + \Psi_6 + \Psi_+, \\ \Psi_{VT4} = \Psi_1 + \Psi_4 + \Psi_5 + \Psi_-, \\ \Psi_{VT5} = \Psi_1 + \Psi_3 + \Psi_5 + \Psi_+, \\ \Psi_{VT6} = \Psi_2 + \Psi_4 + \Psi_6 + \Psi_-. \end{cases} \quad (10)$$

Используя выражения (5)–(8), (10), математическую модель AC/DC-преобразователя по схеме рис. 1 в пространстве состояний ключей схемы можно записать следующим образом:

$$L \frac{di_{L(j)}}{dt} = -U_1(\Psi_4 + \Psi_5 + \Psi_6 + \Psi_+) + U_2(\Psi_1 + \Psi_2 + \Psi_3 + \Psi_-) + (u_{C(j)} - i_{L(j)}R_L), \quad (11)$$

$$\begin{cases} C_d \frac{dU_1}{dt} = i_{L(A)}(\Psi_4 + \Psi_5 + \Psi_6 + \Psi_+) + i_{L(B)}(\Psi_2 + \Psi_3 + \Psi_6 + \Psi_+) + \\ \quad + i_{L(C)}(\Psi_1 + \Psi_3 + \Psi_5 + \Psi_+) - \frac{U_1 + U_2}{R_n}, \\ C_d \frac{dU_2}{dt} = -i_{L(A)}(\Psi_1 + \Psi_2 + \Psi_3 + \Psi_-) - i_{L(B)}(\Psi_1 + \Psi_4 + \Psi_5 + \Psi_-) - \\ \quad - i_{L(C)}(\Psi_2 + \Psi_4 + \Psi_6 + \Psi_-) - \frac{U_1 + U_2}{R_n}. \end{cases} \quad (12)$$

Полученная вторая математическая модель (11), (12) в пространстве состояний ключей схемы может быть полезна при реализации пространственно-векторных методов модуляции.

4. Верификация моделей имитационным моделированием

С целью верификации предложенного подхода к построению математических моделей электромагнитных процессов в схеме на рис. 1 проведено имитационное моделирование статического режима работы преобразователя с активной нагрузкой. Результаты имитационного моделирования сравниваются с результатами математического моделирования на основе системы уравнений (5)–(8). Параметры преобразователя, использованные при верификации, представлены в таблице.

Результаты математического и имитационного моделирования приведены на рис. 5, 6.

Параметры силовой схемы

Действующее значение фазной эдс	115 В
Частота фазной эдс	400 Гц
Частота ШИМ	100 кГц
Мощность нагрузки	1 кВт
Емкость конденсатора C_d	100 мкФ
Сопротивление нагрузки R_H	10 Ом
Сопротивление LC-фильтра R_L	1 мкОм
Индуктивность LC-фильтра L	300 мкГн
Емкость входного LC-фильтра C_f	10 мкФ
Сопротивление фидера переменного напряжения R_f	1 мкОм
Индуктивность фидера переменного напряжения L_f	1 мкГн

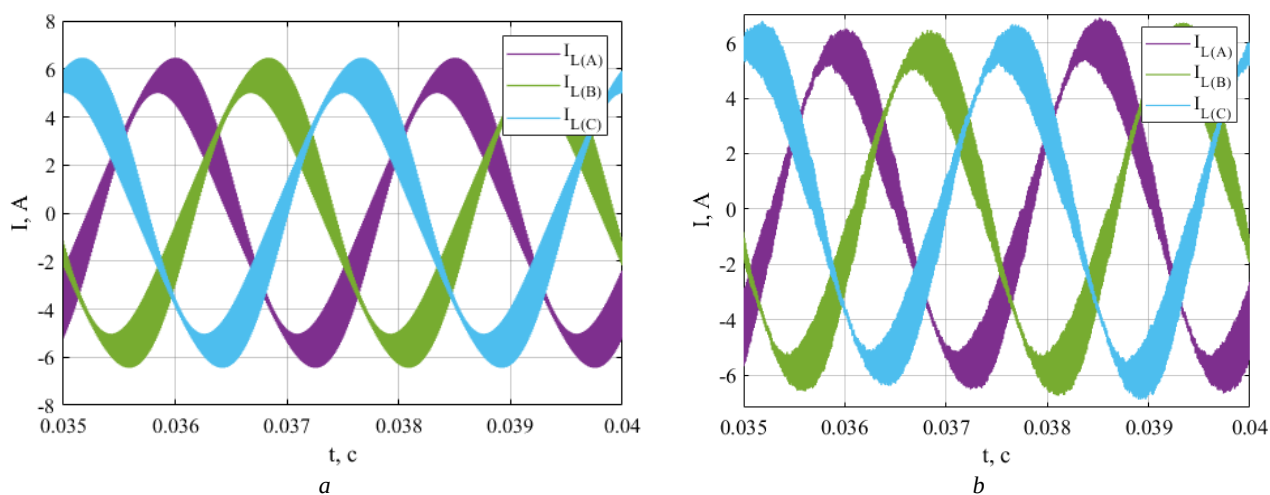


Рис. 5. Входные токи трех фаз AC/DC-преобразователя

Fig. 5. AC/DC converter 3-phase input currents

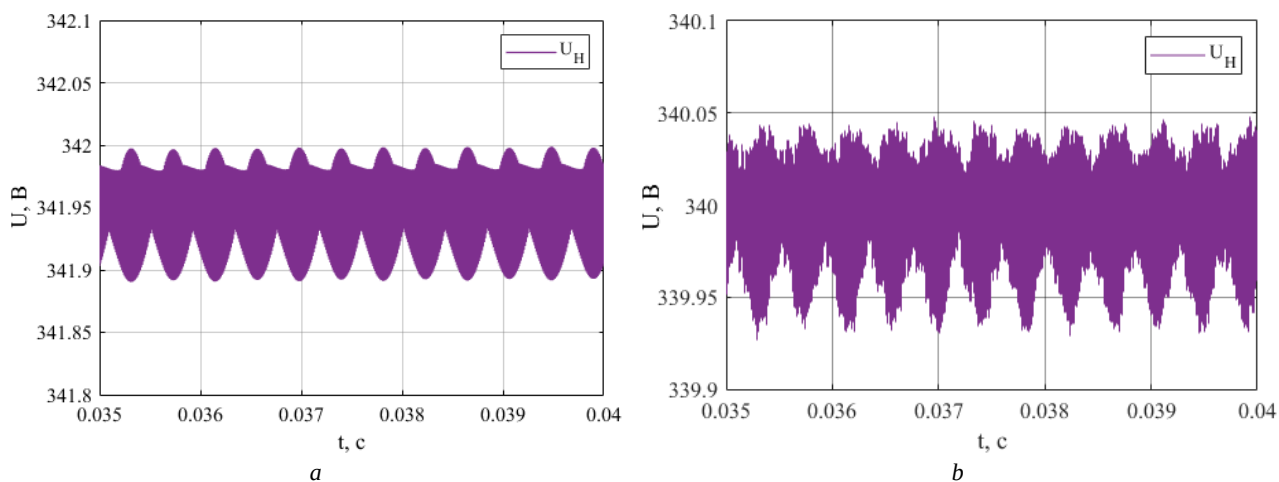


Рис. 6. Выходное выпрямленное напряжение активного выпрямителя

Fig. 6. Active rectifier output voltage

На рис. 5, *a* представлены мгновенные значения фазных переменных токов $i_{L(A)}(t) \dots i_{L(C)}(t)$, полученные с помощью выражений (5) – (8). На рис. 5, *b* – этот же ток, полученный с помощью имитационной модели пакета Matlab/Simulink и SimPowerSystem (лицензия № 355779). На рис. 6 приведены результаты математического (*a*) и имитационного (*b*) моделирования выпрямленного напряжения $U_H(t)$.

Действующее значение фазных входных переменных токов преобразователя составляет 2,94 А (см. рис. 5, *a*) и 2,97 А (см. рис. 5, *b*), среднее значение выпрямленного напряжения – 342 В (см. рис. 6, *a*) и 340 В (см. рис. 6, *b*) для математической и имитационной моделей соответственно. Форма входных токов моделей подобна в обоих моделях. Форма выходных напряжений несколько отличается, что объясняется влиянием дискретного характера численных методов расчета имитационных моделей. Таким образом, по результатам проведенного математического и имитационного моделирования можно сделать вывод, что полученные математические модели достаточно точно отражают процессы функционирования рассматриваемой схемы активного выпрямителя, расхождения в величинах токов и напряжений, полученных двумя способами, составляют менее 1%.

Заключение

В данной работе представлено два варианта математической модели для описания электромагнитных процессов в силовой схеме трехфазного двухуровневого инвертора напряжения с расщепленной емкостью звена постоянного напряжения, работающего в режиме активного выпрямителя (АВ), с входным трехфазным LC -фильтром и четырехпроводным первичным источником переменного тока переменной частоты с активной нагрузкой постоянного тока. Особенность представленных математических моделей состоит в учете наличия нейтрали первичного источника и нагрузки, а также LC -фильтра входного переменного напряжения преобразователя. Первая математическая модель преобразователя разработана в стационарной системе a – b – c -координат и может быть использована для расчета энергетических показателей и режимов работы преобразователя и при синтезе замкнутой системы автоматического управления для улучшения динамических характеристик трехфазного выпрямителя. Вторая математическая модель преобразователя записана в пространстве состояний ключей схемы и может быть применена для выбора оптимального вектора напряжения в пространственно-векторных методах модуляции. Верификация моделей проведена с помощью имитационного моделирования в Matlab инструментами библиотеки SimPowerSystem. Расчеты с помощью математических и имитационной моделей показали высокую сходимость результатов с погрешностью, не превышающей 1%.

Список источников

1. Лёвин А.В., Алексеев И.И., Харитонов С.А., Ковалёв Л.К. Электрический самолёт: от идеи до реализации. М. : Машиностроение, 2010. 288 с.
2. Rakhra P. Modelling and simulation of a MEA twin generator UAV electrical power system // 2011 46th International Universities' Power Engineering Conference (UPEC), Soest, Germany, 05–08 September 2011. P. 1–5.
3. Yanchu Li. Optimal controller design for non-affine nonlinear power systems with static var compensators for hybrid UAVs // Tsinghua Science and Technology, TUP. 2021. V. 27 (1). P.196–206. doi: 10.26599/TST.2020.9010058
4. Zhao H. Energy management strategy for hybrid-electric propulsion UAVs // 2022 IEEE Transportation Electrification Conference and Expo, Asia-Pacific (ITEC Asia-Pacific), Haining, China. 2022. P. 1–6. doi: 10.1109/ITECAsia-Pacific56316.2022.9942171
5. Wang G. A Review of power electronics for grid connection of utility-scale battery energy storage systems // IEEE Transactions on Sustainable Energy. 2016. V. 7 (4). P. 1778–1790. doi: 10.1109/TSTE.2016.2586941
6. Biya T.S., Sindhu M.R. Design and power management of solar powered electric vehicle charging station with energy storage system // 3rd International Conference on Electronics, Communication and Aerospace Technology, ICECA 2019. P. 815–820. doi: 10.1109/ICECA.2019.8821896
7. Brijesh P., Arya G.L., Anoop P., Syam V., Joseph A. Design and development of hybrid power conditioning system for microgrid application // IEEE Texas Power and Energy Conference, TPEC 2023. Memorial Student Center at Texas A&M University, College Station, Texas, USA. 2023. P. 433–438. doi: 10.1109/TPEC56611.2023.10078721
8. Bokopane L., Kusakana K., Vermaak H.J. Energy Management of a Grid-Integrated Hybrid Peer-to-Peer Renewable Charging Station for Electric Vehicles // Open Innovations Conference, OI 2019, Johannesburg, South Africa. 2018. P. 275–280. doi: 10.1109/OI.2018.8535881

9. Коробков Д.В. Использование накопителей энергии в подвижных комплексах связи // Актуальные вопросы энергетики : материалы Всерос. науч.-практ. конф. с междунар. участием. Омск : Изд-во ОмГТУ, 2023. С. 8–12.
10. Divya E., Gnanavadeivel J. Harmonic elimination in three phase PWM rectifier using FPGA control // 2011 International Conference on Emerging Trends in Electrical and Computer Technology, Nagercoil, India. 2011. P. 436–441. doi: 10.1109/ICETECT.2011.5760156
11. Bai H., Wang F., Xing J. Control strategy of combined PWM rectifier/inverter for a high speed generator power system // 2007 2nd IEEE Conference on Industrial Electronics and Applications. Harbin, China. 2007. P. 132–135. doi: 10.1109/ICIEA.2007.4318384
12. Hoang Thi, Thu Giang. Fractional proportional integral controller applied into two parallel 3-phase PWM rectifiers // 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chongqing, China. 2017. P. 9–13. doi: 10.1109/IAEAC.2017.8053966
13. Junjie Ge. Direct power control based on natural switching surface for three-phase PWM rectifiers // IEEE Transactions on Power Electronics. 2014. V. 30 (6). P. 2918–2922. doi: 10.1109/TPEL.2014.2377048
14. Degioanni F., Zurbriggen I.G., Ordonez M. Fast and reliable geometric-based controller for three-phase PWM rectifiers // 2020 IEEE Applied Power Electronics Conference and Exposition (APEC), New Orleans, LA, USA, 2020. P. 1891–1896. doi: 10.1109/APEC39645.2020.9124287
15. Jamma M., Akherraz M., Barar M. ANFIS Based DC-Link voltage control of PWM rectifier-inverter system with enhanced dynamic performance // IECON 2018 – 44th Annual Conference of the IEEE Industrial Electronics Society, 21–23 October 2018, Washington, DC, 2018. P. 2219–2224. doi: 10.1109/IECON.2018.8591620
16. Fekik A. Direct power control of a three-phase PWM-rectifier based on petri nets for the selection of switching states // 2018 7th International Conference on Renewable Energy Research and Applications (ICRERA), Paris, France. 2018. P. 1121–1125. doi: 10.1109/ICRERA.2018.8566936
17. Djabali S., Ait Hamou Ali M., Ammar A. Improved virtual flux-direct power control for PWM rectifier based on second-order generalized integrators // 2020 International Conference on Electrical Engineering (ICEE), Istanbul, Turkey. 2020. P. 267–272. doi: 10.1109/ICEE49691.2020.9249885
18. Fahem K., Chariag D.E., Sbata L. Comparative analysis of model predictive control methods for grid-connected PWM rectifier // 2020 17th International Multi-Conference on Systems, Signals & Devices (SSD), Monastir, Tunisia. 2020. P. 1058–1062. doi: 10.1109/SSD49366.2020.9364162
19. Xinhui Wu, Panda S.K., Jianxin Xu. Supply-side current harmonics control of three phase PWM boost rectifiers under distorted and unbalanced supply voltage conditions // 2007 7th International Conference on Power Electronics and Drive Systems, Bangkok, Thailand. 2007. P. 647–654. doi: 10.1109/PEDS.2007.4487770
20. Chattopadhyay S., Ramanarayanan V. A Voltage-sensorless control method to balance the input currents of a tree-wire boost rectifier under unbalanced input voltages condition // IEEE Transactions on Industrial Electronics. 2005. V. 52 (2). P. 386–398. doi: 10.1109/TIE.2005.843917
21. Bouafia A., Gaubert J.-P., Chaoui A. High performance direct power control of three-phase PWM boost rectifier under different supply voltage conditions // 2013 15th European Conference on Power Electronics and Applications (EPE), Lille, France. 2013. P. 1–8. doi: 10.1109/EPE.2013.6631776
22. Yacoubi L., Al-Haddad K., Fnaiech F., Dessaint L.-A. A DSP-based implementation of a new nonlinear control for a three-phase neutral point clamped boost rectifier prototype // IEEE Transactions on Industrial Electronics. 2005. V. 52 (1). P. 197–205. doi: 10.1109/TIE.2004.837913
23. Bueno A.G., Pomilio J.A. Balancing voltage in the DC bus with split capacitors in three-phase four-wire PWM boost rectifier // 2018 13th IEEE International Conference on Industry Applications (INDUSCON), Sao Paulo, Brazil, 2018. P. 523–529. doi: 10.1109/INDUSCON.2018.8627210
24. He L., En X., Jian X., Xinchun L., Yong K. Modeling and analysis of three-phase four-leg PWM boost-type rectifier for double conversion transformerless UPS // IECON 2011 – 37th Annual Conference of the IEEE Industrial Electronics Society, Melbourne, VIC, Australia. 2011, IEEE. P. 1444–1449.
25. ГОСТ Р 54073–2017. Системы электроснабжения самолетов и вертолетов. Общие требования и нормы качества электроэнергии. Введ. 2018-06-01. М. : Стандартинформ, 2018. 39 с.
26. Vavilov O.A., Korobkov D.V., Yurkevich V.D. Two-level voltage inverter: parametric synthesis of filter and controllers // 2022 IEEE 23rd International Conference of Young Professionals in Electron Devices and Materials (EDM), Altai, Russian Federation, 2022. P. 372–377.
27. Вавилов О.А., Юркевич В.Д., Коробков Д.В. Методика синтеза резонансного регулятора на основе метода разделения движений для инвертора напряжения // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2023. № 63. С. 4–15.
28. Фомин Д.М., Жилина Т.Е. Моделирование в MATLAB/Simulink и SCILAB/Scicos : учеб. пособие. Н. Новгород : Нижегород. гос. техн. ун-т им. Р.Е. Алексеева, 2011. 289 с.

References

1. Levin, A.V., Alekseev, I.I., Kharitonov, S.A. & Kovalev, L.K. (2010) *Elektricheskiy samolet: ot idei do realizatsii* [Electric Aircraft: From Idea to Implementation] Moscow: Mashinostroyeniye.

2. Rakhra, P. (2011) Modelling and Simulation of a MEA Twin Generator UAV Electrical Power System. *2011 46th International Universities' Power Engineering Conference (UPEC)*. Soest, Germany, 05-08 September. pp.1–5.
3. Yanchu, Li. (2022) Optimal controller design for non-affine nonlinear power systems with static var compensators for hybrid UAVs. *Tsinghua Science and Technology, TUP*. 27(1). pp. 196–206. DOI: 10.26599/TST.2020.9010058
4. Zhao, H. (2022) Energy Management Strategy for Hybrid-Electric Propulsion UAVs. *2022 IEEE Transportation Electrification Conference and Expo, Asia-Pacific (ITEC Asia-Pacific)*. Haining, China, October 28–31. pp. 1–6. DOI: 10.1109/ITECAsia-Pacific56316.2022.9942171
5. Wang, G. (2016) A Review of Power Electronics for Grid Connection of Utility-Scale Battery Energy Storage Systems. *IEEE Transactions on Sustainable Energy*. 7(4). pp. 1778–1790. DOI: 10.1109/TSTE.2016.2586941
6. Biya, T.S. & Sindhu, M.R. (2019) Design and Power Management of Solar Powered Electric Vehicle Charging Station with Energy Storage System. *3rd International conference on Electronics, Communication and Aerospace Technology, ICECA 2019*. Coimbatore, India, June 12–14, 2019. pp. 815–820. DOI: 10.1109/ICECA.2019.8821896
7. Brijesh, P., Arya, G.L., Anoop, P., Syam, V. & Joseph, A. (2023) Design and Development of Hybrid Power Conditioning System for Microgrid Application. *IEEE Texas Power and Energy Conference, TPEC 2023*. Memorial Student Center at Texas A&M University, College Station, Texas, USA. February 13 – 14. pp. 433–438. DOI: 10.1109/TPEC56611.2023.10078721
8. Bokopane, L., Kusakana, K. & Vermaak, H.J. (2018) Energy management of a grid-intergrated hybrid peer-to-peer renewable charging station for electric vehicles. *Open Innovations Conference, OI 2019*. Johannesburg, South Africa, October 3–5, 2018. pp. 275–280. DOI: 10.1109/OI.2018.8535881
9. Korobkov, D.V. (2023) Ispol'zovanie nakopiteley energii v podvizhnykh kompleksakh svyazi [Application of energy storage devices in mobile communication complexes]. *Aktual'nye voprosy energetiki* [Current Issues of Energy]. Proc. of the Conference. Omsk, May 25–26, 2023. Omsk: Omsk State Technical University. pp. 8–12.
10. Divya, E. & Gnanavadiel, J. (2011) Harmonic elimination in three phase PWM rectifier using FPGA control. *2011 International Conference on Emerging Trends in Electrical and Computer Technology*. Nagercoil, India. March 23–24, 2011. pp. 436–441. DOI: 10.1109/ICETECT.2011.5760156
11. Bai, H., Wang, F. & Xing, J. (2007) Control strategy of combined PWM rectifier/inverter for a high-speed generator power system. *2007 2nd IEEE Conference on Industrial Electronics and Applications*. Harbin, China. May 23–25, 2007. pp. 132–135. DOI: 10.1109/ICIEA.2007.4318384
12. Hoang Thi Thu Giang. (2017) Fractional proportional integral controller applied into two parallel 3-phase PWM rectifiers. *2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*. Chongqing, China. March 25–26, 2017. pp. 9–13. DOI: 10.1109/IAEAC.2017.8053966
13. Junjie, Ge. (2015) Direct power control based on natural switching surface for three-phase PWM rectifiers. *IEEE Transactions on Power Electronics*. 30(6). pp. 2918–2922. DOI: 10.1109/TPEL.2014.2377048
14. Degioanni, F., Zurbriggen, I.G. & Ordonez, M. (2020) Fast and reliable geometric-based controller for three-phase PWM rectifiers. *2020 IEEE Applied Power Electronics Conference and Exposition (APEC)*. New Orleans, LA, USA. March 15–19, 2020. pp. 1891–1896. DOI: 10.1109/APEC39645.2020.9124287
15. Jamma, M., Akherraz, M. & Barar, M. (2018) ANFIS Based DC-Link voltage control of PWM rectifier-inverter system with enhanced dynamic performance. *IECON 2018 – 44th Annual Conference of the IEEE Industrial Electronics Society*. October 21–23, 2018. pp. 2219–2224. DOI: 10.1109/IECON.2018.8591620
16. Fekik, A. (2018) Direct Power Control of a three-phase PWM-Rectifier based on Petri nets for the selection of Switching States. *2018 7th International Conference on Renewable Energy Research and Applications (ICRERA)*. Paris, France, October 14–17, 2018. pp. 1121–1125. DOI: 10.1109/ICRERA.2018.8566936
17. Djabali, S., Ait Hamou Ali, M. & Ammar, A. (2020) Improved Virtual Flux-Direct Power Control for PWM Rectifier Based on Second-Order Generalized Integrators. *2020 International Conference on Electrical Engineering (ICEE)*. Istanbul, Turkey. September 25–27, 2020. pp. 267–272. DOI: 10.1109/ICEE49691.2020.9249885
18. Fahem, K., Chariag, D.E. & Sbita, L. (2020) Comparative analysis of model predictive control methods for grid-connected PWM rectifier. *2020 17th International Multi-Conference on Systems, Signals & Devices (SSD)*. Monastir, Tunisia, July 20–23, 2020. pp. 1058–1062. DOI: 10.1109/SSD49366.2020.9364162
19. Xinhui, Wu, Panda, S.K. & Jianxin, Xu. (2007) Supply-side current harmonics control of three phase PWM boost rectifiers under distorted and unbalanced supply voltage conditions. *2007 7th International Conference on Power Electronics and Drive Systems*. Bangkok, Thailand. November 27–30, 2007. pp. 647–654. DOI: 10.1109/PEDS.2007.4487770
20. Chattopadhyay, S. & Ramanarayanan, V. (2005) A Voltage-sensorless control method to balance the input currents of a tree-wire boost rectifier under unbalanced input voltages condition. *IEEE Transactions on Industrial Electronics*. 52(2). pp. 386–398. DOI: 10.1109/TIE.2005.843917
21. Bouafia, A., Gaubert, J.-P. & Chaoui, A. (2013) High performance direct power control of three-phase PWM boost rectifier under different supply voltage conditions. *2013 15th European Conference on Power Electronics and Applications (EPE)*. Lille, France. September 2–6, 2013. pp. 1–8. DOI: 10.1109/EPE.2013.6631776
22. Yacoubi, L., Al-Haddad K., Fnaiech, F. & Dessaint, L.-A. (2005) A DSP-based implementation of a new nonlinear control for a three-phase neutral point clamped boost rectifier prototype. *IEEE Transactions on Industrial Electronics*. 52(1). pp. 197–205. DOI: 10.1109/TIE.2004.837913

23. Bueno, A.G. & Pomilio, J.A. (2018) Balancing voltage in the DC bus with split capacitors in three-phase four-wire PWM boost rectifier. *2018 13th IEEE International Conference on Industry Applications (INDUSCON)*. Sao Paulo, Brazil. November 12–14. pp. 523–529. DOI: 10.1109/INDUSCON.2018.8627210
24. He, L., En, X., Jian, X., Xinchun, L. & Yong, K. (2011) Modeling and analysis of three-phase four-leg PWM boost-type rectifier for double conversion transformerless UPS. *IECON 2011 – 37th Annual Conference of the IEEE Industrial Electronics Society*. Melbourne, VIC, Australia. November 7–10, 2011. pp. 1444–1449.
25. GOST R 54073-2017. (2018) *Sistemy elektrosnabzheniya samoletov i vertoletov. Obshchie trebovaniya i normy kachestva elektroenergii. Vved. 2018-06-01* [Power supply systems for aircraft and helicopters. General requirements and standards for power quality]. Moscow: Standartinform.
26. Vavilov, O.A., Korobkov, D.V. & Yurkevich, V.D. (2022) Two-level voltage inverter: parametric synthesis of filter and controllers. *2022 IEEE 23rd International Conference of Young Professionals in Electron Devices and Materials (EDM)*. Altai, Russian Federation, June. pp. 372–377.
27. Vavilov, O.A., Korobkov, D.V. & Yurkevich, V.D. (2023) Resonant controller design by time-scale separation method for a voltage inverter. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 63. pp. 4–15.
28. Fomin, D.M. & Zhilina, T.E. (2011) *Modelirovanie v MATLAB/Simulink i SCILAB/Scicos* [Modeling in MATLAB/Simulink and SCILAB/Scicos]. Nizhny Novgorod: Nizhny Novgorod State Technical University.

Информация об авторах:

Ашурков Никита Владимирович – магистрант Новосибирского государственного технического университета (Новосибирск, Россия). E-mail: 4445@list.ru

Коробков Дмитрий Владиславович – ведущий инженер Института силовой электроники, старший преподаватель Новосибирского государственного технического университета (Новосибирск, Россия). E-mail: korobkov@corp.nstu.ru

Вавилов Олег Андреевич – аспирант Новосибирского государственного технического университета (Новосибирск, Россия). E-mail: vavilov.oleg.rp462@gmail.com

Юркевич Валерий Дмитриевич – профессор, доктор технических наук, профессор Новосибирского государственного технического университета (Новосибирск, Россия). E-mail: yurkev@mail.ru

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

Information about the authors:

Ashurkov Nikita V. (Master's Student, Novosibirsk State Technical University Novosibirsk, Russian Federation). E-mail: 4445@list.ru

Korobkov Dmitry V. (Leading Engineer of the Institute of Power Electronics, Assistant Professor, Novosibirsk State Technical University, Novosibirsk, Russian Federation). E-mail: korobkov@corp.nstu.ru

Vavilov Oleg A. (Post-graduate Student, Novosibirsk State Technical University, Novosibirsk, Russian Federation). E-mail: vavilov.oleg.rp462@gmail.com

Yurkevich Valeriy D. (Doctor of Technical Sciences, Professor, Novosibirsk State Technical University, Novosibirsk, Russian Federation). E-mail: yurkev@mail.ru

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

Поступила в редакцию 14.08.2024; принята к публикации 02.12.2024

Received 14.08.2024; accepted for publication 02.12.2024

Original article

UDC 519.6

doi: 10.17223/19988605/69/3

**Neural network model and software for an information system
to intelligently analyze gas quality****Mais P. Farkhadov¹, Sergei V. Vaskovskii², Ivan A. Brokarev³**^{1,2} V.A. Trapeznikov Institute of Control Sciences of the Russian Academy of Sciences, Moscow, Russian Federation³ National University of Oil and Gas "Gubkin University", Moscow, Russian Federation¹ mais@ipu.ru² vask@ipu.ru³ brokarev.i@gubkin.ru

Abstract. The problem of analyzing the quality of natural gas is solved by traditional methods of gas chromatography. The article proposes an alternative approach using neural networks. An automated information system to determine energy parameters of natural gas and its operation were studied. The system testing was conducted on experimental data obtained from real gas mixtures in laboratory conditions. The gas quality indicators were calculated and the conclusion about system applicability was drawn. The developed mathematics and software allow to provide high performance for the information system in important cases where gas properties can change quickly and constant monitoring is required. Based on experimental results, an algorithmic solution was proposed for natural gas quality analysis, that allows to obtain necessary data with lower time and financial costs.

Keywords: natural gas quality analysis; assessment of analysis system accuracy; automated information system.

For citation: Farkhadov, M.P., Vaskovskii, S.V., Brokarev, I.A. (2024) Neural network model and software for an information system to intelligently analyze gas quality. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika* – Tomsk State University Journal of Control and Computer Science. 69. pp. 22–30. doi: 10.17223/19988605/69/3

Научная статья

doi: 10.17223/19988605/69/3

**Нейросетевая модель и программное обеспечение для информационной системы
интеллектуального анализа качества газа****Маис Паша оглы Фархадов¹, Сергей Владимирович Васьковский²,
Иван Андреевич Брокрев³**^{1,2} Институт проблем управления им. В.А. Трапезникова Российской академии наук, Москва, Россия³ Российский государственный университет нефти и газа (национальный исследовательский университет)

им. И.М. Губкина, Москва, Россия

¹ mais@ipu.ru² vask@ipu.ru³ brokarev.i@gubkin.ru

Аннотация. Проблема анализа качества природного газа решается традиционными методами газовой хроматографии. В статье предлагается альтернативный подход с использованием нейронных сетей. Изучены автоматизированная информационная система определения энергетических параметров природного газа и ее функционирование. Тестирование системы проводилось на экспериментальных данных, полученных на реальных газовых смесях в лабораторных условиях. Рассчитаны показатели качества газа и сделан вывод о применимости системы. Разработанное математическое и программное обеспечение позволяет обеспечить высокую производительность информационной системы в случаях, когда свойства газа могут быстро меняться и требуется постоянный мониторинг. На основе результатов экспериментов предложено алгоритмическое решение для анализа качества природного газа, позволяющее получить необходимые данные с меньшими временными и финансовыми затратами.

Ключевые слова: анализ качества природного газа; оценка точности системы анализа; автоматизированная информационная система.

Для цитирования: Фархадов М.П., Васильковский С.В., Брокарев И.А. Нейросетевая модель и программное обеспечение для информационной системы интеллектуального анализа качества газа // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 69. С. 22–30. doi: 10.17223/19988605/69/3

Introduction

To modernize oil and gas industry, as well as to solve complex industrial problems too difficult for traditional methods, there are a large number of examples of how to use neural networks in this industry [1–5]. One of the promising tasks where artificial neural networks can be used is to analyze natural gas quality. Currently, there are a large number of different methods to analyze the quality indicators of natural gas [6–9].

As a result of the study to analyze advantages and disadvantages of existing methods and automated information systems to determine gas quality indicators, the following conclusions are made. Currently, physicochemical methods of gas quality analysis and such methods-based systems prevail in industry. The main methods are gas chromatography and calorimetry to determine the energy characteristics of natural gas. Physicochemical methods to analyze the component composition and energy parameters of gas have a number of significant disadvantages. In particular, these disadvantages are high cost and large size of the equipment. Also, it takes a lot of time to analyze a single sample.

Currently, alternative analytical methods to analyze gas quality, and systems based on these methods, are being developed. However, these systems are not yet widely used. The characteristic features of systems based on analytical methods are low time costs to analyze, to use models (in particular neural networks), to obtain the required concentrations of components, to use relatively inexpensive and commercially available measuring devices.

This article discusses the problem of calculating the required quality indicators of natural gas using artificial neural networks. The implemented solution based on the information processing method provides analysis of gas quality with low time costs. The performance of the system is expressed in determining the required quality indicators using data obtained from measuring instruments and software for calculating parameters.

1. AIS architecture

The developed architecture of the gas quality analysis system determines the component composition of the equivalent pseudo-gas model and energy parameters of natural gas with low time costs. Also, the architecture uses neural network models, in particular recurrent neural networks, to find unknown concentrations of components of the equivalent pseudo-gas model if the measured properties of the gas are known. Finally, the architecture measures physical parameters of gas with commercially available and relatively inexpensive measuring devices.

The main advantages of the system are as follows. The system analyzes gas mixtures at low time costs - up to several seconds (compared to tens of minutes for a portable gas chromatograph to do the same job), and highly accurately as neural network models are used. The achieved accuracy is up to the third class (± 0.5 MJ/m³) in determining the energy parameters of gas according to the existing regulatory documents. The system measures gas mixture parameters at lesser cost as it uses commercially available measuring devices up to hundreds of thousands of rubles (the cost of a high-precision industrial chromatograph exceeds one million rubles). Finally, the system has smaller size of the measuring equipment compared to a traditional chromatograph.

To accurately analyze natural gas quality, various automated information systems (AIS) based on physical and chemical methods of analysis are currently used in industry. It should be noted that when such AIS is used, the analysis is ensured to be accurate as gas mixtures are measured directly and appropriately process the data. But such systems have the following disadvantages: significant time and economic costs

to analyze data, system development is labor intensive, and the system's equipment is costly to service. Therefore, it is relevant to develop a new method and means to process information, as well as an AIS to implement them to analyze natural gas quality at low development and operation costs.

Currently, neural networks are widely used in oil and gas industry to solve industrial problems too complex for existing methods and algorithms. Neural network models determine operational and cost indicators, control pressure in a gas distribution network, forecast natural gas viscosity, etc. Neural network technology is also relevant to develop new methods and means to process information for intelligent analysis of the quality of natural gas.

As a practical implementation of the proposed solutions, this article experimentally confirms the proposed method of information processing, in particular, how to implement an AIS to analyze gas quality. In the process of studying the system, the results obtained earlier in [10–12] are used. In particular, the developed neural network model is used, apply the multi-criteria assessment to select input parameters, use the algorithm to switch to an equivalent pseudo-gas model, assess reliability by means of a probabilistic method.

The AIS architecture functions as follows. At the process facility where it is planned to apply the developed AIS, in a bypass (a bypass pipeline of the process unit used to transport gas parallel to the shut-off and control valves) from the main process pipe, there are two measuring chambers (main and backup).

The measurement data is sent to a personal computer via a concentrator. The measurements received from the input data module are sent to an algorithm subsystem implemented on a PC in Matlab. The AIS architecture consists of three main units: an information subsystem that implements the developed algorithms, a subsystem to obtain measurement information, including measured physical parameters of gas; a subsystem to calculate energy parameters. The neural network analysis algorithm chooses network architecture and trains the network.

The measurement information subsystem consists of measuring devices and equipment that can be used to obtain values of physical parameters to evaluate how the system operates as a whole. It is worth noting that this subsystem was tested on a model that has all the properties of the proposed system. In particular, the subsystem is able to highly accurately analyze mixtures since the mixtures are prepared by mass flow controllers, and measured using commercially available and relatively inexpensive measuring devices.

The studied subsystem of measurement information also visualizes the main measured parameters, in particular, the speed of sound, thermal conductivity, and carbon dioxide concentration to clearly present the measurement process. To visualize the process, algorithms from the corresponding subsystem are used, in particular, the algorithm to visualize how the parameters of the initial mixtures and the parameters of pseudo-gas mixtures differ from one another. The proposed subsystem to obtain measurement information was tested on available measurement equipment, its main function is to provide measurement information for subsequent processing in the next subsystem.

2. Neural network architecture

The neural network algorithms generate data for the model, develop the model itself, train the model, and test the model on a data array. The data import algorithm is designed to upload data in a convenient format; the format depends on the platform from where the data is imported. The main neural network analysis algorithm includes a number of algorithms implementing the following functions: to skip a number of models if there is no need to consider them; multicriterially assess input parameters and subsequently select the most suitable set of input data; divide data into training, validation, and test samples, then choose sample size, number of splits, and the fraction of data going to each of the samples; select sets of input and output data, round them to values that can be obtained by measuring devices; calculate the matrix of correlation coefficients, normalize and cross-validate the data; implement neural network models.

For the neural network models, set parameters, including model architecture, training algorithm, training termination criterion, number of training cycles. Then calculate training errors for all models, reversely denormalize data, calculate accuracy indicators during training and testing of the neural network models and select a model for subsequent testing.

The architecture of the neural network model is a simple recurrent neural network (RNN) with one hidden layer. The number of neurons in the input layer is three according to the number of input physical parameters (sound speed, thermal conductivity, carbon dioxide concentration). The number of neurons in the output layer is three according to the number of output concentrations of pseudo-gas components except for the concentration of carbon dioxide. The number of neurons in the hidden layer is eleven. A sigmoid function in the form of a hyperbolic tangent for the hidden layer and a linear activation function for the output layer are chosen as the neuron activation functions. The structure of the developed RNN is shown in Figure 1 (n, k, m - the number of neurons in the input, hidden and output layers).

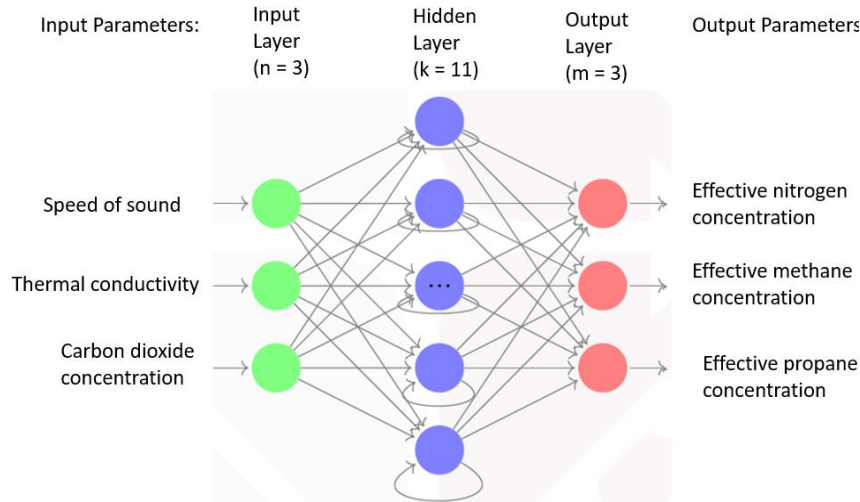


Fig. 1. Neural network model to determine an equivalent pseudogas mixture

The stage of developing a statistical model includes a number of sequential tasks that must be solved to develop the required model, namely, choosing a statistical model, choosing a model architecture, choosing model parameters, training the model, assessing the accuracy of the developed model, testing the developed model. Each of these stages is a multi-stage task, consisting of several subtasks that must be solved for the successful development of the required model. It should be noted that before training the model, the data are cross-validated, i.e. cross-validation is a method of evaluating an analytical model and its behavior on independent data.

The model is tested on a test data sample. This sample includes data that was not used to train the neural network model. The preliminary test sample includes more than four thousand gas mixtures, formed similarly to the training sample, which included approximately one hundred thousand gas mixtures.

The algorithm subsystem gives an important advantage to the proposed architecture of the AIS for gas quality analysis. The algorithms are implemented to prepare the data array and develop the neural network model. At this stage, due to the multifunctionality of the subsystem, it is possible to skip or modify some stages, and so to adjust the subsystem to a specific task. Also in this subsystem it is possible to preliminary train and test the model, multicriterially assess the input parameters, switch to the equivalent pseudo-gas model and other algorithms that make it simpler to further develop the system and provide an opportunity to test some stages of the system's operation.

To assess the accuracy of the model, as in the previous steps, the following parameters are calculated: average absolute error and average relative error. Given the fact that a model can have a satisfactory average error, but at the same time have outliers at individual points, it is also necessary to calculate the maximum absolute error and the maximum relative error.

3. Algorithm subsystem and software

This subsystem includes software with implemented algorithms for the main functions of the subsystem. In the proposed implementation of the system simulation model, the application software package for

solving technical computing problems Matlab with the NIST REFPROP plugin is used as software. As a package of application programs for implementing algorithms and calculating gas parameters in the problem under study, it is possible to use software that operates on modern operating systems, including Linux, macOS and Windows, in which it is possible to implement the given algorithms, which is an advantage of the developed system - the property of the multifunctionality of the system.

The algorithms subsystem includes the following algorithms: algorithms for generating calculated data, algorithms for conducting neural network analysis, algorithms for primary data visualization.

The algorithm for forming the component composition of a gas consists of setting the minimum and maximum (according to standards) concentration values of each component, followed by enumerating all possible variants of the component composition with setting the step for each component. The algorithm for calculating the component composition of an equivalent pseudogas model consists of using the methods described above to calculate, based on the component composition of the source gas, the component composition of a four-, five-component pseudogas model or two types of pseudogas simultaneously, which can be specified in this algorithm. The algorithm for setting temperature and pressure allows you to add pressure and temperature values in the specified range to the component composition to generate calculated data. The algorithm for calculating physical parameters and energy characteristics for a source gas and a model of an equivalent pseudogas consists of choosing the standard by which the calculation will be made, selecting the parameters to be calculated, units of measurement, generating a report on possible calculation errors and generating an array of calculated data. The diagram of the algorithm for calculating physical parameters and quality indicators in NIST REFPROP is shown in the figure 2.

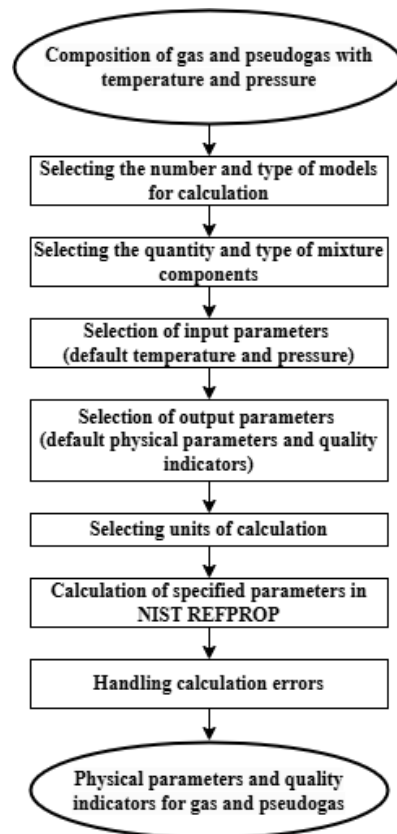


Fig. 2. Scheme of the algorithm for calculating physical parameters and quality indicators in NIST REFPROP

Algorithms that implement neural network analysis consist of generating data for the model, developing the model itself, training it, and testing it on calculated data. The data import algorithm is designed to generate a convenient for presentation type of calculated data, depending on the platform from which the import is carried out. The main algorithm of neural network analysis includes a number of algorithms that implement the following functions: multi-criteria assessment of input parameters; dividing data into training, validation,

test samples with a choice of sample size and number of splits; selection of a set of input and output data; calculation of the matrix of correlation coefficients, normalization and cross-validation of data; implementation of statistical models, calculation of accuracy indicators when training and testing statistical models. Algorithms for simulating a statistical model consist of testing the model selected in the previous step and calculating accuracy indicators at the model testing stage. Algorithms for visualizing calculated data consist of constructing graphs based on differences in the parameters of initial gas mixtures and models of equivalent pseudo-gas mixtures to calculate the deviation and its contribution to the error of the method, as well as graphs on the process of training and testing models to search for possible errors in the course of these processes and eliminate them from the implementation of the algorithms.

The subsystem of the above algorithms is an important advantage of the developed gas quality analysis system, which implements algorithms for preliminary preparation of calculated data and development of a neural network model. At this stage, due to the multifunctionality of the subsystem, it is possible to skip or modify some stages, adjusting the subsystem to a specific task. Also in this subsystem it is possible to implement preliminary training and testing of the model, multi-criteria assessment of input parameters, transition to an equivalent pseudogas model and other algorithms that simplify further development of the system and make it possible to test some stages of the system's operation.

4. Experimental testing

The subsystem of gas analysis was implemented using the same software as the algorithm subsystem. To test the proposed subsystem, a measurement experiment was conducted. In the experiment the physical parameters of gas are measured, in particular, the speed of sound, thermal conductivity coefficient and carbon dioxide concentration for the original gas mixtures and for the corresponding four- and five-component pseudo-gas mixtures. These parameters under the selected thermodynamic standard conditions were measured by two identical measuring devices to assess how reliable the obtained measurement data is. The speed of sound and thermal conductivity were measured for a matrix of gas mixtures. This matrix was built by gradually increasing the content of the components to cover the selected object, namely Russian natural gas. The studied gas mixture matrix corresponds to the ranges of components of Russian natural gas. This, in turn, means that in the experiment, all the natural gas under study was covered; this is another advantage of the system in terms of its testing coverage.

Figure 2 shows the thermal conductivity data for a methane-nitrogen gas mixture. Figure 3 shows the measurement results for the mixture of 96% methane, 4% nitrogen. It is worth noting that the jumps in graphs 3 and 4 are explained by the fact that at the beginning and end of the measurements, a check is made on pure methane to control how reliable the measurements are and to further correct the data on temperature and pressure.

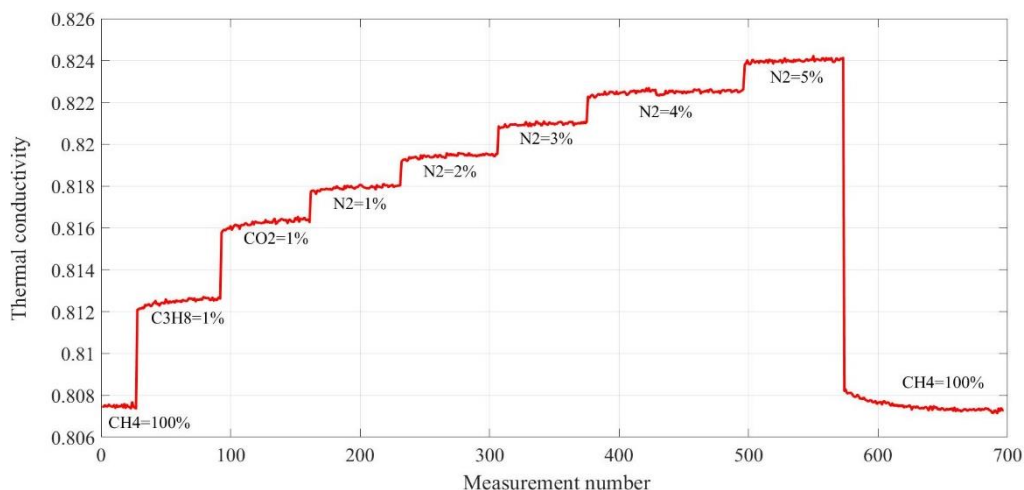


Fig. 2. The results of thermal conductivity measurements in a methane - nitrogen mixtures

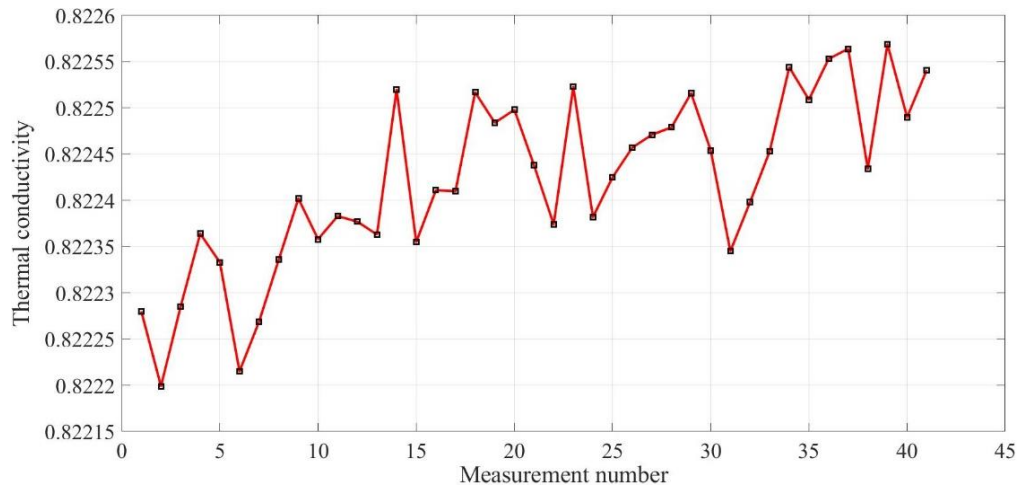


Fig. 3. The results of thermal conductivity measurements for a 96% methane, 4% nitrogen gas mixture

At the next step, the accuracy indicators are calculated of both the component composition of the equivalent pseudo-gas model and of the energy parameters of the gas. This subsystem tests the developed neural network model on the obtained experimental data after the model is corrected and the data is preliminarily processed. The testing results (with the calculated maximum absolute deviation (MAD), average absolute deviation (AAD) and root mean square deviation (RMD) are shown in Table.

Accuracy of component composition of pseudo-gas determined by neural network model on experimental data

Component	MAD, molar fraction, %	AAD, molar fraction, %	RMD
Methane	0,93	0,55	0,63
Propane	0,56	0,38	0,37
Nitrogen	0,48	0,27	0,25

Figures 4 and 5 show how accurately the energy characteristics of the gas are determined (lower volumetric calorific value and Wobbe index).

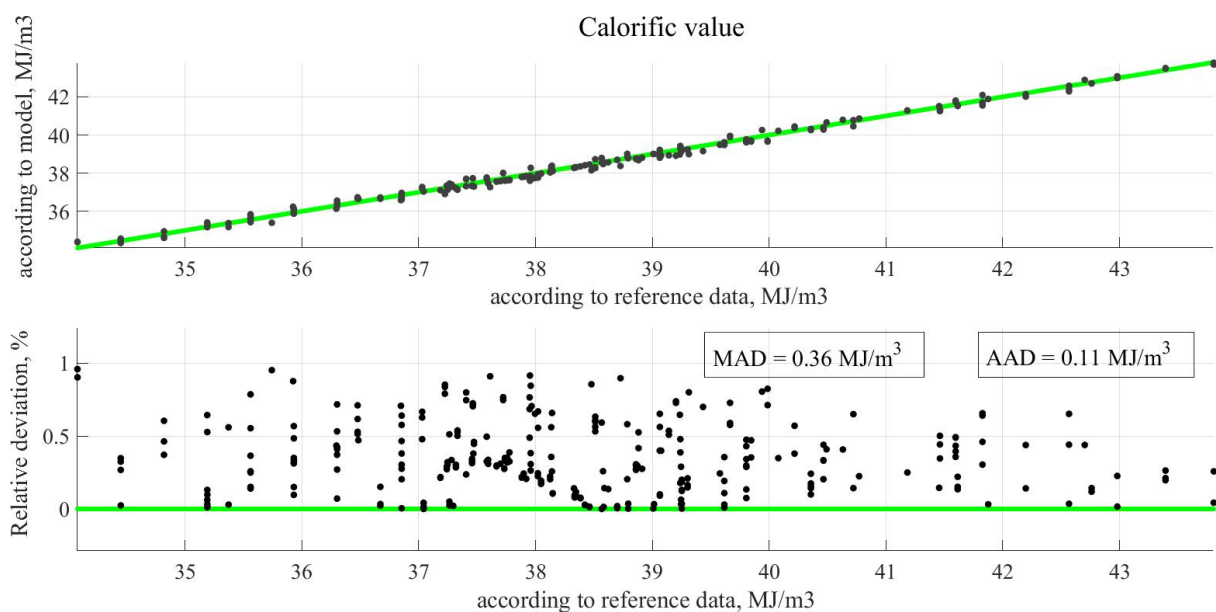


Fig. 4. Accuracy of volumetric calorific value determined by neural network on the testing stage

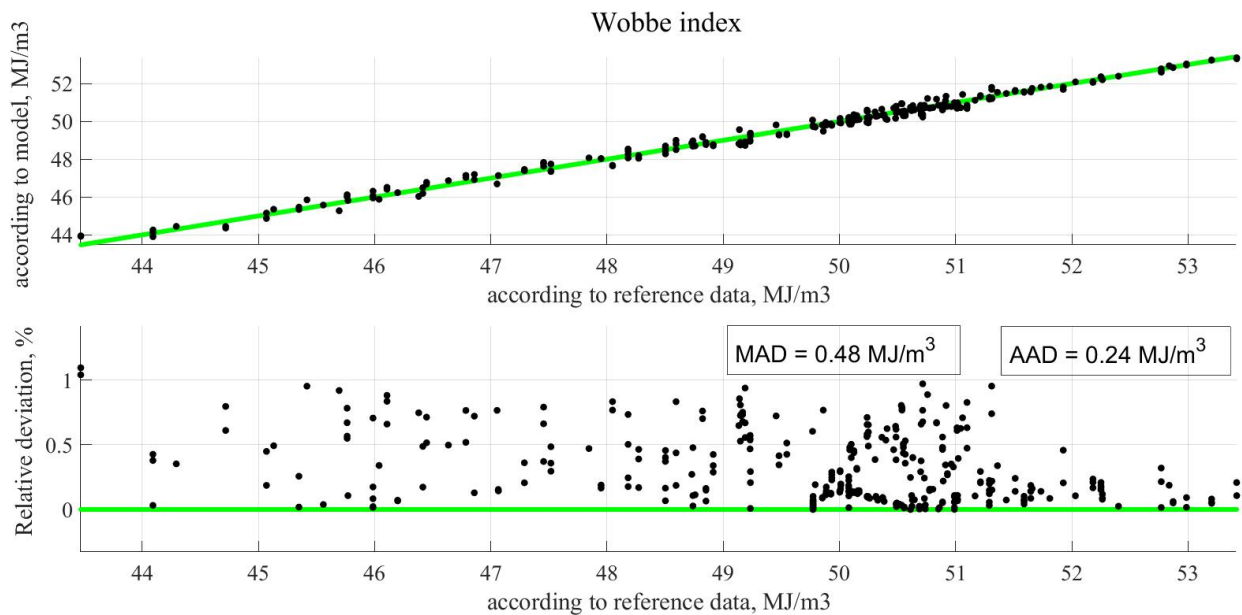


Fig. 5. Accuracy of Wobbe index value determined by neural network on the testing stage.

4. Conclusions

An automated information system is developed to determine the values of energy characteristics of natural gas based on measurements of a given set of gas physical parameters. The AIS architecture consists of an information subsystem that implements the developed algorithms, a measurement information subsystem, and an analysis subsystem. The advantage of the system is its distributed structure, which means that failure of one of the units does not lead the entire system to stop completely. The AIS operates by means of software to perform calculations and measuring devices to provide measurement information.

It is worth noting that high system performance of the system is extremely important for cases where gas properties can change quickly and need to be monitored constantly. For example, the system can monitor how associated gas from oil fields is processed, where the oil fields previously went into flares. Also, the system can monitor gas obtained by hydraulic fracturing, and biogas from various sources [13, 14].

A number of algorithms, methods, and technologies were implemented in the system. These include neural network technologies, multi-criteria assessment of input parameters, an algorithm to switch to an equivalent pseudo-gas model, and reliability assessment using a probabilistic method. All these methods also provide advantages to the system. The developed methods and architecture of the system allow us to study and analyze various modern methods to determine the quality of natural gas, to verify how equipment operates, and to experimentally study how to assess the accuracy of natural gas quality measurements. Based on the research results, an algorithmic solution was proposed, and software was implemented to operate within the existing automated information system to analyze natural gas quality with lower time and financial costs. A crucial part of further work is the calculation and study of the reliability of the automated information system [15].

References

1. Anagnostis, A., Papageorgiou, E. & Bochtis, D. (2020) Application of Artificial Neural Networks for Natural Gas Consumption Forecasting. *Sustainability*. 12. 6409. DOI: 10.3390/su12166409
2. Ashena, R., Thonhauser, G. (2015). Application of Artificial Neural Networks in Geoscience and Petroleum Industry. In: Cranganu, C., Luchian, H. & Breaban, M. (eds) *Artificial Intelligent Approaches in Petroleum Geosciences*. Springer, Cham. DOI: 10.1007/978-3-319-16531-8_4
3. Etesami, D., Zhang, W.J. & Hadian, M. (2021) A formation-based approach for modeling of rate of penetration for an offshore gas field using artificial neural networks. *Journal of Natural Gas Science and Engineering*. 95. pp. 104. DOI: 10.1016/j.jngse.2021.104104

4. Gomaa, I., Elkhatny, S. & Abdurraheem, A. (2020) Real-time determination of rheological properties of high over-balanced drilling fluid used for drilling ultra-deep gas wells using artificial neural network. *Journal of Natural Gas Science and Engineering*. 77. pp. 103. DOI: 10.1016/j.jngse.2020.103224
5. Lötters, J.C., van der Wouden, E.J., Groenesteijn, J., et al. (2014) Real-time composition determination of gas mixtures. *Proc. SENSORS*. Valencia, Spain. pp. 1640–1643.
6. Altfeld, K. & Schley, P. (2012) Development of natural gas qualities in Europe. *Heat Processing*. 3. pp. 77–83.
7. Boersma, A., Sweelssen, J. & Blokland, H. (2018) Multiparameter sensor array for gas composition monitoring. *Proc. Eurosensors*. 2(13). pp. 867.
8. Dörr, H., Koturbash, T. & Kutcherov, V. (2019) Review of impacts of gas qualities with regard to quality determination and energy metering of natural gas. *Measurement Science and Technology*. 30(2). Article ID 022001.
9. Hribar, R., Potočnik, P., Šilc, J. & Papa, G. (2018) A comparison of models for forecasting the residential natural gas demand of an urban area. *Energy*. 167. pp. 511–522.
10. Brokarev, I.A. & Vaskovskii, S.V. (2020) Multi-criteria estimation of input parameters in natural gas quality analysis. *Advances in Systems Science and Applications*. 20(2). pp. 60–69. DOI: 10.17223/19988605/55/2
11. Brokarev, I.A., Farkhadov, M.P. & Vaskovskii, S.V. (2021) Recurrent neural networks to analyze the quality of natural gas. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 55. pp. 11–17. DOI: 10.17223/19988605/55/2
12. Brokarev, I.A., Vaskovskii, S.V. & Farkhadov M.P. (2024) Automated information system for natural gas quality analysis. *Large-Scale Systems Control*. 108. pp. 174–191.
13. Koturbash, T., Bicz, A. & Kutcherov, V.G. (2021) Real-time quality metering of propanated biomethane. *International Journal of Oil, Gas and Coal Technology*. 27(1). pp. 8.
14. Wu, F., Yan, Y. & Yin, C. (2017) Real-time microseismic monitoring technology for hydraulic fracturing in shale gas reservoirs: A case study from the Southern Sichuan Basin. *Natural Gas Industry*. 4(1). pp. 68–71.
15. Kalimulina, E.Y. (2017) Analysis of system reliability with control, dependent failures, and arbitrary repair times. *International Journal of System Assurance Engineering*. 8. pp. 1–9. DOI: 10.1007/s13198-016-0520-5

Information about the authors:

Farhadov Mais Pasha ogly (Doctor of Technical Sciences, Chief Researcher, V.A. Trapeznikov Institute of Control Sciences of the Russian Academy of Sciences, Moscow, Russian Federation). Email: mais@ipu.ru

Vaskovskii Sergei V. (Candidate of Technical Sciences, Senior Researcher, V.A. Trapeznikov Institute of Control Sciences of the Russian Academy of Sciences, Moscow, Russian Federation). E-mail: vask@ipu.ru

Brokarev Ivan A. (Candidate of Technical Sciences, Senior Lecturer of the Department of Automation of Technological Processes of the Russian State University of Oil and Gas (National Research University) named after I.M. Gubkin, Moscow, Russian Federation). E-mail: brokarev.i@gubkin.ru

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

Информация об авторах:

Фархадов Маис Паша оглы – доктор технических наук, главный научный сотрудник, заведующий лабораторией Института проблем управления им. В.А. Трапезникова Российской академии наук (Москва, Россия). E-mail: mais@ipu.ru

Васьковский Сергей Владимирович – кандидат технических наук, старший научный сотрудник лаборатории Института проблем управления им. В.А. Трапезникова Российской академии наук (Москва, Россия). E-mail: vask@ipu.ru

Брокарев Иван Андреевич – кандидат технических наук, старший преподаватель кафедры автоматизации технологических процессов Российского государственного университета нефти и газа (Национальный исследовательский университет) им. И.М. Губкина» (Москва, Россия). E-mail: brokarev.i@gubkin.ru

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

Received 24.09.2024; accepted for publication 02.12.2024

Поступила в редакцию 24.09.2024; принята к публикации 02.12.2024

ОБРАБОТКА ИНФОРМАЦИИ

DATA PROCESSING

Научная статья

УДК 519.862.6

doi: 10.17223/19988605/69/4

**Интеграция ограничений на коэффициенты интеркорреляций
в оптимизационную задачу и условия построения вполне интерпретируемых
неэлементарных линейных регрессий****Михаил Павлович Базилевский***Иркутский государственный университет путей сообщения, Иркутск, Россия, mik2178@yandex.ru*

Аннотация. Сформулировано строгое определение вполне интерпретируемой неэлементарной линейной регрессии, содержащее девять условий. Для контроля мультиколлинеарности в оптимизационную задачу интегрированы линейные ограничения на интеркорреляции как между внешними регрессорами, так и между внутренними переменными из разных регрессоров. Математический аппарат был реализован в специализированной компьютерной программе. С помощью этой программы по реальным данным успешно построена вполне интерпретируемая неэлементарная линейная регрессия, удовлетворяющая всем девяти сформулированным условиям.

Ключевые слова: регрессионный анализ; неэлементарная линейная регрессия; интерпретация; кусочно-заданная функция; метод наименьших квадратов; задача частично-булевого линейного программирования; мультиколлинеарность; железнодорожные перевозки.

Для цитирования: Базилевский М.П. Интеграция ограничений на коэффициенты интеркорреляций в оптимизационную задачу и условия построения вполне интерпретируемых неэлементарных линейных регрессий // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 69. С. 31–41. doi: 10.17223/19988605/69/4

Original article

doi: 10.17223/19988605/69/4

**Integration of constraints on intercorrelation coefficients into the optimization
problem and conditions for constructing quite interpretable non-elementary
linear regressions****Mikhail P. Bazilevskiy***Irkutsk State Transport University, Irkutsk, Russian Federation, mik2178@yandex.ru*

Abstract. The article formulates a rigorous definition of quite interpretable non-elementary linear regression, containing nine conditions. To control multicollinearity, linear constraints on intercorrelations between both external regressors and internal variables from different regressors are integrated into the optimization problem. The mathematical apparatus was implemented in a specialized computer program. Using this program, a quite interpretable non-elementary linear regression has been successfully constructed using real data, satisfying all nine formulated conditions.

Keywords: regression analysis; non-elementary linear regression; interpretation; piecewise-defined function; ordinary least squares method; mixed 0-1 integer linear programming problem; multicollinearity; rail transportation.

For citation: Bazilevskiy, M.P. (2024) Integration of constraints on intercorrelation coefficients into the optimization problem and conditions for constructing quite interpretable non-elementary linear regressions. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 69. pp. 31–41. doi: 10.17223/19988605/69/4

Введение

В настоящее время с помощью моделей машинного обучения решается множество прикладных задач в самых разных областях деятельности человека (см., напр.: [1–3]). К одному из главных свойств моделей, безусловно, относится точность. Любая математическая модель с низкой точностью не имеет никакой практической пользы. При этом высокоточные модели, например в виде многослойных нейронных сетей, зачастую можно охарактеризовать как громоздкие «черные ящики», объяснить механизм функционирования которых не представляется возможным. Если такая «необъяснимая» модель применяется для решения несерьезной задачи, например для прогнозирования числа просмотров видеороликов на некотором интернет-ресурсе, то вряд ли стоит тратить время и пытаться как-то ее объяснить, ведь главное, чтобы она давала точные прогнозы. Но если решается серьезная задача, например с помощью модели машинного обучения ставится медицинский диагноз человеку либо выносится судебное решение, то одной только точности недостаточно. Нужно еще доказать, что модель корректно работает во всех возможных ситуациях и не будет давать сбоев. Иными словами, нужно повысить доверие к модели, т.е. сделать так, чтобы у специалистов не возникло никаких сомнений относительно ее внедрения в производственный процесс. Для этого требуется модель интерпретировать, что означает объяснить механизм ее функционирования в понятных для человека терминах. Интерпретируемое машинное обучение, возникшее относительно недавно [4, 5], является актуальным научным направлением на сегодняшний день [6, 7].

В монографии К. Молнара [5] к интерпретируемым моделям машинного обучения относятся следующие виды регрессионных моделей: линейные [8] и логистические регрессии [9], а также обобщенные линейные [10] и аддитивные модели [11]. При этом универсального строгого математического определения интерпретируемой модели пока не существует. В [12] автором были рассмотрены вопросы построения вполне интерпретируемых линейных регрессий (ВИЛинР), причем идентификация ВИЛинР осуществлялась с помощью метода наименьших квадратов (МНК) [13], а для отбора информативных регрессоров (ОИР) [14, 15] использовался аппарат частично-булевого линейного программирования (ЧБЛП). Ограничения в задаче ЧБЛП ставились на знаки коэффициентов регрессии и абсолютные вклады переменных в общую детерминацию, что позволило эффективно бороться с мультиколлинеарностью [16, 17]. В [18] в задаче построения ВИЛинР появились дополнительные ограничения на коэффициенты интеркорреляций. Строгое математическое определение вполне интерпретируемых квазилинейных регрессий (ВИКЛинР), обобщающих ВИЛинР, появилось в [19].

Процесс поиска новых хорошо интерпретируемых математических форм связи между переменными продолжается. Так, в [20] с использованием бинарных операций $\min$ и $\max$ введены неэлементарные линейные регрессии (НЛинР). Там же задача ОИР в НЛинР была формализована в виде задачи ЧБЛП. Достоинство таких моделей в том, что они легко интерпретируются в виде кусочно-заданных функций. Поэтому в [21] задача ЧБЛП была расширена ограничениями, с помощью которых можно контролировать абсолютные вклады переменных в общую детерминацию. Полученное в результате ее решения уравнение называется вполне интерпретируемой неэлементарной линейной регрессией (ВИНЛинР). Однако строгого математического определения ВИНЛинР пока еще дано не было. К тому же в задаче ЧБЛП для построения ВИНЛинР не были задействованы ограничения на мультиколлинеарность, что все же не позволяет полностью относить такие модели ко вполне интерпретируемым. Цель данной статьи заключается в решении двух указанных проблем.

1. Вполне интерпретируемые неэлементарные линейные регрессии

Пусть имеется выборочная совокупность объема n , содержащая значения зависимой (объясняемой) переменной y и l независимых (объясняющих) переменных $x_1, x_2, \dots, x_l$. Рассмотрим введенную в [20] НЛинР с неизвестными параметрами $\alpha_0, \alpha_j, j = \overline{1, l}, \alpha_j^{\min}, \alpha_j^{\max}, k_j^{\min}, k_j^{\max}, j = \overline{1, C_l^2}$, вида

$$y_i = \alpha_0 + \sum_{j=1}^l \alpha_j x_{i,j} + \sum_{j=1}^{C_l^2} \alpha_j^{\min} \min \{x_{i,\mu_{j,1}}, k_j^{\min} x_{i,\mu_{j,2}}\} + \sum_{j=1}^{C_l^2} \alpha_j^{\max} \max \{x_{i,\mu_{j,1}}, k_j^{\max} x_{i,\mu_{j,2}}\} + \varepsilon_i, \quad i = \overline{1, n}, \quad (1)$$

где C_l^2 – число сочетаний без повторений из l элементов по 2; $\min, \max$ – бинарные операции, возвращающие минимум и максимум двух чисел соответственно; $\mu_{j,1}, \mu_{j,2}, j = \overline{1, C_l^2}$ – элементы матрицы M размера $C_l^2 \times 2$, содержащей в строках в лексикографическом порядке все возможные комбинации пар индексов объясняющих переменных; $\varepsilon_i, i = \overline{1, n}$ – ошибки аппроксимации; $x_{ij} \neq 0, i = \overline{1, n}, j = \overline{1, l}$.

Как отмечено в [20], области возможных значений оценок параметров $k_j^{\min}, k_j^{\max}, j = \overline{1, C_l^2}$, можно записать в виде:

$$k_j^{\min}, k_j^{\max} \in (k_j^{\text{нижн}}, k_j^{\text{верхн}}), \quad j = \overline{1, C_l^2}, \quad (2)$$

$$\text{где } k_j^{\text{нижн}} = \min \left\{ \frac{x_{1,\mu_{j,1}}}{x_{1,\mu_{j,2}}}, \dots, \frac{x_{n,\mu_{j,1}}}{x_{n,\mu_{j,2}}} \right\}, \quad k_j^{\text{верхн}} = \max \left\{ \frac{x_{1,\mu_{j,1}}}{x_{1,\mu_{j,2}}}, \dots, \frac{x_{n,\mu_{j,1}}}{x_{n,\mu_{j,2}}} \right\}.$$

Равномерно разбивая промежутки (2) p точками [20], перейдем от НЛинР (1) к линейной по параметрам регрессии:

$$y_i = \alpha_0 + \sum_{j=1}^l \alpha_j x_{ij} + \sum_{j=1}^{C_l^2} \sum_{k=1}^p \alpha_{jk}^- z_{ijk}^- + \sum_{j=1}^{C_l^2} \sum_{k=1}^p \alpha_{jk}^+ z_{ijk}^+ + \varepsilon_i, \quad i = \overline{1, n}, \quad (3)$$

где $z_{ijk}^- = \min \{x_{i,\mu_{j,1}}, \lambda_{jk} x_{i,\mu_{j,2}}\}, z_{ijk}^+ = \max \{x_{i,\mu_{j,1}}, \lambda_{jk} x_{i,\mu_{j,2}}\}, i = \overline{1, n}, j = \overline{1, C_l^2}, k = \overline{1, p}; \lambda_{jk}, j = \overline{1, C_l^2}, k = \overline{1, p}$ – элементы матрицы Λ размера $C_l^2 \times p$, каждый из которых равен k -й точке разбиения j -го промежутка (2); $\alpha_{jk}^-, \alpha_{jk}^+, j = \overline{1, C_l^2}, k = \overline{1, p}$ – неизвестные параметры.

Используя модель (3), введем строгое определение ВИНЛинР.

Определение. НЛинР (1), оцененная с помощью МНК, называется вполне интерпретируемой, если:

- 1) каждая объясняющая переменная входит в модель не более одного раза;
- 2) знаки всех коэффициентов корреляции $r_{yx_j}, j = \overline{1, l}, r_{yz_{jk}^-}, r_{yz_{jk}^+}, j = \overline{1, C_l^2}, k = \overline{1, p}$, удовлетворяют содержательному смыслу решаемой задачи;
- 3) знаки всех коэффициентов представленной в виде кусочно-заданной функции модели согласуются со знаками соответствующих коэффициентов корреляции r_{yx_j} ;
- 4) все абсолютные вклады регрессоров в общую детерминацию R^2 не ниже выбранного числа $\theta \in [0, 1)$;
- 5) все оценки $\tilde{\alpha}_j, \tilde{\alpha}_j^{\min}, \tilde{\alpha}_j^{\max}$ значимы по t -критерию Стьюдента для выбранного уровня значимости α ;
- 6) все интеркорреляции как между внешними регрессорами, так и между внутренними переменными из разных регрессоров незначимы по t -критерию Стьюдента для выбранного уровня значимости α ;

- 7) все вспомогательные зависимости каждого регрессора от остальных незначимы по F -критерию Фишера для выбранного уровня значимости α ;
- 8) модель в целом значима по F -критерию Фишера для выбранного уровня значимости α ;
- 9) коэффициент детерминации модели $R^2 \geq 0,8$.

Прокомментируем далее некоторые условия сформулированного определения.

Условие 2 можно проверить еще до оценивания НЛинР. Для этого требуется участие экспертов из данной предметной области. Их задача состоит в анализе соответствия содержательному смыслу знаков коэффициентов корреляции между регрессорами и переменной y . Если выявлены несоответствия, то использовать такие регрессоры в процессе оценивания модели не следует.

Заметим, что если эксперты согласовали корреляции между объясняющими переменными и y , то согласование корреляций между y и преобразованными переменными $z_{jk}^-, z_{jk}^+, j = \overline{1, C_l^2}, k = \overline{1, p}$, можно автоматизировать. Например, преобразование $\min\{x_i, k \cdot x_j\}$ не нужно исключать, если для него выполняется одно из четырех условий:

- 1) $k > 0, r_{yx_i} > 0, r_{yx_j} > 0, r_{y, \min\{x_i, kx_j\}} > 0$;
- 2) $k > 0, r_{yx_i} < 0, r_{yx_j} < 0, r_{y, \min\{x_i, kx_j\}} < 0$;
- 3) $k < 0, r_{yx_i} > 0, r_{yx_j} < 0, r_{y, \min\{x_i, kx_j\}} > 0$;
- 4) $k < 0, r_{yx_i} < 0, r_{yx_j} > 0, r_{y, \min\{x_i, kx_j\}} < 0$.

Условие 3 равносильно выполнению неравенств $\alpha_j \cdot r_{yx_j} > 0, j = \overline{1, l}; \alpha_{jk}^- \cdot r_{y, z_{jk}^-} > 0, \alpha_{jk}^+ \cdot r_{y, z_{jk}^+} > 0, j = \overline{1, C_l^2}, k = \overline{1, p}$.

Фигурирующие в условии 4 абсолютные вклады регрессоров в общую детерминацию для модели (3) находятся по формулам $C_{x_j}^{abc} = r_{yx_j} \cdot \beta_j, j = \overline{1, l}; C_{z_{jk}^-}^{abc} = r_{yz_{jk}^-} \cdot \beta_{jk}^-, C_{z_{jk}^+}^{abc} = r_{yz_{jk}^+} \cdot \beta_{jk}^+, j = \overline{1, C_l^2}, k = \overline{1, p}$, где $\beta_j, j = \overline{1, l}; \beta_{jk}^-, \beta_{jk}^+, j = \overline{1, C_l^2}, k = \overline{1, p}$ – неизвестные параметры стандартизованной регрессии вида

$$y_i^* = \sum_{j=1}^l \beta_j x_{ij}^* + \sum_{j=1}^{C_l^2} \sum_{k=1}^p \beta_{jk}^- z_{ijk}^* + \sum_{j=1}^{C_l^2} \sum_{k=1}^p \beta_{jk}^+ z_{ijk}^{**} + \varepsilon_i^*, i = \overline{1, n}, \quad (4)$$

в которой $y_i^* = \frac{y_i - \bar{y}}{\sigma_y}, i = \overline{1, n}; x_{ij}^* = \frac{x_{ij} - \bar{x}_j}{\sigma_{x_j}}, i = \overline{1, n}, j = \overline{1, l}; z_{ijk}^* = \frac{z_{ijk}^- - \bar{z}_{jk}^-}{\sigma_{z_{jk}^-}}, z_{ijk}^{**} = \frac{z_{ijk}^+ - \bar{z}_{jk}^+}{\sigma_{z_{jk}^+}}, j = \overline{1, C_l^2},$

$k = \overline{1, p}; \varepsilon_i^*, i = \overline{1, n}$ – новые ошибки аппроксимации.

При этом справедливо следующее равенство:

$$\sum_{j=1}^l C_{x_j}^{abc} + \sum_{j=1}^{C_l^2} \sum_{k=1}^p C_{z_{jk}^-}^{abc} + \sum_{j=1}^{C_l^2} \sum_{k=1}^p C_{z_{jk}^+}^{abc} = R^2,$$

где R^2 – коэффициент детерминации, оцененной с помощью МНК регрессии (3) или (4).

Условия 6 и 7 отвечают за контроль эффекта мультиколлинеарности в модели.

Поясним условие 6 на простом примере. Пусть оцененная НЛинР имеет вид:

$$\tilde{y} = 2 + 3x_1 + 4 \min\{x_2, 8x_5\} + 6 \max\{x_3, 2x_4\}.$$

Тогда интеркорреляциями между внешними регрессорами являются коэффициенты корреляции между x_1 и $\min\{x_2, 8x_5\}$, x_1 и $\max\{x_3, 2x_4\}$, $\min\{x_2, 8x_5\}$ и $\max\{x_3, 2x_4\}$. А интеркорреляциями между внутренними переменными из разных регрессоров являются коэффициенты корреляции между x_1 и x_2 , x_1 и x_5 , x_1 и x_3 , x_1 и x_4 , x_2 и x_3 , x_2 и x_4 , x_5 и x_3 , x_5 и x_4 .

Если в условии 6 выбран уровень значимости α , то, как отмечено в [21], интеркорреляции будут незначимы по t -критерию Стьюдента тогда, когда все они по абсолютной величине будут

меньше числа $r = \frac{t_{\text{крит}}(\alpha, n-2)}{\sqrt{n-2+t_{\text{крит}}^2(\alpha, n-2)}}$, где $t_{\text{крит}}(\alpha, n-2)$ – критическая точка распределения Стюдента.

Как видно, представленное определение ВИНЛинР носит субъективный характер, поскольку окончательный выбор числа θ в условии 4, уровней значимости α в условиях 5, 6, 7, 8 и даже границы R^2 в условии 9 остается за исследователем. Понятно, что от этого выбора будут зависеть результаты моделирования.

Перейдем к формализации задачи ОИР в стандартизованной регрессии (4) с помощью аппарата математического программирования. Решение этой задачи, если оно существует, должно гарантировать выполнение условий вышеприведенного определения ВИНЛинР. Для этого введем бинарные переменные

$$\delta_j \in \{0,1\}, \quad j = \overline{1, l}; \quad \delta_{jk}^-, \delta_{jk}^+ \in \{0,1\}, \quad j = \overline{1, C_l^2}, \quad k = \overline{1, p}, \quad (5)$$

отвечающие за вхождения регрессоров в модель (4). Например, если $\delta_{jk}^- = 1$, то регрессор z_{jk}^- входит в регрессию, а если $\delta_{jk}^- = 0$, то нет.

Для выполнения условия 1 нужно ввести ограничения

$$\delta_j + \sum_{i \in Y_j} \sum_{k=1}^p \delta_{ik}^- + \sum_{i \in Y_j} \sum_{k=1}^p \delta_{ik}^+ \leq 1, \quad j = \overline{1, l}, \quad (6)$$

где Y_j – множество номеров строк матрицы M , содержащих среди своих элементов число j .

Условие 2, как уже отмечалось, проверяется до решения задачи.

Выполнение условия 3 гарантируют ограничения

$$0 \leq \beta_j \leq M \cdot \delta_j, \quad j \in \{s | r_{yx_s} > 0\}, \quad (7)$$

$$-M \cdot \delta_j \leq \beta_j \leq 0, \quad j \in \{s | r_{yx_s} < 0\}, \quad (8)$$

$$0 \leq \beta_{jk}^- \leq M \cdot \delta_{jk}^-, \quad (j, k) \in \{(s_1, s_2) | r_{yz_{s_1 s_2}^-} > 0\}, \quad (9)$$

$$-M \cdot \delta_{jk}^- \leq \beta_{jk}^- \leq 0, \quad (j, k) \in \{(s_1, s_2) | r_{yz_{s_1 s_2}^-} < 0\}, \quad (10)$$

$$0 \leq \beta_{jk}^+ \leq M \cdot \delta_{jk}^+, \quad (j, k) \in \{(s_1, s_2) | r_{yz_{s_1 s_2}^+} > 0\}, \quad (11)$$

$$-M \cdot \delta_{jk}^+ \leq \beta_{jk}^+ \leq 0, \quad (j, k) \in \{(s_1, s_2) | r_{yz_{s_1 s_2}^+} < 0\}, \quad (12)$$

где M – большое положительное число, процедура выбора которого подробно описана в [20].

Условие 4 контролируется с помощью ограничений

$$r_{yx_j} \cdot \beta_j \geq \theta \cdot \delta_j, \quad j = \overline{1, l}, \quad (13)$$

$$r_{yz_{jk}^-} \cdot \beta_{jk}^- \geq \theta \cdot \delta_{jk}^-, \quad j = \overline{1, C_l^2}, \quad k = \overline{1, p}, \quad (14)$$

$$r_{yz_{jk}^+} \cdot \beta_{jk}^+ \geq \theta \cdot \delta_{jk}^+, \quad j = \overline{1, C_l^2}, \quad k = \overline{1, p}, \quad (15)$$

в которых число $\theta \in [0,1)$ рекомендуется брать не ниже 0,05.

Для контроля интеркорреляций между внешними регрессорами в условии 6 введем ограничения

$$|r_{x_i x_j}| \cdot (\delta_i + \delta_j - 1) \leq r, \quad i = \overline{1, l-1}, \quad j = \overline{i+1, l}, \quad (16)$$

$$|r_{x_j z_{sk}^-}| \cdot (\delta_j + \delta_{sk}^- - 1) \leq r, \quad |r_{x_j z_{sk}^+}| \cdot (\delta_j + \delta_{sk}^+ - 1) \leq r, \quad j = \overline{1, l}, \quad s \in Y \setminus Y_j, \quad k = \overline{1, p}, \quad (17)$$

$$|r_{z_{s_1 k_1}^- z_{s_2 k_2}^-}| \cdot (\delta_{s_1 k_1}^- + \delta_{s_2 k_2}^- - 1) \leq r, \quad |r_{z_{s_1 k_1}^+ z_{s_2 k_2}^+}| \cdot (\delta_{s_1 k_1}^+ + \delta_{s_2 k_2}^+ - 1) \leq r, \quad j_1 = \overline{1, l-1}, \quad j_2 = \overline{j_1+1, l}, \quad (18)$$

$$s_1 \in Y_{j_1} \cap Y_{j_2}, \quad k_1 = \overline{1, p}, \quad s_2 \in Y \setminus (Y_{j_1} \cup Y_{j_2}), \quad s_1 < s_2, \quad k_2 = \overline{1, p},$$

$$\begin{aligned} & \left| r_{z_{s_1 k_1} z_{s_2 k_2}^+} \right| \cdot (\delta_{s_1 k_1}^- + \delta_{s_2 k_2}^+ - 1) \leq r, \quad j_1 = \overline{1, l-1}, \quad j_2 = \overline{j_1+1, l}, \\ & s_1 \in Y_{j_1} \cap Y_{j_2}, \quad k_1 = \overline{1, p}, \quad s_2 \in Y \setminus (Y_{j_1} \cup Y_{j_2}), \quad k_2 = \overline{1, p}, \end{aligned} \quad (19)$$

где Y – множество номеров строк матрицы M .

Подобный прием для линейной регрессии можно найти в [18]. Ограничения (16) предназначены для контроля интеркорреляций между объясняющими переменными, (17) – между объясняющими переменными и бинарными операциями $\min$ и $\max$, (18) – между бинарными операциями $\min$ и $\min$, $\max$ и $\max$, (19) – между бинарными операциями $\min$ и $\max$.

Для контроля интеркорреляций между внутренними переменными из разных регрессоров нужно использовать ограничения

$$\left| r_{x_j x_{\mu s_1}} \right| \cdot (\delta_j + \delta_{sk}^- - 1) \leq r, \quad \left| r_{x_j x_{\mu s_2}} \right| \cdot (\delta_j + \delta_{sk}^- - 1) \leq r, \quad j = \overline{1, l}, \quad s \in Y \setminus Y_j, \quad k = \overline{1, p}, \quad (20)$$

$$\left| r_{x_j x_{\mu s_1}} \right| \cdot (\delta_j + \delta_{sk}^+ - 1) \leq r, \quad \left| r_{x_j x_{\mu s_2}} \right| \cdot (\delta_j + \delta_{sk}^+ - 1) \leq r, \quad j = \overline{1, l}, \quad s \in Y \setminus Y_j, \quad k = \overline{1, p}, \quad (21)$$

$$\begin{aligned} & \left| r_{x_{\mu s_1, 1} x_{\mu s_2, 1}} \right| \cdot (\delta_{s_1 k_1}^- + \delta_{s_2 k_2}^- - 1) \leq r, \quad \left| r_{x_{\mu s_1, 1} x_{\mu s_2, 2}} \right| \cdot (\delta_{s_1 k_1}^- + \delta_{s_2 k_2}^- - 1) \leq r, \\ & \left| r_{x_{\mu s_1, 2} x_{\mu s_2, 1}} \right| \cdot (\delta_{s_1 k_1}^- + \delta_{s_2 k_2}^- - 1) \leq r, \quad \left| r_{x_{\mu s_1, 2} x_{\mu s_2, 2}} \right| \cdot (\delta_{s_1 k_1}^- + \delta_{s_2 k_2}^- - 1) \leq r, \\ & j_1 = \overline{1, l-1}, \quad j_2 = \overline{j_1+1, l}, \quad s_1 \in Y_{j_1} \cap Y_{j_2}, \quad k_1 = \overline{1, p}, \quad s_2 \in Y \setminus (Y_{j_1} \cup Y_{j_2}), \quad s_1 < s_2, \quad k_2 = \overline{1, p}, \end{aligned} \quad (22)$$

$$\begin{aligned} & \left| r_{x_{\mu s_1, 1} x_{\mu s_2, 1}} \right| \cdot (\delta_{s_1 k_1}^+ + \delta_{s_2 k_2}^+ - 1) \leq r, \quad \left| r_{x_{\mu s_1, 1} x_{\mu s_2, 2}} \right| \cdot (\delta_{s_1 k_1}^+ + \delta_{s_2 k_2}^+ - 1) \leq r, \\ & \left| r_{x_{\mu s_1, 2} x_{\mu s_2, 1}} \right| \cdot (\delta_{s_1 k_1}^+ + \delta_{s_2 k_2}^+ - 1) \leq r, \quad \left| r_{x_{\mu s_1, 2} x_{\mu s_2, 2}} \right| \cdot (\delta_{s_1 k_1}^+ + \delta_{s_2 k_2}^+ - 1) \leq r, \\ & j_1 = \overline{1, l-1}, \quad j_2 = \overline{j_1+1, l}, \quad s_1 \in Y_{j_1} \cap Y_{j_2}, \quad k_1 = \overline{1, p}, \quad s_2 \in Y \setminus (Y_{j_1} \cup Y_{j_2}), \quad s_1 < s_2, \quad k_2 = \overline{1, p}, \end{aligned} \quad (23)$$

$$\begin{aligned} & \left| r_{x_{\mu s_1, 1} x_{\mu s_2, 1}} \right| \cdot (\delta_{s_1 k_1}^- + \delta_{s_2 k_2}^+ - 1) \leq r, \quad \left| r_{x_{\mu s_1, 1} x_{\mu s_2, 2}} \right| \cdot (\delta_{s_1 k_1}^- + \delta_{s_2 k_2}^+ - 1) \leq r, \\ & \left| r_{x_{\mu s_1, 2} x_{\mu s_2, 1}} \right| \cdot (\delta_{s_1 k_1}^- + \delta_{s_2 k_2}^+ - 1) \leq r, \quad \left| r_{x_{\mu s_1, 2} x_{\mu s_2, 2}} \right| \cdot (\delta_{s_1 k_1}^- + \delta_{s_2 k_2}^+ - 1) \leq r, \\ & j_1 = \overline{1, l-1}, \quad j_2 = \overline{j_1+1, l}, \quad s_1 \in Y_{j_1} \cap Y_{j_2}, \quad k_1 = \overline{1, p}, \quad s_2 \in Y \setminus (Y_{j_1} \cup Y_{j_2}), \quad k_2 = \overline{1, p}. \end{aligned} \quad (24)$$

Ограничения (20) предназначены для контроля интеркорреляций между внутренними переменными из регрессоров с одной переменной и регрессоров с операцией $\min$, (21) – из регрессоров с одной переменной и регрессоров с операцией $\max$, (22) – из регрессоров с операциями $\min$, (23) – из регрессоров с операциями $\max$, (24) – из регрессоров с операциями $\min$ и $\max$.

В зависимости от бинарных переменных изменение конфигурации системы линейных алгебраических уравнений для нахождения МНК-оценок стандартизованной регрессии можно записать в виде следующих ограничений:

$$-(1-\delta_j)M \leq \sum_{k=1}^l r_{x_j x_k} \cdot \beta_k + \sum_{s=1}^{C_l^2} \sum_{k=1}^p r_{x_j z_{sk}^-} \cdot \beta_{sk}^- + \sum_{s=1}^{C_l^2} \sum_{k=1}^p r_{x_j z_{sk}^+} \cdot \beta_{sk}^+ - r_{yx_j} \leq (1-\delta_j)M, \quad j = \overline{1, l}, \quad (25)$$

$$\begin{aligned} & -(1-\delta_{jk}^-)M \leq \sum_{s=1}^l r_{x_s z_{jk}^-} \cdot \beta_s + \sum_{s_1=1}^{C_l^2} \sum_{s_2=1}^p r_{z_{s_1 s_2}^- z_{jk}^-} \cdot \beta_{s_1 s_2}^- + \\ & + \sum_{s_1=1}^{C_l^2} \sum_{s_2=1}^p r_{z_{s_1 s_2}^+ z_{jk}^-} \cdot \beta_{s_1 s_2}^+ - r_{yz_{jk}^-} \leq (1-\delta_{jk}^-)M, \quad j = \overline{1, C_l^2}, \quad k = \overline{1, p}, \end{aligned} \quad (26)$$

$$\begin{aligned} & -(1-\delta_{jk}^+)M \leq \sum_{s=1}^l r_{x_s z_{jk}^+} \cdot \beta_s + \sum_{s_1=1}^{C_l^2} \sum_{s_2=1}^p r_{z_{s_1 s_2}^- z_{jk}^+} \cdot \beta_{s_1 s_2}^- + \\ & + \sum_{s_1=1}^{C_l^2} \sum_{s_2=1}^p r_{z_{s_1 s_2}^+ z_{jk}^+} \cdot \beta_{s_1 s_2}^+ - r_{yz_{jk}^+} \leq (1-\delta_{jk}^+)M, \quad j = \overline{1, C_l^2}, \quad k = \overline{1, p}. \end{aligned} \quad (27)$$

Например, если $\delta_{jk}^- = 1$, что означает, что регрессор z_{jk}^- входит в регрессию, то соответствующее двойное неравенство (26) трансформируется в строгое равенство. Иными словами, происходит включение уравнения в систему. В противном случае, когда $\delta_{jk}^- = 0$, включения не происходит. Подход к выбору большого числа M в (25)–(27) также обсуждается в [20].

В качестве критерия оптимизации выступает коэффициент детерминации R^2 , который, естественным образом, должен быть максимальным:

$$R^2 = \sum_{j=1}^l r_{yx_j} \cdot \beta_j + \sum_{j=1}^{C_l^2} \sum_{k=1}^p r_{yz_{jk}^-} \cdot \beta_{jk}^- + \sum_{j=1}^{C_l^2} \sum_{k=1}^p r_{yz_{jk}^+} \cdot \beta_{jk}^+ \rightarrow \max. \quad (28)$$

Пусть экспертами успешно выполнена проверка условия 2. Тогда решение оптимизационной задачи ЧБЛП с целевой функцией (28) и линейными ограничениями (5)–(27), если оно существует, гарантирует выполнение условий 1, 3, 4 и 6 в построенной НЛинР. Формализация остальных условий в виде линейных ограничений и последующая их интеграция в предложенную задачу возможна. Например, условие 9 с учетом (28) можно записать в виде простого ограничения $R^2 \geq 0,8$. Однако для обеспечения условий 5, 7 и 8 нужно вводить дополнительные бинарные переменные, что усложнит и без того непростую в вычислительном плане оптимизационную задачу. К тому же условия 5, 7 и 8 по смыслу соотносятся с условиями 4, 6 и 9 соответственно. Так, если абсолютные вклады регрессоров в общую детерминацию высоки, то, скорее всего, и все коэффициенты будут значимы по t -критерию Стьюдента. Если интеркорреляции между внешними регрессорами малы, то, вероятно, и зависимости каждого регрессора от множества остальных будут незначимы. А если коэффициент R^2 высок, то модель в большинстве случаев будет значима в целом по F -критерию Фишера. Таким образом, очень высоки шансы, что решение задачи (5)–(28) будет гарантировать выполнение всех девяти условий ВИНЛинР.

2. Проверка условий ВИНЛинР на примере логистической задачи

Для тестирования разработанного математического аппарата было решено использовать статистические данные, приведенные в [20], по следующим переменным:

- y – отправление грузов ж/д транспортом общего пользования в Иркутской области (млн т);
- x_2 – процент трудоспособного населения;
- x_3 – численность рабочей силы (тыс. чел.);
- x_5 – численность пенсионеров (тыс. чел.);
- x_8 – число собственных легковых автомобилей на 1 000 человек населения (шт.);
- x_{18} – число предприятий и организаций;
- x_{20} – кредиторская задолженность организаций (млн руб.);
- x_{22} – производство электроэнергии (млрд кВт·ч);
- x_{58} – тарифы на грузовые перевозки ж/д транспортом (усл. ед.).

По этим данным в [20] была построена НЛинР вида

$$\begin{aligned} \tilde{y} = & -24,527 + \overset{(0,6427)}{1,189} \min\{x_2; 0,000933x_{18}\} - \overset{(0,1129)}{0,0196} \min\{x_5; 0,006754x_{20}\} - \\ & \overset{(0,0843)}{-0,0323} \min\{x_8; 0,11725x_{58}\} + \overset{(0,1063)}{0,0254} \max\{x_3; 23,079x_{22}\}, \end{aligned} \quad (29)$$

для которой $R^2 = 0,946183$.

В уравнении (29) в скобках под коэффициентами указаны наблюдаемые t -статистики, а над коэффициентами – абсолютные вклады регрессоров в R^2 .

Сначала было решено проверить, можно ли отнести модель (29) к ВИНЛинР, т.е. проверить выполнение всех девяти предложенных выше условий. Будем считать, что число $\theta = 0,05$, а уровень значимости $\alpha = 0,01$.

Очевидно, что для модели (29) выполняются условия 1, 4 и 9.

В работе [20] экспертами проведен подробный содержательный анализ знаков влияния объясняющих переменных на y , поэтому выполнены условие 2 и, как следствие, условие 3.

По критерию Стьюдента для уровня $\alpha = 0,01$ в модели (29) незначим только коэффициент при регрессоре $\min\{x_8; 0,11725x_{58}\}$, но уже для уровня $\alpha = 0,05$ он становится значимым. Поэтому условие 5 можно считать по большому счету выполненным.

Наблюдаемое значение критерия Фишера для модели (29) составляет 70,3619, что гораздо выше критического значения для уровня $\alpha = 0,01$, поэтому модель в целом значима, т.е. условие 8 выполнено.

Проверим наличие мультиколлинеарности в модели (29). Как отмечено выше, пороговое значение r для абсолютных величин интеркорреляций можно найти по формуле

$$r = \frac{t_{\text{крит}}(0,01; 21-2)}{\sqrt{21-2+t_{\text{крит}}^2(0,01; 21-2)}} = \frac{2,8609}{\sqrt{19+2,8609^2}} = 0,5487.$$

Все абсолютные величины интеркорреляций между внешними регрессорами, кроме коэффициента корреляции между $\min\{x_5; 0,006754x_{20}\}$ и $\min\{x_8; 0,11725x_{58}\}$, равного 0,8576, не превышают значения 0,5487. Отсюда следует, что условие 6 нарушено. К тому же абсолютные величины многих интеркорреляций между внутренними переменными превышают значение 0,5487. Например, коэффициент корреляции между x_2 и x_3 равен 0,8199, между x_2 и x_5 равен $-0,9017$, между x_3 и x_5 равен $-0,863$ и др.

Две вспомогательные регрессии из четырех значимы по критерию Фишера для уровня $\alpha = 0,01$: зависимость $\min\{x_5; 0,006754x_{20}\}$ от $\min\{x_2; 0,000933x_{18}\}$, $\min\{x_8; 0,11725x_{58}\}$, $\max\{x_3; 23,079x_{22}\}$ и зависимость $\min\{x_8; 0,11725x_{58}\}$ от $\min\{x_2; 0,000933x_{18}\}$, $\min\{x_5; 0,006754x_{20}\}$, $\max\{x_3; 23,079x_{22}\}$. Тем самым условие 7 нарушено.

Таким образом, модель (29) нельзя отнести к ВИНЛинР из-за присутствия в ней мультиколлинеарности.

После этого с использованием тех же данных была предпринята попытка построить ВИНЛинР. Для автоматической формализации задачи ЧБЛП (5)–(28) по данным о грузовых ж/д перевозках была использована программа построения вполне интерпретируемых элементарных и неэлементарных квазилинейных регрессионных моделей (ВИнтер-2) [18]. Она позволяет автоматически формулировать нужную оптимизационную задачу для пакета LPSolve IDE. Начальные настройки в программе ВИнтер-2 остались те же, которые использовались в [20]. Параметры $\theta = 0,05$, $r = 0,5487$.

В результате с помощью пакета LPSolve была построена НЛинР вида

$$\tilde{y} = -77,893 + \underset{(9,889)}{0,000888} \min\{x_{18}; 1071,164x_{22}\} + \underset{(7,691)}{1,379} \max\{x_2; 0,047x_3\}, \quad (30)$$

для которой $R^2 = 0,923181$.

Далее для модели (30) проверялись 9 условий ВИНЛинР. Нетрудно заметить, что условия 1, 2, 3, 4, 5 и 9 выполняются.

Величина интеркорреляции между внешними регрессорами $\min\{x_{18}; 1071,164x_{22}\}$ и $\max\{x_2; 0,047x_3\}$ составляет 0,283. А величины интеркорреляций между внутренними переменными x_{18} и x_2 , x_{18} и x_3 , x_{22} и x_2 , x_{22} и x_3 составляют 0,0841, 0,0225, 0,3926 и $-0,0013$ соответственно. Все эти величины по модулю меньше, чем 0,5487, поэтому условие 6 выполнено.

Поскольку в модели (30) всего два внешних регрессора, коэффициент корреляции между которыми незначим, то из этого автоматически следует выполнение условия 7.

Наблюдаемое значение критерия Фишера для модели (29) составляет 107,933, что гораздо выше критического значения для уровня $\alpha = 0,01$, поэтому модель в целом значима и условие 8 выполнено.

Таким образом, все 9 условий выполняются, следовательно, модель (30) относится к ВИНЛинР. Это означает, что если представить ее в виде кусочно-заданной функции, то в ней, в отличие от кусочно-заданного представления регрессии (29), можно корректно интерпретировать отдельное влияние абсолютно любой объясняющей переменной на результирующий признак y . В модели (29) этому препятствует мультиколлинеарность, из-за которой различать отдельное влияние объясняющих переменных на y затруднительно.

Заключение

Сформулировано строгое определение ВИНЛинР, содержащее девять условий. Среди них ограничения на количество вхождений объясняющих переменных в модель, на соответствие знаков оценок содержательному смыслу задачи, на значимость коэффициентов по t -критерию Стьюдента и величинам абсолютных вкладов переменных в общую детерминацию, на мультиколлинеарность, на адекватность регрессии по коэффициенту детерминации и критерию Фишера. Выполнение всех этих девяти условий означает, что в поставленной в соответствие регрессии кусочно-заданной функции можно корректно интерпретировать отдельное влияние любой объясняющей переменной на объясняемую переменную y . В ранее предложенную задачу ЧБЛП для выбора оптимальной структурной спецификации ВИНЛинР интегрированы линейные ограничения для контроля мультиколлинеарности как между внешними регрессорами, так и между внутренними переменными из разных регрессоров. Разработанный математический аппарат реализован в программе ВИнтер-2. С помощью программы перестроена полученная ранее модель железнодорожных перевозок в Иркутской области. При этом показано, что все условия ВИНЛинР для нее выполнены. Описанный в статье математический аппарат может успешно применяться в машинном обучении для поиска в статистических данных новых и хорошо интерпретируемых закономерностей.

Список источников

1. Joshi A., Raman B., Mohan C.K., Cenkeramaddi L.R. Application of a new machine learning model to improve earthquake ground motion predictions // *Natural Hazards*. 2024. V. 120 (1). P. 729–753. doi: 10.1007/s11069-023-06230-4
2. Zhou Y., Liu X., Wang D., Liu X., Wang Y. A review on global solar radiation prediction with machine learning models in a comprehensive perspective // *Energy Conversion and Management*. 2021. V. 235. Art. 113960. doi: 10.1016/j.enconman.2021.113960
3. Sonkavde G., Dharrao D.S., Bongale A.M., Deokate S.T., Doreswamy D., Bhat S.K. Forecasting stock market prices using machine learning and deep learning models: A systematic review, performance analysis and discussion of implications // *International Journal of Financial Studies*. 2023. V. 11 (3). Art. 94. doi: 10.3390/ijfs11030094
4. Doshi-Velez F., Kim B. Towards a rigorous science of interpretable machine learning // *arXiv preprint arXiv:1702.08608*. 2017. doi: 10.48550/arXiv.1702.08608
5. Molnar C. Interpretable machine learning. URL: https://originalstatic.aminer.cn/misc/pdf/Molnar-interpretable-machine-learning_compressed.pdf (accessed: 10.06.2024).
6. Chen Z., Xiao F., Guo F., Yan J. Interpretable machine learning for building energy management: A state-of-the-art review // *Advances in Applied Energy*. 2023. V. 9. Art. 100123. doi: 10.1016/j.adapen.2023.100123
7. Esterhuizen J.A., Goldsmith B.R., Linic S. Interpretable machine learning for knowledge generation in heterogeneous catalysis // *Nature catalysis*. 2022. V. 5 (3). P. 175–184. doi: 10.1038/s41929-022-00744-z
8. Maulud D., Abdulazeez A.M. A review on linear regression comprehensive in machine learning // *Journal of Applied Science and Technology Trends*. 2020. V. 1 (2). P. 140–147. doi: 10.38094/jastt1457
9. Schober P., Vetter T.R. Logistic regression in medical research // *Anesthesia & Analgesia*. 2021. V. 132 (2). P. 365–366. doi: 10.1213/ANE.0000000000005247
10. Mejia A.F., Yue Y.R., Bolin D., Lindgren F., Lindquist M.A. A Bayesian general linear modeling approach to cortical surface fMRI data analysis // *Journal of the American Statistical Association*. 2020. V. 115. P. 501–520. doi: 10.1080/01621459.2019.1611582
11. Fasiolo M., Nedellec R., Goude Y., Wood S.N. Scalable visualization methods for modern generalized additive models // *Journal of computational and Graphical Statistics*. 2020. V. 29 (1). P. 78–86. doi: 10.1080/10618600.2019.1629942
12. Базилевский М.П. Построение вполне интерпретируемых линейных регрессионных моделей с помощью метода последовательного повышения абсолютных вкладов переменных в общую детерминацию // *Вестник Воронежского государственного университета. Сер. Системный анализ и информационные технологии*. 2022. № 2. С. 5–16. doi: 10.17308/sait/1995-5499/2022/2/5-16

13. Montgomery D.C., Peck E.A., Vining G.G. Introduction to linear regression analysis. John Wiley & Sons, 2021.
14. Miller A. Subset selection in regression. New York : CRC Press, 2002.
15. Kim S.H., Boukouvala F. Machine learning-based surrogate modeling for data-driven optimization: a comparison of subset selection for regression techniques // *Optimization Letters*. 2020. V. 14 (4). P. 989–1010. doi: 10.1007/s11590-019-01428-7
16. Chan J.Y.L., Leow S.M.H., Bea K.T., Cheng W.K., Phoong S.W., Hong Z.W., Chen Y.L. Mitigating the multicollinearity problem and its machine learning approach: a review // *Mathematics*. 2022. Vol. 10 (8). Art. 1283. doi: 10.3390/math10081283
17. Obite C.P., Olewuezi N.P., Ugwuanyim G.U., Bartholomew D.C. Multicollinearity effect in regression analysis: A feed forward artificial neural network approach // *Asian J. Probab. Stat.* 2020. V. 6 (1). P. 22–33. doi: 10.9734/AJPAS/2020/v6i130151
18. Базилевский М.П. Сравнительный анализ эффективности методов построения вполне интерпретируемых линейных регрессионных моделей // *Моделирование и анализ данных*. 2023. Т. 13, № 4. С. 59–83. doi: 10.17759/mda.2023130404
19. Базилевский М.П. Технология построения вполне интерпретируемых квазилинейных регрессионных моделей // *Прикладная математика и вопросы управления*. 2024. № 1. С. 123–138. doi: 10.15593/2499-9873/2024.1.08
20. Базилевский М.П. Метод построения неэлементарных линейных регрессий на основе аппарата математического программирования // *Проблемы управления*. 2022. № 4. С. 3–14. doi: 10.25728/ru.2022.4.1
21. Базилевский М.П. Построение вполне интерпретируемых неэлементарных линейных регрессионных моделей // *Вестник Югорского государственного университета*. 2022. № 4 (67). С. 105–114. doi: 10.18822/byusu202204105-114

References

1. Joshi, A., Raman, B., Mohan, C.K. & Cenkeramaddi, L.R. (2024) Application of a new machine learning model to improve earthquake ground motion predictions. *Natural Hazards*. 120(1). pp. 729–753. DOI: 10.1007/s11069-023-06230-4
2. Zhou, Y., Liu, Y., Wang, D., Liu, X. & Wang, Y. (2021) A review on global solar radiation prediction with machine learning models in a comprehensive perspective. *Energy Conversion and Management*. 235. p. 113960. DOI: 10.1016/j.enconman.2021.113960
3. Sonkavde, G., Dharrao, D.S., Bongale, A.M., Deokate, S.T., Doreswamy, D. & Bhat, S.K. (2023) Forecasting stock market prices using machine learning and deep learning models: A systematic review, performance analysis and discussion of implications. *International Journal of Financial Studies*. 11(3). p. 94. DOI: 10.3390/ijfs11030094
4. Doshi-Velez, F. & Kim, B. (2017) Towards a rigorous science of interpretable machine learning. *arXiv preprint arXiv:1702.08608*. DOI: 10.48550/arXiv.1702.08608
5. Molnar, C. (2020) *Interpretable machine learning*. [Online] Available from: https://originalstatic.aminer.cn/misc/pdf/Molnar-interpretible-machine-learning_compressed.pdf (Accessed: 10th June 2024).
6. Chen, Z., Xiao, F., Guo, F. & Yan, J. (2023) Interpretable machine learning for building energy management: A state-of-the-art review. *Advances in Applied Energy*. 9. p. 100–123. DOI: 10.1016/j.adapen.2023.100123
7. Esterhuizen, J.A., Goldsmith, B.R. & Linic, S. (2022) Interpretable machine learning for knowledge generation in heterogeneous catalysis. *Nature catalysis*. 5(3). pp. 175–184. DOI: 10.1038/s41929-022-00744-z
8. Maulud, D. & Abdulazez, A.M. (2020) A review on linear regression comprehensive in machine learning. *Journal of Applied Science and Technology Trends*. 1(2). pp. 140–147. DOI: 10.38094/jastt1457
9. Schober, P. & Vetter, T.R. (2021) Logistic regression in medical research. *Anesthesia & Analgesia*. 132(2). pp. 365–366. DOI: 10.1213/ANE.0000000000005247
10. Mejia, A.F., Yue, Y.R., Bolin, D., Lindgren, F. & Lindquist, M.A. (2020) A Bayesian general linear modeling approach to cortical surface fMRI data analysis. *Journal of the American Statistical Association*. 115. pp. 501–520. DOI: 10.1080/01621459.2019.1611582
11. Fasiolo, M., Nedellec, R., Goude, Y. & Wood, S.N. (2020) Scalable visualization methods for modern generalized additive models. *Journal of Computational and Graphical Statistics*. 29(1). pp. 78–86. DOI: 10.1080/10618600.2019.1629942
12. Bazilevskiy, M.P. (2022) Postroenie vpolne interpretiruemykh lineynykh regressiionnykh modeley s pomoshch'yu metoda posledovatel'nogo povysheniya absolyutnykh vkladov peremennykh v obshchuyu determinatsiyu [Construction of quite interpretable linear regression models using the method of successive increase the absolute contributions of variables to the general determination]. *Vestnik Voronezhskogo gosudarstvennogo universiteta. Ser. Sistemnyy analiz i informatsionnye tekhnologii – Proceedings of Voronezh State University. Series: Systems Analysis and Information Technologies*. 2. pp. 5–16. DOI: 10.17308/sait/1995-5499/2022/2/5-16
13. Montgomery, D.C., Peck, E.A. & Vining, G.G. (2021) *Introduction to Linear Regression Analysis*. John Wiley & Sons.
14. Miller, A. (2002) *Subset Selection in Regression*. Chapman and hall/CRC.
15. Kim, S.H. & Boukouvala, F. (2020) Machine learning-based surrogate modeling for data-driven optimization: a comparison of subset selection for regression techniques. *Optimization Letters*. 14(4). pp. 989–1010. DOI: 10.1007/s11590-019-01428-7
16. Chan, J.Y.L., Leow, S.M.H., Bea, K.T., Cheng, W.K., Phoong, S.W., Hong, Z.W. & Chen, Y.L. (2022) Mitigating the multicollinearity problem and its machine learning approach: a review. *Mathematics*. 10(8). p. 1283. DOI: 10.3390/math10081283
17. Obite, C.P., Olewuezi, N.P., Ugwuanyim, G.U. & Bartholomew, D.C. (2020) Multicollinearity effect in regression analysis: A feed forward artificial neural network approach. *Asian Journal of Probability and Statistics*. 6(1). pp. 22–33. DOI: 10.9734/AJPAS/2020/v6i130151
18. Bazilevskiy, M.P. (2023) Sravnitel'nyy analiz effektivnosti metodov postroeniya vpolne interpretiruemykh lineynykh regressiionnykh modeley [Comparative analysis of the effectiveness of methods for constructing quite interpretable linear regression models]. *Modelirovanie i analiz dannykh – Modelling and Data Analysis*. 13(4). pp. 59–83. DOI: 10.17759/mda.2023130404

19. Bazilevskiy, M.P. (2024) Tekhnologiya postroeniya vpolne interpretiruemykh kvazilineynykh regressionnykh modeley [Technology for constructing quite interpretable quasilinear regression models]. *Prikladnaya matematika i voprosy upravleniya – Applied Mathematics and Control Sciences*. 1. pp. 123–138. DOI: 10.15593/2499-9873/2024.1.08
20. Bazilevskiy, M.P. (2022) Metod postroeniya neelementarnykh lineynykh regressiy na osnove apparata matematicheskogo programmirovaniya [A method for constructing nonelementary linear regressions based on mathematical programming]. *Problemy upravleniya*. 4. pp. 3–14. DOI: 10.25728/pu.2022.4.1
21. Bazilevskiy, M.P. (2022) Postroenie vpolne interpretiruemykh neelementarnykh lineynykh regressionnykh modeley [Construction of quite interpretable non-elementary linear regression models]. *Vestnik Yugorskogo gosudarstvennogo universiteta*. no. 4 (67). pp. 105–114. DOI: 10.18822/byusu202204105-114

Информация об авторе:

Базилевский Михаил Павлович – доцент, кандидат технических наук, доцент кафедры «Математика» Иркутского государственного университета путей сообщения (Иркутск, Россия). E-mail: mik2178@yandex.ru

Автор заявляет об отсутствии конфликта интересов.

Information about the author:

Bazilevskiy Mikhail P. (Candidate of Technical Sciences, Associate Professor of the Mathematics Department, Irkutsk State Transport University, Irkutsk, Russian Federation). E-mail: mik2178@yandex.ru

The author declares no conflicts of interests.

Поступила в редакцию 17.07.2024; принята к публикации 02.12.2024

Received 17.07.2024; accepted for publication 02.12.2024

Научная статья

УДК 004.95

doi: 10.17223/19988605/69/5

Элементы системного и вепольного анализа инфраструктуры мониторинга окружающей среды (на примере геомагнитных данных)**Гульнара Равиловна Воробьева¹, Андрей Владимирович Воробьев²,
Эмиль Фанильевич Фарваев³, Глеб Олегович Орлов⁴**^{1, 2, 3, 4} Уфимский университет науки и технологий, Уфа, Россия² Геофизический центр Российской академии наук, Москва, Россия¹ gulnara.vorobeva@gmail.com² geomagnet@list.ru³ farvaev.emil@gmail.com⁴ orlovgleb99@mail.ru

Аннотация. Принятие решений в области обеспечения техносферной безопасности базируется на своевременном получении актуальной и полной информации о состоянии окружающей среды. Геомагнитный мониторинг играет важную роль в сокращении негативного воздействия экстремальных геофизических событий на объекты и системы техносферы (линии электропередач, системы связи, автоматику железных дорог и пр.). В работе геомагнитный мониторинг рассмотрен с точки зрения системного подхода, идентифицированы и сформулированы основные проблемы получения геомагнитных данных, главной из которых является неполное покрытие сетью мониторинга. Предложен и с помощью вепольного представления формализован подход, дополняющий схему получения и использования данных интерполирующей компьютерной моделью. Реализация такой модели использует адаптивную пространственную интерполяцию для определения оптимальных методов интерполяции с учетом географического расположения интерполируемых точек и глобального индекса геомагнитной активности. Эффективность подхода продемонстрирована на примере геомагнитной информации проекта SuperMAG с помощью веб-ориентированного программного решения.

Ключевые слова: системный анализ; вепольный анализ; геомагнитные данные; пространственная интерполяция; поддержка принятия решений.

Благодарности: Исследование выполнено при финансовой поддержке РНФ, проект № 21-77-30010.

Для цитирования: Воробьева Г.Р., Воробьев А.В., Фарваев Э.Ф., Орлов Г.О. Элементы системного и вепольного анализа инфраструктуры данных мониторинга окружающей среды (на примере геомагнитных данных) // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 69. С. 42–53. doi: 10.17223/19988605/69/5

Original article

doi: 10.17223/19988605/69/5

**Elements of system and Su-field analysis of environmental monitoring tools
(using the example of the geomagnetic data)****Gulnara R. Vorobeva¹, Andrei V. Vorobev², Emil F. Farvaev³, Gleb O. Orlov⁴**^{1, 2, 3, 4} Ufa University of Science and Technology, Ufa, Russian Federation² Geophysical Center of RAS, Moscow, Russian Federation¹ gulnara.vorobeva@gmail.com² geomagnet@list.ru

³farvaev.emil@gmail.com

⁴orlovgleb99@mail.ru

Abstract. Decision-making in the field of technosphere safety is based on timely receipt of up-to-date and complete information about the state of the environment. Geomagnetic monitoring plays an important role in reducing the negative impact of extreme geophysical events on objects and systems of the technosphere (power lines, communication systems, railway automation, etc.). In this paper, geomagnetic monitoring is considered from the point of view of system approach, the main problems of obtaining geomagnetic data are identified and formulated, the main of which is incomplete coverage by the monitoring network. Using a Su-field representation, an approach is proposed and formalized, that extends the process of obtaining and using data with an interpolating computer model. The implementation of this model uses adaptive spatial interpolation to determine the optimal interpolation methods, taking into account the geographical location of the interpolated points and the global geomagnetic activity index. The effectiveness of the approach is demonstrated by the example of the geomagnetic information of the SuperMAG project using a web-based software solution.

Keywords: system analysis, Su-field analysis, geomagnetic data, spatial interpolation, decision support.

Acknowledgments: The reported study was funded by RSF, project number 21-77-30010.

For citation: Vorobeve, G.R., Vorobev, A.V., Farvaev, E.F., Orlov, G.O. (2024) Elements of system and Su-field analysis of environmental monitoring tools (using the example of the geomagnetic data). *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 69. pp. 42–53. doi: 10.17223/19988605/69/5

Введение

В настоящее время наблюдается активное развитие систем поддержки принятия решений, базирующихся на применении геопространственных данных различного прикладного назначения. Примером таких данных являются результаты наблюдения геомагнитного поля и его вариаций, необходимые для обеспечения безопасности эксплуатации объектов и систем техносферы в условиях экстремальной геофизической обстановки [1, 2]. Если условия в соответствующем регионе не располагают к размещению магнитометрических измерительных устройств (что часто бывает, к примеру, в суровых условиях арктического региона планеты), то и получить такую информацию и предпринять соответствующее превентивное действие принимающему решения не представляется возможным. Иными словами, наблюдается дефицит необходимых для принятия решения данных.

Представляется целесообразным рассмотреть задачу мониторинга параметров геомагнитного поля и его вариаций с позиций системного анализа, представить отдельные источники данных в их взаимосвязи с единой точкой входа для потребителей геомагнитных данных с варьируемым шагом дискретизации для сети мониторинга. Фактически предполагается, что данные по параметру геомагнитного поля должны быть получены в любой пространственной точке независимо от физического наличия здесь магнитометрической аппаратуры. Мультиаспектный анализ известных решений геомагнитного мониторинга также ставит своей целью выявление основных проблем и перспективных направлений его развития с повышением эффективности соответствующих источников геомагнитных данных для конечных потребителей.

1. Системный анализ задачи геомагнитного мониторинга

В общем виде рассмотрение системы геомагнитного мониторинга как совокупности элементов и подсистем позволяет использовать для ее математического описания теоретико-множественный аппарат (рис. 1) [3, 4]. Если обозначить систему геомагнитного мониторинга как S , то имеем:

$$S = \{m_1, m_2, \dots, m_N\},$$

где m_i – магнитная обсерватория / вариационная станция, на которой осуществляется регистрация параметров геомагнитного поля и его вариаций ($i = 1, \dots, N$).

Множество прошедших обработку данных обозначим как D . Целевой по отношению к системе S будет являться функция, задающая правила отображения множества S в множество D (получение геомагнитных данных по результатам геомагнитного мониторинга):

$$f : S \rightarrow D : D = \{d_1, d_2, \dots, d_N\},$$

где d_i – геомагнитные данные, полученные по результатам обработки измерений i -й магнитной обсерватории / вариационной станции, на которой осуществляется регистрация параметров геомагнитного поля и его вариаций ($i = 1, \dots, N$).

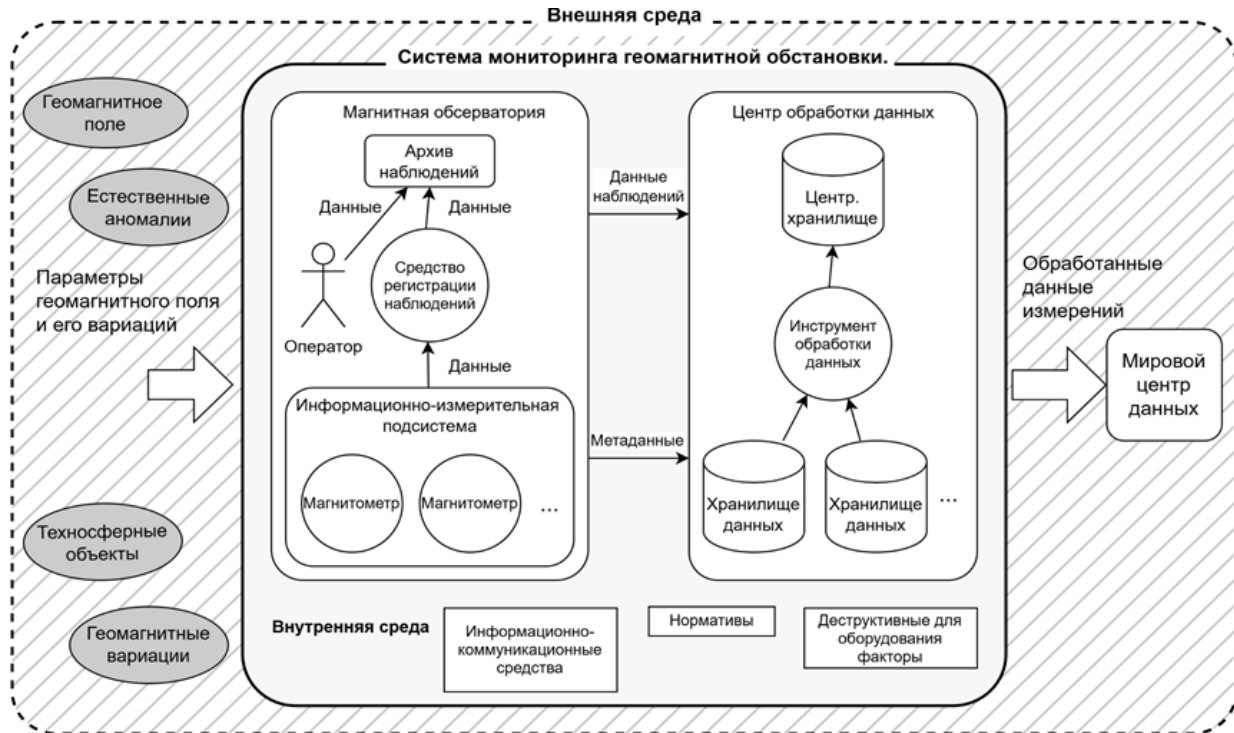


Рис. 1. Представление геомагнитного мониторинга с точки зрения принципов системного анализа
Fig. 1. Presentation of geomagnetic monitoring from the point of view of the principles of system analysis

При этом на результат отображения оказывают воздействие внешние и внутренние факторы, представленные соответственно множествами F_{in} и F_{out} :

$$f : S \rightarrow D : f_i : m_i \rightarrow d_i; f_i = g_i(p_i, t_i, F_{in}, F_{out});$$

$$p_i \supset (lat_i, lng_i), i = 1, \dots, N,$$

где p_i – пространственная точка, где размещается магнитная обсерватория, на которой осуществляется регистрация параметров геомагнитного поля, пространственная точка задается парой географических координат (широта / lat, долгота / lng); g_i – функция, характеризующая воздействие на результат сопоставления i -х элементов множеств внутренних и внешних факторов, времени (t_i) и пространственного расположения (p_i).

Здесь может быть обозначена проблема геомагнитного мониторинга, связанная с тем, что средства регистрации параметров геомагнитного поля и его вариаций распределены по земной поверхности крайне неравномерно. Известно, что наибольшее их количество представлено на территории Европы, в то время как иные регионы планеты охвачены средствами мониторинга весьма поверхностно. Иными словами, потребитель, обратившись к мировому центру данных, может не получить необходимую ему информацию по той причине, что в данной точке p_k (регионе) отсутствуют средства регистрации. Другой вариант – средства регистрации физически присутствуют в искомой пространственной точке p_k (регионе), но в требуемый момент времени не функционировали по тем или иным причинам. В терминах теории множеств сказанное возможно представить следующим образом:

$$\forall m_i \in M (i = 1, \dots, N) \nexists f_k : m_i \rightarrow d_k : f_k = g_k(p_k, t_k, F_{in}, F_{out}); p_k \supset (lat_k, lng_k),$$

где p_k – пространственная точка, для которой необходимо определить параметры геомагнитного поля, но при этом отсутствует магнитная обсерватория, на которой осуществляется регистрация параметров геомагнитного поля и его вариаций.

Соответственно, потребитель геомагнитной информации сталкивается с проблемой дефицита искомых результатов мониторинга параметров геомагнитного поля и его вариаций. В результате невозможно принимать адекватные решения в той или иной области (и особенно в области обеспечения техносферной безопасности), которые напрямую или косвенно зависят от соответствующей геомагнитной обстановки. Если при этом геомагнитные параметры являются одним из компонент информационной поддержки принятия решений, то даже наличие других ее составляющих может привести к серьезным последствиям, в том числе экономическим.

В случае кратковременного отсутствия геомагнитных данных в отдельных случаях возможно их восстановление известными статистическими методами интерполяции временных рядов. Однако в том случае, когда источника данных пространственно не существует, либо пропуски по длительности своей весьма существенны, указанные методы теряют свою эффективность, а полученные на их основе геомагнитные данные могут привести к принятию неверных решений, что особенно критично в области обеспечения техносферной безопасности сложных объектов и систем.

Соответственно необходимо решение, которое позволит оценивать геомагнитную обстановку в условиях дефицита результатов регистрации параметров геомагнитного поля и его вариаций. При этом искомые данные должны учитывать весь комплекс внутренних и внешних факторов, способных оказывать на них негативное искажающее влияние.

2. Вепольный анализ задачи геомагнитного мониторинга

Дальнейшее исследование предполагает рассмотрение потенциальной системы геомагнитного мониторинга с позиций воздействия одних ее компонент на другие (при наличии) для выявления новых / подтверждения уже выявленных проблем при решении соответствующей задачи (рис. 2).

Непосредственно информационно-измерительная система и геомагнитные данные в терминах вепольного анализа рассматриваются как варианты вещества [5, 6]. Преобразование вещества В1 (информационно-измерительная система) в вещество В2 (геомагнитные данные) осуществляется под воздействием геомагнитного поля (П1), параметры которого непосредственно регистрируются приборами и переводятся после некоторых преобразований в геомагнитные данные:

$$B2 \Rightarrow \Pi 1 \rightarrow B1 \rightarrow B2.$$

Фактически имеет место комплексная система, построенная в виде совокупности простых веполей, при этом рассматривается одновременно несколько веществ и энергий в своей взаимосвязи. Данная структура полностью соответствует особенностям пространственно-временной динамики значений параметров геомагнитного поля и его вариаций, что обусловлено комплексным воздействием на измеряемые параметры группы внешних и внутренних факторов (в редких случаях наблюдается воздействие единственного фактора, что является, скорее, исключением).

На более высоком уровне абстракции представляется целесообразным рассмотреть выделенные веполи как особый тип веполей, применяемых для задач на измерение параметров или обнаружение объектов. Если в таком случае представить информационно-измерительную систему как вещество В1, а геомагнитные данные – как вещество В2, то возможно говорить о применении В2 как индикатора существования полей, воздействующих на В1 (непосредственно геомагнитное поле, его вариации, обусловленные различными факторами и пр.).

Здесь представляется целесообразным отметить наличие в рассматриваемой системе так называемых вредных связей. К таковым относятся поля, которые негативно воздействуют на соответствующее вещество, искажая формируемый при этом результат. Причем в рассматриваемом случае негативное воздействие от вредных связей проявляется не только в отношении некоторого вещества, но также и касательно некоторых полей, которые, в свою очередь, оказывают воздействие на исходное вещество в процессе формирования искомого вещества (см. рис. 2).

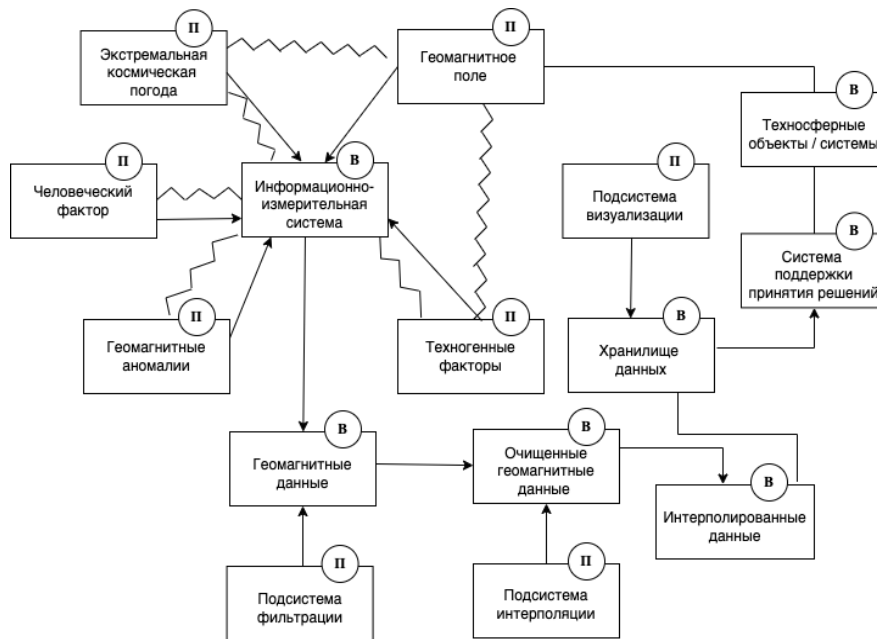


Рис. 2. Вепольная схема геомагнитного мониторинга
Fig. 2. Su-field diagram of geomagnetic monitoring

При этом некоторые из полей могут одновременно и выборочно воздействовать негативным образом на измеряемые значения параметров геомагнитного поля (и на одноименное поле, приведенное в вепольной схеме). Иными словами, имеет место двойственное негативное воздействие ряда полей, проявляющееся в их влиянии как на вещество, так и на другие поля, также влияющие на вещество:

$$B2 \Rightarrow (P1 + P'_{\text{внеш}} + P'_{\text{внутр}}) \rightarrow B1 \rightarrow B2 :$$

$$B2 \Rightarrow ((P1 + P'_{\text{тф}} + P'_{\text{кп}}) + (P'_{\text{чф}} + P'_{\text{кп}} + P'_{\text{тф}}) + P'_{\text{внутр}}) \rightarrow B1 \rightarrow B2,$$

где $P_{\text{внеш}}$ и $P_{\text{внутр}}$ – поля, соответствующие внешним и внутренним факторам; $P_{\text{чф}}$, $P_{\text{кп}}$ и $P_{\text{тф}}$ – человеческий фактор, космическая погода и техногенные факторы, в совокупности составляющие поле $P_{\text{внеш}}$; $P_{\text{га}}$ и $P_{\text{пр}}$ – геомагнитные аномалии и прочие факторы, в совокупности составляющие поле $P_{\text{внутр}}$.

Полностью исключить негативное воздействие вредных связей в контексте геомагнитного мониторинга не представляется возможным. При этом можно его значимо минимизировать посредством введения в вепольную систему дополнительных элементов, частично нейтрализующих последствия искажения соответствующего искомого вещества (геомагнитных данных).

Дополненная схема представляет расширенный и форсированный веполь, поскольку содержит дополнительный элемент, призванный сократить или полностью исключить негативное воздействие внешних и внутренних факторов на результирующие геомагнитные данные. Так, в частности, в обозначенную вепольную структуру добавлена компьютерная модель, которая наряду с информационно-измерительной системой участвует в формировании результирующих геомагнитных данных. Модель представлена двумя ключевыми подсистемами обработки геомагнитных данных, а именно подсистемами фильтрации для очистки данных и интерполяции для их восстановления.

Компьютерная модель с соответствующими настройками может выступать в роли виртуальной информационно-измерительной системы (виртуальной магнитной обсерватории и / или вариационной станции). В этом случае с ее помощью возможно решение проблемы пространственной, временной или пространственно-временной интерполяции геомагнитных данных в условиях отсутствия или неработоспособности реального измерительного устройства в соответствующей пространственной точке или пространственном регионе. Кроме того, компьютерная модель может быть использована для обработки геомагнитных данных, поступивших с физических информационно-измерительных систем. Таким образом, ее применение технически позволит решить проблему, связанную с выраженной пространственной неоднородностью и гетерогенностью источников геомагнитной информа-

ции. Компьютерная модель обеспечивает при этом приведение результирующих геомагнитных данных к единому формату с заданным шагом дискретизации с использованием одних и тех же единиц измерения параметров геомагнитного поля и его вариаций.

3. Описание предлагаемой компьютерной модели

По сути, виртуальная магнитная обсерватория / вариационная станция в общем случае представляет собой компьютерную модель, которая дополняет известные результаты наблюдений и на их основе прогнозирует отсутствующие во времени и / или пространстве значения. Иными словами, в нашем случае речь идет именно об интерполяции геомагнитных данных в некоторой геопространственной точке, по которой по тем или иным причинам (человеческий или техногенный фактор, проявления экстремальной космической погоды, отсутствие средства измерения в точке и пр.) отсутствуют данные измерений параметров геомагнитного поля и его вариаций.

Представляется целесообразным использовать специфику пространственно-временного распределения параметров геомагнитного поля и его вариаций, а также их зависимость от различных внешних и внутренних факторов для получения наилучшего с точки зрения минимизации ошибки восстановления данных метода интерполяции [7, 8]. С этой целью авторами был проведен ряд вычислительных экспериментов, которые при разном сочетании внешних и внутренних факторов обеспечивали восстановление пространственно-временного ряда геомагнитных данных. При этом вычислительные эксперименты были проведены для условных трех пространственных групп восстанавливаемых геомагнитных данных, разделенных по своей широтной принадлежности на высокоширотные, низкоширотные и приэкваториальные области наблюдений. Для достижения поставленной цели определения наиболее эффективного с учетом всех возможных факторов метода пространственной интерполяции были проанализированы результаты геомагнитных наблюдений за 3 календарных года, полученные магнитными обсерваториями и вариационными станциями по всему миру и доступными посредством API, обеспечивающего обращение к геомагнитным данным проекта SuperMAG (URL: <https://supermag.jhuapl.edu/mag/>) [9, 10].

В общем виде для пространственной интерполяции пространственные регионы были разделены на три группы, определенные спецификой пространственного распределения параметров геомагнитного поля: по мере приближения к экватору и полюсам число факторов, влияющих на результаты наблюдений, возрастает (к примеру, кольцевые токи на экваторе и суббури в полярных областях), что, в свою очередь, приводит к росту неопределенности данных о состоянии геомагнитного поля и его вариаций, а также повышает вероятность резкого «скачкообразного» изменения соответствующих уровней временных рядов [3, 5, 9]. В качестве дополнительного фактора был определен индекс геомагнитной активности, характеризующий геомагнитную обстановку в соответствии с параметрами космической погоды. Известно, что в периоды высокой геомагнитной активности, обусловленной экстремальными проявлениями космической погоды, геомагнитные вариации наиболее выражены по сравнению со своими значениями в магнитоспокойные периоды [4, 8, 9].

Анализ эффективности методов пространственной интерполяции был проведен для каждой из выделенных пространственных областей по имеющимся годовым геомагнитным данным, регистрируемым магнитными обсерваториями и станциями с шагом в 1 мин (таблица). Для простоты в работе приведены результаты исследований по одному из компонент вектора геомагнитного поля (горизонтальная составляющая), однако было установлено, что полученные заключения справедливы и для остальных параметров вектора геомагнитного поля. Для проведения вычислительных экспериментов из исходных данных на каждом этапе изымалась одна пространственная точка, атрибутивное значение в которой было рассмотрено как неизвестное и требующее вычисления методами интерполяции. Далее полученный результат сравнивался с оригинальным значением.

При пространственной интерполяции высокоширотных геомагнитных данных наилучшие результаты были получены при применении метода кригинга [11, 12]. В его основе лежит вариограмма, характеризующая пространственную модель известных (опорных) пространственных точек. Каждой

опорной пространственной точке присваивается свой весовой коэффициент, значение которого тем больше, чем ближе известная точка расположена (территориально) по отношению к интерполируемой [11].

В общем виде вычисление атрибутивного значения $Z^*(x_0, y_0)$ в точке (x_0, y_0) по известным n атрибутивным значениям опорных точек $Z(x_i, y_i)$ может быть представлено регрессионным выражением [11]

$$Z^*(x_0, y_0) - m(x_0, y_0) = \sum_{i=1}^n \lambda_i [Z(x_i, y_i) - m(x_i, y_i)],$$

где $m(x_i, y_i)$ и $m(x_0, y_0)$ – математическое ожидание для $Z(x_i, y_i)$ и $Z(x_0, y_0)$ соответственно.

При этом было установлено, что в условиях спокойной магнитосферы наиболее эффективна пространственная интерполяция посредством простого кригинга. В этом случае предполагается, что среднее значение опорных пространственных точек постоянно и известно на всей области исследования ($m(z) = m = \text{const}$ для всех z):

$$Z^*(x_0, y_0) = m + \sum_{i=1}^n \lambda_i [Z(x_i, y_i) - m].$$

**Результаты вычислительных экспериментов
по исследованию методов интерполяции геомагнитных данных**

Пространственные регионы	Геомагнитная активность	Методы интерполяции, значение среднеквадратической ошибки [нТл]									
		Линейная интерполяция	Метод обратно взвешенных расстояний	Полиномиальная регрессия	Триангуляция Делоне	Кригинг	Метод Шепарда	Радиальная интерполяция	Метод градиентных плоскостей	Метод сплайнов	Метод трендов
Высокоширотные	Нет	3,23	5,56	2,98	1,82	1,6	5,56	3,18	4,14	5,12	4,98
	Слабая буря	4,16	6,12	3,12	2,19	1,89	6,12	4,78	6,29	7,15	6,89
	Сильная буря	5,32	8,45	7,15	3,45	2,51	8,45	6,97	7,15	9,12	8,97
Приэкваториальные	Нет	1,2	1,51	1,78	0,17	0,54	1,51	3,21	4,78	5,97	5,3
	Слабая буря	2,5	4,8	2,41	0,65	0,97	4,8	4,86	5,9	5,95	6,3
	Сильная буря	3,1	5,73	5,53	1,31	1,35	5,73	12,1	10,4	11,5	11,9
Низкоширотные	Нет	18,6	17,12	10,15	0,22	7,36	17,12	10,6	11,8	12,1	11,78
	Слабая буря	22,19	20,45	15,92	2,12	10,12	20,45	15,8	16,2	16,98	16,45
	Сильная буря	40,1	28,34	27,6	3,5	12,7	28,34	28,7	28,9	27,6	25,1

В условиях неспокойной магнитосферы наилучшие показатели пространственной интерполяции показал метод универсального кригинга с константной функцией (так называемый кригинг с трендом) [11]. Тренд при этом задается как линейная комбинация известных базисных функций $f_k(x_0, y_0)$ с коэффициентами β_k . Последние, в свою очередь, неизвестны и постоянны внутри окрестности оцениваемой точки (x_0, y_0) :

$$Z^*(x_0, y_0) = \sum_{i=1}^n f_k(x_i, y_i) \beta_k + R(x_0, y_0).$$

Для низкоширотных и приэкваториальных регионов наилучшие результаты пространственной интерполяции были получены с помощью триангуляция Делоне [13, 14]. Для опорных пространственных точек с координатами (X, Y) и атрибутивным значением вида $Z = Z(X, Y)$ построение триангуляции Делоне $DT(S)$ может быть представлено следующим образом. Требуется построить треугольник, образованный опорными пространственными точками S_1, S_2, S_3 и содержащий искомую (интерполируемую) точку. Необходимо определить уравнение плоскости того треугольника, который

включает искомую точку. В общем виде уравнение плоскости содержит четыре неизвестных коэффициента a, b, c, d :

$$a \cdot X + b \cdot Y + c \cdot Z + d = 0.$$

Указанные коэффициенты определяются следующими соотношениями:

$$a = Y_1(Z_2 - Z_3) + Y_2(Z_3 - Z_1) + Y_3(Z_1 - Z_2),$$

$$b = Z_1(X_2 - X_3) + Z_2(X_3 - X_1) + Z_3(X_1 - X_2),$$

$$c = X_1(Y_2 - Y_3) + X_2(Y_3 - Y_1) + X_3(Y_1 - Y_2),$$

$$d = X_1(Y_2Z_3 - Y_3Z_2) + X_2(Y_3Z_1 - Y_1Z_3) + X_3(Y_1Z_2 - Y_2Z_1),$$

где $X_1, Y_1, X_2, Y_2, X_3, Y_3$ – координаты соответствующих точек S_1, S_2, S_3 ; Z_1, Z_2, Z_3 – значения параметров магнитного поля соответствующих точек S_1, S_2, S_3 .

Соответственно, искомая величина $Z^*(X_0, Y_0)$ может быть получена в соответствии с уравнением

$$Z^*(x_0, y_0) = \frac{-a \cdot X_0 - b \cdot Y_0 - d}{c},$$

где a, b, c, d – найденные коэффициенты из уравнения плоскости; X_0, Y_0 – координаты интерполируемой пространственной точки.

С учетом сказанного предлагается следующий алгоритм динамического подбора методов интерполяции (рис. 3):

1. Проведение серии вычислительных экспериментов для определения эффективности разных методов в разных областях (низкие, средние и высокие широты) и при различных внешних факторах (в случае геомагнитных данных – индекс геомагнитной активности).

2. Основанная на результатах экспериментов подготовка пула методов.

3. При поступлении запроса на интерполяцию значения в некоторой точке в заданный момент времени выполнение интерполяции, выбрав метод на основе принадлежности точки определенной пространственной области и степени влияния внешних факторов в указанный момент времени.

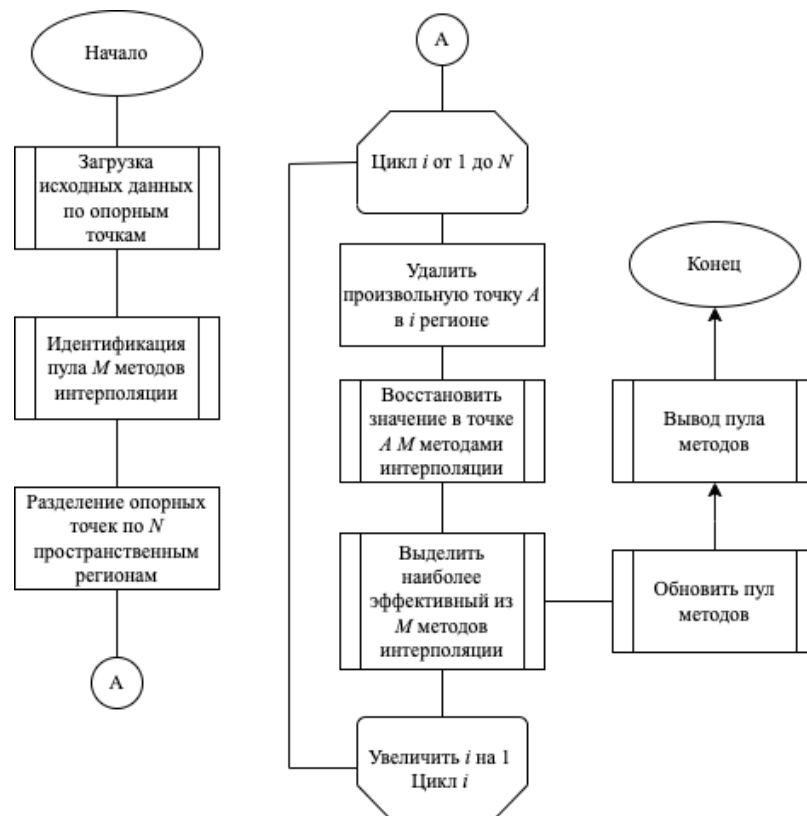


Рис. 3. Алгоритм адаптивной интерполяции
Fig. 3. Adaptive interpolation algorithm

4. Описание и апробация предлагаемого подхода

Предложенный подход был программно реализован в рамках веб-ориентированного исследовательского прототипа приложения, обеспечивающего пространственную интерполяцию и визуализацию геомагнитных данных на основании доступных результатов, полученных наземными магнитными обсерваториями. В общем виде архитектура разработанного веб-приложения и его экранная форма приведены на рис. 4.

В качестве инструмента реализации был выбран Python, что обусловлено опытом авторов в обработке геопространственных данных, а также внушительной базой инструментария для анализа данных, разрабатываемого и поддерживаемого для данного языка программирования. Реализация серверных сценариев использует архитектурный шаблон MVC (Model–View–Controller), который разделяет приложение на три компонента: модель, представление и контроллер. MVC позволяет разработчикам сосредоточиться на отдельных аспектах приложения во время разработки [15, 16].

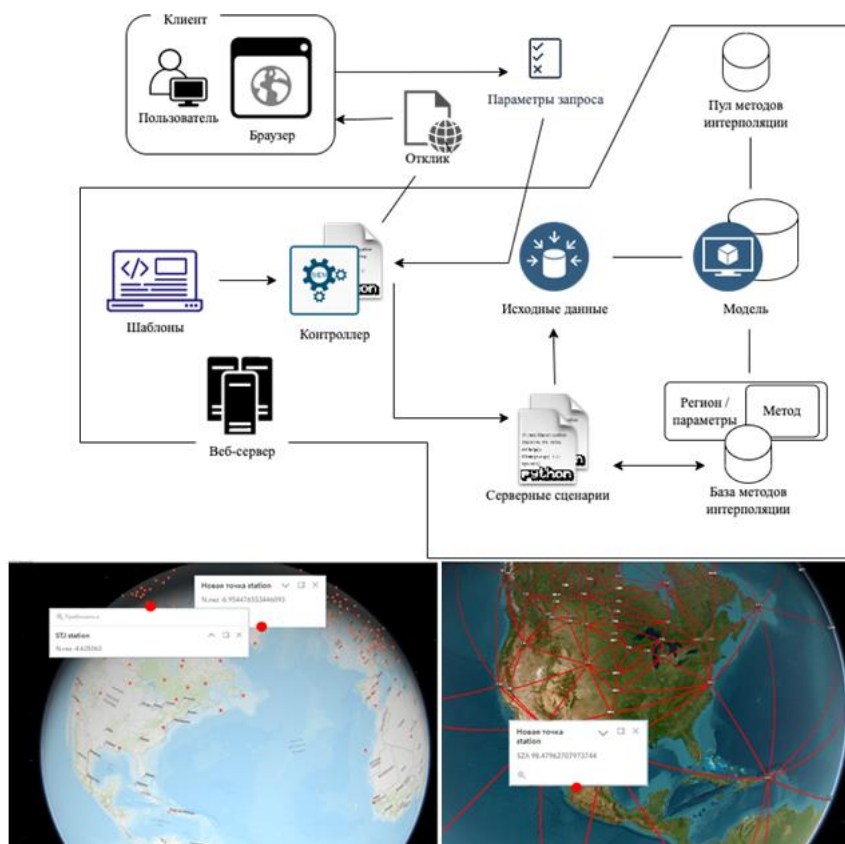


Рис. 4. Архитектура и экранные формы предлагаемого решения
Fig. 4. Architecture and screen forms of the proposed solution

Модель (Model) является объектно-ориентированным отображением данных предметной области и применяется, в частности, в качестве модели ORM при взаимодействии с реляционными базами данных [17]. Контроллер (Controller) взаимодействует с данными (моделью), принимая запросы от клиента, обрабатывая их и формируя результаты. Компонент представления (View) можно рассматривать как опциональный, так как решение представляет наибольший интерес и практическую пользу как удаленная функция для вычисления интерполированного значения.

Функциональность приложения была реализована в двух формах. С одной стороны, представление обеспечивает передачу HTML-разметки с программными компонентами для визуализации результатов в браузере. С другой стороны, предусмотрена возможность взаимодействия с серверной частью через RESTful API, позволяющая сторонним приложениям запрашивать интерполированные данные, отправляя сервису запрос с входными параметрами [18, 19, 20].

Для оценки эффективности предложенного решения проблемы пространственной интерполяции геомагнитных данных был создан исследовательский прототип веб-приложения. Для каждого метода интерполяции был разработан свой серверный модуль, а также создан серверный сценарий, выбирающий наиболее эффективный метод интерполяции.

Эксперименты проводились на клиентской стороне с использованием компьютера с процессором Intel Core i5 10300N 2,5 ГГц, оперативной памятью размером 4 ГБ и скоростью интернет-соединения 52,4 Мбит/с, а на серверной стороне – на базе веб-сервера с процессором 72 * Intel® Xeon® Gold 6140 CPU @ 2,30 ГГц.

На первом этапе оценивалось время получения отклика от серверной стороны. Было обнаружено, что выполнение клиентского запроса с учетом информации о наиболее эффективном методе интерполяции занимает в среднем 12 с, что соответствует требованиям к высокопроизводительным веб-приложениям.

На втором этапе оценивалась ошибка пространственной интерполяции при использовании предложенного адаптивного подхода в сравнении с существующими методами. Исследования проводились на геомагнитных данных в разные периоды геомагнитной активности и для разных пространственных областей: средне-, высоко- и низкоширотных.

Анализ полученных результатов показал, что использование предложенного адаптивного подхода к пространственной интерполяции позволяет снизить среднеквадратическую ошибку восстановления пространственных данных по сравнению с существующими подходами следующим образом: в среднеширотных областях – на 4,71 нТл, в высокоширотных областях – на 4,95 нТл, в низкоширотных областях – на 16,7 нТл.

Заключение

Современные инструменты мониторинга геомагнитного поля включают наземные обсерватории, станции и спутники. Системный подход рассматривает геомагнитный мониторинг как систему элементов, работающих под воздействием внешних и внутренних факторов, которая формирует результирующие магнитные данные, направляемые в надсистему – центр данных, а далее – заинтересованным лицам. Вепольный анализ, проведенный в рамках работы, выявил основные проблемы получения геомагнитных данных, в частности проблему отсутствия данных в некоторых точках земной поверхности.

Разработанная концепция предлагает введение дополнительных компьютерных моделей в систему поддержки принятия решений на основе геомагнитной информации таким образом, чтобы обеспечить получение данных о геомагнитной обстановке в том числе и в тех точках, где отсутствуют или недоступны средства измерения. Предложен подход, комбинирующий методы интерполяции с наименьшей ошибкой для разных условий. На подготовительном этапе для каждого метода оценивается ошибка интерполяции для разных областей и внешних факторов. При дальнейшей интерполяции определяется принадлежность искомой точки одной из пространственных областей, оцениваются внешние факторы и выбирается оптимальный метод интерполяции. Эффективность предложенного решения показана на примере геомагнитной информации для разных широт и уровней активности Кр-индекса.

Список источников

1. Vorobev A., Soloviev A., Pilipenko V., Vorobeva G., Sakharov Y. An Approach to Diagnostics of Geomagnetically Induced Currents Based on Ground Magnetometers Data // Appl. Sci. 2022. V. 12. Art. 1522. doi: 10.3390/app12031522
2. Vorobev A.V., Vorobeva G.R. Approach to Assessment of the Relative Informational Efficiency of Intermagnet Magnetic Observatories // Geomagn. Aeron. 2018. V. 58. P. 625–628. doi: 10.1134/S0016793218050158
3. Кориков А.М., Павлов С.Н. Теория систем и системный анализ. М. : ИНФРА-М, 2024. 288 с.
4. Imboden D., Pfenninger S. Introduction to Systems Analysis Mathematically Modeling Natural Systems. Berlin,; Springer, 2012. 252 p.
5. Петров В. Структурный анализ систем. Вепольный анализ. ТРИЗ. Екатеринбург : Издательские решения, 2018. 280 с.

6. Гареев Р.Т. Эвристические приемы ТРИЗ. М. : МГИУ, 2008. 134 с.
7. Vorobev A.V., Soloviev A.A., Pilipenko V.A., Vorobeve G.R. Interactive computer model for aurora forecast and analysis // *Solar-Terrestrial Physics*. 2022. V. 8 (2). P. 84–90. doi: 10.12737/stp-82202213
8. Vorobev A.V., Pilipenko V.A., Krasnoperov R.I., Vorobeve G.R., Lorentzen D.A. Short-term forecast of the auroral oval position on the basis of the “virtual globe” technology // *Russ. J. Earth Sci.* 2020. V. 20. Art. ES6001. doi: 10.2205/2020ES000721
9. Gjerloev J.W. The SuperMAG data processing technique // *J. Geophys. Res.* 2012. V. 117. Art. A09213.
10. Waters C.L., Gjerloev J.W., Dupont M., Barnes R.J. Global maps of ground magnetometer data // *J. Geophys. Res. Space Physics*. 2015. V. 120. P. 9651–9660. doi:10.1002/2015JA021596
11. Zhang H., Tian Y., Zhao P. Dispersion Curve Interpolation Based on Kriging Method // *Applied Sciences*. 2023. V. 13 (4). Art. 2557. doi: 10.3390/app13042557
12. Lebrezn H., Bardossy A. Geostatistical interpolation by quantile kriging // *Hydrology and Earth System Sciences*. 2019. V. 23. P. 1633–1648. doi: 10.5194/hess-23-1633-2019
13. Alexa M. Conforming weighted Delaunay triangulations // *ACM Transactions on Graphics*. 2020. V. 39. P. 1–16. doi: 10.1145/3414685.3417776
14. Weng Y., Cao J., Chen Zh. Global optimization of optimal Delaunay triangulation with modified whale optimization algorithm // *Engineering with Computers*. 2024. V. 40. P. 1–22. doi: 10.1007/s00366-023-01928-2
15. Nguyen L. et al. Design and Implementation of Web Application Based on MVC Laravel Architecture // *European Journal of Electrical Engineering and Computer Science*. 2022. V. 6. P. 23–29. doi: 10.24018/ejece.2022.6.4.448
16. Rahman M.H., Naderuzzaman M., Kashem M.A. et al. Comparative Study: Performance of MVC Frameworks on RDBMS // *International Journal of Information Technology and Computer Science (IJITCS)*. 2024. V. 16 (1). P. 26–34. doi: 10.5815/ijitcs.2024.01.03
17. Hule K., Ranawat R. Analysis of Different ORM Tools for Data Access Object Tier Generation: A Brief Study // *International Journal of Membrane Science and Technology*. 2023. V. 10. P. 1277–1291. doi: 10.15379/ijmst.v10i1.2842
18. Marculescu B., Zhang M., Arcuri A. On the Faults Found in REST APIs by Automated Test Generation. // *ACM Transactions on Software Engineering and Methodology*. 2022. V. 31. P. 1–43. doi: 10.1145/3491038
19. Golmohammadi A., Zhang M., Arcuri A. Testing RESTful APIs: A Survey // *ACM Transactions on Software Engineering and Methodology*. 2023. V. 33 (1). Art. 27. doi: 10.1145/3617175
20. Bogner J., Kotstein S., Pfaff T. Do RESTful API design rules have an impact on the understandability of Web APIs? // *Empirical Software Engineering*. 2023. V. 28 (6). Art. 132. doi: 10.1007/s10664-023-10367-y

References

1. Vorobev, A., Soloviev, A., Pilipenko, V., Vorobeve, G. & Sakharov, Y. (2022) An Approach to Diagnostics of Geomagnetically Induced Currents Based on Ground Magnetometers Data. *Applied Sciences*. 12. pp. 1522. DOI: 10.3390/app12031522
2. Vorobev, A.V. & Vorobeve, G.R. (2018) Approach to Assessment of the Relative Informational Efficiency of Intermagnet Magnetic Observatories. *Geomagnetism and Aeronomy*. 58. pp. 625–628. DOI: 10.1134/S0016793218050158
3. Korikov, A.M. & Pavlov, S.N. (2024) *Systems theory and systems analysis*. Moscow: INFRA-M.
4. Imboden, D. & Pfenninger, S. (2012) *Introduction to Systems Analysis Mathematically Modeling Natural Systems*. Berlin: Springer.
5. Petrov, V. (2018) *Strukturnyy analiz sistem. Vepol'nyy analiz. TRIZ* [Structural Analysis of Systems. Su-field analysis. TRIZ]. Ekaterinburg: Izdatel'skie resheniya.
6. Gareev, R.T. (2008) *Evristicheskie priemy TRIZ* [Heuristic Techniques of TRIZ]. Moscow: MGIU.
7. Vorobev, A.V., Soloviev, A.A., Pilipenko, V.A. & Vorobeve, G.R. (2022) Interactive computer model for aurora forecast and analysis. *Solar-Terrestrial Physics*. 8(2). pp. 84–90. DOI: 10.12737/stp-82202213
8. Vorobev, A.V., Pilipenko, V.A., Krasnoperov, R.I., Vorobeve, G.R. & Lorentzen, D.A. (2020) Short-term forecast of the auroral oval position on the basis of the “virtual globe” technology. *Russian Journal of Earth Sciences*. 20. P. ES6001. DOI: 10.2205/2020ES000721
9. Gjerloev, J.W. (2012) The SuperMAG data processing technique. *Journal of Geophysical Research*. 117. A09213.
10. Waters, C.L., Gjerloev, J.W., Dupont, M. & Barnes, R.J. (2015) Global maps of ground magnetometer data. *Journal of Geophysical Research: Space Physics*. 120. pp. 9651–9660. DOI: 10.1002/2015JA021596
11. Zhang, H., Tian, Y. & Zhao, P. (2023) Dispersion Curve Interpolation Based on Kriging Method. *Applied Sciences*. 13(4). 2557. DOI: 10.3390/app13042557
12. Lebrezn, H. & Bardossy, A. (2019) Geostatistical interpolation by quantile kriging. *Hydrology and Earth System Sciences*. 23. pp. 1633–1648. DOI: 10.5194/hess-23-1633-2019
13. Alexa, M. (2020) Conforming weighted Delaunay triangulations. *ACM Transactions on Graphics*. 39. pp. 1–16. DOI: 10.1145/3414685.3417776
14. Weng, Y., Cao, J. & Chen, Zh. (2024) Global optimization of optimal Delaunay triangulation with modified whale optimization algorithm. *Engineering with Computers*. 40. pp. 1–22. DOI: 10.1007/s00366-023-01928-2
15. Nguyen, L. et al. (2022) Design and Implementation of Web Application Based on MVC Laravel Architecture. *European Journal of Electrical Engineering and Computer Science*. 6. pp. 23–29. DOI: 10.24018/ejece.2022.6.4.448

16. Rahman, M.H., Naderuzzaman, M., Kashem, M.A. et al. (2024) Comparative Study: Performance of MVC Frameworks on RDBMS. *International Journal of Information Technology and Computer Science (IJITCS)*. 16(1). pp. 26–34. DOI: 10.5815/ijitcs.2024.01.03
17. Hule, K. & Ranawat, R. (2023) Analysis of Different ORM Tools for Data Access Object Tier Generation: A Brief Study. *International Journal of Membrane Science and Technology*. 10. pp. 1277–1291. DOI: 10.15379/ijmst.v10i1.2842
18. Marculescu, B., Zhang, M. & Arcuri, A. (2022) On the Faults Found in REST APIs by Automated Test Generation. *ACM Transactions on Software Engineering and Methodology*. 31. pp. 1–43. DOI: 10.1145/3491038
19. Golmohammadi, A., Zhang, M. & Arcuri, A. (2023) Testing RESTful APIs: A Survey. *ACM Transactions on Software Engineering and Methodology*. 33(1). Art. 27. DOI: 10.1145/3617175
20. Bogner, J., Kotstein, S. & Pfaff, T. (2023) Do RESTful API design rules have an impact on the understandability of Web APIs? *Empirical Software Engineering*. 28(6). Art. 132. DOI: 10.1007/s10664-023-10367-y

Информация об авторах:

Воробьева Гульнара Равилевна – доктор технических наук, профессор кафедры вычислительной математики и кибернетики Института математики, информатики и робототехники Уфимского университета науки и технологий (Уфа, Россия). E-mail: gulnara.vorobeva@gmail.com

Воробьев Андрей Владимирович – доктор технических наук, профессор, заведующий кафедрой информатики Института математики, информатики и робототехники Уфимского университета науки и технологий (Уфа, Россия); научный сотрудник Геофизического центра Российской академии наук (Москва, Россия). E-mail: geomagnet@list.ru

Фарваев Эмиль Фанильевич – аспирант кафедры вычислительной математики и кибернетики Института математики, информатики и робототехники Уфимского университета науки и технологий (Уфа, Россия). E-mail: farvaev.emil@gmail.com

Орлов Глеб Олегович – аспирант кафедры вычислительной математики и кибернетики Института математики, информатики и робототехники 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 Ravilevna (Doctor of Technical Sciences, Professor, Computational Mathematics and Cybernetics Institute of Mathematics, Informatics and Robotics, Ufa University of Science and Technology, Ufa, Russian Federation). E-mail: gulnara.vorobeva@gmail.com

Vorobev Andrei Vladimirovich (Doctor of Technical Sciences, Professor, Head of The Informatics Department of Institute of Mathematics, Informatics and Robotics Ufa University of Science and Technology, Ufa, Russian Federation; Researcher, Geophysical Center of RAS, Moscow, Russian Federation). E-mail: geomagnet@list.ru

Farvaev Emil Fanilevich (Post-graduate Student, Department of Computational Mathematics and Cybernetics, Institute of Mathematics, Informatics and Robotics, Ufa University of Science and Technology, Ufa, Russian Federation). E-mail: farvaev.emil@gmail.com

Orlov Gleb Olegovich (Post-graduate Student, Department of Computational Mathematics and Cybernetics, Institute of Mathematics, Informatics and Robotics, 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 25.09.2024; принята к публикации 02.12.2024

Received 25.09.2024; accepted for publication 02.12.2024

Научная статья

УДК 550.834, 519.6

doi: 10.17223/19988605/69/6

**Вычислительная кинематическая диагностика скорости зондирующего сигнала
с использованием априорной информации об исследуемой среде****Николай Владимирович Гервас¹, Сергей Михайлович Зеркаль²**^{1,2} Новосибирский государственный технический университет, Новосибирск, Россия¹ nik.gervas@mail.ru² zerkal@ngs.ru

Аннотация. Излагается вычислительная технология кинематической диагностики скоростного распределения зондирующего сигнала в сейсмоакустических средах в условиях линеаризации возникающей обратной кинематической задачи. Приводятся постановка задачи, построение решения и вычислительный алгоритм, основанный на методе КТ. Использование КТ обеспечивается организацией томографической системы наблюдений (расположение источников и приемников зондирующего сигнала на окружности). Линеаризация выполняется возле линейной функции глубины (переменная Z пространства XYZ), выделяемой в искомой функции скорости и предполагаемой известной. Основное внимание уделяется организации этой составляющей, что существенно влияет на точность и достоверность получаемого решения. Приводятся результаты компьютерных экспериментов, иллюстрирующих решение рассматриваемой задачи.

Ключевые слова: вычислительный алгоритм; обратная задача; рефрагированные волны; кинематическая диагностика; компьютерная томография.

Для цитирования: Гервас Н.В., Зеркаль С.М. Вычислительная кинематическая диагностика скорости зондирующего сигнала с использованием априорной информации об исследуемой среде // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 69. С. 54–63. doi: 10.17223/19988605/69/6

Original article

doi: 10.17223/19988605/69/6

**Computational kinematic diagnostics of the probing signal speed
using a priori information about the medium under study****Nikolai V. Gervas¹, Sergey M. Zerkal²**^{1,2} Novosibirsk State Technical University, Novosibirsk, Russian Federation¹ nik.gervas@mail.ru² zerkal@ngs.ru

Abstract. A computational technology for kinematic diagnostics of the velocity distribution of a sounding signal in seism acoustic media is presented under the conditions of linearization of the emerging inverse kinematic problem. The problem statement, construction of a solution and a computational algorithm based on the CT method are presented. The use of CT is ensured by the organization of a tomographic observation system (location of sources and receivers of the probing signal on a circle). Linearization is performed near the linear depth function (variable Z of the XYZ space), identified in the desired velocity function and assumed to be known. The main attention is paid to the organization of this component, which significantly affects the accuracy and reliability of the resulting solution. The results of computer experiments illustrating the solution of the problem under consideration are presented.

Keywords: computational algorithm; inverse problem; refracted waves; kinematic diagnostics; computed tomography.

For citation: Gervas, N.V., Zerkal, S.M. (2024) Computational kinematic diagnostics of the probing signal speed using a priori information about the medium under study. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 69. pp. 54–63. doi: 10.17223/19988605/69/6

Введение

Основными характеристиками физических процессов, изучаемых с позиций «кинематики», являются скорость, ускорение, путь и время. Установленные здесь физические законы в своих формулах закрепляют связь между этими характеристиками, что позволяет определять какую-либо из них через измерение других. Это составляет основу решения многих практических проблем, в том числе задач, возникающих при дистанционном измерении внутреннего строения сейсмоакустических и оптических сред на основе теории распространения волн. В этой тематике одной из важных задач считается обратная кинематическая задача (ОКЗ), имеющая математическую формулировку как задача интегральной геометрии, опирающаяся на ариационное исчисление [1, 2]. ОКЗ имеет свою историю, и истоки ее восходят к геофизике. Первая постановка ОКЗ была исследована Г. Герглотцем и Е. Вихертом в 1905–1907 гг. для сферически симметричной модели Земли при наличии роста скорости зондирующего излучения с глубиной (от поверхности к центру Земли). Результаты Г. Герглотца и Е. Вихерта позволили с использованием данных сейсмических наблюдений над землетрясениями сделать первые выводы о глубинном строении Земли. Вскоре нашлось еще одно важное приложение кинематической сейсмики, связанное с лоцированием тяжелой артиллерии противника в ходе Первой мировой войны, при этом использовался феномен поверхностных волн.

ОКЗ исследовалась в одномерной постановке до 60-х гг. прошлого столетия, когда была рассмотрена двумерная постановка (М.М. Лаврентьев, В.Г. Романов), оказавшаяся математически сильно некорректной. ОКЗ, по-видимому, – первая рассмотренная обратная задача для дифференциальных уравнений (уравнение Эйконала) [1, 2]. Известные конструктивные решения ОКЗ получены для линейаризованных постановок. До сих пор решения ОКЗ в полной, нелинейной постановке не получено. Следует отметить, что одномерная ОКЗ имеет важное значение для морской гидроакустики [3–5].

1. Постановка задачи и построение решения

Основным математическим понятием наряду с уравнением Эйконала является функционал Ферма, связывающий время прихода рефрагированной волны из пункта генерации волнового сигнала (источник) в пункт регистрации (приемник), обозначим их x^0 и x^1 соответственно, $x \in R^3$, тогда

$$T(x^0, x^1) = \int_{\gamma(x^0, x^1)} \frac{dS}{V(x)}, \quad (1)$$

где отмеченное время «пробега» рефрагированной волны вдоль лучевой траектории (геодезической) $\gamma(x^0, x^1)$, соединяющей точки x^0 и x^1 , являющейся решением системы уравнений Эйлера для данного функционала, dS – элемент дуги кривой $\gamma(x^0, x^1)$, $V(x)$ – скорость рефрагированной волны в исследуемой среде. Известно, что $\gamma(x^0, x^1)$ доставляет минимум значению данного функционала (принцип Ферма). Вышесказанное находится в соответствии с постановкой задачи геометрической оптики в пространстве [6].

Требуется, зная значения функции $T(x^0, x^1)$, т.е. для различных положений источника и приемника (результаты измерений), найти функцию $V(x)$; очевидно, без знания $\gamma(x^0, x^1)$ это сделать невозможно (на практике эта информация отсутствует). Виды лучевых траекторий определяются скоростным распределением, именно оно и подлежит определению. Таким образом, имеется серьезное обстоятельство, препятствующее решению задачи, т.е. построению обратного оператора; если выражение (1) записать в операторном виде $AV = T$, то $A^{-1}AV = A^{-1}T$, $V = A^{-1}T$.

Однако при наличии определенной априорной информации о скорости распространения используемого волнового сигнала, т.е. о функции V , а именно представлении ее в виде суммы

$$V = V_0 + V_1, V_0 \gg |V_1|, \quad (2)$$

где V_0 – известная функция или константа, становится возможным решение ОКЗ с применением метода линеаризации задачи «возле V_0 ». Систематическое использование линеаризации началось с работ М.М. Лаврентьева и В.Г. Романова [2] в начале второй половины прошлого столетия. Относительно V_0 рабочими вариантами могут быть

$$V_0 = A, A = \text{const}, A > 0, \quad (3)$$

$$V_0 = A + Bx_3, V_1(x), x \in R^3, A = \text{const}, B = \text{const}, A > 0, B > 0. \quad (4)$$

Оба варианта имеют физический смысл и практическое значение. Например, в океане скорость V имеет небольшой градиент по глубине и ее вариации составляют первые проценты от константы 1450, однако такое волновое явление, как «подводный звуковой канал», и некоторые другие заметные колебания возле отмеченной константы делают решение ОКЗ актуальным элементом морской акустики. В этой связи вариант формулы (3) позволяет разрабатывать вычислительные алгоритмы для решения возникающих постановок ОКЗ [3–5]. Вариант формулы (4) является актуальным для сейсмоакустических исследований Земли – как верхней части геологического разреза (инженерные исследования, например подготовка строительной площадки, диагностика состояния фундаментов крупных сооружений, внутреннего пространства туннелей и т.д., разведка и мониторинг состояния эксплуатируемых месторождений полезных ископаемых, в частности углеводородов), так и глубинных сейсмических исследований земной коры и верхней мантии Земли (глубинное сейсмическое зондирование – ГСЗ). Это объясняется, например, тем, что с уплотнением горных пород с глубиной функция V возрастает и имеет тенденцию к квазилинейному росту, что характерно, в частности, для глубоких осадочных «бассейнов», перспективных для залежей углеводородов. На территории России это Южный Урал (Прикаспийская низменность) и Западная Сибирь, известные такими месторождениями; следует отметить в этой связи и океанический шельф.

Подчеркнем, что при организации системы наблюдений (расположение источников и приемников) на «дневной» поверхности (поверхность Земли), т.е. на плоскости $x_3 = 0$, функция $T(x^0, x^1)$ является функцией четырех переменных, поскольку $x^0 = (x_1^0, x_2^0, 0)$ и $x^1 = (x_1^1, x_2^1, 0)$, и задача определения трехмерной функции $V(x_1, x_2, x_3)$ по $T(x^0, x^1)$ есть задача переопределения; кроме того, традиционные профильные системы наблюдений (точки x^0 и x^1 расположены на одной линии) не обеспечивают математическую полноту исходных данных для решения ОКЗ, что отражается на достоверности результата (интерпретации результатов измерений) [1].

Эти проблемы снимаются при выборе системы наблюдений в виде окружности заданного радиуса r с центром в точке $(0, 0, 0)$ на плоскости $x_3 = 0$, имеющей уравнение $x_1^2 + x_2^2 = r^2$, $r = \text{const}$, $r > 0$. Если точки x^0 и x^1 расположены на одной окружности (рис. 1), то время T становится функцией двух полярных углов φ_1, φ_2 и радиуса окружности r , т.е. переопределенность задачи снимается, по функции трех переменных T определяется функция трех переменных V .

Примечательно, что для скорости $V = V_0$, т.е. $V = A + Bx_3$, лучевые траектории $\gamma(x^0, x^1)$, определяются уравнением

$$\left(x_3 + \frac{A}{B}\right)^2 + |x|^2 = R^2, R = \text{const}, R > 0, \quad (5)$$

описывающим сферу с центром в точке $(0, 0, -A/B)$, и уравнением плоскости Q

$$\langle x, v \rangle - P = 0, P = \text{const}, v = (\sin \theta, -\cos \theta), \quad (6)$$

перпендикулярной плоскости $x_3 = 0$ и пересекающей ее по линии, включающей $l(x^0, x^1)$, т.е. $\gamma_0(x^0, x^1)$ – это след пересечения сферы (5) и плоскости (6).

Постоянная R легко вычисляется в рассматриваемой постановке как

$$R = \sqrt{\left(\frac{A}{B}\right)^2 + r^2}. \quad (7)$$

Обозначим через d расстояние между источником и приемником, определяемое как расстояние между двумя точками с заданными координатами. Тогда формула для вычисления времени T_0 имеет вид:

$$T_0(x^0, x^1) = \frac{1}{B} \ln \frac{2R + d}{2R - d}. \quad (8)$$

Итак, в соответствии с приведенными выше распределениями можно заменить $T(x^0, x^1) \approx \approx T_0 + \int_{\gamma_0(x^0, x^1)} n_1 dS$, или $\int_{\gamma_0(x^0, x^1)} n_1 dS = T - T_0$, где $n_1 = \frac{1}{V} - \frac{1}{V_0} = n - n_0$.

Если $T - T_0$ обозначить через T_1 , то

$$\int_{\gamma_0} n_1 dS \approx T_1. \quad (9)$$

Интегральное уравнение (9) может быть получено также в результате линеаризации уравнения Эйконала и с точностью до малых второго порядка относительно T_1 и n_1 , знак приближенного равенства заменим на $=$, т.е. $\int_{\gamma_0} n_1 dS = T_1$.

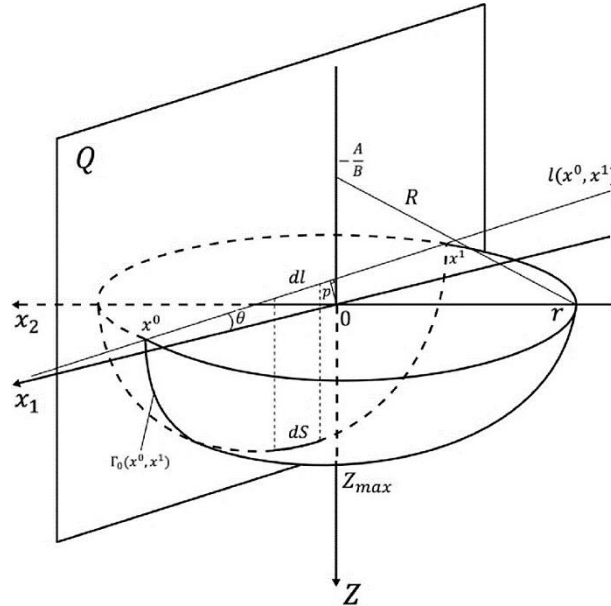


Рис. 1. Иллюстрация геометрических условий линеаризованной постановки ОКЗ

Fig. 1. Illustration of the geometric conditions of the IKP linearized formulation

Далее при переходе от интегрирования по γ_0 к интегрированию по l возникает весовой множитель. Так как $Z = -\frac{A}{B} \pm \sqrt{R^2 - |x|^2} = Z^*$, $x \in R^2$, $Z \equiv x_3$, то с учетом того, что для уравнения $l(x^0, x^1)$ имеем $x_1 \sin \theta - x_2 \cos \theta = P$, и поскольку исследуемый процесс происходит в полупространстве $Z \geq 0$, для dS можно получить $dS = dl \sqrt{1 + (Z'_l)^2}$, или $dS = \frac{\sqrt{R^2 - P^2}}{Z^* + \frac{A}{B}} dl$.

Таким образом, получаем выражения

$$T_1(x^0, x^1) = \sqrt{R^2 - P^2} \int_{l(x^0, x^1)} \frac{n_1(x, Z^*)}{Z^* + \frac{A}{B}} dl, \quad (10)$$

$$\frac{T_1(x^0, x^1)}{\sqrt{R^2 - P^2}} = \int_{l(x^0, x^1)} \frac{n_1(x, Z^*)}{Z^* + \frac{A}{B}} dl. \quad (11)$$

Выражение (11) позволяет, используя интегральное преобразование Радона, получить формулу обращения для вычисления значений $n_1(x, Z^*)$:

$$n_1(x, Z^*) = \left(Z^* + \frac{A}{B} \right) \mathbb{R}^{-1} f(P, \theta), \quad (12)$$

где $\mathbb{R}^{-1}$ – оператор обратного преобразования Радона, а $f(P, \theta) = \frac{T_1(x^0, x^1)}{\sqrt{R^2 - P^2}}$, P и θ – параметры, определяющие прямые $l: P - \langle x, v \rangle = 0, P = \text{const}, v = (\sin \theta, -\cos \theta)$ (см. рис. 1).

Итак, рассматриваемая постановка ОКЗ позволяет свести ее решение к решению классической томографической задачи в круге радиуса r и получить в результате значения скоростного распределения на поверхности второго порядка, образованной множеством лучевых траекторий γ_0 . Изменение радиуса окружности системы наблюдений позволит получить систему вложенных отмеченных поверхностей, на которых искомая скорость V определена, и тем самым возникает объемное скоростное распределение для зондирующего сигнала в исследуемой среде.

Организация томографических измерений (получение проекционных данных) требует задания числа пар источник–приемник при фиксированном их положении (одна томографическая проекция), и количество этих положений, определяемых поворотом системы наблюдений от 0 до 2π с заданным шагом (углом сканирования), соответственно N и M (см. рис. 1). Линеаризация скорости происходит возле линейной функции глубины Z , т.е. возле $V_0 = A + BZ$, $A > 0, B > 0, A = \text{const}, B = \text{const}$. Здесь r – радиус окружности системы наблюдений на плоскости $Z = 0$ с центром в начале системы координат. R – радиус сферической поверхности, образованной лучевыми траекториями (5), Q – плоскость, определенная уравнением (6).

Основным элементом формирования проекционных данных $f(P, \theta)$ является численное решение прямой кинематической задачи (ПКЗ), т.е. вычисление значений $T(x^0, x^1)$ и при этом построение лучей $\gamma(x^0, x^1)$, так как вдоль $\gamma(x^0, x^1)$ распространяется рефрагированный сигнал и фиксируется время его пробега. В общем случае при заданной функции $V(x)$, $x \in R^3$, решение ПКЗ заключается в численном решении начально-краевой задачи для системы трех уравнений второго порядка, разрешенных относительно старших производных (система является следствием из уравнения Эйконала, и ее часто называют «система уравнений луча» [1]), для этого используется разностная схема метода Рунге–Кутты. Приходится решить так называемую «двухточечную задачу», т.е. требуется уметь находить лучевую траекторию, соединяющую две заданные точки, а отмеченная начально-краевая постановка напрямую этого не обеспечивает, поскольку позволяет построить луч, выходящий из заданной точки с заданными направляющими косинусами. Для решения двухточечной задачи применяется вычислительная методика пристрелки. С использованием решения начально-краевой задачи подбираются начальные данные, и в окрестности нужной точки «пристреливаются», например, три точки и делается интерполяция в заданную точку. В результате получаем приближенное значение $T(x^0, x^1)$.

В случае организации томографического численного эксперимента такое вычисление проекционных данных даже при «малоракурсной» томографии, как правило, – весьма трудоемкое задание в связи с большими объемами значений $T(x^0, x^1)$. Однако известен класс сред, для которых значения $T(x^0, x^1)$ вычисляются по явной формуле, что позволяет избежать отмеченных выше трудностей. Это среды, в которых показатель преломления задается формулой

$$n(x) = \sqrt{\lambda_0 + \langle \beta, x \rangle + \alpha^2 |x|^2} \quad (13)$$

при $x \in R^3$, $\lambda_0 = \text{const}$, $\alpha = \text{const}$, $\beta = (\beta_1, \beta_2, \beta_3)$, где $\beta_1 = \text{const}$, $\beta_2 = \text{const}$, $\beta_3 = \text{const}$ [7]. Имеет место для $T(x^0, x^1)$ формула

$$T(\overline{x_0}, \overline{x_1}) = \frac{M\sqrt{M^2 - 1}}{2\alpha^2} (\alpha^2(\lambda_0^2 + \langle \beta, \overline{x_0} \rangle + \alpha^2 |\overline{x_0}|^2) + (\beta + \alpha^2 x_0)) +$$

$$\begin{aligned}
 & + \frac{\ln(M + \sqrt{M^2 - 1})}{2\alpha^3} (\alpha^2(\lambda_0^2 + \langle \beta, \bar{x}_0 \rangle + \alpha^2 |\bar{x}_0|^2)) - (\beta + \alpha^2 \bar{x}_0) + \\
 & + \frac{\sqrt{M^2 - 1}}{\alpha^3} ((\alpha^2 \bar{x}_1 + \beta) - (\alpha^2 \bar{x}_0 + \beta)M)(\beta + \alpha^2 \bar{x}_0),
 \end{aligned} \tag{14}$$

где

$$\begin{aligned}
 M = & \frac{-(\alpha^2 \bar{x}_1 + \beta)(\beta + \alpha^2 \bar{x}_0)}{\alpha^2 \lambda_0^2 - \beta^2} + \\
 & + \frac{\sqrt{[(\alpha^2 \bar{x}_1 + \beta)(\beta + \alpha^2 \bar{x}_0)]^2 + (\alpha^2 \lambda_0^2 - \beta^2)[\alpha^2(\lambda_0^2 + 2\langle \beta, \bar{x}_0 \rangle + \alpha^2 |\bar{x}_0|^2) + (\alpha^2 \bar{x}_1 + \beta)]}}{\alpha^2 \lambda_0^2 - \beta^2}.
 \end{aligned} \tag{15}$$

Использование формулы (14) накладывает соответствующие ограничения на параметры подкоренного выражения и границы исследуемой пространственной области. Выражение (13) представим как $n(x) = \sqrt{\lambda_0 + \varphi(x)}$, где $|\varphi(x)| < \lambda_0$. Пусть для определенности $\lambda_0 \equiv 1$, что удобно для дальнейших рассуждений. Тогда $|\varphi(x)| < 1$. Далее, поскольку ранее использовалось обозначение $Z = x_3$, то область изменения составляет $0 \leq Z \leq Z_{\max}$, $Z_{\max}$ – максимальное заглубление лучевой траектории, а именно точка ее поворота (см. рис. 1), где $Z_{\max} = R - A/B$, границы изменения переменных x_1 и x_2 определяются неравенствами $|x_1| \leq r$ и $|x_2| \leq r$, в дальнейшем $r \equiv 1$.

В нашем случае удобно воспользоваться результатами работы [7], где для сред с показателем преломления $n(\bar{x}) = \sqrt{1 + \langle \beta, \bar{x} \rangle + \alpha^2 |\bar{x}|^2}$ выводятся явные формулы для решения ПКЗ. По проекционным данным $f(P, \theta)$ строится томограмма, содержащая искомое скоростное распределение, и тем самым решение ОКЗ завершается. В нашем случае для построения томограмм могут быть использованы методы и алгоритмы, являющиеся к настоящему времени классическим математическим обеспечением КТ. Без ограничения общности в компьютерных экспериментах использовался метод Ерохина–Шнейдерова [1] с предварительным сглаживанием проекций (измерения, выполненные под определенным ракурсом) бикубическими сплайнами в соответствии с оценочным уровнем случайной помехи.

Поскольку информация о V_0 , т.е. о константах A и B , носит эвристический характер и их точные значения неизвестны, то представляется важным использовать процедуру оптимизации их выбора. Используемый в данных исследованиях класс сред, имеющих скоростное распределение, описываемое формулой (14), где явно заданное выражение для V_0 отсутствует, представляется удобным для наглядного выполнения такой оптимизации.

2. Численное компьютерное моделирование решения ОКЗ

Обычно о параметре A имеется больше информации, и с его выбором больше определенности, чем с параметром B . В этой связи в численных экспериментах, результаты которых представлены на рис. 2–5, оптимизация проводится по одному параметру B , причем полученная тенденция наглядно демонстрирует имеющуюся зависимость точности решения ОКЗ от выбора B .

Для проведения численного компьютерного моделирования рассмотрим среду с показателем преломления $n(x) = \sqrt{1 - x_3 + 0,2x_1 - 2 \cdot 10^{-4} + 0,04|x|^2}$. Здесь существенным является присутствие сравнимого с градиентом по глубине (переменная x_3) изменения по переменной x_1 , это нарушает симметрию, что представляет практический интерес при диагностике неоднородностей такого типа. Относительно скорости $V(x)$ можно записать

$$V(x) = \frac{1}{n(x)} = \frac{1}{\sqrt{1 + \varphi(x)}} \approx 1 - \frac{\varphi(x)}{2} = 1 + 0,5x_3 - 0,1x_1 + 10^{-4}x_2 - 0,02|x|^2 = 1 + 0,5x_3 + V_1(\bar{x}), \tag{16}$$

где $V_1(\bar{x}) = -0,1x_1 + 10^{-4}x_2 - 0,02|x|^2$.

Таким образом, в качестве констант A и B можно выбрать $A = 1$ и $B = 0,5$. Тогда вся необходимая информация для решения ОКЗ по изложенному выше томографическому подходу получена. Действительно, значения параметра B выбирались в соответствии с таблицей.

Зависимость ошибки Q от изменения параметра B

B	0,45	0,5	0,55	0,6	0,65	0,7	0,75
Q	33%	28%	24%	20%	17%	19%	22%

Значение Q является отклонением результата решения ОКЗ n_{1r} от точного значения n_{1e} , которое вычисляется по формуле

$$Q = \sqrt{\frac{\sum_{i,j} [n_{1r}(x_{1i}, x_{2j}, x_3^*) - n_{1e}(x_{1i}, x_{2j}, x_3^*)]^2}{\sum_{i,j} n_{1e}^2(x_{1i}, x_{2j}, x_3^*)}}, \quad (17)$$

где $x_3^* = x_3^*(x_{1i}, x_{2j})$, i и j – текущие параметры сетки, в узлах которой вычисляются значения n_{1r} .

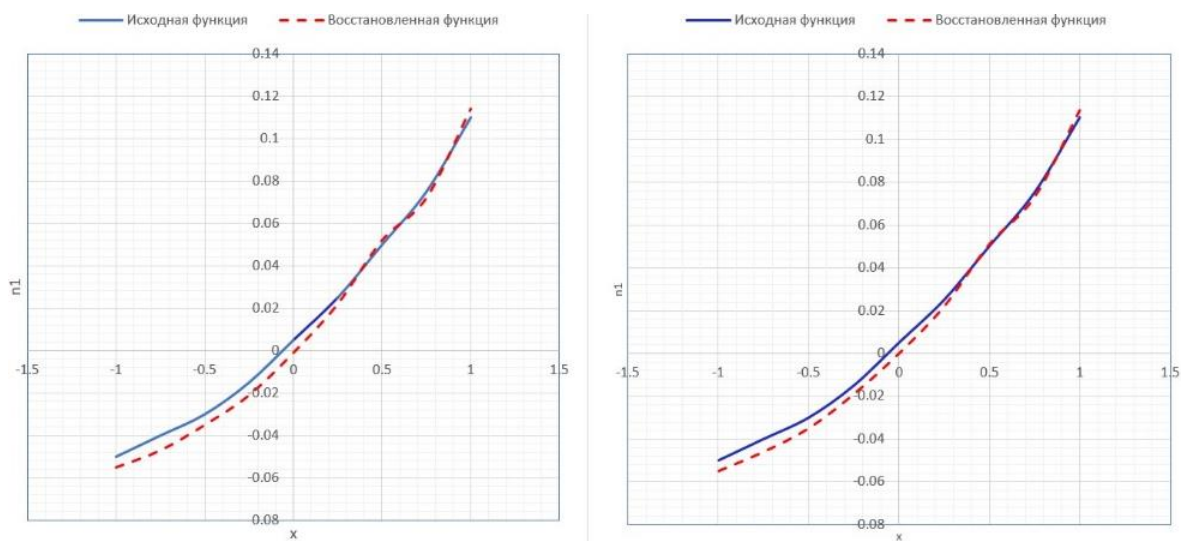


Рис. 2. Графическое изображение сечения плоскостью $x_2 = 0$ пространственного графика функции $n_1(x_1, x_2, Z^*)$ для последовательности значений параметра $B = 0,75$ и $B = 0,7$

Fig. 2. Graphic representation of a section by plane $x_2 = 0$ of the spatial graph of the function $n_1(x_1, x_2, Z^*)$ for a sequence of $B = 0.75$ and $B = 0.7$ values

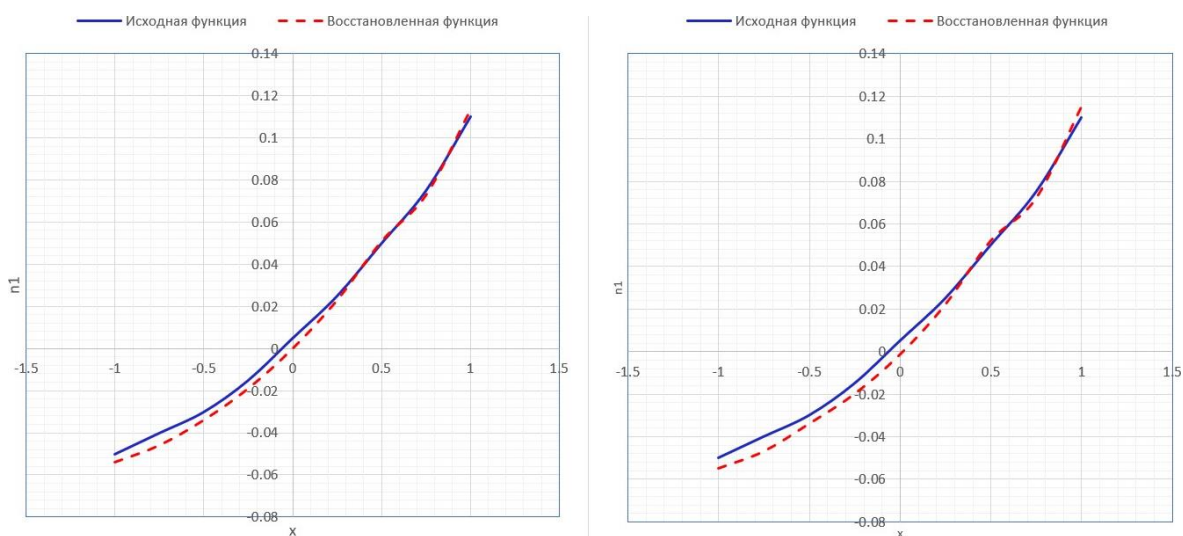


Рис. 3. Графическое изображение сечения плоскостью $x_2 = 0$ пространственного графика функции $n_1(x_1, x_2, Z^*)$ для последовательности значений параметра $B = 0,65$ и $B = 0,6$

Fig. 3. Graphic representation of a section by plane $x_2 = 0$ of the spatial graph of the function $n_1(x_1, x_2, Z^*)$ for a sequence of $B = 0.65$ and $B = 0.6$ values

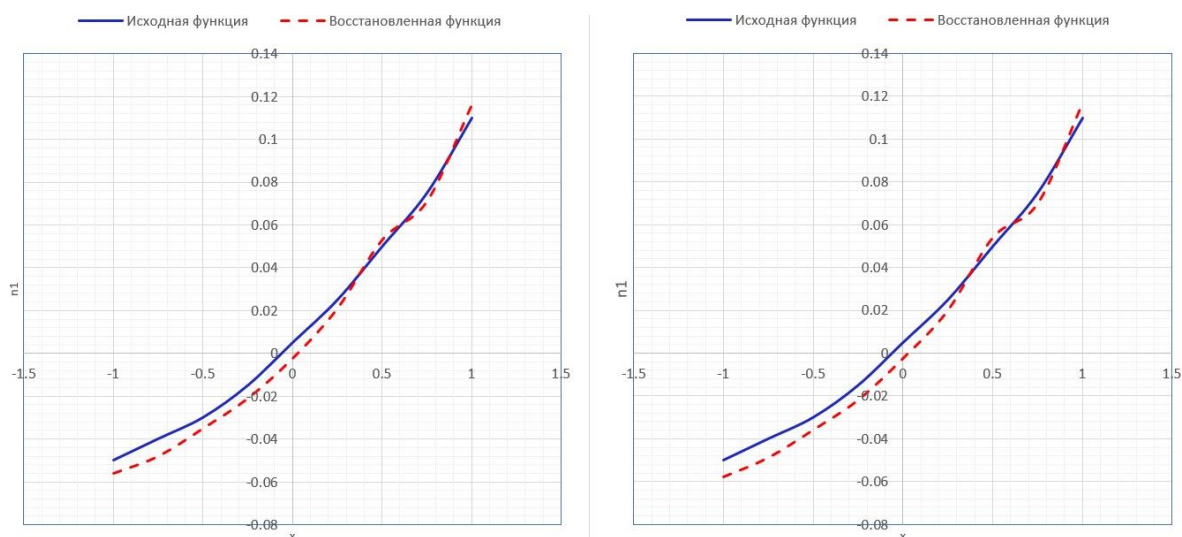


Рис. 4. Графическое изображение сечения плоскостью $x_2 = 0$ пространственного графика функции $n_1(x_1, x_2, Z^*)$ для последовательности значений параметра $B = 0,55$ и $B = 0,5$

Fig. 4. Graphic representation of a section by plane $x_2 = 0$ of the spatial graph of the function $n_1(x_1, x_2, Z^*)$ for a sequence of $B = 0.55$ and $B = 0.5$ values

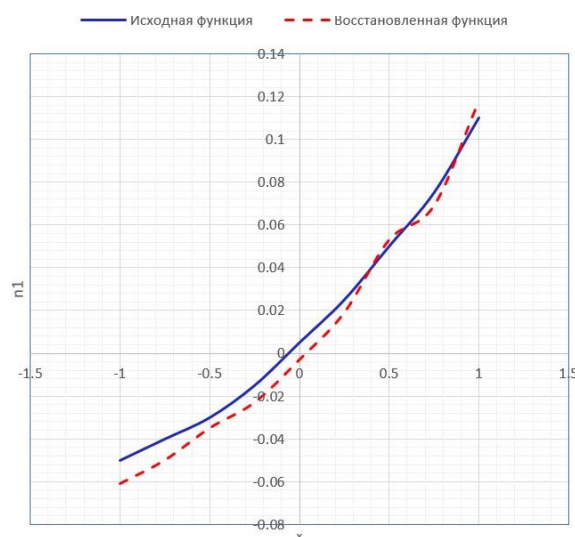


Рис. 5. Графическое изображение сечения плоскостью $x_2 = 0$ пространственного графика функции $n_1(x_1, x_2, Z^*)$ для значения параметра $B = 0,45$

Fig. 5. Graphic representation of a section by plane $x_2 = 0$ of the spatial graph of the function $n_1(x_1, x_2, Z^*)$ for a $B = 0.45$ value

Выполненные компьютерные эксперименты позволяют сделать вывод, что оптимальное в данном случае значение B находится в окрестности $B^* = 0,65$.

Заключение

Данная томографическая постановка ОКЗ в трехмерном случае ($V(x)$, $x \in R^3$) позволяет находить скоростное распределение с точностью, определяемой условиями линейаризации и слабой степенью математической некорректности этой постановки [8]. Точность может быть повышена в результате оптимизации выбора «референтного» скоростного распределения, т.е. функции $V_0 = A + Bx_3$ (x_3 – глубина).

Томографический подход обеспечивает достоверность решения за счет:

- 1) полноты проекционных данных;
- 2) геометрии системы наблюдений, которая снимает некорректность задачи;
- 3) слабой математической некорректности томографической постановки.

Результаты данной работы допускают обобщение на веерную томографию в условиях алгебраической реконструкции, что отвечает интересам кинематической сейсмотомографии, создавая более благоприятные условия при проведении полевых геофизических измерений. Представляется перспективным рассмотренный подход, основанный на априорной информации для организации алгоритмического и программного обеспечения в условиях «эвристического поиска» в тематике искусственного интеллекта с выходом в локальную томографическую реконструкцию, т.е. в ROI-томографию, в том числе при регистрации проекционных данных в полосе и конусе лучей [9].

Следует отметить, что обратная задача скоростной диагностики скоростного распределения зондирующего сигнала в исследуемой среде может возникать как обратная задача для волнового уравнения [10, 11].

Список источников

1. Lavrent'ev M.M., Zerkal S.M., Trofimov O.E. *Computer Modeling in Tomography and Ill-Posed Problems*. Berlin : Walter de Gruyter, 2014. 126 p.
2. Лаврентьев М.М., Романов В.Г., Шишатский С.П. Некорректные задачи математической физики и анализа. М. : Наука, 1980. 286 с.
3. Папкова Ю.И. Звуковое поле в морском волноводе с неоднородной скоростью звука по глубине и трассе // Вестник Московского университета. Сер. 3. Физика. Астрономия. 2021. № 3. С. 27–35.
4. Греков А.Н., Греков Н.А., Сычев Е.Н. Среднечастотные акустические методы и средства для исследования водной среды. Севастополь : ИПТС, 2020. 126 с.
5. Микушин И.И. Метрологическое обеспечение измерений скорости звука в воде. СПб. : Изд-во СПбГЭТУ, 2023. 205 с.
6. Смирнов В.И. Курс высшей математики. М. : Наука, 1974. Т. 4, ч. 1. 336 с.
7. Лагунова К.Ф., Омельченко О.К. Об одном способе определения близких землетрясений и явные формулы для луча и времени // Неклассические проблемы математической физики : сб. науч. тр. / под ред. М. М. Лаврентьева. Новосибирск : ВЦ СО АН СССР, 1981. С. 121–129.
8. Добринский В.И., Лаврентьев М.М. Способ определения профиля скорости звука по записи акустического поля в одной точке // Тезисы докладов Второй Дальневосточной акустической конференции «Человек и океан». Владивосток, 1978. С. 12–14.
9. Peshkov A.V., Zerkal, S.M. Computational-heuristic algorithm for tomographic solution of industrial flaw detection problems // Proc. 16th Int. Scientific and Technical Conference Actual Problems of Electronic Instrument Engineering (APEIE-2023). Novosibirsk. 10–12 Nov. 2023. P. 910–915. doi: 10.1109/APEIE59731.2023.10347879
10. Романов В.Г., Бугуева Т.В. Обратная задача для волнового уравнения с полиномиальной нелинейностью // Сибирский журнал индустриальной математики. 2023. Т. 26, № 1. С. 142–149. doi: 10.33048/SIBJIM.2023.26.113
11. Романов В.Г., Бугуева Т.В. Обратная задача для нелинейного волнового уравнения // Сибирский журнал индустриальной математики. 2022. Т. 25, № 2. С. 83–100. doi: 10.33048/SIBJIM.2021.25.206

References

1. Lavrentiev, M.M., Zerkal, S.M. & Trofimov, O.E. (2014) *Computer Modeling in Tomography and Ill-Posed Problems*. Berlin: Walter de Gruyter.
2. Lavrentiev, M.M., Romanov, V.G. & Shishatsky, S.P. (1980) *Nekorrektnye zadachi matematicheskoy fiziki i analiza* [Ill-posed problems of mathematical physics and analysis]. Moscow: Nauka.
3. Papkova, Yu.I. (2021) Zvukovoe pole v morskoy volnovode s neodnorodnoy skorost'yu zvuka po glubine i trasse [Sound field in a sea waveguide with a non-uniform speed of sound along the depth and path]. *Vestnik Moskovskogo universiteta. Ser. 3. Fizika. Astronomiya*. 3. pp. 27–35.
4. Grekov, A.N., Grekov, N.A. & Sychev, E.N. (2020) *Srednechastotnye akusticheskie metody i sredstva dlya issledovaniya vodnoy sredy* [Mid-frequency acoustic methods and means for studying the aquatic environment]. Sevastopol: IPTS.
5. Mikushin, I.I. (2023) *Metrologicheskoe obespechenie izmereniy skorosti zvuka v vode* [Metrological support for measuring the speed of sound in water]. St. Petersburg: SPbGETU. pp. 193–203.
6. Smirnov, V.I. (1974) *Kurs vysshey matematiki* [Higher Mathematics]. Vol. 4. Moscow: Nauka.
7. Lagunova, K.F. & Omelchenko, O.K. (1981) Ob odnom sposobe opredeleniya blizkikh zemletryaseniy i yavnye formuly dlya luchy i vremeni [About one way to determine nearby earthquakes and explicit formulas for ray and time]. In: Lavrentyev, M. M. (ed.) *Neklassicheskie problemy matematicheskoy fiziki* [Non-classical problems of mathematical physics]. Novosibirsk: USSR AS. pp. 121–129.
8. Dobrinsky, V.I. & Lavrentiev, M.M. (1978) Sposob opredeleniya profilya skorosti zvuka po zapisi akusticheskogo polya v odnoy tochke [Method for determining the sound speed profile from recording the acoustic field at one point]. In: *Chelovek i okean* [Human and Ocean]. Theses of the Second Far Eastern Far Eastern Acoustic Conference. Vladivostok. pp. 12–14.

9. Peshkov, A.V. & Zerkal, S.M. (2023) Computational-heuristic algorithm for tomographic solution of industrial flaw detection problems. *Proceedings 16 International Scientific and Technical Conference Actual Problems of Electronic Instrument Engineering (APEIE-2023)*. Novosibirsk, November 10–12, 2023. pp. 910–915. DOI: 10.1109/APEIE59731.2023.10347879
10. Romanov, V.G. & Bugueva, T.V. (2023) Obratnaya zadacha dlya volnovogo uravneniya s polinomial'noy nelineynost'yu [Inverse problem for the wave equation with polynomial nonlinearity]. *Sibirskiy zhurnal industrial'noy matematiki – Siberian Journal of Industrial Mathematics*. 26(1). pp. 142–149. DOI: 10.33048/SIBJIM.2023.26.113.
11. Romanov, V.G. & Bugueva, T.V. (2022) Obratnaya zadacha dlya nelineynogo volnovogo uravneniya [Inverse problem for a nonlinear wave equation]. *Sibirskiy zhurnal industrial'noy matematiki – Siberian Journal of Industrial Mathematics*. 25(2). pp. 83–100. DOI: 10.33048/SIBJIM.2021.25.206.

Информация об авторах:

Гервас Николай Владимирович – аспирант кафедры вычислительной техники Новосибирского государственного технического университета (Новосибирск, Россия). E-mail: nik.gervas@mail.ru

Зеркаль Сергей Михайлович – доктор технических наук, профессор кафедры вычислительной техники Новосибирского государственного технического университета (Новосибирск, Россия). E-mail: zerkal@ngs.ru

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

Information about the authors:

Gervas Nikolai V. (Post-graduate Student, Novosibirsk State Technical University, Novosibirsk, Russian Federation). E-mail: nik.gervas@mail.ru

Zerkal Sergey M. (Doctor of Technical Sciences, Professor, Novosibirsk State Technical University, Novosibirsk, Russian Federation). E-mail: zerkal@ngs.ru

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

Поступила в редакцию 30.07.2024; принята к публикации 02.12.2024

Received 30.07.2024; accepted for publication 02.12.2024

Научная статья

УДК 004.8

doi: 10.17223/19988605/69/7

Нейросетевое моделирование и автоматическая сегментация корня аорты на снимках ангиографии

Ольга Михайловна Гергет¹, Никита Витальевич Лаптев², Михаил Александрович Чернявский³

¹ Институт проблем управления им. В.А. Трапезникова Российской академии наук, Москва, Россия

² Сибирский государственный медицинский университет, Томск, Россия

³ Национальный медицинский исследовательский центр им. В.А. Алмазова, Санкт-Петербург, Россия

¹ olgagerget@mail.ru

² laptev.nv@ssmu.ru

³ machern@mail.ru

Аннотация. Описан подход к решению задачи сегментации объектов исследования на снимках ангиографии при проведении процедуры транскатетерной имплантации аортального клапана (TAVI). В связи со сложностью анализа большого объема данных для кардиохирургов в статье особое внимание уделено автоматическому анализу медицинских данных, обучению и сравнению современных архитектур нейронных сетей. Подробно рассмотрены и протестированы для прогнозирования масок корня аорты 50 нейронных сетей (энкодеры U-net, U-net++, Linknet, FPN, DeepLabV3+ и декодеры Efficientnet-b0, Efficientnet-b1, Resnext50, Resnet34, Regnetx32). На этапах обучения и тестирования каскад моделей FPN + Efficientnet-b0 позволил получить наилучшую точность прогнозирования по метрикам IOU 0,771, Dice 0,870. Проведенное исследование показывает, что предложенный подход на основе нейросетевых каскадов, позволяющих идти не в сторону детектирования ключевых точек, а в направлении создания сегментационных масок, позволяет прогнозировать в реальном времени расположение аортального клапана и системы доставки, тем самым способствуя правильному позиционированию клапана во время TAVI.

Ключевые слова: сегментация; транскатетерная замена аортального клапана; глубокое обучение – CNN; анализ медицинских изображений; ангиография.

Благодарности: Работа выполнена при финансовой поддержке РНФ в рамках научного проекта № 24-19-00084.

Для цитирования: Гергет О.М., Лаптев Н.В., Чернявский М.А. Нейросетевое моделирование и автоматическая сегментация корня аорты на снимках ангиографии // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 69. С. 64–71. doi: 10.17223/19988605/69/7

Original article

doi: 10.17223/19988605/69/7

Segmentation of the aortic root on angiography images

Olga M. Gerget¹, Nikita V. Laptev², Michael A. Chernyavsky³

¹ V.A. Trapeznikov Institute of Control Sciences, RAS, Moscow, Russian Federation

² Siberian State Medical University, Tomsk, Russian Federation

³ National Medical Research Center named after V.A. Almazov, St. Petersburg, Russian Federation

¹ olgagerget@mail.ru

² laptev.nv@ssmu.ru

³ machern@mail.ru

Abstract. The article describes an approach based on neural network modeling, which is of great interest for segmenting different anatomical structures during TAVI. Given the complexity of analyzing large volumes of data for cardiac surgeons, the article pays special attention to the automatic analysis of medical data, training, and comparison of

modern neural networks. Fifty neural networks were thoroughly examined and tested for predicting aortic root masks (encoders: U-net, U-net++, Linknet, FPN, DeepLabV3+ and decoders: Efficientnet-b0, Efficientnet-b1, Resnext50, Resnet34, Regnetx32). During the training and testing phases, the FPN Efficientnet-b0 cascade architecture demonstrated the best prediction accuracy with metrics IOU 0.771 and Dice 0.870. The conducted study shows that the proposed approach based on neural network cascades, which focuses not on detecting key points but on creating segmentation masks, allows for real-time prediction of the aortic valve and delivery system location, thereby facilitating correct valve positioning during TAVI.

Keywords: segmentation; transcatheter aortic valve replacement; deep learning – CNN; medical image analysis; aortography.

Acknowledgments: The study was financially supported by the Russian Science Foundation under scientific project No. 24-19-00084.

For citation: Gerget, O.M., Laptev, N.V., Chernyavsky, M. A. (2024) Segmentation of the aortic root on angiography images. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 69. pp. 64–71. doi: 10.17223/19988605/69/7

Введение

В связи с развитием визуального сопровождения медицинских процедур интерес к решению задачи сегментации небольших по размеру объектов на зашумленных двумерных и трехмерных данных неуклонно растет. Известные на сегодняшний день классические алгоритмы сегментации медицинских изображений [1] обладают множеством ограничений, в частности алгоритмы, построенные по принципу каскадности, обладают низкой обобщающей способностью и накапливают ошибку вычислений. Трехмерные классические алгоритмы сегментации [1] представлены достаточно малым набором исследований, что обусловлено сложностью и невозможностью сегментации трехмерных медицинских объектов в реальном времени. Большое количество классических методов автоматической сегментации сталкивается со сложностью выделения регионов на изображении в связи с неравномерностью освещения, схожими значениями интенсивности пикселей для разных объектов, малыми объемами выборки, наличием артефактов и т.д.

В настоящее время наибольшее распространение получили нейросетевые подходы при решении задачи сегментации области интереса, что в большей степени обусловлено устойчивостью к шумам, способностью выбирать (активировать) наиболее релевантные характеристические признаки без предварительной обработки, отказоустойчивостью (при неблагоприятных условиях выходная точность нейросетевой модели практически не изменяется) [2–4].

Несмотря на большое разнообразие нейросетевых архитектур (U-net [5], U-net++ [6], Linknet [7], FPN [8], DeepLabV3+ [9] и др.), позволяющих провести сегментации изображения, выделение контура корня аорты на снимках ангиографии при проведении транскатетерной имплантации аортального клапана (TAVI) [10] является нерешенной задачей. Рентгеновская коронарография включает введение контрастных веществ в коронарные сосуды, однако, несмотря на подачу контрастирующих веществ, визуализация сосудов, в частности корня аорты, сильно затруднена наличием различных шумов на изображении, низкой контрастностью, наличием неоднородностей, вызванных воздухом в легких, ограниченностью области видимости, а также быстрым вымыванием контраста из зоны интереса.

В данной работе предложено применение нейросетевого подхода при автоматической сегментации корня аорты для одиночных флюороскопических изображений (снимки ангиографии) с контрастным усилением, что позволяет использовать его как в качестве вспомогательного инструмента для визуализации с целью оптимизации позиционирования клапана, так и в качестве элемента системы для синхронизации данных КТ и снимков ангиографии.

1. Источник данных

В качестве исходных данных для обучения и валидации нейронных сетей использовались оригинальные серии изображений аортографии, собранные во время имплантации 14 саморасширяющихся биопротезов аортального клапана CoreValve пациентам со стенозом аортального клапана

в период с 2015 по 2018 г. [11] и оригинальные серии изображений аортографии, собранные во время имплантации 31 баллонорасширяемого биопротеза аортального клапана Visi-Pro пациентам со стенозом аортального клапана в период с 2020 по 2024 г., выполненные в отделении сердечно-сосудистой хирургии НИИ ФГБУ «Национальный медицинский исследовательский центр имени В.А. Алмазова» Министерства здравоохранения Российской Федерации.

Набор данных представлял собой 50 серий видеопоследовательностей размером $1\,000 \times 1\,000$ пикселей с глубиной 8 бит (шкала от 0 до 255). С целью сегментации корня аорты выбрана серия изображений с разным количеством контрастирующего вещества и в разные временные участки проведения процедуры ТАВИ. Разметка данных осуществлялась с использованием платформы Supervisely AI. Таким образом, итоговая выборка состояла из 2 250 изображений в оттенках серого, из которых 1 800 (80%) изображений использовалось в качестве обучающего набора, а 450 изображений (20%) – в качестве проверочного набора.

Разработка системы сегментации корня аорты состояла из двух основных этапов:

Этап 1. Подготовка данных: маркировка данных для разработки обучающих и проверочных наборов; аннотация изображения интервенционного кардиолога.

Этап 2. Обучение и тестирование нейронных моделей: выбор доступных архитектур нейронных сетей, функции потерь и описательных метрик, оценка качественных и количественных параметров по данным обучения и тестирования.

2. Обучение

В ходе проведения исследования использованы ключевые библиотеки, пакеты и фреймворки, в частности:

– Python версии 3.10.0 (RRID:SCR_008394) и PyCharm версии 2020.1 (RRID:SCR_018221) – в качестве основного языка программирования и интегрированной среды разработки для обработки / обработки данных и разработки нейронных сетей;

– Библиотеки Pytorch версии 2.3.1 (RRID:SCR_018536); Scikit-learn версии 0.20.3 (RRID:SCR_002577); SciPy версии 1.4.1 (RRID:SCR_008058); NumPy версии 1.18.2 (RRID:SCR_008633); Pandas версии 0.24.2 (RRID:SCR_018214); OpenCV версии 4.0.1.23 (RRID:SCR_018214); Seaborn версии 0.10.0 (RRID:SCR_018132) и Matplotlib 3.0.3 (RRID:SCR_008624).

С целью улучшения качества входной информации было предложено и применено несколько последовательных преобразований. Во-первых, реализован алгоритм вычитания из оригинального изображения результата выделения белых объектов на более темном фоне (white top-hat transformation), что позволило повысить контрастность изображения, снизить шум фона и улучшить видимость корня аорты. Необходимым для корректного отображения изображений также стал шаг, связанный с ограничением значений яркости пикселей в диапазоне [0, 255].

Во-вторых, реализован метод адаптивного выравнивания гистограммы (CLAHE) [12], что позволило регулировать контрастность локально в небольших областях изображения. Целесообразность предварительной обработки медицинских данных на основе CLAHE обусловлена многообразием структур, для которых необходимо подбирать разные коэффициенты улучшения качества отображения.

В качестве нейросетевых моделей, позволяющих извлекать признаки, в исследовании использовано несколько разновидностей энкодеров (U-net, U-net++, Linknet, FPN, DeepLabV3+) и декодеров моделей семантической сегментации (Efficientnet-b0, Efficientnet-b1, Resnext50, Resnet34, Regnetx32).

Сравнение моделей осуществлялось на базе основных характеристик эффективности, в частности сглаженного коэффициента Соренсена–Дайса (англ. DiceLoss), IoU, Precision, Recall.

Все нейронные сети обучались на компьютере с характеристиками: процессор Intel(R) Xeon(R) Gold 5218R 2,10 ГГц, 512 ГБ оперативной памяти, NVIDIA GeForce A6000 48 ГБ, Debian 11.

В качестве оптимизатора использовался RMSprop [13] с начальным шагом 0,0001 (англ. learning rate) и множителем угасания 0,9 (англ. learning rate annealing), применяющимся каждые 5 шагов. Batch size для обучения моделей был одинаковым и составлял 16 обучающих примеров.

Для оценки компонентов классификации и регрессии нейронных сетей были выбраны следующие метрики:

$$IOU = \frac{TP}{TP + FP + FN}, \quad (1)$$

$$Dice = \frac{2 \cdot TP}{2 \cdot TP + FP + FN}, \quad (2)$$

$$Presion = \frac{TP}{TP + FP}, \quad (3)$$

$$Recall = \frac{TP}{TP + FN}, \quad (4)$$

где TP – количество истинно положительных результатов, TN – количество истинно отрицательных результатов, FP – количество ложноположительных результатов, FN – количество ложноотрицательных результатов.

3. Полученные результаты

На рис. 1 показан пример обучения исследуемых нейронных сетей на примере энкодера FPN. Графики представляют изменения значений функции потерь, пунктирными линиями выделены изменения функции потерь на тренировочном наборе данных, сплошными – на валидационном.

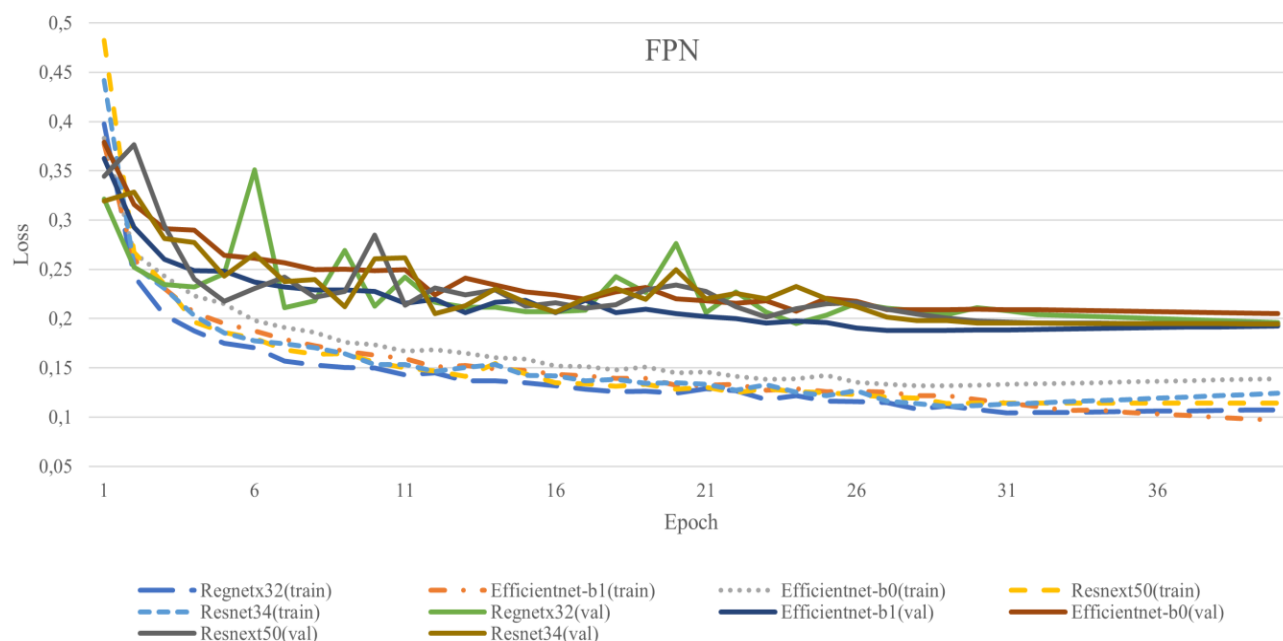


Рис. 1. Функция потерь для исследуемых нейронных сетей
Fig 1. Loss of function for saving neural networks

Результаты обучения моделей нейронных сетей представлены в таблице.

Показатели результативности нейросетевых архитектур семантической сегментации

Энкодер	Декодер	IoU	Dice	Precision	Recall	Скорость
FPN	Resnet34	0,756	0,861	0,862	0,859	0,096
	Resnext50	0,740	0,851	0,857	0,845	0,107
	Efficientnet-b0	0,771	0,870	0,833	0,911	0,101
	Efficientnet-b1	0,743	0,853	0,806	0,905	0,107
	Regnetx32	0,744	0,853	0,866	0,841	0,105

Окончание таблицы

Энкодер	Декодер	IoU	Dice	Precision	Recall	Скорость
Unet	Resnet34	0,730	0,844	0,868	0,822	0,099
	Resnext50	0,766	0,867	0,856	0,879	0,108
	Efficientnet-b0	0,755	0,861	0,848	0,873	0,105
	Efficientnet-b1	0,762	0,865	0,817	0,920	0,108
	Regnetx32	0,763	0,865	0,864	0,867	0,113
Unet++	Resnet34	0,756	0,861	0,862	0,859	0,120
	Resnext50	0,740	0,851	0,857	0,845	0,157
	Efficientnet-b0	0,738	0,849	0,867	0,832	0,117
	Efficientnet-b1	0,743	0,853	0,806	0,905	0,115
	Regnetx32	0,755	0,860	0,867	0,854	0,132
Linknet	Resnet34	0,726	0,841	0,868	0,816	0,097
	Resnext50	0,735	0,847	0,865	0,830	0,111
	Efficientnet-b0	0,753	0,859	0,826	0,894	0,107
	Efficientnet-b1	0,743	0,853	0,803	0,910	0,111
	Regnetx32	0,737	0,849	0,874	0,825	0,111
DeepLabV3+	Resnet34	0,758	0,862	0,867	0,858	0,096
	Resnext50	0,757	0,862	0,866	0,858	0,110
	Efficientnet-b0	0,761	0,864	0,827	0,905	0,106
	Efficientnet-b1	0,751	0,858	0,820	0,899	0,107
	Regnetx32	0,740	0,851	0,862	0,840	0,105
С цветовой предобработкой данных (CLAHE)						
FPN	Resnet34	0,733	0,846	0,876	0,818	0,097
	Resnext50	0,729	0,843	0,858	0,829	0,087
	Efficientnet-b0	0,745	0,854	0,825	0,885	0,096
	Efficientnet-b1	0,760	0,863	0,849	0,879	0,099
	Regnetx32	0,754	0,859	0,870	0,849	0,095
Unet	Resnet34	0,692	0,818	0,879	0,765	0,101
	Resnext50	0,735	0,847	0,885	0,812	0,102
	Efficientnet-b0	0,753	0,859	0,842	0,877	0,095
	Efficientnet-b1	0,734	0,846	0,869	0,825	0,101
	Regnetx32	0,734	0,847	0,875	0,820	0,1
Unet++	Resnet34	0,706	0,827	0,867	0,791	0,145
	Resnext50	0,705	0,827	0,878	0,782	0,148
	Efficientnet-b0	0,758	0,862	0,825	0,903	0,104
	Efficientnet-b1	0,758	0,862	0,881	0,845	0,106
	Regnetx32	0,768	0,869	0,873	0,865	0,117
Linknet	Resnet34	0,729	0,843	0,848	0,838	0,099
	Resnext50	0,703	0,826	0,774	0,886	0,101
	Efficientnet-b0	0,747	0,855	0,834	0,877	0,095
	Efficientnet-b1	0,768	0,869	0,851	0,887	0,101
	Regnetx32	0,732	0,845	0,871	0,821	0,11
DeepLabV3+	Resnet34	0,759	0,863	0,858	0,867	0,11
	Resnext50	0,737	0,848	0,845	0,852	0,124
	Efficientnet-b0	0,753	0,859	0,808	0,917	0,091
	Efficientnet-b1	0,758	0,862	0,844	0,881	0,103
	Regnetx32	0,754	0,860	0,857	0,863	0,099

Согласно результатам, представленным в таблице, наилучшие значения показателей оценки семантической сегментации корня аорты показала модель FPN Efficientnet-b0 с результатами IOU 0,771, Dice 0,870, Recall 0,833, Precision 0,911. Скорость обработки одного изображения 0,101 с. При этом необходимо отметить, что применение предобработки данных не дает значительного увеличения значений метрик качества. Однако, несмотря на отсутствие значимых различий в точности сегментации структур корня аорты, предобработанные изображения значительно улучшают визуальное восприятие изображения и облегчают разметку данных. Примеры изображений с цветовой обработкой и без обработки представлены на рис. 2.

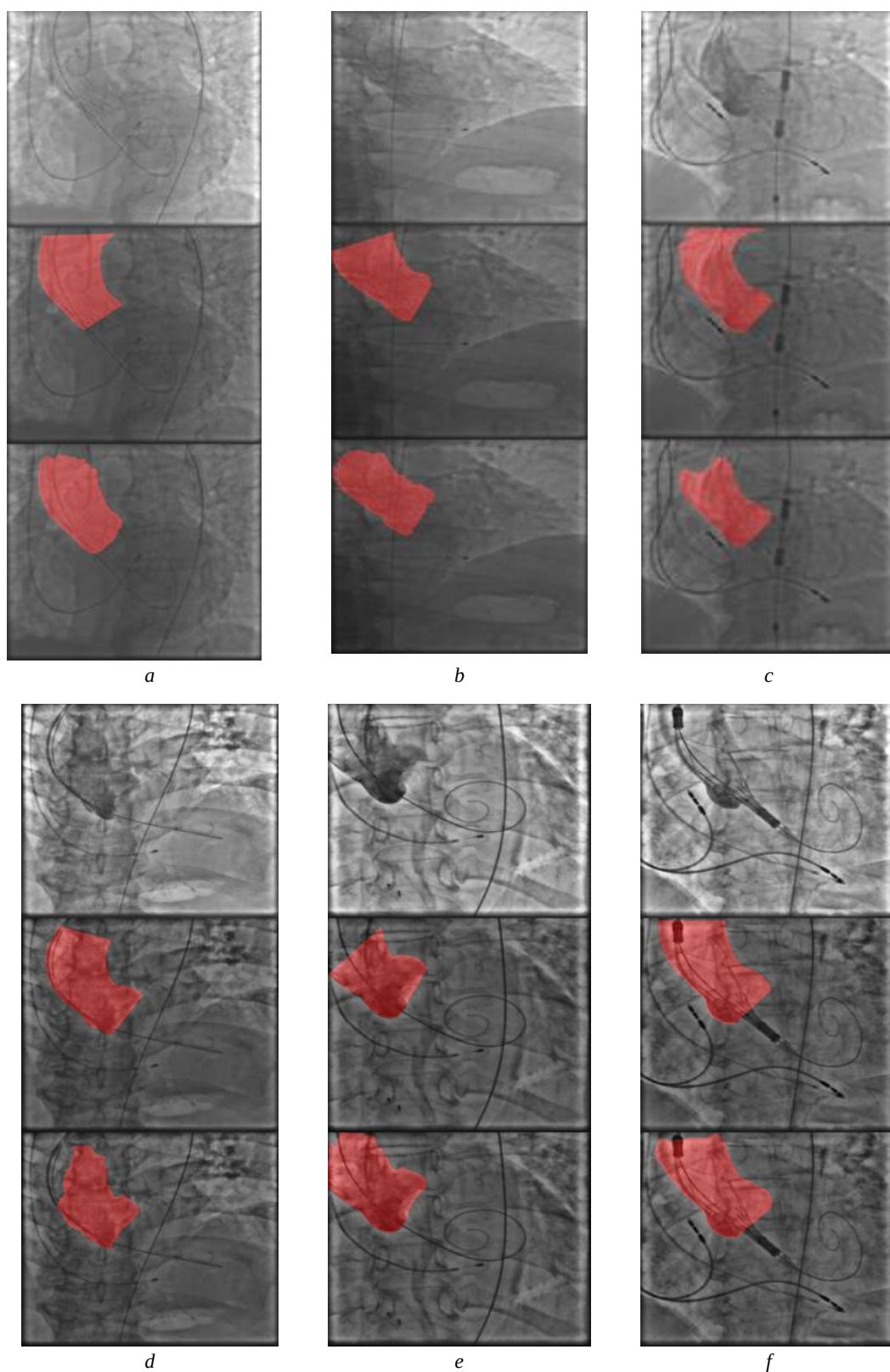


Рис. 2. Примеры сегментации корня аорты: *a–c* – сегментация изображений без цветовой обработки, изображение сверху исходное, изображение в центре с размеченной маской, изображение снизу с предсказанной маской, *d–f* – сегментация изображений с цветовой обработкой (CLANE), изображение сверху исходное, изображение в центре с размеченной маской, изображение снизу с предсказанной маской

Fig. 2. Examples of segmentation of the lower aorta: *a–c* image segmentation without color processing, top image is original, center image with labeled mask, image usually with predicted mask, *d–f* image segmentation with color processing (CLANE), top image is original, center image with labeled mask, the image matches the predicted mask

Заключение

В статье приведен анализ 50 серий видеопоследовательностей, полученных при проведении транскатетерной имплантации аортального клапана, с целью создания системы визуальной помощи для кардиохирурга.

Поскольку реальные данные имеют низкую разрешающую способность, применение известных методов детектирования как с предобработкой входных данных, так и при ее отсутствии не позволяет получить приемлемое качество визуального представления корня аорты на данных с малым количеством контрастирующего вещества. Однако применение каскада нейросетевых архитектур дало возможность получать сегментационные маски и использовать их в качестве ориентира, что позволило визуализировать корень аорты при снижении количества введений контрастирующего вещества во время проведения операции.

Реализация и сравнение различных каскадов архитектур энкодер–декодер, где в качестве энкодера выбраны Unet, Unet++, Linknet, FPN, DeepLabV3+, а в качестве декодера выступают модели EfficientNet-b0, EfficientNet-b1, Resnext50, Resnet34, Regnetx32, позволили выделить с точки зрения точности оптимальную нейросетевую модель для решения задачи распознавания и сегментации контуров корня аорты во время проведения процедуры TAVI. Наилучшим каскадом является FPN Efficientnet-b0. Показатели результативности проведения семантической сегментации в данном случае имеют значения: IOU 0,771, Dice 0,870, Recall 0,833, Precision 0,911, Скорость обработки одного изображения 0,101 с. Модель на основе данной архитектуры (FPN + декодер Efficientnet-b0) может использоваться как в качестве вспомогательного инструмента для оптимизации позиционирования клапана, так и в качестве элемента системы для синхронизации данных КТ и снимков ангиографии.

Список источников

1. Ramesh K.K.D., Kiran Kumar G., Swapna K., Datta D., Suman Rajest S. A review of medical image segmentation algorithms // EAI Endorsed Transactions on Pervasive Health and Technology. 2021. V. 7 (27). Art. e6. doi: 10.4108/eai.12-4-2021.169184
2. Li Y., Wu Y., He J., Jiang W., Wang J., Peng Y., Chen M. Automatic coronary artery segmentation and diagnosis of stenosis by deep learning based on computed tomographic coronary angiography // European Radiology. 2022. V. 32 (9). P. 6037–6045. doi: 10.1007/s00330-022-08761-z
3. Popov M., Amanturdieva A., Zhaksylyk N., Alkanov A., Saniyazbekov A., Aimyshev T. et al. Dataset for Automatic Region-based Coronary Artery Disease Diagnostics Using X-Ray Angiography Images // Scientific Data. 2024. V. 11 (1). Art. 20. doi: 10.1038/s41597-023-02871-z
4. Zhu X., Cheng Z., Wang S., Chen X., Lu G. Coronary angiography image segmentation based on PSPNet // Computer Methods and Programs in Biomedicine. 2021. V. 200. Art. 105897. doi: 10.1016/j.cmpb.2020.105897
5. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation // International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015). 5–9 October, Munich, Germany. 2015. P. 234–241.
6. Zhou Z., Rahman Siddiquee M.M., Tajbakhsh N., Liang J. Unet++: A nested u-net architecture for medical image segmentation // Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop (DLMIA 2018). 20 September, Granada, Spain. 2018. P. 3–11. doi: 10.1007/978-3-030-00889-5_1
7. Chaurasia A., Culurciello E. Linknet: Exploiting encoder representations for efficient semantic segmentation // 2017 IEEE Visual Communications and Image Processing (VCIP). 2017. P. 1–4.
8. Martinsson J., Mogren O. Semantic segmentation of fashion images using feature pyramid networks // Proc. 2019 International Conference on Computer Vision Workshop (ICCVW 2019). 2019. doi: 10.1109/ICCVW.2019.00382
9. Baheti B., Innani S., Gajre S., Talbar S. Semantic scene segmentation in unstructured environment with modified DeepLabV3+ // Pattern Recognition Letters. 2029. V. 138. P. 223–229.
10. Abdelgawad A., Hussein M.A., Naeem H., Abuellata R., Alghamdy S. A comparative study of TAVR versus SAVR in moderate and high-risk surgical patients: Hospital outcome and midterm results // Heart Surgery Forum. 2019. V. 22 (5). P. E331–E339. doi: 10.1532/hfsf.2243.
11. The repository of the Research Laboratory for Processing and Analysis of Big Data (Tomsk Polytechnic University). URL: https://www.dropbox.com/sh/80wpfkdbahuo0l9/AADuysNg3sO00_vjhW8MgZ6Ba?dl=0 (accessed: 10.05.2024).
12. Pisano E.D., Zong S., Hemminger B.M., DeLuca M., Johnston R.E., Muller K. et al. Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms // Journal of Digital Imaging. 1998. V. 11. P. 193–200. doi: 10.1007/BF03178082
13. RMSprop // Keras. URL: <https://keras.io/api/optimizers/rmsprop/> (accessed: 10.05.2024).

References

1. Ramesh, K.K.D., Kiran Kumar, G., Swapna, K., Datta, D. & Suman Rajest, S. (2021) A review of medical image segmentation algorithms. *EAI Endorsed Transactions on Pervasive Health and Technology*. 7(27). DOI: 10.4108/eai.12-4-2021.169184
2. Li, Y., Wu, Y., He, J., Jiang, W., Wang, J., Peng, Y. & Chen, M. (2022). Automatic coronary artery segmentation and diagnosis of stenosis by deep learning based on computed tomographic coronary angiography. *European Radiology*. 32(9). pp. 6037–6045. DOI: 10.1007/s00330-022-08761-z
3. Popov, M., Amanturdieva, A., Zhaksylyk, N., Alkanov, A., Sanlyazbekov, A., Aimyshev, T. et al. (2024) Dataset for Automatic Region-based Coronary Artery Disease Diagnostics Using X-Ray Angiography Images. *Scientific Data*. 11(1). pp. 20. DOI: 10.1038/s41597-023-02871-z
4. Zhu, X., Cheng, Z., Wang, S., Chen, X. & Lu, G. (2021). Coronary angiography image segmentation based on PSPNet. *Computer Methods and Programs in Biomedicine*. 200. DOI: 10.1016/j.cmpb.2020.105897
5. Ronneberger, O., Fischer, P. & Brox, T. (2015) U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015)*. October 5–9, 2015. Munich. Germany. pp 234–241.
6. Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N. & Liang, J. (2018) Unet++: A nested u-net architecture for medical image segmentation. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. 4th International Workshop (DLMIA 2018). September 20. Granada. Spain. pp. 3–11. DOI: 10.1007/978-3-030-00889-5_1
7. Chaurasia, A. & Culurciello, E. (2017) Linknet: Exploiting encoder representations for efficient semantic segmentation. *2017 IEEE Visual Communications and Image Processing (VCIP)*. pp. 1–4.
8. Martinsson, J. & Mogren, O. (2019) Semantic segmentation of fashion images using feature pyramid networks. *Proceedings – 2019 International Conference on Computer Vision Workshop (ICCVW 2019)*. DOI: 10.1109/ICCVW.2019.00382
9. Baheti, B., Innani, S., Gajre, S. & Talbar, S. (2020) Semantic scene segmentation in unstructured environment with modified DeepLabV3+. *Pattern Recognition Letters*. 138. pp. 223–229.
10. Abdelgawad, A., Hussein, M.A., Naeem, H., Abuelata, R. & Alghamdy, S. (2019) A comparative study of TAVR versus SAVR in moderate and high-risk surgical patients: Hospital outcome and midterm results. *Heart Surgery Forum*. 22(5). pp. E331–E339. DOI: 10.1532/hcf.2243
11. Tomsk Polytechnic University. (n.d.) *The dataset presented in this study can be found in the repository of the Research Laboratory for Processing and Analysis of Big Data (Tomsk Polytechnic University)*. [Online] Available from: https://www.dropbox.com/sh/80wpfkabhuo0l9/AADuysNg3sO00_vjhW8MgZ6Ba?dl=0 (Accessed: 10th May 2024).
12. Pisano, E.D., Zong, S., Hemminger, B.M., DeLuca, M., Johnston, R.E., Muller, K., et al. (1998). Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms. *Journal of Digital Imaging*. 11. pp. 193–200. DOI: 10.1007/BF03178082
13. Keras. [Online] Available from: <https://keras.io/api/optimizers/rmsprop/> (Accessed: 10th May 2024).

Информация об авторах:

Гергет Ольга Михайловна – доктор технических наук, ведущий научный сотрудник Института проблем управления им. В.А. Трапезникова Российской академии наук (Москва, Россия). E-mail: olgagerget@mail.ru

Лаптев Никита Витальевич – научный сотрудник отдела медицинских данных и биоинформатики Сибирского государственного медицинского университета (Томск, Россия). E-mail: raa@mail.tsu.ru

Чернявский Михаил Александрович – доктор медицинских наук, доцент кафедры факультетской хирургии с клиникой, заведующий НИО сосудистой и интервенционной хирургии Национального медицинского исследовательского центра им. В.А. Алмазова» Минздрава России (Санкт-Петербург, Россия). E-mail: machern@mail.ru

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

Information about the authors:

Gerget Olga M. (Doctor of Technical Sciences, Leading Researcher Laboratory №80, V.A. Trapeznikov Institute of Control Sciences, RAS, Moscow, Russia). E-mail: olgagerget@mail.ru

Laptev Nikita V. (Researcher at the Department of Medical Data and Bioinformatics of Siberian State Medical University, Tomsk, Russia). E-mail: laptev.nv@ssmu.ru

Chernyavsky Michael A. (Doctor of Medical Sciences, Associate Professor of Department of Faculty Surgery with Clinic, Head of Vascular and Endovascular surgery Department, Almazov National Medical Research Center, Saint-Petersburg, Russia). E-mail: machern@mail.ru

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

Поступила в редакцию 08.07.2024; принята к публикации 02.12.2024

Received 08.07.2024; accepted for publication 02.12.2024

Original article

UDC 004.93'12

doi: 10.17223/19988605/69/8

**Detection of flying objects in images using the YOLOv4-CSP
convolutional neural network model****Stepan G. Nebaba¹, Nikolay G. Markov²**^{1,2} National Research Tomsk Polytechnic University, Tomsk, Russian Federation¹ stepanlfx@tpu.ru² markovng@tpu.ru

Abstract. The effectiveness of the YOLOv4-CSP convolutional neural network model in solving the problem of detecting objects moving in airspace is investigated. Images of flying objects of two classes were used as initial data for training and researching the convolutional neural network model: helicopter-type and aircraft-type unmanned aerial vehicles. Images of such objects were obtained in the optical and infrared wavelength ranges. Two datasets were formed from appropriately labeled source images with objects of these two classes. The first dataset was created from optical images, and the second from images obtained in the infrared wavelength range. The YOLOv4-CSP model was trained using training and validation samples from each dataset. Comprehensive studies of the effectiveness of the trained model were carried out using test samples from datasets. It is shown that the accuracy of detecting flying objects in optical images is higher than in images obtained in the infrared range, and the results for the speed of model calculation when analyzing optical and infrared images are close. Recommendations are given for the use of the YOLOv4-CSP model in computer vision systems for airspace monitoring.

Keywords: computer vision system; convolutional neural network YOLOv4-CSP; detection of flying objects; helicopter-type unmanned aerial vehicle; aircraft-type unmanned aerial vehicle.

For citation: Nebaba, S.G., Markov, N.G. (2024) Detection of flying objects in images using the YOLOv4-CSP convolutional neural network model. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 69. pp. 72–81. doi: 10.17223/19988605/69/8

Научная статья

doi: 10.17223/19988605/69/8

**Детектирование летающих объектов на изображениях
с помощью модели сверточной нейронной сети YOLOv4-CSP****Степан Геннадьевич Небаба¹, Николай Григорьевич Марков²**^{1,2} Национальный исследовательский Томский политехнический университет, Томск, Россия¹ stepanlfx@tpu.ru² markovng@tpu.ru

Аннотация. Исследуется эффективность модели сверточной нейронной сети YOLOv4-CSP при решении задачи детектирования в воздушном пространстве объектов двух классов: беспилотных летательных аппаратов вертолетного типа и самолетного типа. Изображения объектов получены в оптическом и инфракрасном диапазонах длин волн, из них сформировано два соответствующих датасета. Модель YOLOv4-CSP обучена с использованием обучающей и валидационной выборок из каждого датасета. Проведены комплексные исследования эффективности обученной модели с использованием тестовых выборок из датасетов. Показано, что точность детектирования летающих объектов на оптических изображениях выше, чем на изображениях, полученных в инфракрасном диапазоне, а результаты по скорости вычисления модели при анализе оптических и инфракрасных изображений близки. Даны рекомендации по использованию модели YOLOv4-CSP в системах компьютерного зрения для мониторинга воздушного пространства.

Ключевые слова: система компьютерного зрения; сверточная нейронная сеть YOLOv4-CSP; детектирование летающих объектов; беспилотный летательный аппарат вертолетного типа; беспилотный летательный аппарат самолетного типа.

Для цитирования: Небаба С.Г., Марков Н.Г. Детектирование летающих объектов на изображениях с помощью модели сверточной нейронной сети YOLOv4-CSP // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 69. С. 72–81. doi: 10.17223/19988605/69/8

Introduction

Today, research and development of modern computer vision systems (CVS) are intensively conducted all over the world. Such systems are in demand for solving many applied problems in a variety of areas of human activity, such as controlling autonomous vehicles, monitoring the safety of enterprises and hazardous production facilities [1], processing and analyzing medical images, and in a number of other areas of knowledge [2]. The problem of airspace safety above enterprises with hazardous technological objects and processes deserves special attention. To solve this problem, radar systems for detecting flying objects (FO) are widely used. FO usually refers to various manned and unmanned aerial vehicles, birds, etc. In recent years, monitoring of airspace for the purpose of detecting FO has begun to be carried out using various CVS based on deep learning methods and models. Such CVS can analyze images of FO in airspace obtained in optical, infrared (IR) and other wavelength ranges [3]. Moreover, the analysis of images entering the input of such CVS should most often be carried out in real time [4].

The authors of [1–3] believe that such CVSs should be created based on modern deep learning models, primarily based on convolutional neural network (CNN) models [2]. However, on this path it is necessary to solve a number of very complex problems: to select from among the known CNN models the most promising model (models) for implementation in real-time CVS, to conduct a comprehensive analysis of its (their) effectiveness, taking into account the requirements for the accuracy of LO recognition and the speed of calculation this model and, if it meets the requirements of the CVS, solve the problem of software or hardware implementation of such a model as part of the CVS. Moreover, even the preparation of a representative set of images for training and research of a CNN model can be far from a trivial task [5].

The article presents the results of comprehensive studies of the effectiveness of one of the most promising CNN models for creating CVS, YOLOv4-CSP, designed for detection and classification of two classes of aircraft in images: aircraft-type unmanned aerial vehicles (UAVs) and helicopter-type UAVs. At the same time, the problem of recognizing such objects is solved both in optical images (RGB images) and in images obtained in the IR range.

1. The task of detecting flying objects using a CNN model

The task of FO detection in images is the main task solved with the help of CVS when monitoring airspace. It comes down to three subtasks: detecting objects in an image, localizing them, and determining the class of each object. Further, the term «detection» will describe the process of solving all three subtasks. Today, a promising direction in solving such a three-stage detection problem is the use of modern CNN models in CVS [2–4]. When creating a real-time CVS for airspace monitoring, it is necessary to select from among the known CNN models the most promising model(s) for implementation in such a CVS and conduct a comprehensive analysis of its (their) effectiveness, taking into account the requirements for the CVS and, accordingly, to the implemented model CNN.

Let's consider these requirements in more detail. In the images obtained during airspace monitoring using appropriate equipment as part of the CVS, one or more FOs may appear, and in the case of several objects they may belong to different classes. This must be taken into account when initially selecting a CNN model from a set of potentially promising models for solving the problem of FO detection. The main requirements for the CNN model are the contradictory requirements of high speed of model calculations and high accuracy of detection (classification) of FOs in images. It is especially difficult to satisfy the requirement of high computation speed in the case of creating CVS operating in real time [1]. Such CVS must not only detect and classify FOs, but also track their movement in space. The required real-time scale depends on the speed of movement of FO in the airspace. At the same time, as shown in [1, 4], the main computing

resources of the real-time CVS are used to calculate the CNN model, which emphasizes the relevance of creating and using the most efficient CNN models in terms of computation speed in such CVSs.

The speed of detecting FOs in an image (the speed of calculating the CNN model), measured as the number of analyzed images per second in FPS (frames per second), should be at least 25 for most real-time CVS [4].

The accuracy of detecting (classifying) objects in an image is usually assessed using the generally accepted metrics $AP_{0.5}$ (Average Precision) and $mAP_{0.5}$ (mean Average Precision) - the average value of $AP_{0.5}$ for all classes of FOs. In accordance with works [2,3], we will assume that the accuracy of detection (classification) of FOs in images using the CNN model is high (the requirement for detection accuracy is met) if the values of the $AP_{0.5}$ metric for each class of FOs and the $mAP_{0.5}$ metric for all classes exceed the threshold value of 60%.

High algorithmic efficiency of the CNN model can be also highlighted as the third requirement. The algorithmic efficiency of the model is assessed based on the calculation of two indicators: the size of the CNN model and its computational complexity. The CNN model size (or the measure of compactness of the architecture), most often measured in MB (megabytes), is the amount of memory of the CVS computing device required to store the weighting coefficients of convolutional layers and intermediate buffers when calculating the CNN model. In turn, the computational complexity of the CNN model is measured in GFLOPs (Giga-Floating-point Operations per Second) and is determined by the number of multiplication, addition and comparison operations on floating-point numbers when performing convolution and subsampling procedures in all convolutional and subsampling layers of the model. Note that the threshold values of the model size and computational complexity of the CNN are set by the developers of the real-time CVS. They generally consider that a candidate CNN model satisfies their requirement for algorithmic efficiency if it can be implemented on a mid-range GPU or a modern system-on-chip with a programmable integrated circuit to speed up the computation.

The problem solved in this study can be formulated as follows. It is necessary to create real-time CVS based on the CNN model, which should allow detection of two classes of aircraft in images: aircraft-type UAVs and helicopter-type (multi-rotor) UAVs. The task is set to select from the known CNN models the most promising model for detecting FOs in images and the task of its comprehensive study in order to identify the parameters of the model at which it satisfies all three requirements formulated above, and makes it possible to detect not only single FO, but also several objects in the image, including those belonging to various classes.

2. Choosing a CNN model for research

There are two types of CNN models for detecting objects in images: one-stage and two-stage detectors (Figure 1). First, two-stage detectors appeared, the most famous of them is based on the R-CNN model [6]. It provides high detection accuracy, but has a low image analysis speed. The Faster R-CNN model [7] and its extension Mask R-CNN were developed to solve the problem of the low speed of the R-CNN model. However, according to the evaluation of these CNN models from [2], their computational speed is significantly lower than the selected FPS threshold of 25. Moreover, in [2] it is shown that their algorithmic efficiency is quite low. All this does not allow us to consider such two-stage CNN models as a basis for creating real-time CVS.

The analysis of various classes of CNN models carried out in [1, 2] and the above results of our discussion of a number of CNN models suggest that the most suitable for detecting FOs in images, taking into account the specified requirements, are the YOLO (You Only Look Once) class models [8, 9]. CNN models of this class belong to CNN models with a single-stage architecture; they analyze the image in one pass (the stages of detection, localization and classification of objects are performed in parallel, at the same time), which significantly increases the speed of image analysis. Moreover, such CNN models have very high accuracy in object detection [2]. Some YOLO class models have a complex architecture and therefore can exceed threshold values in terms of computational complexity and the size of the memory occupied by the CNN model. Thus, the first models of this family, such as YOLO, YOLOv2 and YOLOv3, are not suitable for CVS with low computing resources [9]. For such CVSs, CNN models with more compact architectures are

needed, including, first of all, a small number of convolutional layers. In this regard, for real-time CVS it is proposed to choose and study the effectiveness of YOLO class CNN models with more modern architectures, starting with the YOLOv4 model [8].

The YOLOv4 CNN model consists of 4 blocks, in contrast to the models of two-stage detectors, consisting of 5 blocks (Figure 1) [8]: Input-input image; Backbone (used to build a deeper CNN in order to increase its accuracy); Neck (used to obtain more detailed spatial and semantic information about an object); Dense Prediction (used to determine bounding box coordinates along with a confidence estimate for a feature class). The last block uses the same principle as the earlier CNN model YOLOv3 [9].

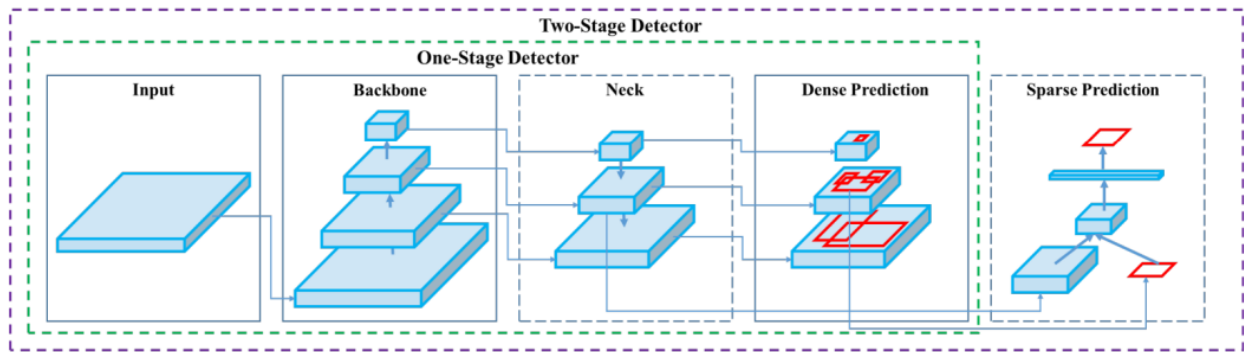


Fig. 1. Single-stage and two-stage detectors

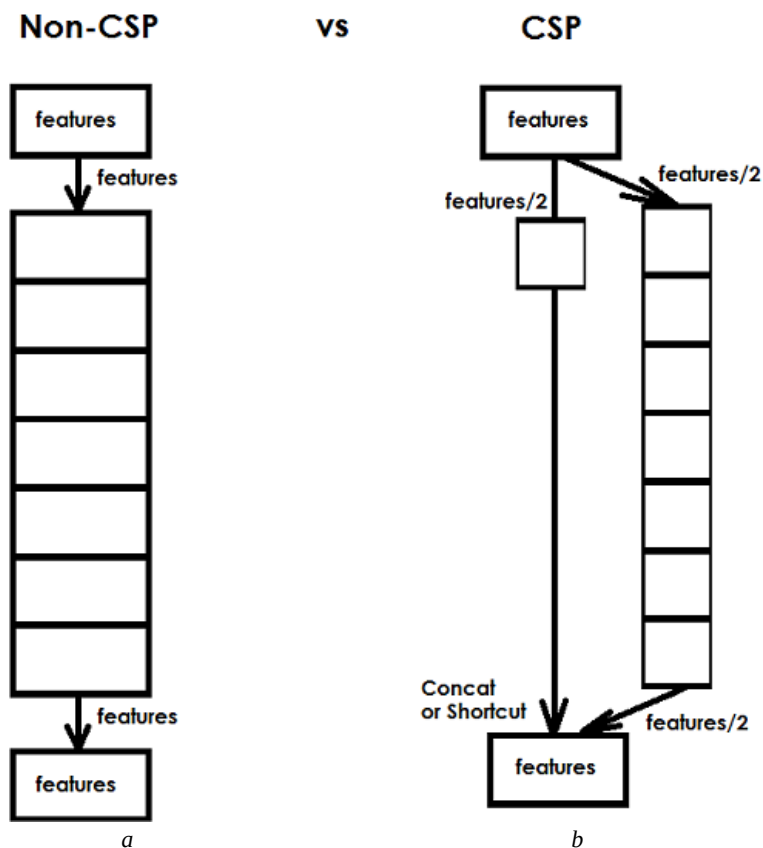


Fig. 2. Backbone block structure: normal connection of layers (a); CSP layer connection (b)

The YOLOv4-CSP CNN model is a modification of the YOLOv4 model and contains a number of important improvements of intermediate detector blocks. The main ones are the following.

1. Backbone. This block uses a regular CSPDarknet53 connection [10] for the YOLOv4 model (Figure 2a) or a CSP connection for the YOLOv4-CSP CNN model (Figure 2b). A CSP connection is more efficient because its basic idea is:

- a half of the output signal goes along the main path (generating more semantic information with a larger receptive field);
- the second half of the signal takes a detour (preserving more spatial information with a small receptive field).

2. Neck. This block uses additional SPP layers [11], as well as the CNN PAN model for path aggregation (Figure 3).

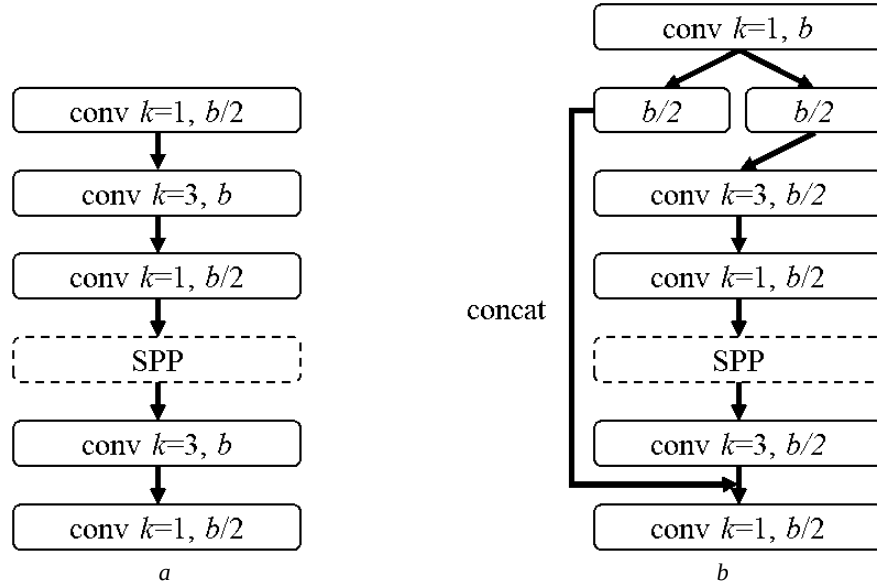


Fig. 3. Neck block structure: YOLOv4 CNN model (a); YOLOv4-CSP CNN model (b)

Moreover, by applying optimal scaling, efficient CNN models can be obtained from the YOLOv4-CSP model for different input image sizes (Fig. 4).

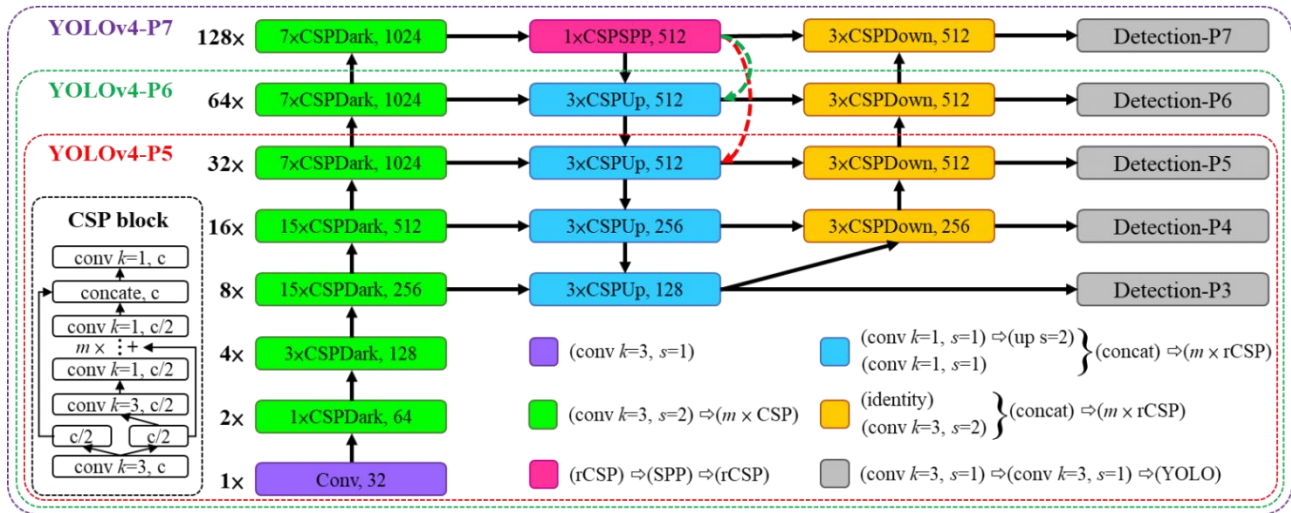


Fig. 4. Examples of YOLOv4-CSP CNN models for input images of various sizes

Compared to previous CNN models of the YOLO class, the YOLOv4-CSP model is more effective for object detection than previous CNN models from this class, which is confirmed by the results of a number of studies on the MS COCO dataset [12]. For example, the average detection accuracy of YOLOv4-CSP (with $\text{mAP}_{0.5}$ metric) is increased from 57,9% to 65,7% compared to the YOLOv3 model. All of the above made it possible to select YOLOv4-CSP for potential use in real-time CVS as a promising CNN model. Next, it is necessary to conduct a study of the effectiveness of this model and determine its compliance with the requirements described above for the CVS and, accordingly, for the CNN model implemented in such a system.

3. Training the YOLOv4-CSP CNN model and studying its effectiveness

3.1. Dataset preparation

To study the effectiveness of the YOLOv4-CSP CNN model in solving the problem of detecting FOs, two datasets were marked and prepared based on the results of shooting FOs of two classes: helicopter-type UAVs and aircraft-type UAVs. The size of each dataset is 6230 labeled images. The first dataset is formed on the basis of 1150 optical images (RGB), the second - on the basis of 1160 images obtained in the IR range. To expand the datasets to 6230 images, data augmentation was carried out in each of them using the following algorithms:

- varying the size of images;
- mirroring images horizontally and vertically;
- rotation of images at an angle from 1 to 15 degrees;
- mosaic placement of objects in the image.

70% of the volume of each of these datasets was used for training the YOLOv4-CSP CNN model, 20% was used for validation, and 10% was used for testing/research. The Python programming language and the PyTorch framework were used for the software implementation of the YOLOv4-CSP CNN. Training of the CNN model and testing/research of its effectiveness were carried out on a computer with the following characteristics: Intel Core i9-11900KF processor, 64 GB RAM, NVIDIA Quadro RTX 6000 video card with 24 GB of video memory.

3.2. Results of FO detection in images

The number of epochs was set equal to 200, and learning rate was set equal to 0.001. These parameters were not changed during training and validation processes for the CNN model. The Adam algorithm was used as an optimizer.

When testing the trained CNN model YOLOv4-CSP and conducting studies of its effectiveness on datasets with optical images and images in the IR range, the following model parameters were changed: input image size (416x416, 512x512, 608x608 pixels), mini-batch size (4, 8) and activation function (Leaky ReLU, Mish).

Figure 5 shows, as an example, the results of detection using the trained and validated CNN model YOLOv4-CSP FO of each class on optical images.

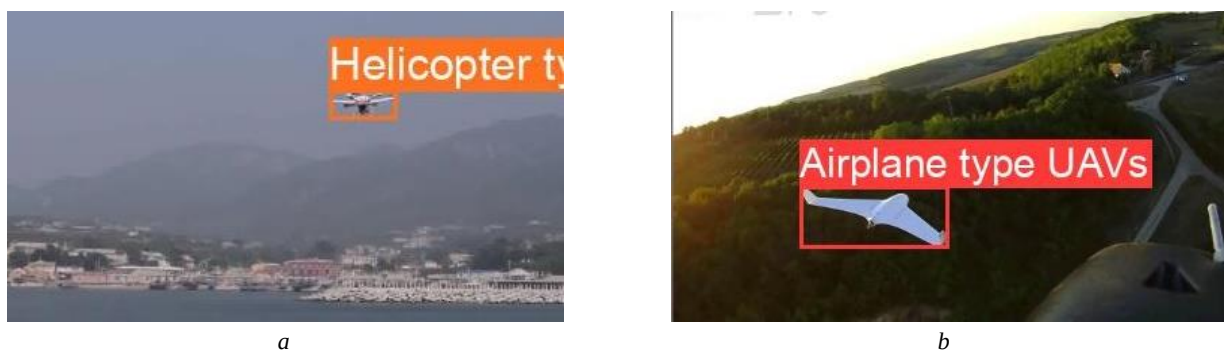


Fig. 5. Results of FO detection on optical images: Helicopter-type UAV (a); Aircraft-type UAV (b)

The first experiment was conducted to determine the effect on the accuracy of FO detection using this model depending on the size of mini-batches, the size of the input optical images, as well as the approaches used to train the model, implemented in the form of BoF (Bag of Freebies) procedures and BoS (Bag of Specials). The Leaky ReLU activation function was used in the experiment. BoF and BoS are procedures that implement methods and techniques that change the training strategy or cost of training a CNN model to improve the accuracy of detecting objects in images with its help. These methods are implemented in the model

in the form of procedures (various plugins and post-processing modules), which can significantly increase the accuracy of object detection at the cost of a slight increase in the cost of training the CNN model [8]. The results of first experiment are shown in table 1.

Table 1

Accuracy of FO detection in optical images for various mini-batch sizes, input images and usage of BoF/BoS procedures

mini-batch size	input image size	BoF/BoS	mAP _{0.5} , %
4	416 × 416	–	61,5
4	416 × 416	+	62,7
8	416 × 416	–	62,3
8	416 × 416	+	63,2
4	512 × 512	–	62,9
4	512 × 512	+	64,7
8	512 × 512	–	63,1
8	512 × 512	+	64,8

From table 1 it follows that the use of BoF/BoS procedures in conducted experiment significantly increases the accuracy of FO detection. In this case, the size of the mini-batch has little effect on it, but the accuracy increases with increasing size. The accuracy of FO detection increases with increasing size of input images.

In the second experiment, the accuracy of FO detection in optical images was determined using the YOLOv4-CSP CNN model using different activation functions and for different sizes of input images. The BoF/BoS procedures were also used, the mini-batch size remained unchanged and was equal to 8. The results of the experiment are shown in table 2.

Table 2

Accuracy of FO detection in optical images for various activation functions and input image sizes

activation function	input image size	mAP _{0.5} , %
Leaky ReLU	416 × 416	63,2
Mish	416 × 416	63,7
Leaky ReLU	512 × 512	64,8
Mish	512 × 512	65,2
Leaky ReLU	608 × 608	65,1
Mish	608 × 608	66,1

In the third experiment, the detection rate of FO was determined using the YOLOv4-CSP CNN model for various sizes of input optical images. In this case, BoF/BoS procedures were used, and the mini-batch size remained unchanged and was equal to 8. The results of the experiment are shown in table 3.

Table 3

Speed of FO detection in optical images for different input image sizes

input image size	Frames per second (FPS)	Average detection time, ms
416 × 416	43	23,2
512 × 512	35	28,6
608 × 608	29	34,5

Here, the detection rate was defined as calculated frames per second and the average time of FO detection per analysis of one image. It was calculated by averaging the results over the detection time of 623 optical images from the test sample of the first dataset.

In experiments №№ 4–6, only the initial data for training, validating and testing the YOLOv4-CSP CNN model was changed: instead of optical images, images obtained in the IR range were fed to the input of the model. The results of these experiments are presented in table 4, table 5 and table 6. The results in table 5 are obtained using a mini-batch size of 8 and using the BoF/BoS procedures. The average FO detection time required to analyze one image was calculated by averaging the results over the detection time of 623 such images from the test sample of the second dataset.

Table 4

Accuracy of FO detection in IR images for various mini-batch sizes, input images and usage of BoF/BoS procedures

mini-batch size	input image size	BoF/BoS	mAP _{0.5} , %
4	416 × 416	–	55,3
4	416 × 416	+	56,9
8	416 × 416	–	56,8
8	416 × 416	+	57,2
4	512 × 512	–	57,6
4	512 × 512	+	59,1
8	512 × 512	–	58,4
8	512 × 512	+	59,5

Table 5

Accuracy of FO detection in IR images for various activation functions and input image sizes

activation function	input image size	mAP _{0.5} , %
Leaky ReLU	416 × 416	56,9
Mish	416 × 416	57,2
Leaky ReLU	512 × 512	59,5
Mish	512 × 512	60,1
Leaky ReLU	608 × 608	60,3
Mish	608 × 608	60,6

Table 6

Speed of FO detection in IR images for different input image sizes

input image size	Frames per second (FPS)	Average detection time, ms
416 × 416	42	23,8
512 × 512	36	27,7
608 × 608	27	37,0

4. Analysis of the obtained results

As it follows from table 1 and table 4, the accuracy of FO detection in images using the YOLOv4-CSP CNN model in the case of using the BoF/BoS procedures changes slightly with increasing mini-batch size. Without the use of these procedures, regardless of the wavelength range in which images of objects are obtained, the size of the mini-batch significantly affects the accuracy of FO detection: the larger its value, the more accurate the detection. This is consistent with the results obtained when detecting various objects of a different physical nature in images [11].

The experimental results in table 2 and table 5 show that the accuracy of FO detection in images using the YOLOv4-CSP CNN model is higher when using the Mish activation function for different input image sizes. This function is smooth, non-monotonic, bounded below and unbounded above, so these properties allow it to obtain better results compared to the Leaky ReLU activation function. However, existing implementations of the Mish activation function depend on support for CUDA hardware instructions, making it less universally usable than the Leaky ReLU activation function. This should be taken into account when developing CVS.

From the experimental results shown in table 3 and table 6, it follows that the best speed of detecting FOs in images using the CNN model under study can be obtained with an input image size of 416x416 pixels. As the size of the input images increases, the detection time per image increases, and the number of frames processed per second decreases. At the same time, the research results indicate that with the considered sizes of input images (416x416, 512x512 and 608x608 pixels), the CVS created on the basis of the YOLOv4-CSP CNN model for FO detection will work in real time (the FPS parameter value in all cases is above 25).

Analysis of the results obtained shows that the accuracy of FO detection using the mAP_{0.5} metric using the YOLOv4-CSP CNN model on optical images is 5–6% higher than on images obtained in the IR range.

This is explained by the fact that the model, when generating feature maps from optical images, also takes into account the characteristics of FO colors. Moreover, exceeding the threshold for the detection accuracy of FO, equal to 60%, occurs in all studied cases of optical image analysis (table 1 and table 2). However, for IR images, exceeding this threshold was detected only for three cases: with input image sizes of 512x512 and 608x608 pixels and the Mish activation function, as well as for input images with sizes of 608x608 pixels and the Leaky ReLU activation function (table 5). It is the results obtained on the accuracy of FO detection that make it possible to formulate recommendations for the use of the YOLOv4-CSP CNN model as the basis of real-time CVS. Thus, for optical images, taking into account the specifics of the created CVS, it can be recommended to select a model with parameters from a fairly wide set. For images obtained in the IR range, it is possible to use in the CVS only three versions of the model, which were identified in the experiments.

Conclusion

The most important task in airspace control is the task of detecting and classifying various FO. In recent years, FO images obtained in the optical and IR wavelength ranges have been increasingly analyzed to solve this problem. For this purpose, CVSs are created based on modern CNN models. We selected the YOLOv4-CSP model, which is part of the YOLO class, as a promising CNN model for potential use in real-time CVS.

Comprehensive studies have been carried out on the effectiveness of the selected CNN model in detecting two classes of FO in images: helicopter-type UAVs and aircraft-type UAVs. It was revealed that the accuracy of detecting such FOs using the YOLOv4-CSP CNN model is positively influenced by the use of BoF/BoS procedures and an increase in the size of input images. To a lesser extent, the accuracy is influenced by the choice of activation function (the best results were obtained using the Mish activation function) and mini-batch size (a larger mini-batch slightly improves detection accuracy). The obtained estimates of the speed of FO detection in images indicate the fundamental possibility of creating a CVS based on the YOLOv4-CSP CNN model, which detects FO in real time.

It has been shown that the accuracy of FO detection using this CNN model in optical images is 5–6% higher than in images obtained in the IR range. Exceeding the threshold for the detection accuracy of FO, equal to 60%, occurs in all studied cases of optical image analysis. However, for IR images, exceeding this threshold was detected only for three variants of the model. The results on the accuracy of FO detection make it possible to formulate recommendations for the use of the YOLOv4-CSP CNN model as the basis for real-time CVS.

References

1. Zoev, I.V., Markov, N.G. & Ryzhova, S.E. (2019) Intelligent computer vision system for unmanned aerial vehicles for monitoring technological objects of oil and gas industry. *Bulletin of the Tomsk Polytechnic University. Geo Assets Engineering*. 330(11). pp. 34–49. DOI: 10.18799/24131830/2019/11/2346
2. Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M.A., Al-Amidie, M. & Farhan, L. (2021) Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*. 8(1). pp. 53. DOI: 10.1186/Fs40537-021-00444-8
3. Makarenko, S.I., Timoshenko, A.V. & Vasilchenko, A.S. (2020) Counter Unmanned Aerial Vehicles. Part 1. Unmanned aerial vehicle as an object of detection and destruction. *Systems of Control, Communication and Security*. 1. pp. 109–146. DOI: 10.24411/2410-9916-2020-10105
4. Klekovkin, V.A. & Markov, N.G. (2023) Model of convolutional neural network LeNet5 for detection and classification of air space objects in images. *Bulletin of the Tomsk Polytechnic University. Industrial Cybernetics*. 1(2). pp. 11–16. DOI: 10.18799/29495407/2023/2/17
5. Nebaba, S.G. & Markov, N.G. (2022) Effectiveness of moving objects detecting and tracking in airspace by images in near-infrared. *Light & Engineering*. 30(2). pp. 62–69. DOI: 10.33383/2021-102
6. Girshick, R., Donahue, J., Darrell, T. & Malik, J. (2014) Rich feature hierarchies for accurate object detection and semantic segmentation. *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 580–587. DOI: 10.48550/arXiv.1311.2524

7. Ren, S., He, K., Girshick, R. & Sun, J. (2015) Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 39(6). pp. 1137–1149. DOI: 10.48550/arXiv.1506.01497
8. Bochkovskiy, A., Wang, C.Y. & Liao, H.Y.M. (2020) YOLOv4: Optimal Speed and Accuracy of Object Detection. *Journal arXiv, preprint arXiv*: 2004.10934v1. DOI: 10.48550/arXiv.2004.10934
9. Redmon, J. & Ali, F. (2018) Yolov3: An incremental improvement. *Journal arXiv, preprint arXiv*: 1804.02767. DOI: 10.48550/arXiv.1804.02767
10. Wang, C.Y., Mark Liao, H.Y., Yeh, I.H., Wu, Y.H., Chen, P.Y. & Hsieh, J.W. (2019) CSPNet: A New Backbone that can Enhance Learning Capability of CNN. *Journal arXiv, preprint arXiv*: 1911.11929v1. DOI: 10.48550/arXiv.1911.11929
11. He, K., Zhang, X., Ren, S. & Sun, J. (2015) Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. *Journal arXiv, preprint arXiv*: 1406.4729v4. DOI: 10.48550/arXiv.1406.4729
12. Wang, C.Y., Bochkovskiy, A. & Liao, H.Y.M. (2021) Scaled-YOLOv4: Scaling Cross Stage Partial Network. *Proceedings of the IEEE/cvf conference on computer vision and pattern recognition*. DOI: 10.48550/arXiv.2011.08036

Information about the authors:

Nebaba Stepan G. (Candidate of Technical Sciences, Associate Professor, National Research Tomsk Polytechnic University, Tomsk, Russian Federation). E-mail: stepanlfx@tpu.ru

Markov Nikolay G. (Doctor of Technical Sciences, Professor, National Research Tomsk Polytechnic University, Tomsk, Russian Federation). E-mail: markovng@tpu.ru

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

Информация об авторах:

Небаба Степан Геннадьевич – кандидат технических наук, доцент отделения информационных технологий Инженерной школы информационных технологий и робототехники Национального исследовательского Томского политехнического университета (Томск, Россия). E-mail: stepanlfx@tpu.ru

Марков Николай Григорьевич – доктор технических наук, профессор отделения информационных технологий Инженерной школы информационных технологий и робототехники Национального исследовательского Томского политехнического университета (Томск, Россия). E-mail: markovng@tpu.ru

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

Received 08.07.2024; accepted for publication 02.12.2024

Поступила в редакцию 08.07.2024; принята к публикации 02.12.2024

Научная статья

УДК 519.2

doi: 10.17223/19988605/69/9

**Обобщенный асинхронный поток событий
с продлевающимся мертвым временем****Людмила Алексеевна Нежелская¹, Валентина Денисовна Пономаренко²**^{1,2} *Национальный исследовательский Томский государственный университет, Томск, Россия*¹ *ludne@mail.ru*² *valya.ponomarenko.00@mail.ru*

Аннотация. Рассматривается дважды стохастический обобщенный асинхронный поток событий с двумя состояниями, функционирующий в стационарном режиме в условиях продлевающегося мертвого времени фиксированной длительности. Решена задача оценивания длительности мертвого времени в рекуррентном обобщенном асинхронном потоке методом моментов. Представлены численные результаты статистических экспериментов, поставленных на имитационной модели потока.

Ключевые слова: дважды стохастический поток событий; обобщенный асинхронный поток событий; инфинитезимальные характеристики; продлевающееся мертвое время; совместная плотность вероятности; условия рекуррентности; преобразование Лапласа; метод моментов; статистические эксперименты.

Для цитирования: Нежелская Л.А., Пономаренко В.Д. Обобщенный асинхронный поток событий с продлевающимся мертвым временем // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 69. С. 82–94. doi: 10.17223/19988605/69/9

Original article

doi: 10.17223/19988605/69/9

Generalized asynchronous flow of events with extending dead time**Lyudmila A. Nezhel'skaya¹, Valentina D. Ponomarenko²**^{1,2} *National Research Tomsk State University, Tomsk, Russian Federation*¹ *ludne@mail.ru*² *valya.ponomarenko.00@mail.ru*

Abstract. We consider a double stochastic generalized asynchronous flow of events with two states, operating in a stationary mode under conditions of an extended dead time of a fixed duration. The problem of estimating the duration of dead time in a recurrent generalized asynchronous flow of events using the method of moments has been solved. Numerical results of statistical experiments performed on a simulation model are presented.

Keywords: double stochastic flow of events; generalized asynchronous flow of events; infinitesimal characteristics; extending dead time; joint probability density; low recurrence conditions; Laplace transform; method of moments; statistical experiments.

For citation: Nezhel'skaya, L.A., Ponomarenko, V.D. (2024) Generalized asynchronous flow of events with extending dead time. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 69. pp. 82–94. doi: 10.17223/19988605/69/9

Введение

Математическая модель дважды стохастического потока является наиболее приближенной к реальности моделью потоков заявок, запросов, сообщений, протекающих в компьютерных сетях, телекоммуникационных сетях и сетях связи [1]. Для дважды стохастического потока характерно сле-

дующее: сопровождающий случайный процесс является принципиально ненаблюдаемым и моменты наступления событий случайны. Дважды стохастические потоки делятся на два класса: к первому классу относятся потоки, интенсивность которых есть непрерывный случайный процесс [2, 3], ко второму – потоки с интенсивностью в виде кусочно-постоянного случайного процесса с конечным числом состояний [4–8].

В большом количестве работ по теории массового обслуживания решаются задачи, когда все события входящего в систему массового обслуживания потока доступны для наблюдения. Однако такая ситуация далеко не всегда соответствует реальности, так как в реальных системах входящее событие способно породить период мертвого времени для регистрирующего прибора. В течение этого периода другие события входящего потока становятся ненаблюдаемыми – теряются [9]. Мертвое время, порождаемое наступившим событием, является в некотором роде искажающим фактором при решении различного рода задач оценивания состояний потока или его параметров по моментам наступления событий в наблюдаемом потоке.

В работах [10–14] решены задачи оценивания длительности мертвого времени, когда наступившие в течение периода мертвого времени события не вызывают его продления (непродлевающееся мертвое время).

Поток событий, рассматриваемый в данной работе, относится ко второму классу дважды стохастических потоков. Предполагается, что событие, наступившее в течение периода мертвого времени, хотя и не наблюдается, все же способно продлить общий период ненаблюдаемости исходного потока (продлевающееся мертвое время). При этом принимается, что мертвое время имеет фиксированную длительность.

Отличие настоящего исследования от проведенного в [10] в том, что в данной работе обобщенный асинхронный поток рассматривается в условиях продлевающегося мертвого времени фиксированной длительности.

В настоящей работе, являющейся продолжением исследования, начатого в [15], выводится и решается уравнение моментов в задаче оценивания длительности продлевающегося мертвого времени, проводятся статистические эксперименты с целью установления качества получаемых оценок.

1. Математическая модель потока

Рассматривается обобщенный асинхронный поток событий, обладающий следующими свойствами: стационарность, ординарность, наличие последствия.

Сопровождающий процесс данного потока $\lambda(t)$ является принципиально ненаблюдаемым кусочно-постоянным случайным процессом с двумя состояниями S_1 и S_2 ; будем говорить, что имеет место состояние S_i процесса $\lambda(t)$, если $\lambda(t) = \lambda_i$, $i = 1, 2$, $\lambda_1 > \lambda_2 \geq 0$. Длительность пребывания процесса $\lambda(t)$ в состоянии S_i является случайной величиной с функцией распределения $F_i(t) = 1 - e^{-\lambda_i t}$, $t \geq 0$, $i = 1, 2$. В течение времени пребывания процесса $\lambda(t)$ в состоянии S_1 имеет место пуассоновский поток событий с параметром λ_1 ; в течение времени пребывания процесса $\lambda(t)$ в состоянии S_2 – пуассоновский поток событий с параметром λ_2 .

В момент перехода сопровождающего случайного процесса $\lambda(t)$ из состояния S_1 в состояние S_2 с вероятностью p ($0 \leq p \leq 1$) инициируется дополнительное событие потока в состоянии S_2 (сначала переход из S_1 в S_2 , затем наступление дополнительного события в S_2). Аналогично при переходе сопровождающего случайного процесса $\lambda(t)$ из состояния S_2 в состояние S_1 с вероятностью q ($0 \leq q \leq 1$) инициируется дополнительное событие потока в состоянии S_1 (сначала переход из S_2 в S_1 , затем наступление дополнительного события в S_1).

В момент наступления каждого события потока наступает период ненаблюдаемости фиксированной длительности T (мертвое время), так что другие события, наступившие в течение времени T ,

недоступны наблюдению. Каждое событие, ненаблюдаемое в течение мертвого времени, вызывает продление периода ненаблюдаемости на величину T ; наблюдаться будет лишь то событие, которое наступило после окончания последнего периода ненаблюдаемости.

На рис. 1 приведены схема формирования наблюдаемого потока событий и одна из реализаций сопровождающего процесса $\lambda(t)$, где $S_i, i=1,2, \dots$ – состояния процесса $\lambda(t)$; $t_1, t_2, \dots$ – моменты времени наступления событий в наблюдаемом потоке. Наблюдаемые события обозначены незакрашенными кружками, а ненаблюдаемые события, т.е. недоступные наблюдению из-за наличия мертвого времени, – закрашенными кружками. Штриховкой обозначен период ненаблюдаемости. Длительность общего периода ненаблюдаемости ξ – случайная величина.

Матрицы инфинитезимальных характеристик сопровождающего процесса $\lambda(t)$ имеют вид [16]:

$$\mathbf{D}_0 = \begin{bmatrix} -(\lambda_1 + \alpha_1) & (1-p)\alpha_1 \\ (1-q)\alpha_2 & -(\lambda_2 + \alpha_2) \end{bmatrix}, \quad \mathbf{D}_1 = \begin{bmatrix} \lambda_1 & p\alpha_1 \\ q\alpha_2 & \lambda_2 \end{bmatrix}.$$

Элементами матрицы $\mathbf{D}_1$ являются интенсивности переходов процесса $\lambda(t)$ из состояния в состояние с наступлением события потока. Недиagonalные элементы матрицы $\mathbf{D}_0$ – интенсивности переходов процесса $\lambda(t)$ из состояния в состояние без наступления события. Диагональные элементы матрицы $\mathbf{D}_0$ – интенсивности выхода процесса $\lambda(t)$ из своих состояний, взятые с противоположным знаком.

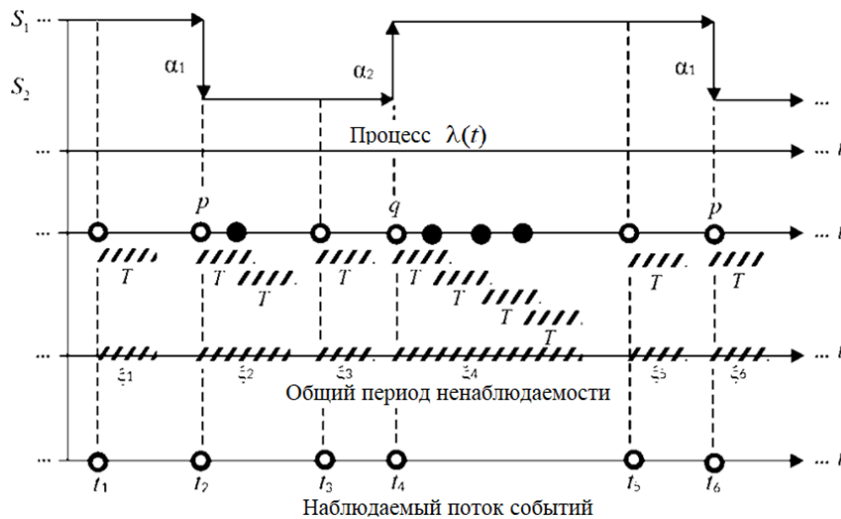


Рис. 1. Реализация наблюдаемого потока событий

Fig. 1. Realization of the observed event flow

Цель данного исследования заключается в получении и решении уравнения моментов в задаче оценивания длительности продлевающегося мертвого времени в рекуррентном обобщенном асинхронном потоке событий, а также в проведении на имитационной модели потока ряда статистических экспериментов с целью установления качества полученных оценок.

2. Двумерная плотность вероятности длительностей двух соседних интервалов.

Условия рекуррентности

Лемма 1. Для обобщенного асинхронного потока событий, функционирующего в условиях полной наблюдаемости, сопровождающий кусочно-постоянный случайный процесс $\lambda(t)$ является марковским процессом [16].

Пусть $t_1, t_2, \dots, t_k, \dots$ – моменты времени наступления событий в потоке.

Лемма 2. Последовательность $\{\lambda(t_k)\}$, порождаемая совокупностью моментов наступления событий $t_1, t_2, \dots, t_k, \dots$, является вложенной цепью Маркова [16].

Предполагается, что обобщенный асинхронный поток событий функционирует в условиях полной наблюдаемости ($T = 0$) в стационарном режиме.

Обозначим $\tau_k = t_{k+1} - t_k$, $\tau_k \geq 0$, – значение длительности k -го интервала между соседними событиями t_k и t_{k+1} , $k = 1, 2, \dots$, наблюдаемого потока. В силу стационарного режима функционирования потока для плотности вероятности значений τ_k справедливо $p(\tau_k) = p(\tau)$, $\tau_k \geq 0$, $\forall k$.

Это позволяет без ограничения общности считать момент наступления события t_k равным нулю, или, что то же самое, момент наступления события есть $\tau = 0$.

В общем случае исследуемый поток в условиях полной наблюдаемости является коррелированным потоком. Для достижения поставленной цели необходимо выписать условия рекуррентности потока и для каждого из выписанных условий определить явный вид плотности вероятности значений длительности интервала между соседними событиями потока.

Воспользуемся следующей теоремой.

Теорема 1. Плотность вероятности значений длительности интервала между соседними событиями в обобщенном асинхронном потоке в случае $z_2 - z_1 \neq 0$ имеет вид [17]:

$$\begin{aligned} p(\tau) &= \gamma z_1 e^{-z_1 \tau} + (1 - \gamma) z_2 e^{-z_2 \tau}, \tau \geq 0, \\ \gamma &= \frac{1}{z_2 - z_1} \left[z_2 - \lambda_1 - \lambda_2 + \frac{(\alpha_1 + \alpha_2)(\lambda_1 \lambda_2 - p q \alpha_1 \alpha_2)}{\alpha_1 \lambda_2 + \alpha_2 \lambda_1 + (p + q) \alpha_1 \alpha_2} \right], \\ z_1 &= \frac{1}{2} \left[\lambda_1 + \alpha_1 + \lambda_2 + \alpha_2 - \sqrt{(\lambda_1 + \alpha_1 - \lambda_2 - \alpha_2)^2 + 4 \alpha_1 \alpha_2 (1 - p)(1 - q)} \right], \\ z_2 &= \frac{1}{2} \left[\lambda_1 + \alpha_1 + \lambda_2 + \alpha_2 + \sqrt{(\lambda_1 + \alpha_1 - \lambda_2 - \alpha_2)^2 + 4 \alpha_1 \alpha_2 (1 - p)(1 - q)} \right], \\ z_2 &> z_1 > 0, z_2 - z_1 \neq 0. \end{aligned} \quad (1)$$

Обозначим $p(\tau_1, \tau_2)$ – совместная плотность вероятности длительностей двух соседних интервалов (t_1, t_2) , (t_2, t_3) ; $\tau_1 = t_2 - t_1$, $\tau_2 = t_3 - t_2$ – значения длительностей, $\tau_1 \geq 0$, $\tau_2 \geq 0$. В силу стационарного режима расположение одного интервала (t_k, t_{k+1}) либо двух соседних интервалов (t_k, t_{k+1}) , (t_{k+1}, t_{k+2}) между моментами наступления событий $t_1, t_2, \dots, t_k, \dots$, $k = 1, 2, \dots$, на временной оси произвольно.

Справедлива следующая теорема.

Теорема 2. Совместная плотность вероятности $p(\tau_1, \tau_2)$ длительностей двух соседних интервалов для обобщенного асинхронного потока событий в случае $z_2 - z_1 \neq 0$ имеет вид [17]:

$$\begin{aligned} p(\tau_1, \tau_2) &= p(\tau_1) p(\tau_2) + \gamma (1 - \gamma) \frac{\lambda_1 \lambda_2 - p q \alpha_1 \alpha_2}{z_1 z_2} (z_1 e^{-z_1 \tau_1} - z_2 e^{-z_2 \tau_1}) (z_1 e^{-z_1 \tau_2} - z_2 e^{-z_2 \tau_2}), \\ \tau_1 &\geq 0, \tau_2 \geq 0, \\ \gamma (1 - \gamma) &= \frac{\alpha_1 \alpha_2 z_1 z_2 (\lambda_1 - \lambda_2 + p \alpha_1 - q \alpha_2) (\lambda_1 - \lambda_2 + q \alpha_1 - p \alpha_2)}{(z_2 - z_1)^2 (\lambda_1 \alpha_2 + \lambda_2 \alpha_1 + (p + q) \alpha_1 \alpha_2)^2}, \\ z_2 &> z_1 > 0, z_2 - z_1 \neq 0, \end{aligned} \quad (2)$$

где z_1, z_2 определены в (1); $p(\tau_k)$ определены в (1) для $\tau = \tau_k$, $k = 1, 2$.

Из анализа (2) вытекают условия рекуррентности и соответствующие плотности $p(\tau)$.

1. Если $\lambda_1 - \lambda_2 + p \alpha_1 - q \alpha_2 = 0$, то $\gamma (1 - \gamma) = 0$ и $p(\tau_1, \tau_2)$ факторизуется.

Подставляя условие $\lambda_1 - \lambda_2 + p \alpha_1 - q \alpha_2 = 0$ в (1), находим плотность вероятности значений длительности интервала между соседними событиями $p(\tau)$ в виде:

$$p(\tau) = (p \alpha_1 + \lambda_1) e^{-(p \alpha_1 + \lambda_1) \tau}, \tau \geq 0.$$

В этом случае обобщенный асинхронный поток вырождается в простейший с параметром $p \alpha_1 + \lambda_1$.

2. Если $\lambda_1 - \lambda_2 + q\alpha_1 - p\alpha_2 = 0$, то $\gamma(1 - \gamma) = 0$ и $p(\tau_1, \tau_2)$ факторизуется.

Подставляя условие $\lambda_1 - \lambda_2 + q\alpha_1 - p\alpha_2 = 0$ в (1), находим плотность вероятности значений длительности интервала между соседними событиями $p(\tau)$ в виде:

$$p(\tau) = (q\alpha_1 + \lambda_1)e^{-(q\alpha_1 + \lambda_1)\tau}, \tau \geq 0.$$

Как и в предыдущем случае обобщенный асинхронный поток вырождается в простейший с параметром $q\alpha_1 + \lambda_1$.

3. Если $\lambda_1\lambda_2 - pq\alpha_1\alpha_2 = 0$, то совместная плотность вероятности $p(\tau_1, \tau_2)$ факторизуется. Подставляя условие $\lambda_1\lambda_2 - pq\alpha_1\alpha_2 = 0$ в (1), находим плотность вероятности значений длительности интервала между соседними событиями $p(\tau)$ в виде:

$$p(\tau) = \gamma z_1 e^{-z_1\tau} + (1 - \gamma) z_2 e^{-z_2\tau}, \tau \geq 0, \quad (3)$$

$$\gamma = \frac{z_2 - \lambda_1 - \lambda_2}{z_2 - z_1}, \quad z_2 - z_1 \neq 0.$$

Далее везде рассматривается рекуррентный поток в случае 3.

3. Преобразование Лапласа плотности вероятности общего периода ненаблюдаемости

Пусть ξ – длительность общего периода ненаблюдаемости в рекуррентном обобщенном асинхронном потоке, функционирующем в условиях продлевающегося мертвого времени фиксированной длительности $T \neq 0$. Последовательность $t_1, t_2, \dots$ моментов наступления событий в наблюдаемом потоке образует вложенную цепь Маркова, и рекуррентность наблюдаемого потока сохраняется.

Рассмотрим функцию Пальма $\phi_0(T) = \int_T^\infty \tilde{p}(x) dx$ – вероятность того, что на интервале $(0, T)$ событий рекуррентного потока не наступит при условии, что в начальный момент времени интервала $(0, T)$ событие наступило [18]; здесь $\tilde{p}(x)$ – плотность вероятности длительности интервала между соседними событиями в дважды стохастическом рекуррентном потоке.

Теорема 3. Преобразование Лапласа плотности вероятности значений длительности общего периода ненаблюдаемости в рекуррентном обобщенном асинхронном потоке с продлевающимся мертвым временем в случае $z_2 - z_1 \neq 0$ имеет вид:

$$g_\xi(s) = \frac{\phi_0(T)}{e^{sT}} \left[1 - \left[(1 - e^{-(s+z_1)T}) z_1 \frac{z_2 - \lambda_1 - \lambda_2}{(s+z_1)(z_2-z_1)} + (1 - e^{-(s+z_2)T}) z_2 \frac{-z_1 + \lambda_1 + \lambda_2}{(s+z_2)(z_2-z_1)} \right] \right]^{-1}, \quad (4)$$

где $\phi_0(T) = \gamma e^{-z_1 T} + (1 - \gamma) e^{-z_2 T}$; величина γ определена в (3); z_1, z_2 – в (1).

Доказательство. Воспользуемся результатом, полученным в [17]. Преобразование Лапласа плотности вероятности значений длительности общего периода ненаблюдаемости в рекуррентном дважды стохастическом потоке событий, функционирующем в условиях продлевающегося мертвого времени, имеет вид:

$$g_\xi(s) = \frac{\phi_0(T)}{e^{sT}} \left[1 - \int_0^T e^{-sx} \tilde{p}(x) dx \right]^{-1}. \quad (5)$$

Переобозначая плотность $p(\tau)$, определенную в (3), на $\tilde{p}(x)$, подставляя в (5) и выполняя необходимые преобразования, приходим к (4). Теорема доказана.

Лемма 3. Математическое ожидание длительности общего периода ненаблюдаемости ξ в рекуррентном обобщенном асинхронном потоке событий с продлевающимся мертвым временем в случае $z_2 - z_1 \neq 0$ имеет вид:

$$M\xi = \frac{1}{z_1 z_2 \varphi_0(T)} \left[\alpha_1 + \alpha_2 - \gamma z_2 e^{-z_1 T} - (1 - \gamma) z_1 e^{-z_2 T} \right], \quad (6)$$

где z_1, z_2 определены в (1); γ – в (3); $\varphi_0(T)$ – в (4).

Доказательство. Можно показать, что с использованием вида $g_\xi(s)$, определенного в (5), математическое ожидание ξ запишется в виде:

$$M\xi = -g'_\xi(s) \Big|_{s=0} = T + \frac{1}{\varphi_0(T)} \int_0^T x \tilde{p}(x) dx. \quad (7)$$

Подставляя $\tilde{p}(x)$, определенную в (3), в (7) и выполняя необходимые преобразования, находим (6).

4. Преобразование Лапласа плотности вероятности длительности интервала между соседними событиями в наблюдаемом потоке

Рассмотрим интервал времени (t_k, t_{k+1}) , значение длительности которого есть $\tau_k = t_{k+1} - t_k$. С другой стороны, длительность этого интервала равна $\tau = \xi + \eta$, где η – длительность интервала между моментом окончания общего периода ненаблюдаемости и моментом t_{k+1} . Индекс k опущен в силу произвольного расположения интервала на временной оси. Случайные величины η и ξ являются независимыми. Тогда плотность вероятности $p(\tau)$ длительности интервала между событиями в наблюдаемом потоке запишется в виде:

$$p(\tau) = \int_0^\tau p(\xi) p(\eta|\xi) d\xi = \int_0^\tau p(\xi) p(\tau - \xi|\xi) d\xi. \quad (8)$$

Теорема 4. Преобразование Лапласа плотности вероятности значений длительности интервала между соседними событиями в рекуррентном обобщенном асинхронном потоке с продлевающимся мертвым временем в случае $z_2 - z_1 \neq 0$ имеет вид:

$$g_\tau(s) = \left[(z_1 + s)(z_2 + s)(\alpha_1 + \alpha_2) \right]^{-1} \left[z_1 z_2 (s + \alpha_1 + \alpha_2) g_\xi(s) + \right. \\ \left. + \alpha_1 \alpha_2 (\lambda_1 - \lambda_2 + q \alpha_1 - p \alpha_2) (\lambda_1 - \lambda_2 + p \alpha_1 - q \alpha_2) [z_1 z_2]^{-1} s g_\xi(\alpha_1 + \alpha_2 + s) \right], \quad (9)$$

где функция $g_\xi(s)$ определена формулой (4); $g_\xi(\alpha_1 + \alpha_2 + s)$ определена формулой (4), в которой нужно заменить аргумент s на $\alpha_1 + \alpha_2 + s$.

Доказательство. Найдём выражение для $p(\tau - \xi|\xi)$. Пусть момент наступления события в наблюдаемом потоке есть $\tau = 0$. Рассмотрим временной интервал $(0, \tau) = (0, \xi + \eta)$ и зафиксируем ξ . Пусть $p_{ij}(\tau - \xi)$ – условные вероятности того, что в интервале длительности $\eta = \tau - \xi$ не наступит событий наблюдаемого потока и $\lambda(\xi + \eta) = \lambda_j$ при условии, что $\lambda(\xi) = \lambda_i$, $i, j = 1, 2$. По смыслу введенные условные вероятности не отличаются от вероятностей $p_{ij}(\tau)$ для обобщенного асинхронного потока, полученных в [17]:

$$p_{11}(\tau) = -\frac{\lambda_1 + \alpha_1 - z_2}{z_2 - z_1} e^{-z_1 \tau} + \frac{\lambda_1 + \alpha_1 - z_1}{z_2 - z_1} e^{-z_2 \tau}, \\ p_{12}(\tau) = \frac{\alpha_1(1-p)}{z_2 - z_1} (e^{-z_1 \tau} - e^{-z_2 \tau}), \quad p_{21}(\tau) = \frac{\alpha_2(1-q)}{z_2 - z_1} (e^{-z_1 \tau} - e^{-z_2 \tau}), \\ p_{22}(\tau) = \frac{\lambda_1 + \alpha_1 - z_1}{z_2 - z_1} e^{-z_1 \tau} - \frac{\lambda_1 + \alpha_1 - z_2}{z_2 - z_1} e^{-z_2 \tau}, \quad \tau \geq 0, \quad z_2 - z_1 \neq 0; \quad (10)$$

z_1, z_2 определены в (1).

Тогда условные вероятности $p(\tau - \xi | \xi)$, $i, j = 1, 2$, определяются формулами (10), в которых аргумент τ заменён на $\tau - \xi$.

Введем $P_i(\tau - \xi) = p_{i1}(\tau - \xi) + p_{i2}(\tau - \xi)$ – условную вероятность того, что в интервале (ξ, τ) событий наблюдаемого потока не произойдет при условии, что $\lambda(\xi) = \lambda_i$, $i = 1, 2$. Тогда условная плотность вероятности длительности интервала (ξ, τ) по определению есть $p_i(\tau - \xi) = -P'_i(\tau - \xi)$, $i = 1, 2$. Учитывая (10), находим

$$\begin{aligned} p_1(\tau - \xi) &= \frac{1}{z_2 - z_1} \left\{ z_1(-\lambda_1 - \alpha_1 p + z_2) e^{-z_1(\tau - \xi)} + z_2(\lambda_1 + \alpha_1 p - z_1) e^{-z_2(\tau - \xi)} \right\}, \\ p_2(\tau - \xi) &= \frac{1}{z_2 - z_1} \left\{ z_1[\lambda_1 + \alpha_1 + \alpha_2(1 - q) - z_1] e^{-z_1(\tau - \xi)} + z_2[-(\lambda_1 + \alpha_1 + \alpha_2(1 - q)) + z_2] e^{-z_2(\tau - \xi)} \right\}, \tau \geq \xi. \end{aligned} \quad (11)$$

Введем вероятность $\pi_i(\tau | \xi)$ – условную вероятность того, что $\lambda(\tau) = \lambda_i$, $i = 1, 2$, при условии, что в момент $\tau = 0$ событие наступило и наступило мертвое время длительности ξ . Тогда

$$p(\tau - \xi | \xi) = \pi_1(\tau = \xi | \xi) p_1(\tau - \xi) + \pi_2(\tau = \xi | \xi) p_2(\tau - \xi). \quad (12)$$

Вероятности $\pi_i(\tau = \xi | \xi)$ по смыслу совпадают с вероятностями $\pi_i(T)$, $i = 1, 2$, определенными в [17], в которых вместо периода ненаблюдаемости T следует рассматривать ξ :

$$\begin{aligned} \pi_2(\xi) &= \pi_2(\tau = \xi | \xi) = \pi_2 - [\pi_2 - \pi_2(0 | \xi)] e^{-(\alpha_1 + \alpha_2)\xi}, \\ \pi_2(0 | \xi) &= \frac{p_{12} + \delta \pi_2 [1 - e^{-(\alpha_1 + \alpha_2)\xi}]}{1 - \delta e^{-(\alpha_1 + \alpha_2)\xi}}, \delta = (\lambda_1 \lambda_2 - p q \alpha_1 \alpha_2) / z_1 z_2, \\ p_{12} &= \alpha_1(\lambda_2 + p \alpha_2) / z_1 z_2, \pi_2 = \alpha_1 / (\alpha_1 + \alpha_2); \pi_1(\tau = \xi | \xi) = 1 - \pi_2(\tau = \xi | \xi). \end{aligned} \quad (13)$$

Принимая во внимание условие рекуррентности $\lambda_1 \lambda_2 - p q \alpha_1 \alpha_2 = 0$ и подставляя (11), (13) в (12), находим

$$\begin{aligned} p(\tau - \xi | \xi) &= \begin{cases} 0, & 0 \leq \tau < \xi, \\ \Gamma(\xi) z_1 e^{-z_1(\tau - \xi)} + (1 - \Gamma(\xi)) z_2 e^{-z_2(\tau - \xi)}, & \tau \geq \xi, \end{cases} \\ \Gamma(\xi) &= \frac{1}{z_2 - z_1} \left[z_2 - (\lambda_1 + p \alpha_1) \pi_1(\xi) - (\lambda_2 + q \alpha_2) \pi_2(\xi) \right], \\ \pi_2(\xi) &= \pi_2 \left[1 - \alpha_2 \frac{\lambda_1 - \lambda_2 + q \alpha_1 - p \alpha_2}{z_1 z_2} e^{-(\alpha_1 + \alpha_2)\xi} \right], \pi_1(\xi) = 1 - \pi_2(\xi). \end{aligned} \quad (14)$$

Найдем преобразование Лапласа плотности $p(\tau)$, используя (8) и (14). Имеем

$$\begin{aligned} g_\tau(s) &= \int_0^\infty e^{-s\tau} p(\tau) d\tau = \\ &= \int_0^\infty e^{-s\tau} \left[\int_0^\tau p(\xi) p(\tau - \xi | \xi) d\xi \right] d\tau = \int_0^\infty p(\xi) \left[\int_\xi^\infty e^{-s\tau} p(\tau - \xi | \xi) d\tau \right] d\xi = \\ &= \int_0^\infty p(\xi) \left[\int_\xi^\infty e^{-s\tau} \Gamma(\xi) z_1 e^{-z_1(\tau - \xi)} d\tau + \int_\xi^\infty e^{-s\tau} (1 - \Gamma(\xi)) z_2 e^{-z_2(\tau - \xi)} d\tau \right] d\xi. \end{aligned}$$

В выражении для $g_\tau(s)$ выполним замену переменных: $\tau - \xi = \eta$, $\tau = \xi + \eta$. Тогда

$$g_\tau(s) = \int_0^\infty p(\xi) \left[\int_0^\infty e^{-s(\eta + \xi)} z_1 \Gamma(\xi) e^{-z_1 \eta} d\eta + \int_0^\infty e^{-s(\eta + \xi)} z_2 (1 - \Gamma(\xi)) e^{-z_2 \eta} d\eta \right] d\xi.$$

Подставляя явный вид $\Gamma(\xi)$, определенный в (14), и выполняя достаточно трудоемкие преобразования, приходим к (9).

5. Оценка длительности мёртвого времени

Преобразование Лапласа (9) позволяет получить начальные моменты

$$M(\tau^l) = (-1)^l g_{\tau}^{(l)}(s) \Big|_{s=0}, \quad l=1, 2, \dots$$

Поэтому будем решать задачу оценивания длительности мертвого времени T методом моментов. Введем статистики $C_l = \frac{1}{n} \sum_{k=1}^n \tau_k^l$, $l=1, 2, \dots$, где $\tau_k = t_{k+1} - t_k$ – значение длительности интервала между моментами t_k и t_{k+1} наступления событий в рекуррентном обобщенном асинхронном потоке с продлевающимся мертвым временем.

Предположим, что параметры потока $\alpha_1, \alpha_2, \lambda_1, \lambda_2, p, q$ являются известными. При количестве наблюдений $n \rightarrow \infty$ выборочный момент C_1 стремится к теоретическому моменту $M(\tau)$ [19]. Тогда для оценки длительности мертвого времени T имеем уравнение моментов $M(\tau) = C_1$.

Вычислив производную от (9) по s в точке $s=0$ и взяв её со знаком минус, получим уравнение для нахождения значения $\hat{T}$ оценки $\hat{T}$:

$$\begin{aligned} & \frac{1}{z_1 z_2} (z_1 + z_2 - \frac{z_1 z_2}{\alpha_1 + \alpha_2}) - \frac{\alpha_1 \alpha_2 (\lambda_1 - \lambda_2 + q \alpha_1 - p \alpha_2) (\lambda_1 - \lambda_2 + p \alpha_1 - q \alpha_2)}{(\alpha_1 + \alpha_2) (z_1 z_2)^2} \times \\ & \times \frac{\gamma e^{-(z_1 + \alpha_1 + \alpha_2)T} + (1 - \gamma) e^{-(z_2 + \alpha_1 + \alpha_2)T}}{\left[1 - \frac{\gamma z_1}{z_1 + \alpha_1 + \alpha_2} (1 - e^{-(z_1 + \alpha_1 + \alpha_2)T}) - \frac{(1 - \gamma) z_2}{z_2 + \alpha_1 + \alpha_2} (1 - e^{-(z_2 + \alpha_1 + \alpha_2)T}) \right]} + M \xi = C_1, \end{aligned} \quad (15)$$

где z_1, z_2 определены в (1); $M \xi$ – в (6).

Обозначим левую часть уравнения (15) через $f(T)$. Можно показать, что функция $f(T)$, $T \geq 0$, является возрастающей. Уравнение $f(T) = C_1$ решается численно на интервале $0 \leq T \leq \tau_{\min}$, $\tau_{\min} = \min\{\tau_k\}$, $k = \overline{1, n}$. Возможные ситуации:

1) $f(T=0) \geq C_1$; в качестве значения оценки $\hat{T}$ выбирается $\hat{T}^* = 0$;

2) $f(T=0) < C_1$, тогда возможны случаи:

а) корень уравнения моментов (15) попадает в полуинтервал $(0, \tau_{\min}]$, тогда он и выбирается в качестве значения оценки $\hat{T} = \hat{T}^*$;

б) корень уравнения (15) больше, чем $\tau_{\min}$, тогда в качестве значения оценки $\hat{T}$ выбирается $\hat{T}^* = \tau_{\min}$.

6. Результаты статистических экспериментов

Для получения статистики C_1 построена имитационная модель, выходом которой является последовательность $t_1, t_2, \dots$ моментов наступления событий.

Имитационная модель реализована на языке программирования C# с использованием технологии WPF (Windows Presentation Foundation) и библиотеки Math.NET Numerics. Результат работы имитационной модели приведен на рис. 2.

Для установления стационарного режима проведен следующий эксперимент.

ЭКСПЕРИМЕНТ 1. В данном эксперименте фиксируется число опытов N , которые производятся над имитационной моделью, а время моделирования T_m изменяется. Значения исходных данных для имитационной модели представлены в табл. 1.

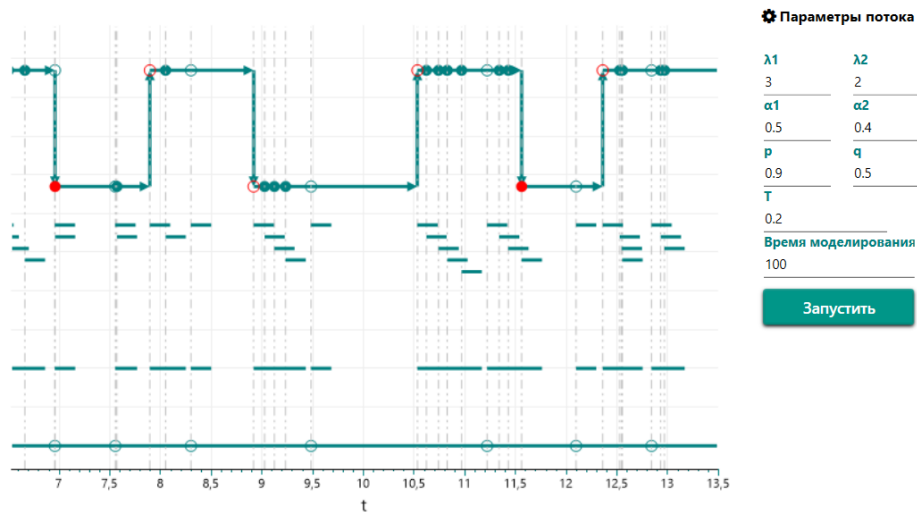


Рис. 2. Скриншот работы программы
Fig. 2 Screenshot of the program

Таблица 1

Исходные данные для эксперимента 1

$\lambda_1 = 0,8$	$\lambda_2 = 0,7$
$\alpha_1 = 0,1$	$\alpha_2 = 0,3$
$p = 0,5$	$q = 0,5$
$N = 1\,000$	$T = 0,1$

Численные результаты эксперимента 1, полученные в результате работы имитационной модели, приведены в табл. 2. В первой строке задается значение длительности времени моделирования T_m ; во второй строке – выборочное среднее значение длительности интервала между событиями в исходном потоке $\hat{M}(\hat{\tau}^\circ)$; в третьей строке – выборочное среднее значение длительности интервала между событиями в наблюдаемом потоке $\hat{M}(\hat{\tau})$; в четвертой строке – выборочное среднее значение длительности общего периода ненаблюдаемости $\hat{M}(\hat{\xi})$.

Таблица 2

Зависимость $\hat{M}(\hat{\tau})$, $\hat{M}(\hat{\tau}^\circ)$, $\hat{M}(\hat{\xi})$ от времени моделирования T_m

T_m	100	150	200	...	700	750	800	850	900	950	1 000
$\hat{M}(\hat{\tau}^\circ)$	1,1611	1,1802	1,1747	...	1,1792	1,1807	1,1767	1,1796	1,1742	1,1723	1,1758
$\hat{M}(\hat{\tau})$	1,2607	1,2851	1,2810	...	1,2858	1,2853	1,2806	1,2850	1,2888	1,2819	1,2825
$\hat{M}(\hat{\xi})$	0,1042	0,1043	0,1045	...	0,1045	0,1044	0,1043	0,1044	0,1044	0,1045	0,1043

Данные из табл. 2 демонстрируют сходимость оценок $\hat{M}(\hat{\tau})$, $\hat{M}(\hat{\tau}^\circ)$, $\hat{M}(\hat{\xi})$ к некоторым постоянным величинам. При времени моделирования $T_m \geq 800$ ед. времени устанавливается стационарный режим.

Для получения значения $\hat{T}$ оценки $\hat{T}$ длительности мёртвого времени T разработан следующий алгоритм оценивания:

- 1) задаются параметры потока с учетом условий $\lambda_1 > \lambda_2$, $\lambda_1 \lambda_2 = p q \alpha_1 \alpha_2$, $0 < p \leq 1$, $0 < q \leq 1$;
- 2) в течение T_m ед. времени работает имитационная модель потока для получения выборки $\tau_1, \tau_2, \dots, \tau_n$, находится $\tau_{\min} = \min\{\tau_k\}$, вычисляется статистика C_1 ;
- 3) табулируется функция $f(T)$ и численно устанавливается, что $f(T)$ является возрастающей функцией;

4) численно решается уравнение (15) методом Ньютона; корень уравнения моментов (15) является единственным;

5) в результате шагов 1–4 получим $\hat{T}_1^*$ – значение оценки $\hat{T}$ в первом опыте;

6) шаги 2–4 повторяются для N опытов и находятся значения оценок $\hat{T}_1^*, \hat{T}_2^*, \dots, \hat{T}_N^*$.

Выборочное среднее оценки и выборочная вариация оценки вычисляются по следующим формулам: $\hat{M}(\hat{T}) = \frac{1}{N} \sum_{j=1}^N \hat{T}_j^*$, $\hat{V}(\hat{T}) = \frac{1}{N} \sum_{j=1}^N (\hat{T}_j^* - T)^2$, где T – значение длительности мертвого времени, известное из имитационной модели потока.

Выборочное среднее количество событий исходного потока и выборочное среднее количество событий наблюдаемого потока вычисляются по формулам $\bar{n}_{\text{исх}} = \frac{1}{N} \sum_{j=1}^N n_{\text{исх}}^{(j)}$, $\bar{n}_{\text{набл}} = \frac{1}{N} \sum_{j=1}^N n_{\text{набл}}^{(j)}$, где $n_{\text{исх}}^{(j)}$ – количество событий исходного потока в j -м опыте, $n_{\text{набл}}^{(j)}$ – количество событий наблюдаемого потока в j -м опыте.

Для установления качества полученной оценки проведен следующий эксперимент.

ЭКСПЕРИМЕНТ 2. В данном эксперименте при одних и тех же параметрах потока меняется значение длительности мертвого времени T от $T = 0,1$ до $T = 1$ ед. времени.

Значения исходных данных представлены в табл. 3. Из эксперимента 1 время моделирования полагается равным $T_m = 800$ ед. времени.

Таблица 3

Исходные данные для эксперимента 2

$\lambda_1 = 5$	$\lambda_2 = 0,02$
$\alpha_1 = 1$	$\alpha_2 = 5$
$p = 0,1$	$q = 0,2$
$N = 1\,000$	$T_m = 800$

Табулируется функция $f(T)$. На рис. 3 представлен график функции $f(T)$ при значениях параметров, приведенных в табл. 3.

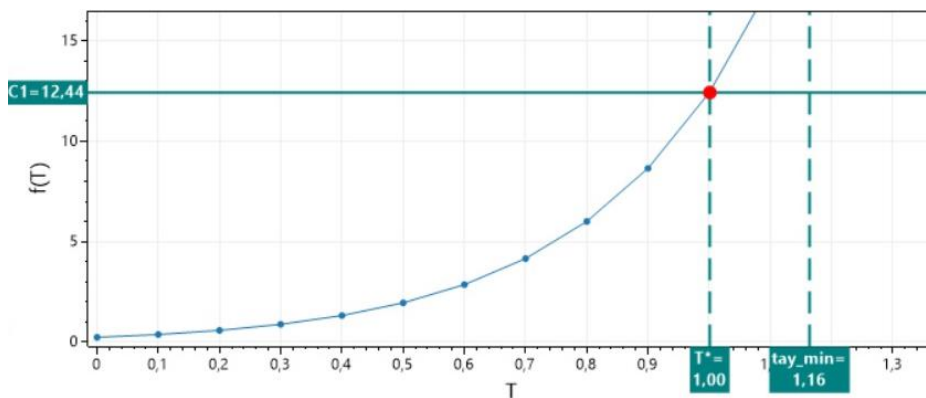


Рис. 3. График функции $f(T)$
Fig. 3. Graph of the function $f(T)$

Численные результаты работы алгоритма оценивания приведены в табл. 4. В первой строке табл. 4 задается значение длительности мертвого времени T ; во второй и третьей строках – выборочное среднее оценки $\hat{M}(\hat{T})$ и выборочная вариация $\hat{V}(\hat{T})$ соответственно; в четвертой и пятой строках – выборочное среднее количество событий исходного потока $\bar{n}_{\text{исх}}$ и выборочное среднее количество событий наблюдаемого потока $\bar{n}_{\text{набл}}$; в шестой строке – процент наблюдаемых событий $p = (\bar{n}_{\text{набл}} / \bar{n}_{\text{исх}}) 100\%$.

Зависимость $\hat{M}(\hat{T}), \hat{V}(\hat{T}), \bar{n}_{\text{исх}}, \bar{n}_{\text{набл}}, p$ от значения длительности мёртвого времени T

T	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1
$\hat{M}(\hat{T})$	0,0977	0,1995	0,2999	0,4001	0,4995	0,5996	0,6993	0,7987	0,8971	0,9735
$\hat{V}(\hat{T})$	$6 \cdot 10^{-6}$	$3 \cdot 10^{-6}$	$5 \cdot 10^{-6}$	$1,4 \cdot 10^{-5}$	$4,5 \cdot 10^{-5}$	$9,7 \cdot 10^{-5}$	$1,89 \cdot 10^{-4}$	$3,54 \cdot 10^{-4}$	$5,93 \cdot 10^{-4}$	$10,83 \cdot 10^{-4}$
$\bar{n}_{\text{исх}}$	3 470	3 470	3 469	3 466	3 479	3 470	3 469	3 464	3 471	3 469
$\bar{n}_{\text{набл}}$	2 144	1 372	904	608	415	290	201	141	100	71
p	61,79	39,53	26,05	17,55	11,97	8,36	5,8	4,08	2,87	2,04

Анализируя результаты, представленные в табл. 4, видим, что с ростом значения длительности мертвого времени число событий в наблюдаемом потоке уменьшается, выборочная вариация $\hat{V}(\hat{T})$ растёт, что является естественным.

Замечание. Оценки $\hat{T}$ являются состоятельными, так как: а) случайные величины $\tau_1, \tau_2, \dots, \tau_k$ являются независимыми и одинаково распределёнными; б) теоретический момент $M(\tau)$ существует; в) численно показано, что уравнение моментов (15) имеет единственное решение [19].

Заключение

В данной работе рассмотрен дважды стохастический обобщенный асинхронный поток случайных событий с двумя состояниями, функционирующий в стационарном режиме в условиях продлевающегося мертвого времени фиксированной длительности. Построена математическая модель данного потока. Получено преобразование Лапласа плотности вероятности значений длительности интервалов между соседними событиями в рекуррентном обобщенном асинхронном потоке (9), с использованием которого выведено уравнение моментов (15) для решения задачи оценивания длительности мертвого времени T . Поставлены статистические эксперименты на разработанной имитационной модели потока. Анализ полученных численных результатов позволяет утверждать, что полученные значения $\hat{T}$ оценок $\hat{T}$ длительности мертвого времени T имеют приемлемое качество в смысле выборочной вариации $\hat{V}(\hat{T})$; оценки $\hat{T}$ являются состоятельными.

Список источников

1. Вишневецкий В.М., Дудин А.Н., Клименок В.И. Стохастические системы с коррелированными потоками. Теория и применение в телекоммуникационных сетях. М. : Техносфера, 2018. 564 с.
2. Cox D.R. The analysis of non-Markovian stochastic processes by the inclusion of supplementary variables // Proc. of the Cambridge Philosophical Society. 1955. V. 51 (3). P. 433–441.
3. Kingman Y.F.C. On doubly stochastic Poisson process // Proc. of the Cambridge Philosophical Society. 1964. V. 60 (4). P. 923–930.
4. Башарин Г.П., Кокотушкин В.А., Наумов В.А. О методе эквивалентных замен расчёта фрагментов сетей связи. Ч. 1 // Известия АН СССР. Техническая кибернетика. 1979. № 6. С. 92–99.
5. Башарин Г.П., Кокотушкин В.А., Наумов В.А. О методе эквивалентных замен расчёта фрагментов сетей связи. Ч. 2 // Известия АН СССР. Техническая кибернетика. 1980. № 1. С. 55–61.
6. Neuts M.F. A versatile Markov point process // Journal of Applied Probability. 1979. V. 16. P. 764–779.
7. Lucantoni D.M. New results on the single server queue with a bath markovian arrival process // Communication in Statistics Stochastic Models. 1991. V. 7. P. 1–46.
8. Нежелская Л.А. Совместная плотность вероятностей длительности интервалов модулированного МАР-потока событий и условия рекуррентности потока // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2015. № 1 (30). С. 57–67.
9. Апанасович В.В., Коляда А.А., Чернявский А.Ф. Статистический анализ случайных потоков в физическом эксперименте. Минск : Университетское, 1988. 256 с.
10. Горцев А.М., Леонова М.А., Нежелская Л.А. Сравнение МП- и ММ-оценок длительности мертвого времени в обобщенном асинхронном потоке событий // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2013. № 4 (25). С. 32–42.

11. Горцев А.М., Калягин А.А., Нежелская Л.А. Оценка максимального правдоподобия длительности мертвого времени в обобщенном полусинхронном потоке // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2015. № 1 (30). С. 27–37.
12. Василевская Т.П., Горцев А.М., Нежелская Л.А. Оценивание длительности мертвого времени и параметров синхронного альтернирующего потока с проявлением либо не проявлением событий // Вестник Томского государственного университета. 2004. № S9-2. С. 129–138.
13. Горцев А.М., Ниссенбаум О.В. Оценивание длительности мертвого времени и параметров асинхронного альтернирующего потока событий с иницированием лишнего события // Вестник Томского государственного университета. 2004. № 284. С. 137–145.
14. Горцев А.М., Веткина А.В. Оценивание методом моментов параметра равномерного распределения длительности непродлевающегося случайного мертвого времени в рекуррентном полусинхронном потоке событий в общем и особом случаях // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2022. № 61. С. 47–60. doi: 10.17223/19988605/61/5
15. Нежелская Л.А., Пономаренко В.Д. Имитационное моделирование обобщенного асинхронного потока событий с продлевающимся мертвым временем // Труды Томского государственного университета. Сер. физико-математическая. 2023. Т. 308. С. 27–35.
16. Нежелская Л.А. Оценка состояний дважды стохастических потоков событий : учеб. пособие. Томск : Изд-во Том. гос. ун-та, 2020. 210 с.
17. Нежелская Л.А. Оценка состояний и параметров дважды стохастических потоков событий : дис. ... д-ра физ.-мат. наук. Томск, 2016. 341 с.
18. Хинчин А.Я. Работы по математической теории массового обслуживания. М. : Физматгиз, 1963. 236 с.
19. Малинковский Ю.В. Теория вероятностей и математическая статистика. Гомель : ГГУ им. Ф. Скорины, 2004. Ч. 2: Математическая статистика. 146 с.

References

1. Vishnevskiy, V.M., Dudin, A.N. & Klimenok, V.I. (2018) *Stokhasticheskie sistemy s korrelirovannymi potokami. Teoriya i primeneniye v telekommunikatsionnykh setyakh* [Stochastic Systems with Correlated Flows. Theory and Application in Telecommunication Networks]. Moscow: Tekhnosfera.
2. Cox, D.P. (1955) The analysis of non-Markovian stochastic process by the inclusion of supplementary variables. *Proceeding of the Cambridge Philosophical Society*. 51(3). pp. 433–441.
3. Kingman, Y.F.C. (1964) On double stochastic Poisson process. *Proceeding of the Cambridge Philosophical Society*. 60(4). pp. 923–930.
4. Basharin, G.P., Kokotushkin, V.A. & Naumov, V.A. (1979) O metode ekvivalentnykh zamen rascheta fragmentov setey svyazi. Ch. 1 [On the equivalent substitutions method for computing fragments of communication networks. Part 1]. *Izvestiya AN SSSR. Tekhnicheskaya kibernetika*. 6. pp. 92–99.
5. Basharin, G.P., Kokotushkin, V.A. & Naumov, V.A. (1980) O metode ekvivalentnykh zamen rascheta fragmentov setey svyazi. Ch. 2 [On the equivalent substitutions method for computing fragments of communication networks. Part 2]. *Izvestiya AN SSSR. Tekhnicheskaya kibernetika*. 1. pp. 55–61.
6. Neuts, M.F. (1979) A versatile Markov point process. *Journal of Applied Probability*. 16. pp. 764–779.
7. Lucantoni, D.M. (1991) New results on the single server queue with a bath Markovian arrival process. *Communication in Statistics Stochastic Models*. 7. pp. 1–46.
8. Nezhe'skaya, L.A. (2015) Joint probability density of the intervals' duration in modulated MAP event flows and its recurrence conditions. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 1(30). pp. 57–67.
9. Apanasovich, V.V., Kolyada, A.A. & Chernyavskiy, A.F. (1988) *Statisticheskiy analiz sluchaynykh potokov v fizicheskom eksperimente* [Statistical Analysis of Random Flows in Physical Experiment]. Minsk: Universitetskoe.
10. Gortsev, A.M., Leonova, M.A. & Nezhe'skaya, L.A. (2013) The comparison of maximum likelihood estimation and method of moments estimation of dead time value in a generalized asynchronous flow of events. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 4(25). pp. 32–42.
11. Gortsev, A.M., Kalyagin, A.A. & Nezhe'skaya, L.A. (2015) Maximum likelihood estimation of dead time at a generalized semisynchronous flow of events. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 1(30). pp. 27–37.
12. Vasilevskaya, T.P., Gortsev, A.M. & Nezhe'skaya, L.A. (2004) Estimation of dead time duration and parameters of synchronous alternating flow with manifestation or non-manifestation of events. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. S9–2. pp. 129–138.
13. Gortsev, A.M. & Nissenbaum, O.V. (2004) Estimation of the dead time duration and parameters of the asynchronous alternating stream of events with initialization of an extra event. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 284. pp. 137–145.

14. Gortsev, A.M. & Vetkina, A.V. (2022) Estimation by the method of moments of the parameter of the uniform distribution of the duration of non-transient random dead time in the recurrent semi-synchronous flow of events in general and special cases. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 61. pp. 47–60. DOI: 10.17223/19988605/61/5
15. Nezhel'skaya, L.A. & Ponomarenko, V.D. (2023) Simulation model of the generalized asynchronous flow of events with prolonging dead time. *Trudy Tomskogo gosudarstvennogo universiteta. Ser. fiziko-matematicheskaya*. 308. pp. 27–35.
16. Nezhel'skaya, L.A. (2020) *Otsenka sostoyaniy dvazhdy stokhasticheskikh potokov sobytiy* [Estimation of States of Doubly Stochastic Flows of Events]. Tomsk: Tomsk State University.
17. Nezhel'skaya, L.A. (2016) *Otsenka sostoyaniy i parametrov dvazhdy stokhasticheskikh potokov sobytiy* [Evaluation of states and parameters of doubly stochastic flows of events]. Physics and Mathematics Dr. Diss. Tomsk.
18. Khinchin, A.Ya. (1963) *Raboty po matematicheskoy teorii massovogo obsluzhivaniya* [Works on the Mathematical Theory of Queuing]. Moscow: Fizmatgiz
19. Malinkovsky, Yu.V. (2004) *Teoriya veroyatnostey i matematicheskaya statistika* [Theory of Probabilities and Mathematical Statics]. Vol. 2. Gomel: GSU.

Информация об авторах:

Нежелская Людмила Алексеевна – профессор, доктор физико-математических наук, заведующая кафедрой прикладной математики Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: ludne@mail.ru

Пonomarenko Валентина Денисовна – магистрант кафедры прикладной математики Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: valya.ponomarenko.00@mail.ru

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

Information about the authors:

Nezhel'skaya Lyudmila A. (Doctor of Physical and Mathematical Sciences, Professor, Institute of Applied Mathematics and Computer Science, National Research Tomsk State University, Tomsk, Russia Federation). E-mail: ludne@mail.ru

Ponomarenko Valentina D. (Master's Student, Department of Applied Mathematics, Institute of Applied Mathematics and Computer Science, National Research Tomsk State University, Tomsk, Russia Federation). E-mail: valya.ponomarenko.00@mail.ru

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

Поступила в редакцию 08.09.2024; принята к публикации 02.12.2024

Received 08.09.2024; accepted for publication 02.12.2024

Научная статья

УДК 519.246.5

doi: 10.17223/19988605/69/10

**Интервальное распознавание больших выбросов в экосистеме
по наблюдениям с фотоловушек****Гурами Шалвович Цициашвили¹, Владимир Николаевич Бочарников²**¹ *Институт прикладной математики Дальневосточного отделения Российской академии наук,
Владивосток, Россия, gurat@iam.dvo.ru*² *Тихоокеанский институт географии Дальневосточного отделения Российской академии наук,
Владивосток, Россия, vbocharnikov@mail.ru*

Аннотация. Исследуется связь временного ряда появления у фотоловушек неразличимых животных с временным рядом характеристик внешней среды. Задача решается интервальным распознаванием образов и оценкой коэффициента корреляции между рядами при наличии мешающего параметра, характеризующего число подходов животного к фотоловушке. Устранение мешающего параметра при анализе этой связи производится с помощью усиленного закона больших чисел.

Ключевые слова: корреляция; интервальное распознавание; закон больших чисел; мешающий параметр.

Благодарности: Работа выполнена в рамках государственного задания ИПМ ДВО РАН (№ 075-00459-24-00).

Для цитирования: Цициашвили Г.Ш., Бочарников В.Н. Интервальное распознавание больших выбросов в экосистеме по наблюдениям с фотоловушек // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 69. С. 95–102. doi: 10.17223/19988605/69/10

Original article

doi: 10.17223/19988605/69/10

**Interval detection of large emissions in the ecosystem
based on observations from camera traps****Gurami Sh. Tsitsiashvili¹, Vladimir N. Bocharnikov²**¹ *Institute for Applied Mathematics, Far Eastern Branch of the Russian Academy of Sciences,
Vladivostok, Russian Federation, gurat@iam.dvo.ru*² *Pacific Institute of Geography, Far Eastern Branch of the Russian Academy of Sciences,
Vladivostok, Russian Federation, vbocharnikov@mail.ru*

Abstract. The relationship between the time series of the appearance of indistinguishable animals in camera traps is investigated with a time series of environmental characteristics. The problem is solved by interval pattern recognition and estimation of the correlation coefficient between the rows in the presence of an interfering parameter characterizing the number of approaches of the animal to the camera trap. The elimination of the interfering parameter in the analysis of this relationship is carried out using the strong law of large numbers.

Keywords: correlation; interval recognition; law of large numbers; interfering parameter.

Acknowledgments: The study was performed within the framework of the state assignment of the Institute for Applied Mathematics FEB RAS (No. 075-00459-24-00).

For citation: Tsitsiashvili, G.Sh., Bocharnikov, V.N. (2024) Interval detection of large emissions in the ecosystem based on observations from camera traps. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 69. pp. 95–102. doi: 10.17223/19988605/69/10

Введение

В работе рассматривается задача интервального распознавания больших выбросов некоторых характеристик X в экосистеме по наблюдениям Y с фотоловушек. Здесь ряд $X = \{x_1, \dots, x_s\}$ характеризует, например, частоту подходов копытных животных к солонцам, необходимым для коррекции биохимического состава организмов оленей, лосей, косуль и других видов. Можно брать данное обозначение в более широком смысле, например как визуальную регистрацию крупных хищников (медведей, тигров, леопардов, волков) при их посещении населенных пунктов, при встречах вблизи дорог или фиксацию следов жизнедеятельности и останков жертв при нападениях на домашних животных. Есть и другие причины возникновения конфликтов между человеком и дикими животными. Такие подходы могут сигнализировать о нехватке пищи для крупных хищников и, как следствие, нарушении некоторых балансов в экосистеме, приводящих к ее неустойчивости [1–3]. Задача данной работы была предопределена также имеющимися работами по математическому моделированию состояния лесных экосистем [4], динамике численности популяций кабана [5] и исследованию биологии кабарги [6].

Что касается наблюдений с фотоловушек $Y = \{y_1, \dots, y_s\}$, то они свидетельствуют о среднемесечном подходе к фотоловушке животных (изюбрей, оленей, косуль, кабанов), служащих кормом для крупных хищников. Однако подсчет общего числа подошедших животных не может быть сделан без серьезных ошибок, если выполнять его по частоте регистраций фотоловушкой. Причина в том, что невозможно доказать различие копытных животных, следовательно, один и тот же зверь может регистрироваться несколько раз и учитываться как разные животные.

Поэтому между определенными с помощью фотоловушек частотой (средним числом наблюдаемых подходов указанных животных) и средним числом $Z = \{z_1, \dots, z_s\}$ самих подошедших животных существует некоторая связь, включающая в себя неизвестный (назовем его мешающим) параметр. Не помогают в этих расчетах и фиксированные сведения по дате и времени пребывания, что также фиксируется фотоловушкой. Необходимо убирать мешающий параметр и строить данные по встречам животных как относительное число, фактически совпадающее с общим числом регистраций всех животных по месяцам. Эти параметры выражают средние значения суточных и сезонных посещений животными солонцов [7–9]. Отсутствие такого параметра ставит задачу интервального распознавания больших выбросов в ряду наблюдений X по реальным наблюдениям с фотоловушек Z . Кроме того, если имеются дополнительные наблюдения за экосистемой $W = \{w_1, \dots, w_s\}$, то возникает задача интервального распознавания в наблюдениях Z по наблюдениям W . В частности, такая задача появляется, если в качестве наблюдений W берутся широко распространенные в метеорологии ряды наблюдений за давлением H_{500} на высоте 5 км.

1. Интервальное распознавание больших выбросов в наблюдениях X по наблюдениям Y

1.1. Алгоритм интервального распознавания по одному признаку

Пусть задано положительное число p , $0 < p < 1$, по которому определяется $r = [ps]$ (здесь $[a]$ – целая часть вещественного числа a). Обозначим $\bar{x} = x'_r$. Рассмотрим теперь алгоритм интервального распознавания образов больших выбросов ($x_j \geq \bar{x}$) в ряду X по наблюдениям Y [10]. Для этого сначала упорядочим по величине, начиная с наименьшего, ряд X , преобразовав его в соответствующий вариационный ряд, где $x'_1 \leq \dots \leq x'_s$,

$$x'_1 = \min_{1 \leq j \leq s} x_j, \quad x'_s = \max_{1 \leq j \leq s} x_j, \quad x'_j = \max_{1 \leq i_1 < i_2 < \dots < i_{n-j+1} \leq s} \min(x_{i_1}, x_{i_2}, \dots, x_{i_{n-j+1}}), \quad j = 2, \dots, s-1. \quad (1)$$

Зададим множество индексов $K = \{k: x_k \geq \bar{x}\}$. Определим $|K|$ – число элементов множества K , и вычислим $\underline{y}$, $\bar{y}$ по формулам

$$\underline{y} = \min_{k \in K} y_k, \quad \bar{y} = \max_{k \in K} y_k. \quad (2)$$

Будем распознавать принадлежность $j \in K$ по следующему правилу:

$$\underline{y} \leq y_j \leq \bar{y}, \quad j = 1, \dots, s. \quad (3)$$

Очевидно, что все $k \in K$ будут распознаваться правильно. Однако, если $j \notin K$, то могут быть ошибки, когда $j \notin K$ распознается как $j \in K$. Обозначим $|\bar{K}|$ число значений $j \notin K$, которые правильно распознаются, т.е. распознаются как $j \in K$. Тогда отношение

$$\rho = \frac{|K| + |\bar{K}|}{s} \quad (4)$$

характеризует качество интервального распознавания.

1.2. Алгоритм интервального распознавания по нескольким признакам

Предположим, что имеются наблюдения с нескольких фотоловушек $Y_i = \{y_{1i}, \dots, y_{si}\}$, $i = 1, \dots, I$. Для каждого из этих наблюдений вычислены

$$\underline{y}_i = \min_{k \in K} y_{ki}, \quad \bar{y}_i = \max_{k \in K} y_{ki}. \quad (5)$$

Будем распознавать принадлежность $j \in K$ по правилу

$$\underline{y}_i \leq y_{ji} \leq \bar{y}_i, \quad i = 1, \dots, I. \quad (6)$$

Если по-прежнему $|\bar{K}|$ является числом правильно распознаваемых индексов $j \notin K$, то формула (4) будет характеризовать качество интервального распознавания по I признакам. Очевидно, что с расширением множества признаков качество распознавания будет увеличиваться / не уменьшаться.

Такой способ интервального распознавания больших выбросов требует числа арифметических операций, пропорциональных sI , т.е. линейно зависит от длины ряда s и от числа признаков I . Метод интервального распознавания образов многократно проверялся на реальных наблюдениях зависимости больших выбросов основного признака X от сопутствующих признаков I . Причем проверка производилась на коротких рядах наблюдений, когда длина ряда наблюдений n невелика, а число сопутствующих признаков I велико [11].

2. Статистическая оценка интенсивности входного потока при наличии мешающего параметра

Рассмотрим теперь связь между наблюдениями за признаком Y , характеризующим среднее число подходов животных к фотоловушке, и признаком Z , характеризующим среднее число животных, подошедших к фотоловушке. Для этого сначала предположим, что на полуинтервале $[0, n)$ задан пуассоновский поток точек интенсивностью λ . Эти точки определяют моменты прихода заявок в систему. Разобьем полуинтервал $[0, n)$ на полуинтервалы $[0, 1), [1, 2), \dots, [n-1, n)$. Обозначим $\xi_1, \dots, \xi_n$ независимые и одинаково распределенные случайные величины, определяющие число точек потока в выделенных полуинтервалах. Пусть теперь каждая точка входного пуассоновского потока, попавшая в полуинтервал $[i-1, i)$, порождает в полуинтервалах $[i, i+1), \dots, [i+m-1, i+m)$ по одной новой точке (моменты фиксации прибором заявки). Тогда в полуинтервале $[0, n)$ случайное число точек пуассоновского потока и случайное число точек, порождаемых точками пуассоновского потока, определяются равенствами

$$C_n(\lambda) = \sum_{t=1}^n \xi_t, \quad D_n(\lambda) = \sum_{t=1}^m t \xi_{n-t+1} + \sum_{t=m+1}^n m \xi_{n-t+1}. \quad (7)$$

Значком $\Rightarrow$ обозначаем сходимость с вероятностью единица, а значком $\xrightarrow{P}$ обозначаем сходимость по вероятности при $n \rightarrow \infty$.

Утверждение 1. Пусть при некотором положительном M выполняется неравенство $m \leq M$, тогда

$$R_n(\lambda) = \frac{C_n(\lambda)}{n} \Rightarrow \lambda, \quad S_n(\lambda) = \frac{D_n(\lambda)}{n} \Rightarrow m\lambda, \quad n \rightarrow \infty. \quad (8)$$

Доказательство. Используя усиленный закон больших чисел (см., напр.: [12. Гл. 8, 4]), получаем сходимость с вероятностью единица при $n \rightarrow \infty$:

$$\frac{C_n(\lambda)}{n} \Rightarrow \lambda, \quad \frac{1}{n} \sum_{k=1}^m k \xi_{n-k+1} \Rightarrow 0, \quad \frac{1}{n-m} \sum_{k=m+1}^n m \xi_{n-k+1} \Rightarrow m\lambda,$$

откуда автоматически следуют формулы (8).

Замечание 1. Наличие линейной зависимости между средним (за n отрезков времени, например дней) числом животных, подошедших к фотоловушке, и средним числом подходов животных к фотоловушкам даже при отсутствии информации о коэффициенте пропорциональности m может быть использовано при анализе всплесков численности животных. Для этого удобно воспользоваться методом интервального распознавания образов.

3. Соединение интервального распознавания с устранением мешающего параметра

Для интервального распознавания можно воспользоваться формулами (3) (при наличии нескольких признаков – формулами (5), (6)). Предположим, что имеется s пуассоновских потоков неизвестной нам интенсивности y_j , $j = 1, \dots, s$, и пусть $z_j = S_n(y_j)$. Тогда из утверждения 1 и формул (1), (8) следуют предельные соотношения

$$\min_{k \in K} z_k \Rightarrow m \min_{k \in K} y_k, \quad \max_{k \in K} z_k \Rightarrow m \max_{k \in K} y_k, \quad n \rightarrow \infty. \quad (9)$$

В этом смысле замена $y_j \in Y$ на $z_j \in Z$ приводит с вероятностью единица к тому же результату интервального распознавания при $n \rightarrow \infty$. Действительно, вследствие формулы (9) принадлежность $j \in K$ означает, что с вероятностью единица

$$\min_{k \in K} \lim_{n \rightarrow \infty} z_k = m \min_{k \in K} y_k \leq \lim_{n \rightarrow \infty} z_j = m y_j \leq \max_{k \in K} \lim_{n \rightarrow \infty} z_k = \max_{k \in K} m y_k.$$

Аналогичное соотношение можно привести для интервального распознавания при наличии нескольких признаков. В свою очередь, для определения погрешностей этой замены можно воспользоваться вместо утверждения 1 следующим утверждением.

Утверждение 2. Пусть при некоторых положительных M, Λ выполняются неравенства $m \leq M, \lambda \leq \Lambda$, тогда для любого $\varepsilon > 0$ выполняется неравенство

$$P(|S_n(\lambda) - m\lambda| > \varepsilon) \leq \frac{M^2 \Lambda}{n \varepsilon^2} \left(1 + \frac{2 + \Lambda^2 M^2}{4n} \right). \quad (10)$$

Доказательство. Обозначим

$$A(m, n) = \frac{\alpha(m)}{n}, \quad \alpha(m) = \frac{1-m}{2}; \quad B(m, n) = \frac{\beta(m)}{n}, \quad \beta(m) = \frac{1}{2} + \frac{1}{6m} - \frac{2m}{3}.$$

Тогда при $m \geq 1$ выполняются неравенства

$$\alpha^2(m) \leq \frac{m^2}{4}, \quad \beta(m) \leq \frac{1}{2}.$$

Вычислим у случайной величины $D_n(\lambda)$ математическое ожидание и дисперсию:

$$M(D_n(\lambda)) = mn\lambda(1 + A(m, n)), \quad D(D_n(\lambda)) = m^2 n \lambda (1 + B(m, n)).$$

Отсюда следуют равенства

$$M(S_n(\lambda)) = m\lambda(1 + A(m, n)), \quad D(S_n(\lambda)) = \frac{m^2\lambda}{n}(1 + B(m, n)).$$

Используя эти равенства, получаем неравенство

$$M(S_n(\lambda) - m\lambda)^2 = D(S_n(\lambda)) + (M(S_n(\lambda)) - m\lambda)^2 = \frac{m^2\lambda}{n}(1 + B(m, n) + \lambda^2 n A^2(m, n)) \leq \frac{m^2\lambda}{n} \left(1 + \frac{2 + \Lambda^2 m^2}{4n}\right),$$

с помощью которого получается формула (10).

Замечание 2. Конечно, такая оценка является достаточно грубой, так как величины m, λ заменяются на верхние оценки M, Λ . Но такова цена за устранение мешающего параметра m .

4. Интервальное распознавание больших выбросов в наблюдениях Z по наблюдениям W

Предполагается, что может быть несколько, как правило, неизвестное число причин, которые определяют количественные показатели, полученные в таких наблюдениях. Например, при посещениях копытными зверями существенное влияние могут оказывать различные погодные условия, а также качественные характеристики тех биотопов, где расположены фотоловушки, воздерживаются ли животные от посещения таких мест, если обнаружат присутствие человека или опасных для них хищников [13, 14].

Изложенный в работе метод интервального распознавания предполагает, что наблюдение X характеризует всю экосистему и потому может быть названо основным признаком. В свою очередь, наблюдение Z естественно назвать сопутствующим признаком. Однако изложенный в предыдущем разделе метод соединения интервального распознавания с устранением мешающего параметра может быть рассмотрен и по-другому. Например, можно в качестве основного параметра взять наблюдение Z за частотой прихода животных к фотоловушкам, а в качестве сопутствующего признака рассмотреть, например, метеорологические наблюдения. В этом случае необходимо задаться частотой превышения некоторого критического уровня [15] и по ней определить множество K тех наблюдений, в которых этот критический уровень превышает. А далее следует воспользоваться оценками интервального распознавания из предыдущего раздела.

Для того чтобы построить такую процедуру интервального распознавания, нужно сначала рассмотреть интервальное распознавание больших выбросов в наблюдениях X по наблюдениям W . Для этого по заданному p можно определить аналог L множества K , но уже для ряда Z . Преобразуем ряд $Z = \{z_1, \dots, z_s\}$ в соответствующий ему вариационный ряд $z'_1 \leq \dots \leq z'_s$. Пусть $\bar{z} = z'_r$, и зададим множество индексов $L = \{l : z_l \geq \bar{z}\}$. Тогда справедливо предельное соотношение

$$\min_{l \in L} z_l \Rightarrow m \min_{k \in K} y_k, \quad \max_{l \in L} z_l \Rightarrow m \max_{k \in K} y_k, \quad n \rightarrow \infty,$$

причем $z_l \Rightarrow m y_l$, $l \in L$, и значит в пределе при $n \rightarrow \infty$ можно полагать $L = K$. Далее выделим в ряду Z множество элементов $\{z_l : l \in L\}$ и уже по ним построим процедуру интервального распознавания по наблюдениям W , полагая

$$\underline{w} = \min_{j \in L} w_j, \quad \bar{w} = \max_{j \in L} w_j$$

и определяя принадлежность $j \in L$ условием $\underline{w} \leq w_j \leq \bar{w}$. В этом смысле замена $y_j \in Y$ на $z_j \in Z$ приводит с вероятностью единица к тому же результату интервального распознавания по наблюдениям Z при $n \rightarrow \infty$. Аналогичная процедура может быть построена для распознавания по нескольким признакам $(w_{1i}, \dots, w_{si})$, $i = 1, \dots, I$.

5. Оценка коэффициента корреляции по неточным данным

Наряду с методом интервального распознавания можно также использовать оценки коэффициента корреляции, который часто применяется при обработке метеорологических данных [16]. Предположим, что рассматривается коэффициент корреляции между случайными величинами X, Z . Однако вместо случайной величины X берется случайная величина $Y = XK(1 + \varepsilon + \gamma)$. Здесь ε – случайная величина с нулевым средним и конечной дисперсией: $M(\varepsilon) = 0$, $D(\varepsilon) = \sigma^2$, не зависящая от случайного вектора (X, Y) , а детерминированная величина $|\gamma| < 1$. Далее будем предполагать, что $|\gamma(T)| \rightarrow 0$, $\sigma^2(T) \rightarrow 0$, $T \rightarrow \infty$. Это предположение оправдывается соотношениями, полученными в [10], если отрезок времени, по которому определяются случайные величины X, Y, Z достаточно велик. Причем коэффициент K неизвестен и может считаться мешающим параметром. Вычислим теперь коэффициент корреляции между случайными величинами Y, Z :

$$\rho(Y, Z) = \frac{\text{cov}(Y, Z)}{\sqrt{D(Y) \cdot D(Z)}}, \quad \text{cov}(Y, Z) = M(YZ) - M(Y)M(Z).$$

Пересчитаем дисперсию $D(Y)$ и ковариацию $\text{cov}(Y, Z)$ через дисперсию $D(X)$ и ковариацию $\text{cov}(X, Z)$:

$$\text{cov}(Y, Z) = \text{cov}(XK(1 + \varepsilon + \gamma), Z) = M(XK(1 + \varepsilon + \gamma)Z) - M(XK(1 + \varepsilon + \gamma))M(Z) = K(1 + \gamma)\text{cov}(X, Z),$$

$$\begin{aligned} D(Y) &= K^2 D(X(1 + \varepsilon + \gamma)) = K^2 \left[M(X^2(1 + \varepsilon + \gamma)^2) - (M(X(1 + \varepsilon + \gamma)))^2 \right] = \\ &= K^2 (1 + \gamma)^2 D(X) \left[1 + \frac{\sigma^2 M(X^2)}{(1 + \gamma)^2 D(X)} \right]. \end{aligned}$$

Отсюда следует формула

$$\rho(Y, Z) = \rho(X, Z)(1 + \delta)^{-1/2}, \quad \delta = \frac{\sigma^2 M(X^2)}{(1 + \gamma)^2 D(X)} > 0.$$

Для дальнейших оценок можно воспользоваться рядом Тейлора, справедливым при $|\delta| < 1$:

$$(1 + \delta)^{-1/2} = 1 + \sum_{k=1}^{\infty} (-1)^k \delta^k \frac{(2k-1)!!}{2^k k!}, \quad (2k-1)!! = 1 \cdot 3 \cdot \dots \cdot (2k-1), \quad k = 1, \dots$$

Отсюда при $0 < \delta < 1$ следует неравенство

$$1 > \frac{\rho(Y, Z)}{\rho(X, Z)} > 1 - \delta/2.$$

Пользуясь разложением в ряд Тейлора функции $(1 + \delta)^{-1/2}$, можно получать при $\delta \ll 1$ и более точные оценки. Точность оценки $\rho(Y, Z)$ тем выше, чем больше отрезок длины T , и определяется условиями $|\gamma(T)|, \sigma^2(T) \rightarrow 0$, $T \rightarrow \infty$. Все эти оценки базируются на предположении независимости случайной величины ε от случайного вектора (X, Z) и позволяют избавиться от неизвестного параметра K , который можно полагать мешающим. Предположение о независимости случайной величины ε и случайного вектора (X, Y) можно некоторым образом ослабить.

Заключение

Основная идея данной работы состоит в соединении интервального распознавания с устранением мешающего параметра при определении интенсивности потока животных, приходящих к фотоловушке / фотоловушкам. Это позволяет предложить использование математической процедуры для устранения мешающего параметра при обработке больших данных, полученных дистанционно с фотоловушек или иных средств автоматического обнаружения, а также при обработке данных по конфликтам человека и опасных животных, во многих исследовательских процедурах зоологов, использующих количественные данные.

Список источников

1. Разжевайкин В.Н. Принцип эволюционной оптимальности как инструмент моделирования структурированных биологических систем // Журнал общей биологии. 2010. Т. 71, № 1. С. 75–84.
2. Разжевайкин В.Н. Об экологической стабильности и принципе Гаузе // Исследование операций (модели, системы, решения). М. : ВЦ РАН, 2011. С. 22–29.
3. Разжевайкин В.Н. Анализ моделей динамики популяций : учеб. пособие. М. : МФТИ (ГУ), 2010. 174 с.
4. Апалькова Т.Г., Белотелов Н.В., Ольчев А.В. Использование моделей временных рядов для анализа взаимосвязи эксергии и продукционных процессов в лесных экосистемах // Моделирование, декомпозиция и оптимизация сложных динамических процессов. 2018. Т. 33, № 1. С. 191–202.
5. Белотелов Н.В., Назарова В.М., Зайцев В.А. Модель популяции кабана // Моделирование, декомпозиция и оптимизация сложных динамических процессов. 2020. Т. 35, № 1. С. 152–165.
6. Зайцев В.А., Максимова Д.А., Смирнов Ю.В., Белотелов Н.В. Использование участка обитания самцом кабарги (*Moschus moschiferus* L.) в Центральном Сихотэ-Алине // Зоологический журнал. 2021. Т. 100, № 4. С. 462–480.
7. Лукаревский В.С., Лукаревский С.В. Оценка численности дальневосточного леопарда (*PANTHERA PARDUS*) в России // Зоологический журнал. 2019. Т. 98, № 5. С. 567–577.
8. Огурцов С.С. Обзор программного обеспечения для обработки данных с фотоловушек: последние новинки, работа с видео и ГИС // Nature Conservation Research. Заповедная наука. 2019. Т. 4, № 2. С. 95–124.
9. Калинин Ю.Н. Суточная активность копытных на солонцах Алтайского заповедника. Полевые исследования в Алтайском биосферном заповеднике // Зоология. 2023. Т. 5. С. 6–14.
10. Цициашвили Г.Ш., Бочарников В.Н. Две естественнонаучные задачи статистической оценки при наличии мешающего параметра // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2023. Т. 65, № 4. С. 89–94. doi: 10.17223/19988605/65/9
11. Tsitsiasvili G.Sh. Processing Large Outliers in Arrays of Observations // Mathematics. 2022. V. 10 (18). Art. 3399.
12. Боровков А.А. Курс теории вероятностей. М. : Наука, 1972. 288 с.
13. Данилкин А.А. Динамика населения диких опытных России: гипотезы, факторы, закономерности. М. : Т-во научных изданий КМК, 2009. 310 с.
14. Калинин Ю.Н. Опыт использования данных с автоматических фотокамер на солонцах для оценки состояния группировки благородного оленя *Cervus elaphus sibiricus* в Алтайском заповеднике // Вестник охотоведения. 2019. Т. 16, № 2. С. 111–118.
15. Shatilina T.A., Moroz V.V., Tsitsiasvili G.Sh., Radchenkova T.V. Formation of Large Anomalies in the Thermal Conditions of Waters on the Western and Eastern Shelf of Sakhalin Island // Physical Oceanography. 2024. V. 31 (1). P. 33–45.
16. Малинин В.Н. Статистические методы анализа гидрометеорологической информации : учебник. СПб. : Изд-во РГГМУ, 2008. 408 с.

References

1. Razzhevaykin, V.N. (2010) Printsip evolyutsionnoy optimal'nosti kak instrument modelirovaniya strukturirovannykh biologicheskikh sistem [The principle of evolutionary optimality as a tool for modeling structured biological systems]. *Zhurnal obshchestvennoy biologii – Journal of General Biology*. 71(1). pp. 75–84.
2. Razzhevaykin, V.N. (2011) Ob ekologicheskoy stabil'nosti i printsipe Gauze [On environmental stability and the Gause principle]. In: Abramov, A.P. (ed.) *Issledovanie operatsiy (modeli, sistemy, resheniya)* [Operations research (models, systems, solutions)]. Moscow: RAS. pp. 22–29.
3. Razzhevaykin, V.N. (2010) *Analiz modeley dinamiki populyatsiy* [Analysis of models of population dynamics]. Moscow: MFTI (GU).
4. Apalkova, T.G., Belotelov, N.V. & Olchev, A.V. (2018) Ispol'zovanie modeley vremennykh ryadov dlya analiza vzaimosvyazi ek-sergii i produktsionnykh protsessov v lesnykh ekosistemakh [Using time series models to analyze the relationship between exergy and production processes in forest ecosystems]. *Modelirovanie, dekompozitsiya i optimizatsiya slozhnykh dinamicheskikh protsessov – Modeling, Decomposition and Optimization of Complex Dynamic Processes*. 33(1). pp. 191–202.
5. Belotelov, N.V., Nazarova, V.M. & Zaitsev, V.A. (2020) Model' populyatsii kabana [Boar population model]. *Modelirovanie, dekompozitsiya i optimizatsiya slozhnykh dinamicheskikh protsessov – Modeling, Decomposition and Optimization of Complex Dynamic Processes*. 35(1). pp. 152–165.
6. Zaytsev, V.A., Maksimova, D.A., Smirnov, Yu.V. & Belotelov, N.V. (2021) Ispol'zovanie uchastka obitaniya samtsom kabargi (*Moschus Moschiferus* L.) v Tsentral'nom Sikhote-Aline [The use of a habitat by a male musk deer (*Moschus Moschiferus* L.) in the Central Sikhote-Alin]. *Zoologicheskii zhurnal – Zoological Journal*. 100(4). pp. 462–480.
7. Lukarevsky, V.S. & Lukarevsky S.V. (2019) Otsenka chislenosti dal'nevostochnogo leoparda (*PANTHERA PARDUS*) v Rossii [Estimation of the number of the Far Eastern leopard (*PANTHERA PARDUS*) in Russia]. *Zoologicheskii zhurnal – Zoological Journal*. 98(5). pp. 567–577.
8. Ogurtsov, S.S. (2019) Obzor programmnoy obespecheniya dlya obrabotki dannykh s fotolovushek: poslednie novinki, rabota s video i GIS [Review of software for processing data from camera traps: The latest innovations, working with video and GIS]. *Nature Conservation Research*. 4(2). pp. 95–124.

9. Kalinkin, Yu.N. (2023) Sutochnaya aktivnost' kopytnykh na solontsakh Altayskogo zapovednika. Polevye issledovaniya v Al-tayskom biosfernom zapovednike [Daily activity of ungulates in the salt lakes of the Altai Reserve. Field studies in the Altai Biosphere Reserve]. *Zoologiya – Zoology*. 5. pp. 6–14.
10. Tsitsiashvili, G.Sh. & Bocharnikov, V.N. (2023) Two natural science tasks of statistical assessment in the presence of an interfering parameter. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 62(4). pp. 89–94. DOI: 10.17223/19988605/65/9
11. Tsitsiashvili, G.Sh. (2022) Processing Large Outliers in Arrays of Observations. *Mathematics*. 10(18). Art. 3399.
12. Borovkov, A.A. (1972) *Kurs teorii veroyatnostey* [Probability Theory]. Moscow: Nauka.
13. Danilkin, A.A. (2009) *Dinamika naseleniya dikikh opytnykh Rossii: gipotezy, faktory, zakonomernosti* [Dynamics of the population of wild animals in Russia: Hypotheses, factors, patterns]. Moscow: T-vo nauchnykh izdaniy KMK.
14. Kalinkin, Yu.N. (2019) Opyt ispol'zovaniya dannykh s avtomaticheskikh fotokamer na solontsakh dlya otsenki sostoyaniya gruppirovki blagorodnogo olenya Cervus elaphus sibiricus v Altayskom zapovednike [The experience of using data from automatic cameras on salt lakes to assess the condition of the red deer group Cervus elaphus sibiricus in the Altai Reserve]. *Vestnik okhotovedeniya*. 16(2). pp. 111–118.
15. Shatilina, T.A., Moroz, V.V., Tsitsiashvili, G.Sh. & Radchenkova, T.V. (2024) Formation of Large Anomalies in the Thermal Conditions of Waters on the Western and Eastern Shelf of Sakhalin Island. *Physical Oceanography*. 31(1). pp. 33–45.
16. Malinin, V.N. (2008) *Statisticheskie metody analiza gidrometeorologicheskoy informatsii* [Statistical Methods of Analysis of Hydrometeorological Information]. St. Petersburg: RGG-MU.

Информация об авторах:

Цициашвили Гурами Шалвович – профессор, доктор физико-математических наук, главный научный сотрудник Института прикладной математики Дальневосточного отделения Российской академии наук (Владивосток, Россия). E-mail: guram@iam.dvo.ru

Бочарников Владимир Николаевич – доктор биологических наук, ведущий научный сотрудник Тихоокеанского института географии Дальневосточного отделения Российской академии наук (Владивосток, Россия). E-mail: vboc_harnikov@mail.ru

Вклад авторов: Цициашвили Г.Ш. построил вероятностные модели и оценки их параметров. Бочарников В.Н. дал детальное описание работы фотоловушек при наблюдениях за животными, неразличимыми по снимкам, и сформулировал содержательную задачу обработки этих данных. Авторы заявляют об отсутствии конфликта интересов.

Information about the authors:

Tsitsiashvili Gurami Sh. (Doctor of Physical and Mathematical Sciences, Main Researcher of Institute for Applied Mathematics, Far Eastern Branch of the Russian Academy of Sciences, Vladivostok, Russian Federation). E-mail: guram@iam.dvo.ru

Bocharnikov Vladimir N. (Doctor of Biological Sciences, Leading Researcher of Institute of Pacific Ocean Geography, Far Eastern Branch of the Russian Academy of Sciences, Vladivostok, Russian Federation). E-mail: vbocharnikov@mail.ru

Contribution of the authors: Tsitsiashvili G.Sh. built probabilistic models and estimates of their parameters. Bocharnikov V.N. gave a detailed description of the work of camera traps when observing animals indistinguishable from images and formulated a meaningful task of processing these data. The authors declare no conflicts of interests.

Поступила в редакцию 23.08.2024; принята к публикации 02.12.2024

Received 23.08.2024; accepted for publication 02.12.2024

ИНФОРМАТИКА И ПРОГРАММИРОВАНИЕ

INFORMATICS AND PROGRAMMING

Original article

UDC 004.023

doi: 10.17223/19988605/69/11

Quantum circuit optimization via local qubit reordering by quantum annealing

Mariia A. Maltseva

University of Trento, Trento, Italy;

Institute of Applied Mathematical Research, Karelian Research Centre of the Russian Academy of Science,
Petrozavodsk, Russian Federation, mariia.maltseva@unitn.it

Abstract. Mapping quantum circuits to the nearest-neighbor topology architecture requires additional gates. In this paper, we propose a hybrid approach to optimize quantum circuits for the two-dimensional nearest-neighbor architecture based on graph partitioning and solving SAT problem using quantum annealing. The technique's applicability and solution quality are studied via real experiments.

Keywords: circuit optimization; graph partitioning; quantum annealing; Boolean satisfiability.

For citation: Maltseva, M.A. (2024) Quantum circuit optimization via local qubit reordering by quantum annealing. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 69. pp. 103–111. doi: 10.17223/19988605/69/11

Научная статья

doi: 10.17223/19988605/69/11

Оптимизация квантовой цепи путем локального переупорядочивания кубитов
методом квантового отжига

Мария Алексеевна Мальцева

Университет Тренто, Тренто, Италия;

Институт прикладных математических исследований Карельского научного центра
Российской академии наук, Петрозаводск, Россия, mariia.maltseva@unitn.it

Аннотация. Отображение квантовых цепей на архитектуру с топологией ближайшего соседа требует введения дополнительных вентилях. В статье предложен гибридный подход к оптимизации квантовых цепей для двумерной архитектуры ближайшего соседа на основе построения разрезов графа и решения SAT-задачи на системе квантового отжига. Применимость метода и качество решения исследованы с помощью натуральных экспериментов.

Ключевые слова: оптимизация квантовой цепи; разрез графа; квантовый отжиг; выполнимость булевых формул.

Для цитирования: Мальцева М.А. Оптимизация квантовой цепи путем локального переупорядочивания кубитов методом квантового отжига // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 69. С. 103–111. doi: 10.17223/19988605/69/11

Introduction

Quantum computing is a cutting-edge field that harnesses the principles of quantum mechanics to perform calculations exponentially faster than classical counterpart and solve complex problems that are intractable to it. Similar to classical circuits in traditional computing, quantum circuits (QCs) are the fundamental blocks of quantum algorithms and consist of quantum gates placed sequentially and operating on quantum bits (qubits). However, the noisy nature of current quantum systems, together with hardware constraints and limited available resources, poses the significant challenges to the implementation of QCs. QC optimization is therefore a crucial task, and various techniques have been proposed to perform it.

In general, there is a wide variety of QC optimization criteria. In [1], the authors aim to reduce the number of two-qubit controlled-NOT (CNOT) gates because of their proneness to errors. For that, the problem is converted to tree search problem and then solved by A^* graph-traversal heuristic search allowing to find optimal path between any two nodes of the graph.

QC depth is another popular QC optimization criteria, and it means the largest number of single-qubit and two-qubit gates on any path from input to output of a circuit [2]. In [3], the authors propose to construct depth-optimal QC via meet-in-the-middle algorithm being a heuristic graph-traversal path-finding search.

There is a separate group of topology-aware QC optimization techniques. As QC is a sequence of quantum gates operating on qubits, logical qubits used in a gate description should be mapped to physical qubits embedded into the architecture. Commonly, architecture should satisfy the Nearest Neighbor (NN) condition implying that physical qubits can interact only with their neighbors [4] and making the process of logical qubits mapping non-trivial. However, it allows to reduce impact of noise on quantum operations [5], but introduces significant overhead because of auxiliary SWAP gates (SWAPs) applied to make interacting qubits adjacent. Thus, QC optimization in NN architecture is significant to minimize the number of additional cost-expensive SWAPs (SWAP count) [6].

Several approaches to optimize QC by minimizing SWAP count have been proposed. They differ by the way of QC representation and optimization technique used. For instance, circuit is commonly studied as a NN-compliant circuit, and heuristics are used to minimize SWAP count. In [6], graph bipartitioning applied to qubit line adjacency graph is proposed as a SWAP count minimization technique in linear (one-dimensional) NN architecture. In [7], QC is represented as a graph of interactions between each control-target pair as an edge with weight equal to the number of times this pair appears in the circuit and SWAP count is minimized by qubit lines rearrangement in two-dimensional (2D) NN architecture using Harmony Search algorithm being a sort of gradient-based metaheuristic.

Despite global qubit lines rearrangement, qubit lines can be rearranged locally, implying that qubits used in the description of only the current gate are reordered. This kind of qubit lines rearrangement has been proposed in [8, 9] with SWAP count minimization performed by heuristics.

We are focusing on optimizing QCs for the 2D NN architecture intensively studied recently [9–12] and follow the approach proposed in [9] where the authors consider a circuit as a gate dependency graph and divide it into NN-compliant subcircuits by local reordering of qubits. For that, they propose to partition the graph using Boolean satisfiability. Further, to connect 2D placements of qubits obtained for each subcircuit, A^* search is used. The novelty of the present paper is that we propose hybrid quantum-classical QC optimization approach with quantum computing techniques replacing the classical counterpart. The choice of such a revolutionary type of computing is the main advantage of our approach. More specifically, quantum annealing implemented in D-Wave quantum machine is proposed to solve the problem under study. Being a hardware implemented heuristic, it returns the solution in constant time, opposed to traditional classical heuristics. It makes quantum annealing beneficial for solving QC optimization problem. The details are described further.

The structure of the paper is as follows. In Section 1, the problem is formulated accurately. Section 2 is devoted to the theoretical fundamentals of quantum theory. In Section 3, the approach of QC optimization is proposed. Section 4 illustrates encouraging optimization results of several QCs.

1. Problem statement

In this paper, we study possibilities of QC optimization via local qubit line mapping in 2D NN architecture. For such an architecture, gate set is commonly restricted and consists of single-qubit gates and CNOT. In combination, the aforementioned gates form a universal gate set meaning that any unitary operation can be approximated by a QC composed of only these gates [13].

To introduce a 2D architecture, a 2D grid is defined and each qubit has four neighbors (two horizontal

$$\begin{bmatrix} q_1 & q_2 & q_3 \\ q_4 & q_5 & q_6 \end{bmatrix}$$

and two vertical) at most here. For instance, in grid qubit q_1 has two neighboring qubits which are q_2 and q_4 .

In 2D NN architecture, two-qubit gates act only on adjacent qubits, and SWAPs are used to rearrange non-adjacent interacting qubit lines to perform given operations. Count of additional SWAPs depends on the initial qubit lines placement. Therefore, finding a 2D placement of qubit lines that minimizes the SWAP count is essential for optimizing the circuit.

More precisely, we have to find a sequence of qubit lines placements $\{A^k\}_{k=1}^m$ meaning that qubit lines placement is changed for, say, m times, where A^k is a 2D grid (matrix) with $A^k_{ij} = q$, if qubit q is positioned at (i, j) . For convenience, we also define inverse mapping $q \rightarrow (i_q^k, j_q^k)$ as a position of qubit q in 2D placement A^k . Such a sequence consists of qubit line placements each of which is valid only for one subcircuit constructed from a given circuit by gates recombination and becoming NN-compliant after qubit line reordering.

To map each intermediate qubit line placement A^k to the next one, additional SWAPs are needed. For a 2D grid, its count is calculated via the Manhattan distance and should be minimized, i.e.

$$s(k) := \sum_{q=1}^n |i_q^k - i_q^{k+1}| + |j_q^k - j_q^{k+1}| \rightarrow \min. \quad (1)$$

Then, the overall SWAP count within the whole circuit is as follows:

$$\sum_{k=1}^m s(k) = \sum_{k=1}^m \sum_{q=1}^n |i_q^k - i_q^{k+1}| + |j_q^k - j_q^{k+1}| \rightarrow \min, \quad (2)$$

where n is the number of qubits.

However, quantum gates composing a given circuit should be recombined according to the gates dependency graph $G = (V, E)$, where V is a set of gates of the circuit and E is the set of dependencies of the gates with Boolean weight $e_{g_1, g_2} = 1$, if g_2 follows g_1 and they are not commutative (meaning that interchange of them causes different result of circuit implementation), and $e_{g_1, g_2} = 0$, if g_1 and g_2 are commutative.

As such, the problem is a graph partitioning problem with goal function (2).

2. Scientific background

2.1 Quantum theory basic properties

Qubit is a basic unit of information in QC [14]. A qubit has the two basic Boolean states. Because of the quantum theory nature, qubit can also store a superposition (linear combination) of the two basic Boolean

states $|0\rangle \equiv \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ or $|1\rangle \equiv \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, i.e. the qubit state $|\varphi\rangle$ can be expressed as follows:

$$|\varphi\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \alpha|0\rangle + \beta|1\rangle,$$

where $\alpha, \beta \in \mathbb{C}$, such that $\alpha^2 + \beta^2 = 1$.

The next fundamental quantum property is entanglement [13]. It allows qubits to interact in pairs regardless of the distance between them and makes them dependent on each other that way.

To perform computations in quantum computing, quantum gates are used. They act on qubits and change their states correspondingly in that way [13]. Quantum gates are used as basic blocks to compose more complex functionality that is represented by a QC [15].

Algebraically a quantum gate is represented as a unitary operator [13] meaning that its inverse matrix is equal to its adjoint matrix. Unitarity conserves being of a quantum system in one of the possible states [13].

There is a great variety of quantum gates. Some of them are single-qubit, i.e. aforementioned NOT gate inverting the qubit. Other gates operate on several qubits. For example, CNOT (controlled-NOT) gate operates on 2 qubits and entangles them by inverting the target qubit if the control qubit is in state $|1\rangle$. For example, if control qubit is the first one and target is the second one (CNOT(q_1, q_2)), then it is as follows

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Three CNOT gates being in combination construct a two-qubit SWAP gate that interchanges input qubits [13].

2.2. Quantum Annealing

Quantum Annealing (QA) is analogous to traditional Simulated Annealing (SA) that is a local search optimization heuristic resembling the physical annealing process. Similar to the physical cooling, cooling scheme manages SA process and transition to the next candidate solution is done with temperature-dependent probability decreasing over time. The difference between SA and QA is that the latter is able to find global optimum thanks to the quantum tunneling enabling annealing process to pass through the energetic barriers [16].

The QA system aims to decrease the cost function of the problem, which can be represented in an Ising model [17]:

$$\varepsilon(\bar{z}) = \sum_{i \in V} h_i z_i + \sum_{(i,j) \in E} J_{ij} z_i z_j, \quad (3)$$

where $\bar{z} = \{-1, 1\}^{|V|}$ is a set of spins describing qubits from V , $G = (V, E)$ is an undirected graph of allowed interactions between qubits, $J_{ij} = J_{ji}$ is interaction strength of the qubits i and j connected by an edge and h_i is the on-site energy of qubit i .

Note that QA is not a universal quantum computer. QA can solve only Quadratic Unconstrained Binary Optimization (QUBO) problems defined as follows [18]:

$$\min_{\bar{x} \in \{0,1\}^{|V|}} \bar{x}^T Q \bar{x}, \quad (4)$$

where $Q_{ii} = h_i$, $Q_{ij} = J_{ij}$ from the Ising model (3).

As for the hardware implementation, QA is developed by D-Wave and can be run online for a limited access time. IT-company Fujitsu provides an alternative quantum-inspired Digital annealer.

3. Solution approach

We introduce a hybrid quantum-classical SWAP count optimization of a QC to be implemented in a 2D NN architecture following [9]. The key ingredients of the original approach involves Boolean satisfiability (that checks if there is at least one set of the variables' assignments under which the given Boolean formula evaluates to TRUE [19]) and A* heuristic graph-traversal path-finding search [20].

Algorithmic description of our approach is presented below. It differs from [9] by solving SAT-problem in QUBO format using QA (see Step 3.2-Step 3.4 below).

Step 1. Require: 2D grid with initial qubits placement; gate dependency graph G ; set of gates Γ ; partition $P = \emptyset$.

Step 2. While Γ not empty.

Step 3. Binary search to construct a subcircuit G^k (taking into account gate dependency graph G):

Step 3.1. Convert G^k to a CNF CNF_k ;

Step 3.2. Convert a SAT-problem with CNF_k to a QUBO problem with coefficient matrix Q_k (by *qubovert* Python package);

Step 3.3. Solve QUBO problem defined by Q_k using QA;

Step 3.4. Extract the first solution (with the lowest energy) - qubit lines placement A^k ;

Step 3.5. If CNF_k is TRUE;

Step 3.5.1. Partition $P = P \cup G^k$, and $\Gamma = \Gamma \setminus G^k$, and return to Step 2;

Step 3.6. Return to Step 3;

Step 4. Apply A^* search to connect $\{A^k\}_{k=1}^m$ and calculate number of SWAP gates needed for that.

The approach is iterative and is based on binary search technique used at Step 3 to construct a subcircuit taking into account gate dependencies. To obtain CNF at Step 3.1, the following conditions should be formulated:

Condition 1. Interacting bits of all gates are adjacent.

Condition 2. Each qubit is assigned to only one cell on a 2D grid.

Condition 3. At most one qubit is assigned to each cell on a 2D grid.

We introduce Boolean variables x_{ijr} that are 1 if qubit q_r is assigned to cell (i, j) and 0 otherwise. For simplicity, 2D grid of 3 rows and 3 columns where cells are placed rowly is considered.

To express Condition 1 as a Boolean function, we consider a CNOT(q_1, q_2) as an example. If q_1 is assigned to (1,1), then q_2 should be assigned to either one of (0,1), (1,0), (1,2), or (2,1). This condition can be expressed as

$$\neg x_{111} \vee x_{012} \vee x_{102} \vee x_{122} \vee x_{212}. \quad (5)$$

Then, applying disjunction to the Boolean formulas of such conditions for the other cells of the grid, we obtain a formula for the condition such that q_1 and q_2 should be adjacent on the grid. After that, we apply conjunction of the obtained Boolean formulas for all the pairs of interacting qubits.

As for Condition 2, each qubit is assigned to at least one cell and at most one cell. The former part of the condition can be expressed as

$$\bigvee_{(i,j)} x_{ijr} = 1 \quad (6)$$

for each qubit q_r . The latter part prohibits assigning the qubit to two different cells:

$$\neg x_{i_1 j_1 r} \vee \neg x_{i_2 j_2 r} = 1 \quad (7)$$

for each qubit q_r and each pair of cells $(i_1, j_1) \neq (i_2, j_2)$.

Condition 3 prohibits assigning two different qubits to one cell meaning that

$$\neg x_{ijr_1} \vee \neg x_{ijr_2} = 1 \quad (8)$$

for each cell (i, j) and each pair of qubits q_{r_1} and q_{r_2} , $r_1 \neq r_2$.

At Step 3.3, QA solves QUBO problem constructed from CNF by *qubovert* automatically. If the solution is satisfiable, i.e. there is such a qubit placement A^k satisfied to (5)–(8) then G^k forms a NN-compliant subcircuit with qubit placement A^k , and binary search on the rest of circuit is continued.

After partitioning the gate dependency graph G into NN-compliant subcircuits with qubit placements $\{A^k\}_{k=1}^m$, A^* search is used to connect them optimally (Step 4). A^* search is a heuristic that allows to find optimal way from the starting node S of a graph to the target one T based on a cost function $f(N) = g(N) + h(N)$, where N is a current node during the search, $g(N)$ is a cumulative cost from S to N [9], and $h(N)$ is a heuristic that estimates the cost from the current qubit placement N to T . In our problem, a qubit placement A^k is

a node of a graph with A^1 as the starting node and A^m as the target one. Cumulative cost $g(A^k)$ is SWAP count needed to obtain A^k from A^1 and it is the sum of the Manhattan distances between the positions of each qubit in the A^k and A^1 , i.e.

$$g(A^k) = \sum_{i=1}^{k-1} s(i),$$

where $s(i)$ is defined in (1). Heuristic cost $h(A^k)$ is calculated analogously.

4. Numerical experiments

Recall that in order to solve a SAT problem in QA, at first it should be converted into QUBO format (4). It can be done using the technique from [19] which requires preliminary setup to encode a SAT-problem to the Ising model. Alternative way is to use Python package *qubovert* that performs such a conversion automatically by using *sat* library. In this paper, we propose to follow the latter approach. To obtain a SAT-problem, each of the Conditions 1, 2, and 3 ((5) – (8)) are added to the resulting Boolean formula with some positive Lagrangian parameter λ tuned by hand. For the experiments performed, $\lambda=100$ for (5), (7) and (8), and $\lambda=10$ for (6).

To verify the applicability of the approach, we optimize one quantum gate from *Reversible logic synthesis benchmark page* by D. Maslov (available at <https://reversiblebenchmarks.github.io/>) and another one created by us. To perform the experiments using QA, we use a real quantum annealer, namely, *D-Wave Advantage_system5.4* with 5614 qubits and 40050 couplers, fixing the annealing time to $20\mu s$ and taking the best out of 2000 simulation runs. Due to the limited access time on it, we also optimize the circuits via SA and classical SAT-solver *PicoSAT* run on local Ubuntu machine (Ubuntu 23.04 with 31 GiB RAM using 11th Gen Intel(R) Core(TM) i7-11700 @ 2.50GHz). Note that numerical experiments are performed only for the circuits with no more than 5 qubits because of the high computational demand on *qubovert*. Along the lines, we also benchmark the resulting quantum part of the algorithm in terms of logical and physical qubits to demonstrate its computational complexity of QA.

4.1. Fredkin gate optimization

Fredkin gate is a gate from *Reversible logic synthesis benchmark page* by D. Maslov. It is frequently used in QC because of its small cost in some QC technologies. The circuit realization at elementary quantum gate set is illustrated in Fig. 1 and consists of 7 two-qubit quantum gates (enumerated as $C_1 \dots C_7$) operating on different pairs of 3 qubits. In case when qubits are placed in 2×2 grid rowly with numbering starting at the left side, i.e. NN topology defined as $\begin{bmatrix} 0 & 1 \\ 2 & - \end{bmatrix}$, 6 SWAPs are needed to implement the circuit (as 1 SWAP

is needed before each gate to make it NN-compliant, and 1 SWAP returns the qubits to their initial order after the gate). In Fig. 2, the dependency of the gates of Fredkin gate from Fig. 1 is illustrated.

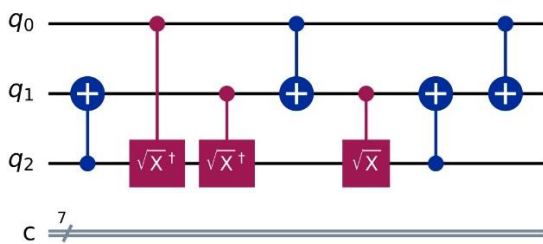


Fig. 1. Fredkin gate with initial SWAP count = 6

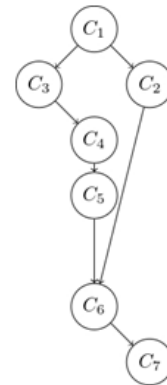


Fig. 2. Dependency of gates in Fredkin gate from Fig. 1

At first, we use SA due to the limited access time of QA on D-Wave. After applying the proposed approach, the circuit is divided into 2 NN-compliant subcircuits as follows from Fig. 3. The qubits are placed following $A^1 = \begin{bmatrix} 2 & 1 \\ 0 & - \end{bmatrix}$ and $A^2 = \begin{bmatrix} 1 & 2 \\ 0 & - \end{bmatrix}$ within the first and the second subcircuits, respectively. 1 SWAP is needed to make them consistent.

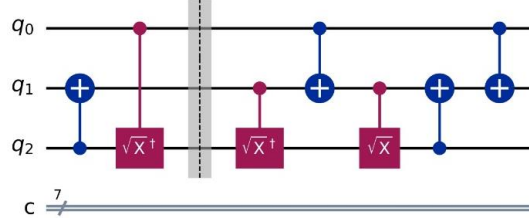


Fig. 3. The optimized Fredkin gate from Fig. 1

QA divides the circuit into subcircuits in the same way as SA does. The qubits are placed following $A^1 = \begin{bmatrix} - & 0 \\ 1 & 2 \end{bmatrix}$ and $A^2 = \begin{bmatrix} - & 0 \\ 2 & 1 \end{bmatrix}$ within the first and the second subcircuits, respectively. Similar to SA, 1 SWAP is needed to connect them. We should note that SA always returns positive result in case when a subcircuit is satisfiable unlike QA. Thus, we should repeat QA for several times (it is 15 in our case) to reduce the probability to exclude satisfiable subcircuit accidentally. In order to convert the subcircuit to QUBO, 19 logical qubits are used and 36 physical qubits are used for the first subcircuit whereas the second one holds 33 physical qubits. It implies that QA hardware (with more than 5000 qubits) used can solve bigger QUBO problems.

To validate the result obtained above, we replicate the classical approach proposed in [9] and use the classical SAT solver *PicoSAT* via *pyeda* Python package. In this experiment the circuit is divided in the same way as both in QA and SA (see Fig. 3). And the qubits are placed following $A^1 = \begin{bmatrix} - & 1 \\ 0 & 2 \end{bmatrix}$ and $A^2 = \begin{bmatrix} - & 2 \\ 0 & 1 \end{bmatrix}$ within the first and the second subcircuits, respectively. 1 SWAP is required to make them consistent.

4.2. Optimization of CNOT-based circuit

To verify the applicability of the proposed approach for a bigger QC, we construct one illustrated in Fig. 4. It contains 6 CNOT gates (enumerated as $C_1 \dots C_6$) operating on some pairs of 5 qubits. Initially, it requires 6 SWAPs in order to implement the circuit in 2D NN architecture that is a 2D grid where qubits are placed rowly, i.e. $\begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & - \end{bmatrix}$. In Fig. 5, the dependency of the gates of circuit from Fig. 4 is depicted.

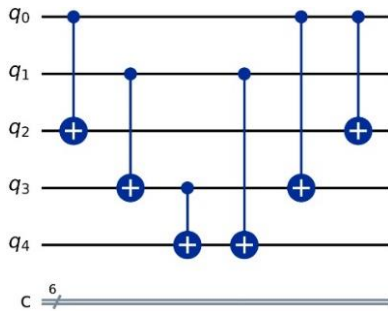


Fig. 4. The circuit to be optimized. Initial SWAP count = 6

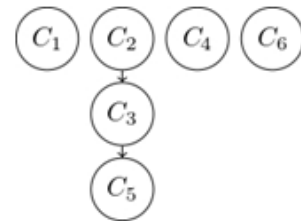


Fig. 5. Dependency of gates in circuit from Fig. 4

After applying SA, the number of SWAPs needed is 2 at best. During the optimization, the circuit was divided into 2 NN-compliant subcircuits: the first one contains C_1, C_2, C_3, C_5 , and C_6 , and the second subcir-

cuit is formed by C_4 gate (see Fig. 5). Note that C_4 , C_5 , and C_6 gates are commutative and it allows to place C_4 after C_6 . The qubits are placed following $A^1 = \begin{bmatrix} - & 4 & 2 \\ 1 & 3 & 0 \end{bmatrix}$ and $A^2 = \begin{bmatrix} 3 & - & 2 \\ 1 & 4 & 0 \end{bmatrix}$ within the first and the second subcircuits, respectively.

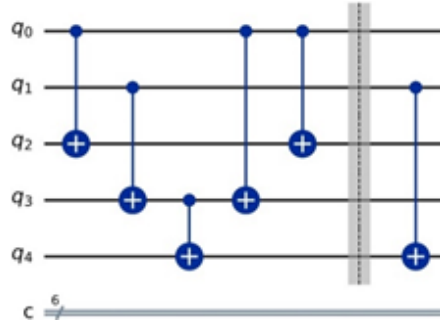


Fig. 6. The optimized circuit from Fig. 4

QA also divides the circuit from Fig. 4 in the same way as SA (see Fig. 6) where optimal qubits placements follow $A^1 = \begin{bmatrix} 4 & 3 & 0 \\ - & 1 & 2 \end{bmatrix}$ and $A^2 = \begin{bmatrix} 0 & 1 & 4 \\ - & 2 & 3 \end{bmatrix}$ within the first and the second subcircuits, respectively. In order to make these optimal qubits placements consistent, 5 SWAPs are needed.

To encode the first subcircuit into QUBO, 103 logical and 354 physical qubits are used and 111 logical and 304 physical qubits are used for the encoding of the second subcircuit into QUBO.

Here, QA returns suboptimal solution due to the dimension of the problem. As this CNOT-based circuit requires more logical and physical qubits than Fredkin gate investigated in Section 4.1, it seems that more simulation runs are needed to obtain better solution. However, its number is bounded by the access time provided by D-Wave machine that is also limited. We decided to increase the number of simulation runs up to 4000 and the result is encouraging. The proposed technique partitions the circuit into 2 subcircuits: the 1st one consists of the first four CNOT gates (C_1 , C_2 , C_3 , C_4) with qubit placement $A^1 = \begin{bmatrix} 2 & 0 & - \\ 4 & 3 & 1 \end{bmatrix}$

and the other two gates construct the 2nd subcircuit with qubit placement $A^2 = \begin{bmatrix} 0 & - & 4 \\ 2 & 3 & 1 \end{bmatrix}$. Only 3 SWAPs are needed to make the subcircuits consistent.

By optimizing the circuit using *PicoSAT*, the circuit is divided in the same way as both in QA and SA (see Fig. 6). Optimal qubits placements of the subcircuits follows $A^1 = \begin{bmatrix} - & 4 & 2 \\ 1 & 3 & 0 \end{bmatrix}$ and $A^2 = \begin{bmatrix} - & 4 & 0 \\ 3 & 1 & 2 \end{bmatrix}$ within the first and the second subcircuits, respectively. 2 SWAPs are required to make them consistent.

Conclusion

In this work, we introduce hybrid QC optimization approach by local reordering of qubits using QA and Boolean satisfiability. Its promising nature is explained by the use of a breakthrough computing for solving such a problem, which provides a solution in constant time. Numerical experiments show that QA performs well on small scale: the solution obtained is comparable with one returned by classical SAT-solver. However, while solving a larger problem, QA may return suboptimal results, and making the simulation process longer improves the solution.

Author express thanks to Rumyantsev Alexander S. (Dr. Sci. in Physics and Mathematics) for the significant comments and advice while writing the paper.

References

1. Davis, M.G., Smith, E., Tudor, A., Sen, K., Siddiqi, I. & Iancu, C. (2020) Towards Optimal Topology Aware Quantum Circuit Synthesis. *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)*. pp. 223–234. DOI: 10.1109/QCE49297.2020.00036
2. Arabzadeh, M., Saheb Zamani, M., Sedighi, M. & Saeedi, M. (2013) Depth-optimized reversible circuit synthesis. *Quantum Information Processing*. 12(4). pp. 1677–1699. DOI: 10.1007/s11128-012-0482-8
3. Amy, M., Maslov, D., Mosca, M. & Roetteler, M. (2013) A Meet-in-the-Middle Algorithm for Fast Synthesis of Depth-Optimal Quantum Circuits. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*. 32(6). pp. 818–830. DOI: 10.1109/tcad.2013.2244643
4. Wille, R., Lye, A. & Drechsler, R. (2014) Optimal SWAP gate insertion for nearest neighbor quantum circuits. *2014 19th Asia and South Pacific Design Automation Conference (ASP-DAC)*. pp. 489–494. DOI: 10.1109/ASPDAC.2014.6742939
5. Datta, K., Kole, A., Sengupta, I. & Drechsler, R. (2022) Nearest Neighbor Mapping of Quantum Circuits to Two-Dimensional Hexagonal Qubit Architecture. *2022 IEEE 52nd International Symposium on Multiple-Valued Logic (ISMVL)*. DOI: 10.1109/ismvl52857.2022.00013
6. Chakrabarti, A., Sur-Kolay, S. & Chaudhury, A. (2011) Linear Nearest Neighbor Synthesis of Reversible Circuits by Graph Partitioning. *arXiv*. DOI: 10.48550/ARXIV.1112.0564
7. Alfaiakawi, M.G., Ahmad, I. & Hamdan, S. (2016) Harmony-Search Algorithm for 2D Nearest Neighbor Quantum Circuits Realization. *Expert Systems With Applications*. 61(C). pp. 16–27. DOI: 10.1016/j.eswa.2016.04.038
8. Bhattacharjee, A., Bandyopadhyay, C., Wille, R., Drechsler, R. & Rahaman, H. (2019) Improved Look-Ahead Approaches for Nearest Neighbor Synthesis of 1D Quantum Circuits. *2019 32nd International Conference on VLSI Design and 2019 18th International Conference on Embedded Systems (VLSID)*. pp. 203–208. DOI: 10.1109/VLSID.2019.00054
9. Wakaki, H. & Yamashita, S. (2019) Mapping a Quantum Circuit to 2D Nearest Neighbor Architecture by Changing the Gate Order. *IEICE Transactions on Information and Systems*. E102.D. pp. 2127–2134. DOI: 10.1587/transinf.2018EDP7439
10. Shafaei, A., Saeedi, M. & Pedram, M. (2014) Qubit placement to minimize communication overhead in 2D quantum architectures. *Proceedings of the Asia and South Pacific Design Automation Conference (ASP-DAC)*. pp. 495–500. DOI: 10.1109/ASPDAC.2014.6742940
11. Wille, R., Keszocze, O., Walter, M., Rohrs, P., Chattopadhyay, A. & Drechsler, R. (2016) Look-ahead schemes for nearest neighbor optimization of 1D and 2D quantum circuits. *2016 21st Asia and South Pacific Design Automation Conference (ASP-DAC)*. pp. 292–297. DOI: 10.1109/ASPDAC.2016.7428026
12. Farghadan, A. & Mohammadzadeh, N. (2017) Quantum circuit physical design flow for 2D nearest-neighbor architectures: Quantum Circuit Design Flow. *International Journal of Circuit Theory and Applications*. 45. pp. 989–1000. DOI: 10.1002/cta.2335
13. Nielsen, M.A. & Chuang, I.L. (2010) *Quantum computation and quantum information*. Cambridge University Press. [Online] Available from: <https://profmcrz.wordpress.com/wp-content/uploads/2017/08/quantum-computation-and-quantum-information-nielsen-chuang.pdf>. (Accessed: 27th September 2024).
14. Pereira da Silva, R. (2023) Gate-based Quantum Computing: An Overview. *SSRN Electronic Journal*. DOI: 10.2139/ssrn.4347584
15. Racorean, O. (2015) Quantum Gates and Quantum Circuits of Stock Portfolio. *arxiv*. DOI: 10.2139/ssrn.2630341
16. Yarkoni, S., Raponi, E., Bäck, T. & Schmitt, S. (2022) Quantum annealing for industry applications: introduction and review. *Reports on Progress in Physics*. 85(10). pp. 1–43. DOI: 10.1088/1361-6633/ac8c54
17. Su, J. (2018). *Towards Quantum Computing: Solving Satisfiability Problem by Quantum Annealing*. UCLA Electronic Theses and Dissertations. [Online] Available from: <https://escholarship.org/uc/item/8qp5200s>. (Accessed: 26th August 2024).
18. Bonomi, A., De Min, T., Zardini, E., Blanzieri, E., Cavecchia, V. & Pastorello, D. (2022) Quantum annealing learning search implementations. *Quantum Information and Computation*. 22(3 & 4). pp. 181–208. DOI: 10.26421/qic22.3-4-1
19. Bian, Z., Chudak, F., Macready, W., Roy, A., Sebastiani, R. & Varotti, S. (2018) Solving SAT and MaxSAT with a Quantum Annealer: Foundations, Encodings, and Preliminary Results. *arXiv*. [Online] Available from: <https://arxiv.org/abs/1811.02524>. (Accessed: 26th August 2024).
20. Foad, D., Ghifari, A., Kusuma, M.B., Hanafiah, N. & Gunawan, E. (2021) A Systematic Literature Review of A* Pathfinding. *Procedia Computer Science*. 179. pp. 507–514. DOI: 10.1016/j.procs.2021.01.034

Information about the author:

Maltseva Mariia A. (Post-graduate Student, Department of Information Engineering and Computer Science, University of Trento, Trento, Italy; Junior research assistant, Institute of Applied Mathematical Research, Karelian Research Centre of the Russian Academy of Science, Petrozavodsk, Russian Federation). E-mail: mariia.maltseva@unitn.it

The author declares no conflicts of interests.

Информация об авторе:

Мальцева Мария Алексеевна – аспирант факультета информационной инженерии и информатики Университета Тренто (Тренто, Италия); младший научный сотрудник Института прикладных математических исследований Карельского научно-го центра Российской академии наук (Петрозаводск, Россия). E-mail: mariia.maltseva@unitn.it

Автор заявляет об отсутствии конфликта интересов.

Received 28.08.2024; accepted for publication 02.12.2024

Поступила в редакцию 28.08.2024; принята к публикации 02.12.2024

Научная статья

УДК 004.048

doi: 10.17223/19988605/69/12

**Программный комплекс нейросетевой классификации данных
газоаналитического обследования дыхательной системы**

**Артём Викторович Обходский¹, Денис Евгеньевич Кульбакин²,
Елена Владимировна Обходская³, Александр Сергеевич Попов⁴,
Евгений Олегович Родионов⁵, Виктор Иванович Сачков⁶,
Владимир Иванович Чернов⁷, Евгений Лхамацыренович Чойнзон⁸**

¹ *Национальный исследовательский Томский политехнический университет, Томск, Россия*

^{2, 5, 7, 8} *НИИ онкологии ТНИМЦ Российской академии наук, Томск, Россия*

^{3, 4, 6} *Национальный исследовательский Томский государственный университет, Томск, Россия*

^{5, 8} *Сибирский государственный медицинский университет, Томск, Россия*

¹ *art707@yandex.ru*

² *kulbakin_d@mail.ru*

³ *fil330a@yandex.ru*

⁴ *asptomsktpu@gmail.com*

⁵ *rodionov_eo@oncology.tomsk.ru*

⁶ *vicsachkov@gmail.com*

⁷ *chernov@tnimc.ru*

⁸ *choynzonov@tnimc.ru*

Аннотация. Предложены алгоритм и архитектура нейронной сети для классификации сигнальных паттернов, формируемых в устройствах для анализа состава газовой смеси в выдыхаемом воздухе. В устройствах применяется набор неселективных полупроводниковых газовых датчиков, синхронизированных между собой и работающих в режиме термодиклирования. Реализованный алгоритм и нейронная сеть обеспечивают нормировку значений сигналов одновременно со всего набора датчиков и дифференциацию сигнальных паттернов. Программный комплекс реализован на алгоритмическом языке программирования C++ в среде Qt и позволяет проводить дообучение нейронной сети за счет интеграции инструментов системы управления базой данных и нейросетевого анализа данных при увеличении объемов базы. В исследовании применялись данные от проб выдыхаемого воздуха, отобранные у здоровых добровольцев и пациентов со злокачественными новообразованиями орофарингеальной области и гортани. Полученные показатели чувствительности и специфичности нейросетевого классификатора данных сопоставимы с современными высокоточными рентгенологическими методами диагностики опухолей дыхательных путей.

Ключевые слова: программный комплекс; классификация данных; искусственная нейронная сеть; параметры обучения; анализ эффективности классификатора.

Благодарности: Исследование выполнено при финансовой поддержке гранта РНФ в рамках научного проекта № 23-15-00177 «Теоретическое и клиническое обоснование молекулярной оценки состава выдыхаемого воздуха для диагностики онкологических заболеваний», <https://rscf.ru/project/23-15-00177/>

Для цитирования: Обходский А.В., Кульбакин Д.Е., Обходская Е.В., Попов А.С., Родионов Е.О., Сачков В.И., Чернов В.И., Чойнзон Е.Л. Программный комплекс нейросетевой классификации данных газоаналитического обследования дыхательной системы // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 69. С. 112–123. doi: 10.17223/19988605/69/12

Neural network classification software for the gas analytical survey data of respiratory system

Artem V. Obkhodskiy¹, Denis E. Kulbakin², Elena V. Obkhodskaya³, Aleksandr S. Popov⁴, Evgeniy O. Rodionov⁵, Victor I. Sachkov⁶, Vladimir I. Chernov⁷, Evgeny L. Choyazonov⁸

¹ National Research Tomsk Polytechnic University, Tomsk, Russian Federation

^{2, 5, 7, 8} Cancer Research Institute – branch of the Tomsk National Research Medical Center of the Russian Academy of Science, Tomsk, Russian Federation

^{3, 4, 6} National Research Tomsk State University, Tomsk, Russian Federation

^{5, 8} Siberian State Medical University, Tomsk, Russian Federation

¹ art707@yandex.ru

² kulbakin_d@mail.ru

³ fil330a@yandex.ru

⁴ asptomsktpu@gmail.com

⁵ rodionov_eo@oncology.tomsk.ru

⁶ vicsachkov@gmail.com

⁷ chernov@tnimc.ru

⁸ choyazonov@tnimc.ru

Abstract. An algorithm and neural network architecture are proposed for classifying signal patterns generated in devices for analyzing the composition of a gas mixture in exhaled air. The devices use a set of non-selective semiconductor gas sensors synchronized with each other and operating in the thermal cycling mode. The implemented algorithm and neural network provide for the normalization of signal values simultaneously from the entire set of sensors and the differentiation of signal patterns. The software package is implemented in the C++ algorithmic programming language in the Qt environment and allows for additional training of the neural network by integrating tools from the database management system and neural network data analysis when increasing the database volume. The studies used data from exhaled air samples taken from healthy volunteers and patients with squamous cell carcinoma of the larynx and oropharynx. The obtained sensitivity and specificity indicators of the neural network data classifier are comparable with modern high-precision X-ray methods for diagnosing respiratory tract tumors.

Keywords: software package; data classification; artificial neural network; training parameters; classifier efficiency analysis.

Acknowledgments: The study was supported by the Russian Science Foundation grant № 23-15-00177 «Theoretical and clinical substantiation of the molecular assessment of the composition of exhaled air for the diagnosis of oncological diseases», <https://rscf.ru/en/project/23-15-00177>.

For citation: Obkhodskiy, A.V., Kulbakin, D.E., Obkhodskaya, E.V., Popov, A.S., Rodionov, E.O., Sachkov, V.I., Chernov, V.I., Choyazonov, E.L. (2024) Neural network classification software for the gas analytical survey data of respiratory system. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 69. pp. 112–123. doi: 10.17223/19988605/69/12

Введение

Уровень развития медицинских технологий, обеспечивающих успешное лечение онкологических заболеваний, напрямую влияет на показатель средней продолжительности жизни и, как следствие, на качество жизни значительной части населения страны. С увеличением возраста растет и риск развития онкопатологий [1]. В связи с этим для улучшения качества жизни необходимо уделять внимание развитию технологий, направленных на профилактику и предварительную диагностику онкологических заболеваний на ранних стадиях [2].

С развитием технологий полупроводниковых материалов, чувствительных к компонентам газовых смесей, микропроцессорной техники и методов цифровой обработки данных интерес к сенсорным технологиям диагностики рака по анализу выдыхаемого воздуха значительно возрос, о чем свидетельствует быстро растущее число публикаций на эту тему. Основными характеристиками раковой

опухоли являются неконтролируемый рост клеток внутри организма, изменение обмена веществ в организме и характерный для этого патологического процесса выброс летучих органических соединений. В зависимости от того, где расположены раковые клетки, они имеют разные характерные профили выделяемых летучих органических соединений, которые используются для дифференциации по сравнению с людьми, у которых онкопатология отсутствует. Из-за низкой растворимости в крови характеристические летучие органические соединения легко диффундируют в альвеолярный воздух и впоследствии выделяются с выдыхаемым воздухом, что позволяет их обнаруживать. Технологии анализа летучих органических соединений в выдыхаемом воздухе в последнее время привлекают повышенное внимание в связи с их возможным применением для раннего скрининга рака благодаря сравнительно высокой скорости проведения анализа и неинвазивности метода [3].

Ранняя предварительная диагностика в случаях злокачественных новообразований головы и шеи позволяет проводить более эффективное и своевременное лечение, обеспечивает меньшую потерю функциональности и снижение уровня смертности [4]. В результате своевременного выявления и терапии больных с местно-распространенными злокачественными опухолями орофарингеальной области и гортани, которые входят в структуру локализаций рака головы и шеи, может быть исключен целый комплекс негативных последствий: инвалидизация, нарушение физиологических функций, тяжелые косметические потери, а также интенсивное возникновение психоэмоциональных травм [5]. Современные методики обследования больных со злокачественными новообразованиями орофарингеальной области и гортани основываются на использовании эндоскопических, рентгенологических и морфологических методов диагностики [6]. Эффективность таких методов зависит как от доступности инструментальной и лабораторной базы, так и от опыта специалистов диагностического профиля. В связи с этим к новым методам предварительной диагностики предъявляются требования низкой себестоимости, независимости от человеческого фактора, а также возможности использования в неспециализированных клиниках врачами первичного звена (ЛОР-врачами, стоматологами и др.).

В качестве неинвазивного метода скрининга злокачественных новообразований головы и шеи может выступать технология, предусматривающая анализ газовой смеси в выдыхаемом воздухе. В устройствах, использующих такой подход, применяются различные наборы газовых датчиков и интеллектуальные методы обработки данных [7]. Применяемое оборудование для этой технологии относительно недорогое. Оно обладает высокой мобильностью, не требует дополнительного переобучения персонала, является оперативным. Газовые полупроводниковые датчики обладают инерционностью, но при этом обеспечивают получение данных без существенных задержек.

В рамках настоящего исследования устройство-газоанализатор используется для неселективной идентификации летучих органических соединений в выдыхаемом воздухе больных злокачественными новообразованиями головы и шеи и здоровых добровольцев [8]. Полученные данные структурируются для дальнейшего установления их принадлежности классу проб от пациентов с патологией или здоровых добровольцев. Идентифицируемая совокупность летучих органических соединений будет считаться биомаркером заболевания, если она статистически отличается у привлекаемых для исследования групп людей.

Злокачественные новообразования изменяют состав и содержание некоторых летучих органических соединений, присутствующих в дыхании человека, которые могут рассматриваться в качестве маркера заболевания [9, 10]. Качественно и количественно идентифицировать химические вещества в газовой смеси на основе большого количества данных позволяют методы машинного обучения. Совместно с устройствами-газоанализаторами часто применяются метод опорных векторов, метод ближайших соседей и искусственные нейронные сети [11]. Большинство работ, связанных с анализом летучих органических соединений и их взаимосвязью с онкологическими заболеваниями, так или иначе используют искусственные нейронные сети.

В работе [12] исследовали применение различных методик классификации для определения рака легких. В эксперименте приняли участие 100 добровольцев с практически одинаковым распределением по количеству здоровых и больных. Для классификации применялись модифицированный k-NN-метод, классификаторы на основе линейных и квадратичных дискриминантных функций и ис-

кусственная нейронная сеть. Точность классификации добровольцев во всех трех случаях оказалась на уровне 89%. Такой показатель сопоставим с современными рентгенологическими методами диагностики злокачественных новообразований. Например, применение ПЭТ / КТ обладает чувствительностью и специфичностью на уровне 90,1 и 96,2% соответственно [13]. При этом показатели точности классификатора данных на основе искусственной нейронной сети могут быть улучшены путем увеличения количества эпох обучения и внутренних слоев, учитывая небольшое количество статистических данных. В работе [12] количество эпох обучения составило 20 и использовалась нейронная сеть с одним внутренним слоем.

Цель настоящего исследования – формирование базовой архитектуры, алгоритма и нейронной сети для построения программного комплекса нейросетевой классификации данных газоаналитического обследования дыхательной системы, обеспечивающего обработку данных, формируемых с помощью специализированных устройств анализа состава газовой смеси в выдыхаемом воздухе. К создаваемому программному комплексу предъявлялись требования выполнения всех операций и стадий по обработке данных преимущественно в автоматическом режиме, возможности оценки частных и обобщенных критериев эффективности формируемых нейронных сетей, структурирования разнородных наборов данных, получаемых от устройств анализа состава газовых смесей, и реализации механизмов дообучения нейронных сетей. Наряду с этим программный комплекс должен осуществлять подбор обучающих выборок данных для формирования нейросетевых классификаторов патологий дыхательной системы разного типа.

1. Формат представления данных газоаналитического обследования дыхательной системы

В нашем исследовании измерение совокупных сигнальных паттернов газовой смеси в выдыхаемом воздухе осуществляется с помощью устройства-газоанализатора, обеспечивающего единый унифицированный набор действий для отбора проб у всех исследуемых добровольцев при соблюдении нормальных внешних условий [8]. Датчики во время функционирования устройства-газоанализатора работают в режиме термоциклирования и за время сбора данных формируют нескольких полных циклов от точки начала нагрева до точки окончания остывания.

В результате оцифровки состава газовой смеси в выдыхаемом воздухе по каждому обследуемому добровольцу формируется массив значений из 24 числовых рядов, каждый из которых включает в среднем 7 периодов термоциклирования. Автоматизация процесса пробоотбора в устройстве-газоанализаторе обеспечивает всегда одинаковый объем вводимого воздуха для проб всех добровольцев, равный 250 мл. Программное обеспечение устройства-газоанализатора запускает процесс сбора данных на 90 с. Аналого-цифровое преобразование сигналов с 24 полупроводниковых металлооксидных неселективных газовых датчиков выполняется с частотой дискретизации 50 Гц. После завершения измерения сигнальных паттернов включается продувка устройства очищенным воздухом.

На рис. 1 представлен пример графика с усредненными значениями величин откликов S с одного датчика для двух групп обследуемых добровольцев. В группе со злокачественными новообразованиями головы и шеи усреднение проводилось по 23 пробам, в контрольной группе здоровых добровольцев – по 25 пробам.

Параметр S вычислялся по формуле

$$S = k \frac{U_i - U_i^0}{U_i^0}, \quad (1)$$

где k – коэффициент масштабирования, равный 100 000, U_i – уровень напряжения на выходе датчика в 4-м периоде одного цикла измерения, U_i^0 – уровень напряжения на выходе датчика в 1-м периоде одного цикла измерения, соответствующего по времени U_i и характеризующего фоновый воздух. Для расчетов при обработке данных из одного периода термоциклирования используется только интервал от начала до окончания нагрева датчиков с длительностью 3,5 с как более информативный с точки зрения преобразования измеряемого параметра.

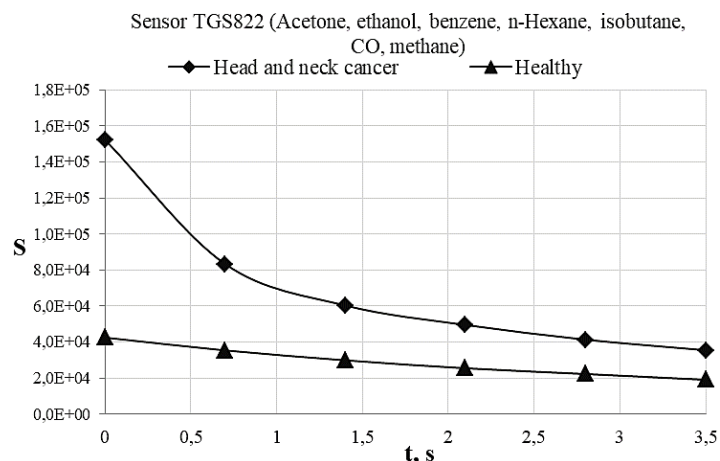


Рис. 1. Усредненные отклики датчика TGS822 на выдыхаемый воздух
Fig. 1. Averaged responses of the TGS822 sensor to exhaled air

Отличие форм кривых и соответствующих значений сигналов на графиках показывает изменение в составе выдыхаемого воздуха, в частности в концентрациях активных молекул. Такие отличия в усредненных данных для пациентов с патологией и здоровых добровольцев являются явным признаком дифференциации добровольцев для искусственной нейронной сети при ее обучении и верификации.

В ходе подготовки данных для последующей классификации нейронной сетью с помощью программного комплекса формируется массив-строка содержащий последовательно выстроенные ряды значений от всех 24 датчиков, каждый из которых включает соотношение 4 и 1 периодов термоциклирования от точки начала нагрева до точки начала остывания на периоде 3,5 с. С частотой дискретизации 50 Гц формируется 175 значений для каждого отдельного ряда. В ходе исследования экспериментальным путем выявлено, если брать каждое 10-е значение отдельного ряда, то погрешность не увеличивается; таким образом, был существенно уменьшен размер входного массива данных со 175 значений для каждого датчика до 17. В результате на вход нейросетевого классификатора по каждому отдельному обследуемому добровольцу подавался массив-строка из 17×24 значений, полученных с газовых датчиков. Общая размерность входного слоя нейронной сети составила 408 значений. На текущем этапе работ не проводились исследования, направленные на улучшение точности нейросетевого классификатора за счет применения информации о возрасте, гендерной принадлежности, локализации опухолевого процесса, факте курения и употребления алкоголя, а также длительного приема лекарственных препаратов по поводу возможных имеющихся хронических заболеваний и других факторов.

2. Нейросетевой классификатор данных

В ходе исследования проведено несколько серий экспериментов, при которых определялись возможности нейронной сети по классификации данных от здоровых добровольцев и пациентов со злокачественными новообразованиями головы и шеи. Такой подход позволяет использовать характерные преимущества бинарных классификаторов. В перспективе этот метод может использоваться в предварительной диагностике разных заболеваний путем последовательного применения отдельных нейросетевых бинарных классификаторов.

На верхнем функциональном уровне алгоритм программного комплекса нейросетевой классификации данных проводит предварительный отбор проб выдыхаемого воздуха у пациентов с наличием патологии и добровольцев без патологического процесса. Далее рассчитывается показатель порога разделения положительных и отрицательных проб, а затем с использованием программного комплекса предварительно диагностируется злокачественное новообразование головы и шеи при определенных значениях выходного нейрона искусственной нейронной сети. Значения формируются предваритель-

но обученной нейронной сетью на соответствующем наборе добровольцев, включающем 48 человек в соотношении примерно 50:50 – с наличием и без патологии.

Архитектура нейронной сети синтезируется для достижения определенной цели. Учитывается большое количество параметров, таких как законы соответствия выборок исходным образам, размерности, природа изменчивости входных и выходных данных и т.д. Существуют общие рекомендации для выбора архитектуры нейронной сети, для начального определения ее размерности, передаточных функций нейронов, форматов данных и т.д. Если отсутствует возможность подобрать архитектуру под конкретную цель, то она выбирается экспериментальным путем из наиболее близких вариантов [14].

Исходя из поставленной задачи предварительной диагностики заболевания, наиболее подходящим вариантом архитектуры нейронной сети для классификации здоровых добровольцев и пациентов с патологическим процессом является нейронная сеть прямого распространения типа многослойный перцептрон. Наличие одного или нескольких скрытых слоев в такой сети позволяет осуществлять нелинейные преобразования разнообразных сигналов. При этом вопрос выбора размера скрытого слоя в каждом конкретном случае, как правило, решается опытным путем. Количество и размерность скрытого слоя в ходе нашего исследования варьировали для достижения максимальной точности и обеспечения работоспособности классификатора на имеющемся статистическом наборе данных. Общая размерность входного слоя нейронной сети соответствует количеству предварительно обработанных выходных сигналов с датчиков и составляет 408. Выходной слой – один нейрон, принимающий значение в диапазоне от -1 до 1 и разграниченный порогом разделения положительных и отрицательных проб. Нормировка на такой диапазон сделана для всех функций активации программного комплекса для удобства анализа диаграмм перекрестной проверки. Значения нейронов выходного слоя представляют собой вероятность. Значение, близкое к 1 выше порога разделения на выходном нейроне, означает соответствие обрабатываемых данных набору пациентов с патологией, ниже порога до -1 – здоровым добровольцам. Рекуррентные нейронные сети не подходят для поставленной задачи из-за наличия обратных связей, что делает их ближе к задачам предсказания трендов, чем к задачам бинарной классификации. Радиальные нейронные сети имеют сложности в обучении, особенно для используемых размерностей, что впоследствии приведет к значительному снижению точности классификации.

Обучение искусственной нейронной сети осуществлялось с помощью алгоритма обратного распространения сигнала ошибки. Этот метод представляет собой разновидность градиентного метода, оптимизированного с учетом архитектуры многослойного перцептрона. С его помощью производилась минимизация функции ошибки, которая определена в пространстве параметров нейронной сети. Параметрами сети в этом случае являются весовые коэффициенты всех нейронов. Успешное завершение процесса обучения сети определялось по минимуму ошибки на тестовом множестве примеров обработки данных. Тестовое множество выделялось из исходного множества имеющегося набора данных проб выдыхаемого воздуха, из которого также выделялось и обучающее множество. Тестовые и обучающие множества при этом не пересекались.

При определении оптимальных параметров обучения нейронной сети на имеющемся наборе данных варьировали гиперпараметры. Для базовой архитектуры нейронной сети применена нормализация – алгоритм, согласно которому некоторым слоям нейронной сети на вход подаются данные, предварительно обработанные и имеющие нулевое математическое ожидание и единичную дисперсию. В ходе исследования осуществлялся подбор подходящих функций активации. Они выбираются в зависимости от задачи и входных данных. Также возможно использование разных функций активации на разных слоях нейронной сети. При исследовании влияния типа функций активации в скрытых слоях на качество работы нейронной сети с имеющимся набором данных удалось установить то, что при использовании линейных функций активации `linear` в каждом скрытом слое нейронная сеть сохраняет свое качество с незначительным отклонением.

В ходе исследования проверялась зависимость эффективности нейронной сети от параметра скорости обучения `learning_rate` с оптимизацией по алгоритму глубокого обучения `adam`. В результате

оптимальными являются значения в диапазоне $\text{learning_rate} = 0,0008 \div 0,001$. Исследование зависимости эффективности классификатора от количества скрытых слоев показало, что уменьшение их количества с 3 до 1 с шагом 1 не приводит к увеличению точности классификатора, однако необходимо наличие как минимум одной нормализации между скрытыми и выходным слоями для сохранения его точности.

3. Структура и функции программного комплекса

Все компоненты программного комплекса нейросетевой классификации данных газоаналитического обследования дыхательной системы взаимодействуют между собой двумя способами: через системные переменные и вызов функций с параметрами или без них. Структура программного комплекса представлена на рис. 2.

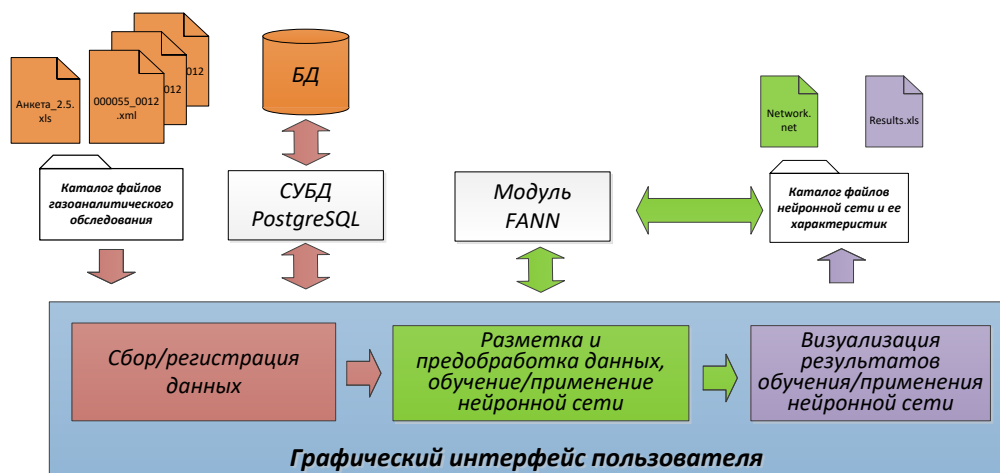


Рис. 2. Структура программного комплекса нейросетевой классификации данных газоаналитического обследования дыхательной системы

Fig. 2. The structure of the neural network classification software for the gas analytical examination data of respiratory system

Графический интерфейс пользователя обеспечивает функции переключения экранов по стадиям обработки данных, переключения режимов работы комплекса и отображение процесса их выполнения, защиты от неправильных действий пользователя, а также индикации текущего состояния процесса обработки данных. В нем реализован экран для выполнения функций сбора и регистрации данных. С помощью него данные из файла «Анкета_X.X.xls» с метаданными по добровольцам и многочисленным файлам «Номер пробы_серийный номер устройства.xml», в которых содержатся соответствующие оцифрованные сигнальные паттерны по каждому отдельному добровольцу, предварительно помещенных в каталог файлов газоаналитического обследования, заносятся и структурируются в базу данных. Перенос данных в базу осуществляется для их последующей автоматизированной обработки и анализа. Для взаимодействия программного комплекса с базой данных применяется система управления PostgreSQL. Связь программного комплекса с базой данных осуществляется при помощи стандартизированных SQL запросов.

Задача проектирования логической структуры базы данных газоаналитического обследования дыхательной системы пациентов с исследуемыми патологическими процессами (плоскоклеточный рак гортани, рак орофарингеальной области) и добровольцев контрольной группы без опухолевых патологий состояла в определении состава таблиц и атрибутов (полей) таблиц базы. Для каждого поля определены тип, название и сформировано его содержание. С целью минимизации объема отдельных таблиц каждому полю присвоено уникальное обозначение (классификатор). При этом обозначения выбирались с учетом сохранения логической связи с названиями полей: например, поле Medical address обозначено символами MED_ADDRESS, полю Type treatment присвоено обозначение TYPE_TREAT и т.д.

Разработанная логическая и физическая структура базы данных позволяет сохранять неструктурированные метаданные, а также осуществлять оперативный поиск необходимой информации по наборам задаваемых атрибутов средствами системы управления. Каждая сущность логической структуры базы данных имеет уникальное имя, которое отражает ее смысловое значение. Для каждой сущности заданы ее подробное описание и особенности взаимодействия с другими таблицами. Регистрация данных в базе выполняется периодически после накопления очередной порции данных в устройствах анализа состава газовой смеси в выдыхаемом воздухе, размещаемых в медицинских учреждениях.

Этап обработки данных выполняется с помощью экрана графического интерфейса, в котором реализованы функции разметки и предобработки данных, обучения и применения нейронной сети. До проведения нейросетевого анализа данных сигнальные паттерны, зарегистрированные в базе, могут подвергаться предварительной обработке (нормировке, приведению к диапазону, выполнению калибровочной функции, численному сокращению, преобразованию типа и т.д.), после чего они становятся доступными для нейросетевого анализа.

Функция нейросетевого анализа данных выполняется в двух режимах. Первый режим предусматривает выбор записей в базе для обучения искусственной нейронной сети, настройку гиперпараметров, обучение и синтез нейронной сети в файл Network.net. Во втором режиме осуществляется применение существующей сети для классификации отдельных записей из базы данных.

Разработка программного комплекса осуществлялась в среде Qt под управлением операционной системы Linux Ubuntu. Для программной реализации искусственной нейронной сети использовались встраиваемые библиотеки. После анализа существующих библиотек, написанных на языке C/C++, был сделан выбор в пользу FANN (Fast Artificial Neural Network). Данная библиотека создавалась с использованием окружения операционной системы GNU/Linux, языка C и компилятора gcc, что совместимо с конфигурацией используемых программных модулей и компонентов в программном комплексе.

После запуска экрана обработки данных происходит проверка того, какое действие необходимо выполнить. Если необходимо обучить нейронную сеть, то запускается функция, в результате работы которой формируется выходной файл Network.net, содержащий параметры обученной нейронной сети: количество слоев, количество нейронов в каждом слое, тип активационной функции, алгоритм обучения, количество рассчитанных весов и их значения. При обучении нейронной сети задаются параметры с помощью функции `fann_set_{training_algorithm, learning_rate, activation_function_hidden, activation_function_output, train_error_function, quickprop_decay}` и др.}. Если программный комплекс работает в режиме применения существующей нейронной сети, обученной ранее, то ее загрузка из файла Network.net происходит с помощью функции `fann_create` с параметрами. Файл, содержащий параметры и весовые коэффициенты нейронной сети, находится в рабочем каталоге программного комплекса.

После того как нейронная сеть обучена и все ее выходные файлы сгенерированы в рабочий каталог, с помощью функции визуализации результатов обучения и применения нейронной сети в графическом интерфейсе происходит формирование отчета об эффективности бинарного классификатора на обучающем наборе с выбранными гиперпараметрами. Данные отчета заносятся в файл Results.xls. При этом, если до этого выполнялось действие по обучению нейронной сети, файл будет содержать зависимости потерь и точности при обучении и верификации от количества эпох для обучающих и тестовых наборов данных с визуализацией t-SNE в виде таблиц значений и графиков [15], диаграмму распределения классов при выполнении перекрестной проверки, ROC-характеристику, показатель AUC [16] и т.д. Если же выполнялось действие по применению нейронной сети, то в файл отчета добавятся сведения о значении выходного нейрона и классификации обрабатываемого набора данных.

4. Применение программного комплекса

В тестовом эксперименте рассматривался набор данных от 23 здоровых добровольцев и 25 пациентов с плоскоклеточным раком гортани и орофарингеальной области. После обучения нейронной

сети и перекрестной проверки были получены результаты, представленные на рис. 3, 4. Точность предварительной диагностики составила 79,17%, чувствительность – 80,12%, специфичность – 78,26%.

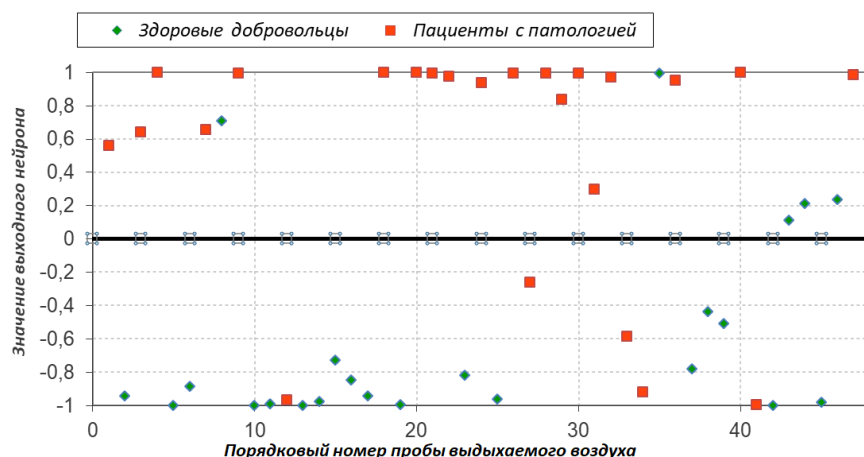


Рис. 3. Диаграмма классификации проб при перекрестной проверке для всех здоровых добровольцев и пациентов с плоскоклеточным раком гортани и орофарингеальной области

Fig. 3. The samples classification diagram of cross-validation for all healthy volunteers and patients with squamous cell carcinoma of the larynx and oropharynx

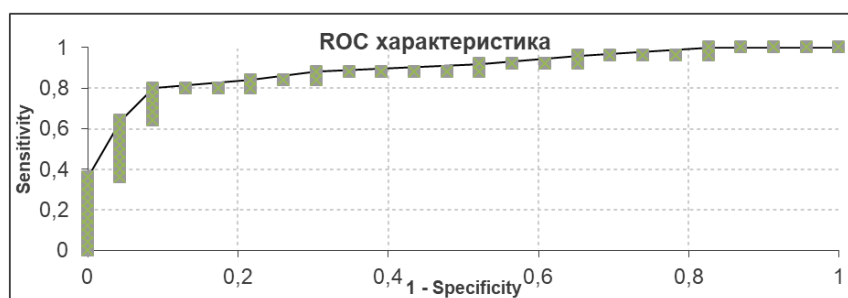


Рис. 4. ROC-характеристика для всех здоровых добровольцев и пациентов с плоскоклеточным раком гортани и орофарингеальной области

Fig. 4. ROC curve for all healthy volunteers and patients with squamous cell carcinoma of the larynx and oropharynx

Порог разделения положительных и отрицательных тестов был найден с помощью анализа ROC-характеристики и составил 0,071. Величина AUC-ROC составила 0,873, что говорит о хорошем качестве классификатора.

В проведенном исследовании состав добровольцев по группам участников сопоставим по полу и возрасту. На начальном этапе исследования изучалась возможность использования сигнальных паттернов выдыхаемого воздуха для дифференциации здоровых добровольцев и пациентов с патологией. Поэтому весь набор данных о добровольцах при обучении и применении нейросетевого классификатора не применялся. В дальнейшем учет дополнительных сведений, таких как факт курения, пол, возраст, локализация опухолевого процесса, стадия заболевания, факт употребления алкоголя, а также длительный прием лекарственных препаратов по поводу возможных имеющихся хронических заболеваний и др., при обработке данных позволит улучшить работу нейронной сети при классификации добровольцев и методику предварительной диагностики в целом.

Заключение

В работе предложены алгоритм и архитектура нейронной сети для создания программного комплекса нейросетевой классификации данных газоаналитического обследования дыхательной системы. Программный комплекс обеспечивает обработку данных, формируемых с помощью специализированных устройств анализа состава газовой смеси в выдыхаемом воздухе. В качестве чувствительных

элементов в каждом отдельном устройстве применяется набор неселективных полупроводниковых газовых датчиков, синхронизированных между собой и работающих в режиме термоциклирования. Преимущества созданного программного комплекса заключаются в автоматизации операций по обработке данных за счет интеграции инструментов системы управления базой данных и нейросетевого анализа данных. Это позволяет структурировать записи в базе, переобучать и дообучать нейронные сети при увеличении объемов данных с пробами выдыхаемого воздуха, что, в свою очередь, обеспечивает формирование рабочих нейронных сетей для целевых медицинских приборов предварительной диагностики заболеваний дыхательной системы. В нашем исследовании программный комплекс применялся для обработки данных от проб выдыхаемого воздуха, отобранных у здоровых добровольцев и пациентов со злокачественными новообразованиями орорингеальной области и гортани.

Вследствие полносвязности искусственной нейронной сети алгоритмическая сложность вычислений пропорциональна квадрату числа нейронов, что влечет за собой значительные временные затраты на обучение сетей большого размера [14]. Поскольку обучение нейронной сети включает в себя большое количество необходимых математических операций, которые поддаются распараллеливанию, при дальнейшем развитии программного комплекса нейросетевой классификации данных газоаналитического обследования дыхательной системы необходимо заложить функции проведения расчетов с применением вычислительных ресурсов графических ускорителей и программной библиотеки CUDANN.

На текущем этапе исследования полученные показатели чувствительности и специфичности сопоставимы с современными высокоточными рентгенологическими методами диагностики опухолей дыхательных путей (МРТ, СКТ, ПЭТ) [13, 17]. Достигнутые показатели эффективности нейросетевого классификатора данных газоаналитического обследования дыхательной системы отражают его перспективность в предварительной диагностике опухолевых процессов у больных раком орорингеальной области и гортани.

Список источников

1. Shah J.P., Patel S.G. Head and neck surgery and oncology. Mosby, 2013. 713 p.
2. Zhuikova L.D., Choyazonov E.L., Ananina O.A., Odintsova I.N. Cancer incidence in Siberia and Russian Far East // *Siberian Journal of Oncology*. 2019. V. 18, is. 6. P. 5–11. doi: 10.21294/1814-4861-2019-18-6-5-11
3. Krilaviciute A., Stock C., Leja M., Brenner H. Potential of non-invasive breath tests for preselecting individuals for invasive gastric cancer screening endoscopy // *J. Breath Res.* 2018. V. 12. Art. 036009. doi: 10.1088/1752-7163/aab5be
4. Hirsch F.R., Franklin W.A., Gazdar A.F., Bunn P.A. Early detection of lung cancer: Clinical perspectives of recent advances in biology and radiology // *Clinical Cancer Research*. 2001. V. 7. P. 5–22.
5. Harris A., Lyu L., Wasserman-Winko T., George S., Johnson J.T., Nilsen M.L. Neck Disability and Swallowing Function in Post-treatment Head and Neck Cancer Patients // *Otolaryngology – Head and Neck Surgery*. 2020. V. 163 (4). P. 763–770. doi: 10.1177/0194599820923630
6. Clinical Practice Guidelines in Oncology // The National Comprehensive Cancer Network. URL: https://www.nccn.org/professionals/physician_gls/default.aspx
7. Behera B., Joshi R., Anil Vishnu G.K., Bhalerao S., Pandya H.J. Electronic nose: A non-invasive technology for breath analysis of diabetes and lung cancer patients // *Journal of Breath Research*. 2019. V. 13 (2). Art. 024001.
8. Chernov V.I., Choyazonov E.L., Kulbakin D.E., Obkhodskaya E.V., Obkhodskiy A.V., Popov A.S., Sachkova A.S., Sachkov V.I. Cancer Diagnosis by Neural Network Analysis of Data from Semiconductor Sensors // *Diagnostics*. 2020. V. 10 (9). Art. 677. doi: 10.3390/diagnostics10090677
9. Gordon S.M., Szidon J.P., Krotoszynski B.K., Gibbons R.D., O'Neill H.J. Volatile organic compounds in exhaled air from patients with lung cancer // *Clinical Chemistry*. 1985. V. 31 (8). P. 1278–1282.
10. Phillips M., Cataneo R.N., Cummin A.R.C., Gagliardi A.J., Gleeson K., Greenberg J., Maxfield R.A., Rom W.N. Detection of lung cancer with volatile markers in the breath // *Chest*. 2003. V. 123 (6). P. 2115–2123.
11. Kesavaraj G., Sukumaran S. A study on classification techniques in data mining // 4th International Conference on Computing, Communications and Networking Technologies. ICCCNT 2013. Art. 6726842.
12. Blatt R., Bonarini A., Calabró E., Della Torre M., Matteucci M., Pastorino U. Lung cancer identification by an electronic nose based on an array of MOS sensors // *IEEE International Conference on Neural Networks – Conference Proc.* 2007. P. 1423–1428.
13. Rivera M.P., Mehta A.C., Wahidi M.M. Establishing the Diagnosis of Lung Cancer // *Chest*. 2013. V. 143 (5). P. e142S–e165S.
14. Сорокоумова Д.А., Корелин О.Н., Сорокоумов А.В. Построение и обучение нейронной сети для решения задачи распознавания речи // *Информатика и системы управления*. 2015. № 3 (110). С. 77–83.

15. van der Maaten L.J.P., Hinton G.E. Visualizing High-Dimensional Data Using t-SNE // *Journal of Machine Learning Research*. 2008. V. 9. P. 2579–2605.
16. Sokolova M., Japkowicz N., Szpakowicz S. Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation // *Advances in Artificial Intelligence*. 2006. V. 4304. P. 1015–1021. URL: <https://cdn.aaai.org/Workshops/2006/WS-06-06/WS06-06-006.pdf>
17. Becker M., Zaidi H. Imaging in head and neck squamous cell carcinoma: the potential role of PET/MRI // *Br. J. Radiol.* 2014. V. 87 (1036). Art. 20130677.

References

1. Shah, J.P. & Patel, S.G. (2013) *Head and Neck Surgery and Oncology*. Mosby.
2. Zhuikova, L.D., Choyznzonov, E.L., Ananina, O.A. & Odintsova, I.N. (2019) Cancer incidence in Siberia and Russian Far East. *Siberian Journal of Oncology*. 18(6). pp. 5–11. DOI: 10.21294/1814-4861-2019-18-6-5-11
3. Krilaviciute, A., Stock, C., Leja, M. & Brenner, H. (2018) Potential of non-invasive breath tests for preselecting individuals for invasive gastric cancer screening endoscopy. *Journal of Breath Research*. 12. Art. 036009. DOI: 10.1088/1752-7163/aab5be
4. Hirsch, F.R., Franklin, W.A., Gazdar, A.F. & Bunn, P.A. (2001) Early detection of lung cancer: Clinical perspectives of recent advances in biology and radiology. *Clinical Cancer Research*. 7. pp. 5–22.
5. Harris, A., Lyu, L., Wasserman-Winko, T., George, S., Johnson, J.T. & Nilsen, M.L. (2020) Neck Disability and Swallowing Function in Posttreatment Head and Neck Cancer Patients. *Otolaryngology – Head and Neck Surgery*. 163(4). pp. 763–770.
6. The National Comprehensive Cancer Network. (n.d.) *Clinical Practice Guidelines in Oncology*. [Online] Available from: https://www.nccn.org/professionals/physician_gls/default.aspx
7. Behera, B., Joshi, R., Anil Vishnu, G.K., Bhalariao, S. & Pandya, H.J. (2019) Electronic nose: A non-invasive technology for breath analysis of diabetes and lung cancer patients. *Journal of Breath Research*. 13(2). Art. 024001.
8. Chernov, V.I., Choyznzonov, E.L., Kulbakin, D.E., Obkhodskaya, E.V., Obkhodskiy, A.V., Popov, A.S., Sachkova, A.S. & Sachkov, V.I. (2020) Cancer Diagnosis by Neural Network Analysis of Data from Semiconductor Sensors. *Diagnostics*. 10(9). Art. 677. DOI: 10.3390/diagnostics10090677.
9. Gordon, S.M., Szidon, J.P., Krotoszynski, B.K., Gibbons, R.D. & O'Neill, H.J. (1985) Volatile organic compounds in exhaled air from patients with lung cancer. *Clinical Chemistry*. 31(8). pp. 1278–1282.
10. Phillips, M., Cataneo, R.N., Cummin, A.R.C., Gagliardi, A.J., Gleeson, K., Greenberg, J., Maxfield, R.A. & Rom, W.N. (2003) Detection of lung cancer with volatile markers in the breath. *Chest*. 123(6). pp. 2115–2123.
11. Kesavaraj, G. & Sukumaran, S. (2013) A study on classification techniques in data mining. *4th International Conference on Computing, Communications and Networking Technologies, ICCCNT*. Art. 6726842.
12. Blatt, R., Bonarini, A., Calabró, E., Della Torre, M., Matteucci, M. & Pastorino, U. (2007) Lung cancer identification by an electronic nose based on an array of MOS sensors. *IEEE International Conference on Neural Networks – Conference Proceedings*. pp. 1423–1428
13. Rivera, M.P., Mehta, A.C. & Wahidi, M.M. (2013) Establishing the Diagnosis of Lung Cancer. *Chest*. 143(5). e142S–e165S.
14. Sorokoumova, D., Korelin, O. & Sorokoumov, A. (2015) Postroenie i obuchenie neyronnoy seti dlya resheniya zadachi raspoznavaniya rechi [Development and training of the neural network for voice recognition solution]. *Informatika i sistemy upravleniya – Computer Science and Management Systems*. 3(110). pp. 77–83.
15. van der Maaten, L.J.P. & Hinton, G.E. (2008) Visualizing High-Dimensional Data Using t-SNE. *Journal of Machine Learning Research*. 9. pp. 2579–2605.
16. Sokolova, M., Japkowicz, N. & Szpakowicz, S. (2006) Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation. *Advances in Artificial Intelligence*. 4304. pp. 1015–1021.
17. Becker, M. & Zaidi, H. (2014) Imaging in head and neck squamous cell carcinoma: the potential role of PET/MRI. *British Journal of Radiology*. 87(1036). Art. 20130677.

Информация об авторах:

Обходский Артём Викторович – кандидат технических наук, доцент инженерной школы ядерных технологий Национального исследовательского Томского политехнического университета (Томск, Россия). E-mail: art707@yandex.ru

Кульбакин Денис Евгеньевич – доктор медицинских наук, заведующий отделением Научно-исследовательского института онкологии филиала Томского национального исследовательского медицинского центра Российской академии наук (Томск, Россия). E-mail: kulbakin_d@mail.ru.

Обходская Елена Владимировна – кандидат технических наук, старший научный сотрудник лаборатории химических технологий Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: fil330a@yandex.ru

Попов Александр Сергеевич – младший научный сотрудник лаборатории химических технологий Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: asptomsktpu@gmail.com

Родионов Евгений Олегович – кандидат медицинских наук, старший научный сотрудник Научно-исследовательского института онкологии филиала Томского национального исследовательского медицинского центра Российской академии наук; ассистент кафедры онкологии Сибирского государственного медицинского университета (Томск, Россия). E-mail: rodionov_eo@oncology.tomsk.ru

Сачков Виктор Иванович – доктор химических наук, заведующий лабораторией химических технологий Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: vicsachkov@gmail.com

Чернов Владимир Иванович – доктор медицинских наук, профессор, член-корреспондент РАН, заместитель директора по научной и инновационной работе Томского НИМЦ; заведующий отделением радионуклидной терапии и диагностики Научно-исследовательского института онкологии филиала Томского национального исследовательского медицинского центра Российской академии наук (Томск, Россия). E-mail: chernov@tnimc.ru

Чойнзон Евгений Лхамцыренович – доктор медицинских наук, профессор, академик РАН, директор Научно-исследовательского института онкологии филиала Томского национального исследовательского медицинского центра Российской академии наук; заведующий кафедрой онкологии Сибирского государственного медицинского университета (Томск, Россия). E-mail: choynzonov@tnimc.ru

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

Information about the authors:

Obkhodskiy Artem V. (Candidate of Technical Sciences, Associate Professor, School of Nuclear Technology, National Research Tomsk Polytechnic University, Tomsk, Russian Federation). E-mail: art707@yandex.ru

Kulbakin Denis E. (Doctor of Medical Sciences, Head of Department of Head and Neck Tumors, Cancer Research Institute, Tomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, Russian Federation). E-mail: kulbakin_d@mail.ru

Obkhodskaya Elena V. (Candidate of Technical Sciences, Senior Researcher, Laboratory of Chemical Technologies, Chemical faculty, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: fil330a@yandex.ru

Popov Aleksandr S. (Junior Researcher, Laboratory of Chemical Technologies, Chemical faculty, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: asptomsktpu@gmail.com

Rodionov Evgeniy O. (Candidate of Medical Sciences, Senior Researcher, Department of Thoracic Oncology, Cancer Research Institute, Tomsk National Research Medical Center of the Russian Academy of Sciences; Assistant, Department of Oncology, Siberian State Medical University, Tomsk, Russian Federation). E-mail: rodionov_eo@oncology.tomsk.ru

Sachkov Victor I. (Doctor of Chemistry Sciences, Associate Professor, Head of the Laboratory of Chemical Technologies, Chemical faculty, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: vicsachkov@gmail.com

Chernov Vladimir I. (Doctor of Medical Sciences, Professor, Deputy Director for Science and Innovation, Head of Nuclear Medicine Department, Cancer Research Institute, Tomsk National Research Medical Center of the Russian Academy of Sciences, Tomsk, Russian Federation). E-mail: chernov@tnimc.ru

Choynzonov Evgeny L. (Doctor of Medical Sciences, Professor, Member of the Russian Academy of Sciences, Director of the Cancer Research Institute, Tomsk National Research Medical Center of the Russian Academy of Sciences; Head of Oncology Department of Siberian State Medical University, Tomsk, Russian Federation). E-mail: choynzonov@tnimc.ru

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

Received 18.07.2024; accepted for publication 02.12.2024

Поступила в редакцию 18.07.2024; принята к публикации 02.12.2024

Научная статья

УДК 621.382

doi: 10.17223/19988605/69/13

Реализация на современных ПЛИС вычислителя сигмоидной функции активации нейронных сетей табличным методом**Инна Владимировна Ушенина***Пензенский государственный технологический университет, Пенза, Россия, ivl23@yandex.ru*

Аннотация. Вычисление функции сигмоида реализуется методом поразрядного отображения. В рамках этого метода аргументы и значения сигмоида представляются в двоичном коде в формате с фиксированной запятой. Каждый разряд значения сигмоида отделен от других и представляется в виде булевой функции от разрядов аргумента или ее таблицы истинности. Оцениваются возможности реализации вычислителей разрядов значений функции сигмоида на блоках программируемой логики ПЛИС. Анализируется два способа реализации: на основе таблиц истинности и на основе минимизированных булевых функций. Во всех реализованных схемах аргументы и значения функции сигмоида имеют равную друг другу разрядность. Схемы, реализованные по таблицам истинности, имеют разрядности от 6 до 11 бит. Показано, что вычислители отдельных разрядов значений функции сигмоида при 7- и 8-разрядном представлении аргумента могут размещаться всего на одном блоке ПЛИС и выполняют вычисления за наименьшее время. Предложенный вариант реализации вычислителя сигмоидной функции может использоваться в составе обученных нейронных сетей, реализуемых аппаратно.

Ключевые слова: нейронная сеть; сигмоид; ПЛИС; табличный метод.

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 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

Original article

doi: 10.17223/19988605/69/13

Realization of the sigmoid activation function for neural networks on current FPGAs by the table-driven method**Inna V. Ushenina***Penza State Technological University, Penza, Russian Federation, ivl23@yandex.ru*

Abstract. In the work, the sigmoid function is implemented using the bit-level mapping method. Within this method, inputs and outputs of the sigmoid function are represented in binary code in fixed-point format. Each output bit is separated from others and represented by a Boolean function of the input bits or its truth table. The possibilities of implementing sigmoid function output bit calculators on FPGA programmable logic blocks are assessed. Two implementation ways are analyzed: on the base of truth tables and on the base of minimized Boolean functions. All implemented circuits have equal bit widths of inputs and outputs to each other. The circuits based on truth tables have bit widths in the range of 6 to 11 bits. It is shown that the sigmoid output bit calculators of 7- and 8-bit inputs occupy just a single programmable logic block and make calculations in the shortest time. The proposed variant of the sigmoid function calculator can be used as a part of trained neural networks implemented in hardware.

Keywords: neural network; sigmoid function; FPGA; table-driven method.

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. (2024) Realization of the sigmoid activation function for neural networks on current FPGAs by the table-driven method. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 69. pp. 124–133. doi: 10.17223/19988605/69/13

Введение

Обучение и развертывание нейронных сетей на аппаратной платформе с возможностью массивно-параллельных вычислений значительно экономит время вычислений, а также позволяет выполнять их в реальном масштабе времени. Доступной платформой, позволяющей осуществлять массивно-параллельные вычисления, являются ПЛИС.

В нейронных сетях типа «многослойный персептрон», обучающихся по алгоритму обратного распространения ошибки, в качестве функции активации нейронов часто используется сигмоид

$$s(x) = \frac{1}{1 + e^{-x}}. \quad (1)$$

Эта функция имеет подходящую форму, непрерывно дифференцируема и требует небольших дополнительных аппаратных затрат на вычисление производной, нужной при обучении.

В настоящее время чаще всего используется три метода вычисления функции сигмоида на ПЛИС, каждый из которых позволяет обойти вычисления частного и экспоненты, напрямую не реализуемые на современных ПЛИС.

Наиболее распространена аппроксимация сигмоида, в частности кусочно-линейная [1–4], когда сигмоид заменяется набором аппроксимирующих отрезков. Недостатком метода является дополнительная погрешность вычислений; преимуществом при реализации на ПЛИС – небольшое количество умножителей или даже их отсутствие [1, 2].

Метод CORDIC (COordinate Rotation DIgital Computer) заключается в постепенном приближении к вычисляемому значению за счет последовательности поворотов вектора и уточнения его координат. Вычислитель такого типа содержит только регистры сдвига и устройства сложения / вычитания и не требует умножителей [5]. Метод CORDIC может обеспечить высокую точность результатов вычислений, но за счет большой латентности схемы, так как каждое уточнение результата требует дополнительной итерации вычислений.

В рамках табличного метода вычислитель представляет собой блок оперативной памяти, хранящий заранее рассчитанные значения функции [6–8]. На его адресные входы подаются значения аргумента. Преимущества метода – меньшая погрешность вычислений по сравнению с аппроксимацией, а также высокая скорость, связанная как с отсутствием латентности, так и с тем, что время доступа к памяти, как правило, меньше времени вычисления. Однако с увеличением разрядности вычислений требуется все больший объем памяти, которой ПЛИС имеют весьма ограниченные запасы.

Разновидность табличного метода – метод поразрядного отображения [9]. В рамках этого метода аргументы и значения сигмоида представляются в двоичном коде в формате с фиксированной запятой. Каждый разряд значения сигмоида отделяется от других и представляется в виде булевой функции от разрядов аргумента или ее таблицы истинности. Разрядность аргумента влияет на размеры таблиц истинности и сложность булевых функций. Применительно к реализации на ПЛИС поразрядное отображение дает преимущество по сравнению со стандартным табличным методом: при работе с отдельными разрядами можно освободить блочную память ПЛИС и реализовать вычисления на распределенной памяти – табличных преобразователях (Look-Up Tables, LUT), которые имеются на ПЛИС в большом количестве в составе блоков программируемой логики.

В [9] и последовавших за ней работах [10–12] отдельные разряды значений функций активации вычисляются по булевым функциям, представленным в минимальной дизъюнктивной нормальной форме (МДНФ). Схемы вычислителей имеют разрядность не более 7 бит, при которой булевы функции достаточно просты и реализуются на малом количестве ресурсов. В [13] показано, что в обученной нейронной сети разрядность вычислителя функции сигмоида должна быть не менее 7–8 бит, так

как именно при этих разрядностях происходит резкое снижение среднеквадратической ошибки выходных сигналов в слоях нейронной сети. Там же показано, что обучение нейронной сети и достижение «жесткой» сходимости, то есть условия

$$|t_j - x_{L,j}| \leq 2^{-4}. \quad (2)$$

требует разрядности вычислений функции сигмоида не менее 9–10 бит. Здесь t_j – целевое значение j -го нейрона выходного слоя сети, $x_{L,j}$ – значение выходного сигнала j -го нейрона выходного слоя сети, L – количество слоев нейронной сети.

Очевидно, что повышение разрядности вычислений значительно усложняет булевы функции и может сделать схемы вычислителей практически нереализуемыми из-за возросшей ресурсоемкости.

Для нейронной сети, развернутой или обучающейся аппаратно, важны также время вычислений и анализ вклада в это время различных влияющих факторов: разрядности вычислителя, времени работы логических элементов, задержек распространения сигналов по соединениям между ними. В работах [9, 11, 12] ресурсоемкость и время вычислений приводятся для 7-битных вычислителей; кроме того, функцией активации не всегда является сигмоид, а целевым устройством – ПЛИС.

В данной статье представлены и проанализированы схемы вычислителей отдельных разрядов значений функции сигмоида (далее – вычислителей) разрядностью от 6 до 11 бит, работающие методом поразрядного отображения и реализованные на блоках программируемой логики ПЛИС. Анализируется два способа реализации: на основе таблиц истинности и на основе минимизированных булевых функций в МДНФ. Для каждой схемы получены характеристики ресурсоемкости, такие как требуемое количество блоков программируемой логики, табличных преобразователей, мультиплексоров, и временные характеристики, включая общее время вычисления и задержку распространения сигналов по соединениям между логическими элементами.

Реализация и анализ схем вычислителей выполнены для двух ПЛИС – XC7A200 семейства Artix7 от Xilinx и GW2AR семейства Arora от Gowin Semiconductor. Ресурсоемкость и время вычисления оценивались с использованием встроенных средств пакетов Vivado IDE и Gowin EDA.

1. Ресурсы современных ПЛИС и вычисление функции сигмоида

Основной ресурс ПЛИС – массив блоков программируемой логики. Название таких блоков зависит от производителя: у Xilinx/AMD это конфигурируемые логические блоки (Configurable Logic Block, CLB), у Gowin – конфигурируемые блоки вычисления функций (Configurable Function Unit, CFU) и т.д. Для реализации комбинационных схем эти блоки содержат следующие ресурсы:

1. LUT – небольшие блоки оперативной памяти, позволяющие реализовать таблицу истинности или имитировать работу комбинационной логики при вычислении булевых функций. Входные сигналы LUT являются адресом нужной ячейки, где хранится 1 бит информации. В современных ПЛИС количество входов у LUT обычно равно четырем или шести.

2. 2-входовые мультиплексоры, которые соединены с выходами LUT.

3. Логика переноса: 2-входовые мультиплексоры и 2-входовые элементы «исключающее ИЛИ».

Поскольку вычисления выполняются отдельно для каждого разряда значения функции сигмоида, 1-битная распределенная память, т.е. LUT, и примыкающие к ней мультиплексоры могут удобно использоваться для реализации схем вычислителей.

Так как сигмоид – симметричная функция, здесь рассматривается только вычисление разрядов ее значений при неотрицательных значениях аргумента, находящихся в диапазоне $[0; 8)$. Вычисления при значениях аргумента в диапазоне $(-8; 0]$ могут быть выполнены аналогично или с использованием схемы вычитания, работающей согласно формуле

$$s(x) = 1 - s(|x|). \quad (3)$$

В настоящей работе этот вопрос не рассматривается.

В работе представлены схемы вычислителей разрядностью от 6 до 11 бит. Во всех схемах аргументы и значения функции имеют равную друг другу разрядность. В представлении аргументов из

всех имеющихся разрядов три старших представляют целую часть, остальные – дробную часть. В представлении значений функции сигмоида, находящихся при неотрицательных аргументах в диапазоне $[0,5; 1)$, все имеющиеся разряды представляют дробную часть.

Таблицы истинности для реализации схем вычислителей первым способом составлялись по таблицам аргументов сигмоида и соответствующих им значений, представленных в двоичном коде с нужной разрядностью. По полученным таблицам истинности определялись кодовые слова, записываемые в LUT. Булевы функции в МДНФ для реализации схем вычислителей вторым способом были получены по соответствующим им таблицам истинности с использованием алгоритма Куайна–МакКласки. Описание всех схем выполнено на языке VHDL.

2. Реализация вычислителей по таблицам истинности (первый способ)

Во многих ПЛИС LUT и мультиплексоры могут быть соединены, как показано на рис. 1, с получением структур, удобных для реализации вычислителей по таблицам истинности. При этом адресными входами LUT будут 4 или 6 младших разрядов аргумента функции сигмоида, а управляющими входами мультиплексоров – оставшиеся старшие разряды.

CFU ПЛИС GW2AR содержит восемь 4-входовых LUT (см. рис. 1, а), значит на одном CFU может быть размещена не более чем 7-разрядная таблица истинности. Для реализации 8-разрядной таблицы истинности понадобится два CFU (рис. 2), на которых размещается примитив 8-разрядного табличного преобразователя LUT8. Это возможно, так как каждый CFU содержит 8 мультиплексоров.

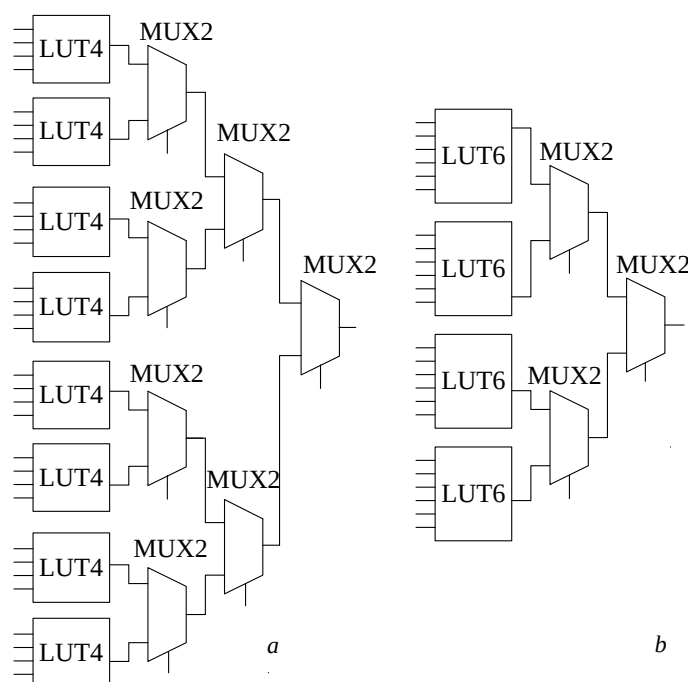


Рис. 1. LUT и мультиплексоры в блоках программируемой логики: 4-входовые LUT и 2-входовые мультиплексоры в CFU ПЛИС GW2AR (а) и 6-входовые LUT и 2-входовые мультиплексоры в CLB ПЛИС XC7A200 (б)

Fig. 1. LUTs and multiplexers in programmable logic blocks: 4-input LUTs and 2-input multiplexers in CFUs of GW2AR FPGAs (a) and 6-input LUTs and 2-input multiplexers in CLBs of XC7A200 FPGAs (b)

Для вычислителей большей разрядности нужно несколько примитивов LUT8, выходы которых должны поступать на мультиплексор. В табл. 1 представлены характеристики схем вычислителей разрядностью 6–11 бит, реализованных на GW2AR по таблицам истинности. Вычислители разрядностью 6 и 7 бит умещаются в один CFU, и задержки распространения сигнала по внутренним соединениям CFU оказываются равными нулю. 8-разрядный вычислитель реализован на двух CFU и, вероятно, соединение между ними имеет ненулевую задержку, но для примитивов (в данном случае для LUT8)

временной анализатор сообщает только общее время задержки (в данном случае 1,638 нс). Вычислители с разрядностью 9, 10 и 11 бит содержат соответственно 2, 4 и 8 примитивов LUT8, выходы которых поступают соответственно на 2-, 4- и 8-входной мультиплексор, реализованный еще на одном CFU.

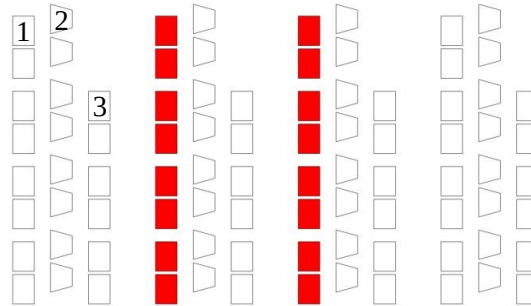


Рис. 2. Результат размещения на двух CFU 8-разрядной таблицы истинности. Фрагмент изображения кристалла GW2AR взят из редактора FloorPlanner среды Gowin EDA. Показаны 4 CFU. 1 – LUT4, 2 – MUX2, 3 – регистры. Занятые LUT выделены редактором. Занятые мультиплексоры редактором не выделяются

Fig. 2. Result of mapping an 8-bit truth table on two CFUs. The piece of GW2AR device image has been taken from Gowin EDA FloorPlanner. Four CFUs are shown. 1 – LUT4, 2 – MUX2, 3 – registers. The occupied LUTs are colored by FloorPlanner. The occupied multiplexers remain uncolored in FloorPlanner

Таблица 1

Характеристики вычислителей, реализованных по таблицам истинности на ПЛИС GW2AR

Разрядность, бит	Кол-во CFU	Общее время вычисления, нс	Задержка по соединениям, нс
6	1	0,780	0
7	1	0,864	0
8	2	1,638	–
9	5	1,887	0,403
10	9	2,369	0,679
11	17	3,092	0,645

CLB ПЛИС XC7A200 содержат четыре 6-входных LUT и три мультиплексора (см. рис. 1, b), значит один CLB позволит реализовать не более чем 8-разрядную таблицу истинности. На рис. 3 представлен результат размещения на ресурсах CLB ПЛИС XC7A200 8-разрядной таблицы истинности, которая реализована на четырех примитивах LUT6 и трех мультиплексорах. Для реализации 9-, 10- и 11-разрядных таблиц истинности потребуется соответственно 3, 5 и 9 CLB, из которых на одном CLB реализован 2-, 4- и 8-разрядный мультиплексор, а остальные реализуют 8-разрядные таблицы истинности от младших разрядов, как показано на рис. 3.

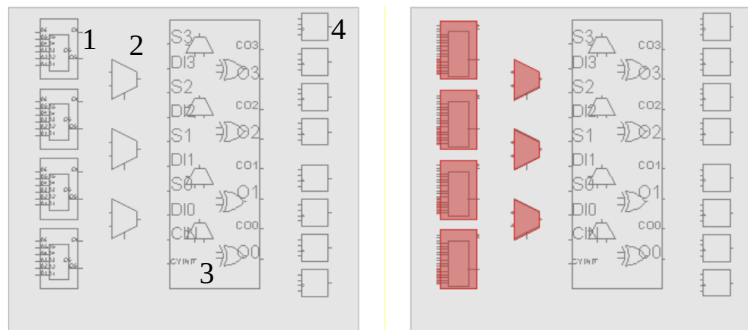


Рис. 3. Результат размещения на одном CLB 8-разрядной таблицы истинности. Фрагмент изображения кристалла XC7A200 взят из редактора FloorPlanner Vivado IDE. Показано 2 CLB. 1 – LUT6, 2 – MUX2, 3 – логика переноса, 4 – регистры.

Занятые LUT и мультиплексоры выделены редактором

Fig. 3. Result of mapping an 8-bit truth table on a single CLB. The piece of XC7A200 device image has been taken from Vivado IDE FloorPlanner. Two CLB are shown. 1 – LUT6, 2 – MUX2, 3 – carry logic, 4 – registers. The occupied LUTs and multiplexers are colored by FloorPlanner

Таблица 2

Характеристики вычислителей, реализованных по таблицам истинности на ПЛИС XC7A200

Разрядность, бит	Кол-во CLB	Общее время вычисления, нс	Задержка по соединениям, нс
6	1	0,199	0
7	1	0,257	0
8	1	0,324	0
9	3	1,134	0,55
10	5	1,294	0,710
11	9	1,231	0,5

Характеристики схем вычислителей с разрядностью 6–11 бит, реализованных на ПЛИС XC7A200 по таблицам истинности, представлены в табл. 2. Вычислители разрядностью 6–8 бит занимают один CLB, причем выходы LUT соединяются только с входами мультимплексоров. Задержки, вносимые этими соединениями в работу вычислителя, равны нулю. Задержки распространения сигналов по соединениям, указанные в табл. 2 для вычислителей разрядностью более 8 бит, соответствуют соединениям между CLB.

3. Реализация вычислителей по минимизированным булевым функциям (второй способ)

Из табл. 1 и 2 видно, что, когда вычислители перестают помещаться в одном блоке программируемой логики, увеличение точности на один разряд приводит к увеличению ресурсоемкости более чем в два раза. На примере 9-разрядных вычислителей проверим, удастся ли снизить ресурсоемкость схем, если реализовывать их на основе минимизированных булевых функций в МДНФ, представленных в табл. 3. Старший, 9-й разряд аргумента функции сигмоида в табл. 3 обозначен буквой a , 8-й разряд – буквой b и т.д. Первый после запятой разряд значения функции сигмоида обозначен $s1$, второй – $s2$ и т.д. При неотрицательных значениях аргумента $s1$ всегда равен единице. Разряд $s2$ представляют 6 импликант; разряд $s3$ – 12; далее количество импликант постепенно нарастает и доходит до 64 для разряда $s9$.

Таблица 3

МДНФ булевых функций для вычисления разрядов значений сигмоида с точностью 9 бит

Разряд	Функция
$s2$	$a + b + cd + ce + cf + cghi$
$s3$	$a + b + \bar{c}di + \bar{c}dh + \bar{c}dg + \bar{c}df + \bar{c}de + \bar{c}d\bar{e}\bar{f}\bar{g} + \bar{c}d\bar{e}\bar{f}\bar{h} + \bar{c}d\bar{e}\bar{f}\bar{i} + defgi + defgh$
$s4$	$a + bc + bde + \bar{b}\bar{c}\bar{d}\bar{e}\bar{i} + \bar{b}\bar{c}\bar{d}\bar{e}\bar{h} + \bar{b}\bar{c}\bar{e}\bar{g} + \bar{b}\bar{c}\bar{e}\bar{f} + \bar{b}\bar{d}\bar{e}\bar{f}\bar{g}\bar{h}\bar{i} + \bar{c}\bar{e}\bar{f}\bar{g} + \bar{c}\bar{e}\bar{f}\bar{h} + \bar{c}\bar{e}\bar{f}\bar{i} + \bar{c}\bar{d}\bar{e} +$ $+ bdfgh + \bar{c}\bar{d}\bar{f} + \bar{c}\bar{d}\bar{g} + \bar{b}\bar{d}\bar{e}\bar{f}\bar{g}\bar{h} + \bar{c}\bar{d}\bar{h}\bar{i} + \bar{b}\bar{c}\bar{e}\bar{h}\bar{i}$
...	...
$s9$	$abc + abd + abe + ac\bar{d}\bar{e} + \bar{a}\bar{b}\bar{c}\bar{d}\bar{f}\bar{g}\bar{h}\bar{i} + \bar{a}\bar{b}\bar{c}\bar{d}\bar{e}\bar{f}\bar{h} + \bar{a}\bar{b}\bar{d}\bar{e}\bar{f}\bar{g}\bar{i} + \bar{a}\bar{b}\bar{d}\bar{e}\bar{h}\bar{i} + \bar{b}\bar{c}\bar{d}\bar{e}\bar{g}\bar{h}\bar{i} + \bar{a}\bar{b}\bar{c}\bar{d}\bar{e}\bar{f}\bar{g}\bar{i} + \bar{b}\bar{c}\bar{d}\bar{e}\bar{f}\bar{g}\bar{h} + \bar{a}\bar{c}\bar{d}\bar{e}\bar{f}\bar{g}\bar{h} +$ $+ \bar{a}\bar{c}\bar{d}\bar{e}\bar{f}\bar{g}\bar{h}\bar{i} + \bar{a}\bar{c}\bar{d}\bar{e}\bar{g}\bar{h}\bar{i} + \bar{b}\bar{c}\bar{d}\bar{e}\bar{f}\bar{h}\bar{i} + \bar{b}\bar{c}\bar{d}\bar{e}\bar{f}\bar{i} + \bar{b}\bar{c}\bar{e}\bar{g}\bar{h}\bar{i} + \bar{b}\bar{c}\bar{d}\bar{e}\bar{g}\bar{h} + \bar{b}\bar{c}\bar{d}\bar{e}\bar{h}\bar{i} + \bar{b}\bar{c}\bar{d}\bar{f}\bar{h} + \bar{b}\bar{c}\bar{d}\bar{f}\bar{g} + \bar{a}\bar{b}\bar{c}\bar{d}\bar{e}\bar{f}\bar{g}\bar{i} +$ $+ \bar{a}\bar{c}\bar{d}\bar{e}\bar{f} + \bar{a}\bar{c}\bar{d}\bar{f}\bar{g} + \bar{a}\bar{c}\bar{d}\bar{e}\bar{f}\bar{g}\bar{i} + \bar{a}\bar{c}\bar{e}\bar{f}\bar{g}\bar{h} + \bar{b}\bar{c}\bar{d}\bar{e}\bar{h} + \bar{a}\bar{b}\bar{d}\bar{g}\bar{h}\bar{i} + \bar{b}\bar{c}\bar{d}\bar{e}\bar{f}\bar{g} + \bar{a}\bar{c}\bar{f}\bar{g}\bar{h}\bar{i} + \bar{a}\bar{b}\bar{c}\bar{d}\bar{e}\bar{f} + \bar{a}\bar{b}\bar{c}\bar{d}\bar{e}\bar{f}\bar{h}\bar{i} + \bar{a}\bar{b}\bar{c}\bar{d}\bar{e}\bar{f}\bar{i} +$ $+ \bar{b}\bar{e}\bar{f}\bar{g}\bar{h}\bar{i} + \bar{a}\bar{c}\bar{d}\bar{e}\bar{f}\bar{g} + \bar{b}\bar{c}\bar{e}\bar{f}\bar{g}\bar{h} + \bar{a}\bar{c}\bar{d}\bar{f}\bar{g}\bar{i} + \bar{c}\bar{d}\bar{e}\bar{f}\bar{g}\bar{h}\bar{i} + \bar{a}\bar{c}\bar{d}\bar{f}\bar{i} + \bar{c}\bar{d}\bar{e}\bar{f}\bar{g}\bar{i} + \bar{b}\bar{e}\bar{f}\bar{g}\bar{h}\bar{i} + \bar{b}\bar{c}\bar{e}\bar{f}\bar{g}\bar{h} + \bar{a}\bar{b}\bar{d}\bar{e}\bar{f}\bar{g}\bar{h} + \bar{a}\bar{b}\bar{d}\bar{e}\bar{g}\bar{h}\bar{i} +$ $+ \bar{a}\bar{c}\bar{e}\bar{f}\bar{g}\bar{h} + \bar{a}\bar{b}\bar{e}\bar{f}\bar{g}\bar{h}\bar{i} + \bar{a}\bar{b}\bar{d}\bar{e}\bar{f}\bar{g}\bar{i} + \bar{a}\bar{b}\bar{d}\bar{e}\bar{f}\bar{h}\bar{i} + \bar{a}\bar{c}\bar{d}\bar{e}\bar{g}\bar{h}\bar{i} + \bar{b}\bar{d}\bar{e}\bar{f}\bar{g}\bar{h}\bar{i} + \bar{b}\bar{d}\bar{e}\bar{f}\bar{g}\bar{i} + \bar{a}\bar{b}\bar{d}\bar{e}\bar{f}\bar{g}\bar{h}\bar{i} + \bar{b}\bar{c}\bar{d}\bar{e}\bar{f}\bar{h}\bar{i} +$ $+ \bar{a}\bar{b}\bar{c}\bar{e}\bar{f}\bar{g}\bar{h}\bar{i} + \bar{a}\bar{b}\bar{d}\bar{e}\bar{f}\bar{g}\bar{h}\bar{i} + \bar{a}\bar{c}\bar{d}\bar{f}\bar{h} + \bar{c}\bar{d}\bar{e}\bar{f}\bar{h}\bar{i} + \bar{a}\bar{d}\bar{e}\bar{g}\bar{h}\bar{i} + \bar{b}\bar{d}\bar{e}\bar{f}\bar{g}\bar{h} + \bar{a}\bar{c}\bar{e}\bar{f}\bar{h}\bar{i} + \bar{a}\bar{b}\bar{c}\bar{e}\bar{f}\bar{g}\bar{i} + \bar{a}\bar{c}\bar{d}\bar{f}\bar{g}\bar{h}\bar{i} + \bar{b}\bar{c}\bar{d}\bar{e}\bar{f}\bar{g}\bar{h} + \bar{a}\bar{d}\bar{e}\bar{f}\bar{g}\bar{h}\bar{i}$

Vivado IDE синтезирует схемы по булевым функциям, используя минимальное количество ресурсов. На рис. 4, a – c показаны синтезированные схемы вычисления разрядов $s2$ – $s5$, которые помещаются в один CLB ПЛИС XC7A200 и реализованы только на LUT. Так, вычисление $s2$ может быть выполнено всего на двух LUT: LUT6, расположенный на рис. 4, a слева, вычисляет функцию $d + e + f + ghi$, а LUT6, расположенный справа, умножает эту функцию на c и складывает ее с a и b . Вычисление $s3$ требует трех LUT; вычисление $s4$ – четырех. Вычисление $s5$ выполняется такой же схемой, что и $s4$.

Схема, синтезированная для разряда $s7$, представлена на рис. 4, d . Схема для разряда $s6$ отличается от представленной на рис. 4, d тем, что два из шести LUT обрабатывают пять, а не шесть входных сигналов. Разряды $s8$ и $s9$ вычислить, используя менее трех CLB, не удалось.

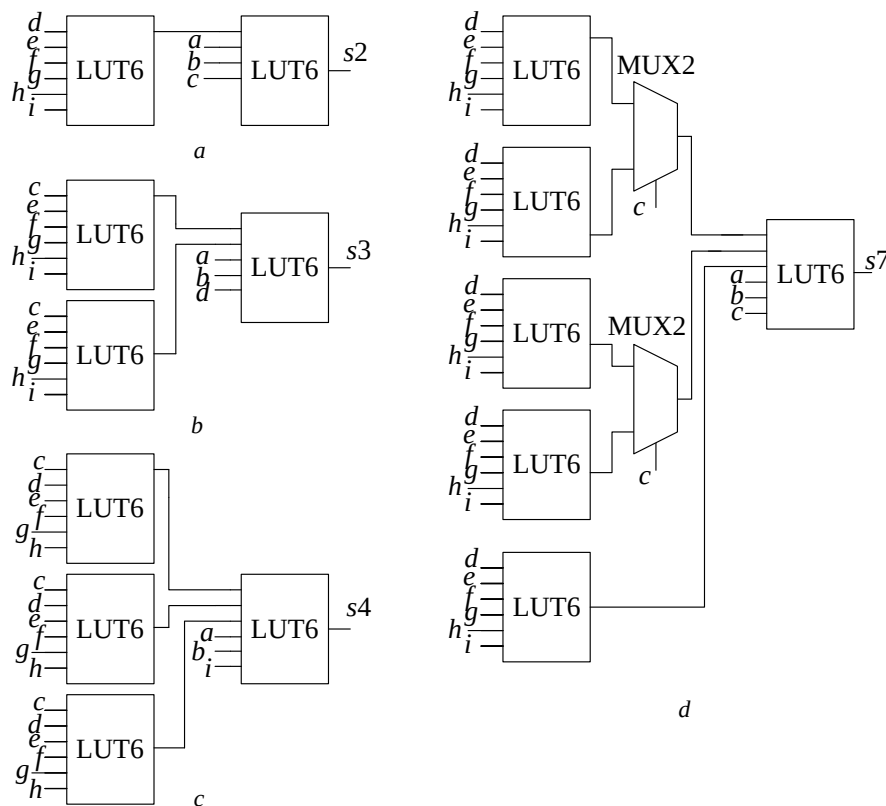


Рис. 4. Схемы вычисления разрядов $s2-s7$ значений функции сигмоида по минимизированным булевым функциям, синтезированные для ПЛИС XC7A200

Fig. 4. Circuits for calculating the sigmoid $s2-s7$ bits using the minimized Boolean functions, synthesized for the XC7A200 FPGA

Характеристики вычислителей разрядов $s2-s7$, синтезированных по булевым функциям, представлены в табл. 4. Общую ресурсоемкость вычислителей разрядов $s2-s9$ удалось снизить с 24 до 14 CLB, т.е. на 42% по сравнению со схемами, выполненными по таблицам истинности.

Таблица 4

Характеристики вычислителей, реализованных по булевым функциям в МДНФ на XC7A200

Разряд	Ресурсоемкость			Время вычисления	
	LUT6	MUX2	CLB	Общее	Задержка распространения по соединениям, нс
$s2$	2	0	1	0,509	0,315
$s3$	3	0	1	0,509	0,315
$s4$	4	0	1	0,695	0,501
$s5$	4	0	1	0,695	0,501
$s6$	6	2	2	0,682	0,214
$s7$	6	2	2	0,779	0,285

Соединения между LUT внутри CLB ПЛИС XC7A200 могут быть реализованы только с использованием внешних трассировочных ресурсов. Поэтому задержки по соединениям для вычислителей разрядов $s2-s5$, размещающихся в одном CLB, отличны от нуля. Тем не менее общее время вычисления в схемах, представленных на рис.4, оказывается на 31–55% меньше, чем в схемах, реализованных по таблицам истинности, за счет снижения задержек, вносимых логическими элементами.

Для ПЛИС GW2AR реализация вычислителей по булевым функциям возможна менее чем на пяти CFU для разрядов $s2-s7$. Схемы вычислителей разрядов $s2$ и $s3$ на ПЛИС GW2AR (рис. 5) отли-

чаются от тех, что были получены для XC7A200. Вычислители $s4$ – $s7$ реализованы в соответствии со схемами, представленными на рис. 4, с, d.

Результаты реализации и временного анализа вычислителей $s2$ – $s7$ на ПЛИС GW2AR приведены в табл. 5. По сравнению с первым способом удалось добиться снижения количества используемых CFU на 42%, что совпадает с результатом, достигнутым на ПЛИС XC7A200. Заметим, что для вычисления $s6$ требуется три CFU, а для $s7$ – четыре, так как в схеме для $s6$ два из пяти LUT обрабатывают пять, а не шесть переменных, а значит, требуют меньше ресурсов.

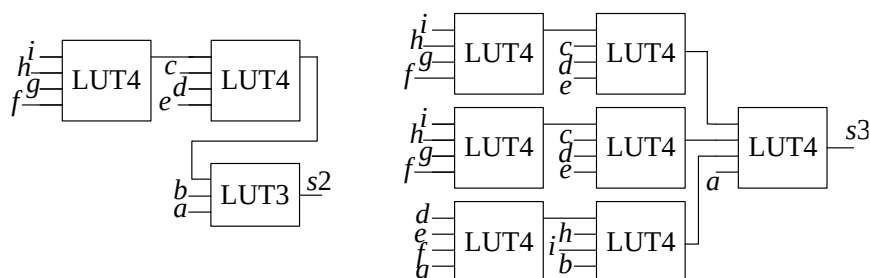


Рис. 5. Схемы вычисления разрядов $s2$ и $s3$ значений функции сигмоида по минимизированным булевым функциям, полученные для ПЛИС GW2AR

Fig. 5. Circuits for calculating the sigmoid $s2$ and $s3$ bits using the minimized Boolean functions, received for the GW2AR FPGA

Таблица 5

Характеристики вычислителей, реализованных по булевым функциям в МДНФ на GW2AR

Разряд	Ресурсоемкость			Время вычисления	
	LUT4	MUX2	CFU	Общее	Задержка распространения по соединениям, нс
$s2$	3	0	1	1,378	0,005
$s3$	7	0	1	1,623	0,148
$s4$	16	12	2	1,690	0,270
$s5$	16	12	2	1,690	0,270
$s6$	22	14	3	2,312	0,598
$s7$	26	18	4	2,288	0,617

Что касается временных характеристик, достигнуть существенного выигрыша при переходе ко второму способу реализации вычислителей не удалось. Для разрядов $s6$ и $s7$ время вычисления оказалось даже больше, чем при реализации вычислителей первым способом. Это частично объясняется тем, что при реализации по таблицам истинности пути, по которым проходят данные, пролегают через три CFU, но на двух из них размещен примитив LUT8, и вносимая им задержка минимизирована производителем. То есть произвольно относительно друг друга могут быть размещены LUT8 (на двух CFU) и мультиплексор (на одном CFU), между которыми – одно соединение. При переходе ко второму способу часть путей проходит через три CFU, которые могут располагаться произвольно относительно друг друга, и задержка распространения по соединениям оказывается выше. Также задержка распространения через цепочку из трех примитивов LUT6–мультиплексор–LUT6 при втором способе (см. рис. 4, d) оказывается больше, чем через LUT8–мультиплексор при первом.

Реализация вторым способом вычислителей с разрядностью 10 бит позволяет добиться меньшей экономии ресурсов и времени. Из-за усложнения булевых функций общее снижение ресурсоемкости при переходе ко второму способу вычисления разрядов $s2$ – $s6$ составляет около 30%. Время вычисления оказывается меньше при использовании второго способа только для разрядов $s2$ и $s3$; вычисление остальных разрядов обоими способами требует примерно равного времени.

Заключение

Вычислители отдельных разрядов значений функции сигмоида при 7- и 8-разрядном представлении аргумента могут размещаться всего на одном блоке программируемой логики и выполняют

вычисления за наименьшее время. Это создает хорошую перспективу для их использования в составе обученных нейронных сетей, реализуемых аппаратно.

Вычисления с точностью 9–10 бит, нужные при обучении нейронных сетей, требуют больше ресурсов и времени. При этом вычисление старших разрядов значений функции сигмоида по представляющим их минимизированным булевым функциям, а не по таблицам истинности, требует меньше логических ресурсов и может сократить время вычислений.

Список источников

1. Alippi C., Storti-Gajani G. Simple approximation of sigmoidal functions: realistic design of digital neural networks capable of learning // Proc. 1991 IEEE International Symposium on Circuits and Systems (ISCAS). Singapore, 1991. V. 3. P. 1505–1508.
2. Amin H., Curtis K.M., Hayes-Gill B.R. Piecewise linear approximation applied to nonlinear function of a neural network // IEE Proc. – Circuits, Devices and Systems. 1997. V. 144 (6). P. 313–317. doi: 10.1049/ip-cds:19971587
3. Tatas K., Gemenaris M. High-Performance and Low-Cost Approximation of ANN Sigmoid Activation Functions on FPGAs // Proc. 12th International Conference on Modern Circuits and Systems Technologies (MOCAS). Athens, Greece, 2023. P. 1–4.
4. 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
5. Chen H., Jiang L., Luo Y., Lu Z., Fu Y., Li L., Yu Z. A CORDIC-Based Architecture with Adjustable Precision and Flexible Scalability to Implement Sigmoid and Tanh Functions // Proc. 2020 IEEE International Symposium on Circuits and Systems (ISCAS). Seville, Spain, 2020. P. 1–5.
6. Sartin M.A., da Silva A.C.R. Approximation of hyperbolic tangent activation function using hybrid methods // Proc. 8th International Workshop on Reconfigurable and Communication-Centric Systems-on-Chip (ReCoSoC). Darmstadt, Germany, 2013. P. 1–6.
7. Saranya S., Elango B. Implementation of PWL and LUT based approximation for hyperbolic tangent activation function in VLSI // Proc. 2014 International Conference on Communication and Signal Processing. Melmaruvathur, India, 2014. P. 1778–1782.
8. Xie Y., Raj A.N.J., Hu Z., Huang S., Fan Z., Joler M. A twofold lookup table architecture for efficient approximation of activation functions // IEEE Transactions on Very Large Scale Integration (VLSI) Systems. 2020. V. 28 (12). P. 2540–2550. doi: 10.1109/TVLSI.2020.3015391
9. Tommiska M.T. Efficient digital implementation of the sigmoid function for reprogrammable logic // IEE Proc. – Computers and Digital Techniques. 2003. V. 150 (6). P. 403–411. doi: 10.1049/ip-cdt:20030965
10. 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.433-440.5647
11. 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
12. 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
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

References

1. Alippi, C. & Storti-Gajani, G. (1991) Simple approximation of sigmoidal functions: realistic design of digital neural networks capable of learning. *Proceedings 1991 IEEE International Symposium on Circuits and Systems (ISCAS)*. 3. pp. 1505–1508.
2. Amin, H., Curtis, K.M. & Hayes-Gill, B.R. (1997) Piecewise linear approximation applied to nonlinear function of a neural network. *IEE Proceedings-Circuits, Devices and Systems*. 144(6). pp. 313–317. DOI: 10.1049/ip-cds:19971587
3. Tatas, K. & Gemenaris, M. (2023) High-Performance and Low-Cost Approximation of ANN Sigmoid Activation Functions on FPGAs. *Proceedings of the 12th International Conference on Modern Circuits and Systems Technologies (MOCAS)*. Athens, Greece. pp. 1–4.
4. 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
5. Chen, H., Jiang, L., Luo, Y., Lu, Z., Fu, Y., Li, L. & Yu, Z. (2020) A CORDIC-Based Architecture with Adjustable Precision and Flexible Scalability to Implement Sigmoid and Tanh Functions. *Proceedings 2020 IEEE International Symposium on Circuits and Systems (ISCAS)*. Seville, Spain. pp. 1–5.
6. Sartin, M.A. & da Silva, A.C.R. (2013) Approximation of hyperbolic tangent activation function using hybrid methods. *Proceedings 8th International Workshop on Reconfigurable and Communication-Centric Systems-on-Chip (ReCoSoC)*. Darmstadt, Germany. pp. 1–6.

7. Saranya, S. & Elango, B. (2014) Implementation of PWL and LUT based approximation for hyperbolic tangent activation function in VLSI. *Proceedings 2014 International Conference on Communication and Signal Processing*. Melmaruvathur, India. pp. 1778–1782.
8. Xie, Y., Raj, A.N.J., Hu, Z., Huang, S., Fan, Z. & Joler, M. (2020) A twofold lookup table architecture for efficient approximation of activation functions. *IEEE Transactions on Very Large-Scale Integration (VLSI) Systems*. 28(12). pp. 2540–2550. DOI: 10.1109/TVLSI.2020.3015391
9. Tommiska, M.T. (2003) Efficient digital implementation of the sigmoid function for reprogrammable logic. *IEE Proceedings-Computers and Digital Techniques*. 150(6). pp. 403–411. DOI: 10.1049/ip-cdt:20030965
10. 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
11. 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
12. 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
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

Информация об авторе:

Ушенина Инна Владимировна – доцент, кандидат технических наук, доцент кафедры «Программирование» Пензенского государственного технологического университета (Пенза, Россия). E-mail: ivl23@yandex.ru

Авторы заявляют 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.

Received 20.07.2024; accepted for publication 02.12.2024

Поступила в редакцию 20.07.2024; принята к публикации 02.12.2024

**ПРОЕКТИРОВАНИЕ И ДИАГНОСТИКА
ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ****DESIGNING AND DIAGNOSTICS OF COMPUTER SYSTEMS**

Научная статья

УДК 519.7

doi: 10.17223/19988605/69/14

**Тестирование многовыходных комбинационных схем
в условиях известных реакций на тестовые наборы****Анжела Юрьевна Матросова¹, Марина Максимовна Савенкова²,
Виктор Алексеевич Провкин³, Ксения Алексеевна Сухорученко⁴**^{1, 2, 3, 4} *Национальный исследовательский Томский государственный университет, Томск, Россия*¹ *mau11@yandex.ru*² *marina1412_11@mail.ru*³ *prowkan@mail.ru*⁴ *ksenia_s.a@bk.ru*

Аннотация. При построении тестовых последовательностей для многовыходных логических схем обычно используются тестовые наборы на множестве входных переменных схемы, такие что реакции на каждый из наборов отличаются в исправном и неисправном состоянии схемы хотя бы на одном из ее выходов. Знание реакций на каждый из наборов может оказаться полезным при построении тестовых последовательностей с целью включения подходящих наборов, ориентированных как на снижение потребляемой при тестировании мощности, так и на обнаружение неисправностей. Предлагается точный алгоритм поиска всех тестовых наборов с соответствующими им реакциями исправной и неисправной схем. Алгоритм основан на выполнении операций над ROBDD-графами, зависящими только от входных переменных, только от выходных переменных и от тех и других переменных вместе. Обсуждаются возможности применения тестовых наборов совместно с их реакциями с целью дебаггинга.

Ключевые слова: комбинационные схемы; тестовые наборы; ROBDD-графы.

Для цитирования: Матросова А.Ю., Савенкова М.М., Провкин В.А., Сухорученко К.А. Тестирование многовыходных комбинационных схем в условиях известных реакций на тестовые наборы // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 69. С. 134–143. doi: 10.17223/19988605/69/14

Original article

doi: 10.17223/19988605/69/14

**Testing of multi-output combinational circuits when reactions
on test patterns are known****Anzhela Yu. Matrosova¹, Marina M. Savenkova², Viktor A. Provkin³, Ksenia A. Sukhoruchenko⁴**^{1, 2, 3, 4} *National Research Tomsk State University, Tomsk, Russian Federation*¹ *mau11@yandex.ru*² *marina1412_11@mail.ru*

³ *prowkan@mail.ru*

⁴ *ksenia_s.a@bk.ru*

Abstract. Deriving test sequences for multi-output combinational circuits is usually connected with applying test patterns on input circuit variables, so that the reaction on each test pattern differs at least by one output from correct one. Finding the suitable reactions on test input patterns may be used when deriving test sequences is oriented to cut power dissipation or to facilitate finding the circuit faults. The precise algorithm of deriving all input test patterns together with their reactions that differs fault and fault free circuits is suggested. It is based on applying operations on ROBDDs depending on only input circuit variables, only output circuit variables and both input and output circuit variables. Some possibilities of using input test patterns together with their reactions for debugging are discussed.

Keywords: combinational circuits; test sequences; ROBDDs.

For citation: Matrosova, A.Yu., Savenkova, M.M., Provkina, 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'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 69. pp. 134–143. doi: 10.17223/19988605/69/14

Введение

Традиционные подходы к тестированию неисправностей многовыходных комбинационных схем связаны с нахождением тестовых наборов для заданной неисправности в виде подходящего множества ее входных наборов. Различия реакций исправной и неисправной схем на каждом из наборов могут быть установлены двоичным моделированием. В случае, когда находятся все тестовые наборы и число их велико, выполнить двоичное моделирование для каждого из наборов практически невозможно. В этой ситуации предлагается воспользоваться операциями над ROBDD-графами, зависящими от различных подмножеств переменных схемы, с тем чтобы получить компактное представление тестовых наборов в виде совокупности ROBDD-графов от входных переменных. Каждому графу сопоставляется пара реакций исправной и неисправной схем (пара наборов значений выходных переменных этих схем). Из полученных наборов и их реакций далее можно выбирать наборы с полезными для нас свойствами. например с целью построения тестовых последовательностей, ориентированных на снижение потребляемой при тестировании мощности или для ускорения процесса поиска неисправности в предъявленной неисправной схеме. Предполагается, что возможны неисправности элементов схемы при условии, что связи между элементами сохраняются.

В данной работе предлагается точный алгоритм поиска тестовых наборов с соответствующими им реакциями на основе представления исправной и неисправной схем однокоренными ROBDD-графами, построенными на множестве входных и выходных переменных [1], ROBDD-графами, построенными на множестве выходных переменных схемы [2], и традиционными ROBDD-графами [3] на множестве входных переменных схемы. Обсуждаются возможности применения тестовых наборов совместно с их реакциями с целью обнаружения неисправностей в предъявленной схеме.

1. Представление многовыходной схемы однокоренным ROBDD-графом

Логическая функция многовыходной схемы может быть представлена в виде однокорневого, направленного, ациклического графа [2], который состоит из внутренних вершин и концевых вершин, называемых 0-терминальными и 1-терминальными вершинами. Каждая внутренняя вершина помечена логической переменной, входной или выходной, и имеет две дочерние вершины. Ребро от внутренней вершины до дочерней вершины задает присвоение переменной, которой помечена рассматриваемая внутренняя вершина, значения 0 или 1 для функции, сопоставляемой дочерней вершине.

Пусть задана многовыходная комбинационная схема C , реализующая систему булевых функций $y_1 = f_1(x_1, x_2, \dots, x_n)$, $y_2 = f_2(x_1, x_2, \dots, x_n)$, ..., $y_m = f_m(x_1, x_2, \dots, x_n)$. На ее основе строится функция f , которая зависит от входных и выходных переменных этой схемы:

$$f(x_1, x_2, \dots, x_n, y_1, y_2, \dots, y_m) = \bigwedge_{i=1}^m [y_i \sim f_i(x_1, x_2, \dots, x_n)], \quad (1)$$

где $\sim$ – логическая операция «эквивалентность».

В дальнейшем будем называть эту функцию входо-выходной функцией схемы. Данная функция обладает следующими свойствами:

1. Если вместо входных переменных схемы C подставить в (1) константы из $\{0, 1\}$, то получим функцию, которая принимает единичное значение на единственном выходном наборе схемы. Этот набор является реакцией схемы C на данный входной набор.
2. Если вместо выходных переменных схемы C подставить в (1) константы из $\{0, 1\}$, то получим функцию, которая принимает единичное значение на всех входных наборах схемы C , обеспечивающих проявление заданной реакции на ее выходах. В частном случае, если заданная реакция недостижима, то полученная функция тождественно равна 0.

Для получения всех достижимых реакций схемы можно построить ROBDD-граф входо-выходной функции схемы C . Если известна структура схемы C , то ROBDD-граф для входо-выходной функции может быть получен непосредственно по ее структуре.

Рассмотрим множество B_N булевых векторов длины $N = n + m$. Каждая из первых n компонент соответствует значениям входных переменных схемы C , m последних компонент вектора соответствуют значениям выходных переменных этой схемы.

Если построить ROBDD-граф для входо-выходной функции схемы C , выбрав порядок разложения от выходных переменных к входным [1], то каждая простая цепь, начинающаяся в корне этого графа и заканчивающаяся в 1-терминальной вершине, соответствует достижимому выходному набору схемы и входному набору либо множеству входных наборов, обеспечивающих его достижение.

Набор значений выходных переменных схемы называется недостижимым, если его невозможно получить на выходах схемы никакой комбинацией значений входных переменных. Недостижимому набору или множеству недостижимых наборов соответствует простая цепь ROBDD-графа для входо-выходной функции схемы C , заканчивающаяся в 0-терминальной вершине и не содержащая внутренних вершин, помеченных входными переменными схемы.

Пусть реакция $\beta = \beta^1 \beta^2 \dots \beta^m$ схемы C является достижимой. Тогда она обеспечивается некоторым множеством $A = \{\alpha_1, \alpha_2, \dots, \alpha_k\}$ входных наборов длины n , то есть входо-выходная функция схемы C принимает единичные значения на множестве булевых векторов длины N вида

$$\begin{aligned} & \alpha_1^1 \alpha_1^2 \dots \alpha_1^n \beta^1 \beta^2 \dots \beta^m \\ & \alpha_2^1 \alpha_2^2 \dots \alpha_2^n \beta^1 \beta^2 \dots \beta^m \\ & \dots \\ & \alpha_k^1 \alpha_k^2 \dots \alpha_k^n \beta^1 \beta^2 \dots \beta^m \end{aligned}$$

Последние m компонент каждого набора не меняют свои значения, поскольку они соответствуют реакции $\beta = \beta^1 \beta^2 \dots \beta^m$ (выходному набору) схемы C на соответствующие входные наборы.

Пусть $\Gamma = \{\gamma_1, \gamma_2, \dots, \gamma_l\}$ – множество входных наборов схемы C , которые не обеспечивают реакцию β . Пространство B_n размерности n разбивается на два подмножества A и Γ , причем на наборах множества Γ

$$\begin{aligned} & \gamma_1^1 \gamma_1^2 \dots \gamma_1^n \beta^1 \beta^2 \dots \beta^m \\ & \gamma_2^1 \gamma_2^2 \dots \gamma_2^n \beta^1 \beta^2 \dots \beta^m \\ & \dots \\ & \gamma_l^1 \gamma_l^2 \dots \gamma_l^n \beta^1 \beta^2 \dots \beta^m \end{aligned}$$

входо-выходная функция схемы C принимает нулевое значение, поскольку реакция β не может быть обеспечена ни одним набором этого множества. Следовательно, пути, соответствующие векторам из множества Γ , ведут в 0-терминальную вершину ROBDD-графа схемы C .

Итак, входо-выходная функция схемы C принимает значение 0 на недостижимых выходных наборах схемы. В этом случае входные переменные N -мерного вектора на соответствующих троичных векторах имеют неопределенные значения.

Кроме того, входо-выходная функция принимает значение 0 на N -мерных векторах, представляющих входные наборы, которые не обеспечивают приписанные им выходные наборы схемы C . Следовательно, входо-выходная функция схемы C принимает нулевые значения на недостижимых входо-выходных состояниях схемы C . Обозначим это множество состояний как I (impossible). Функция схемы C принимает единичные значения на достижимых входо-выходных состояниях схемы. Назовем это множество E (expected). Объединение множеств E, I образует булево пространство размерности 2^N .

Проиллюстрируем сказанное примером. Пусть поведение схемы C описывается системой из трех булевых функций. Достижимые выходные наборы вместе с условиями их достижимости представлены таблицей истинности (табл. 1) входо-выходной функции $f(x_1, x_2, x_3, x_4, y_1, y_2, y_3)$.

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

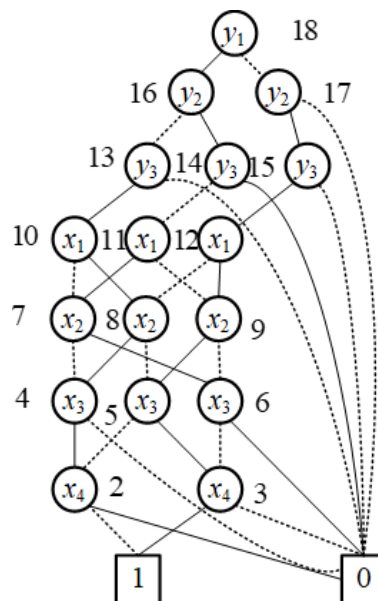


Рис. 1. ROBDD-граф функции $f(x_1, x_2, x_3, x_4, y_1, y_2, y_3)$

Fig. 1. ROBDD of function $f(x_1, x_2, x_3, x_4, y_1, y_2, y_3)$

Достижимым состояниям соответствуют единичные значения входо-выходной функции схемы C . Нулевые значения этой функции в таблице не представлены. На рис. 1 задан ROBDD-граф входо-выходной функции схемы C , зависящий от входных и выходных переменных системы булевых функций, описывающий поведение схемы C .

Простая цепь $y_1 y_2 y_3 \overline{x_1} \overline{x_2} \overline{x_3} x_4$ в ROBDD-графе, ведущая в 1-концевую вершину, представляет входной набор 0101, обеспечивающий реакцию 101 выходных переменных схемы.

Простая цепь $y_1 y_2 y_3 \overline{x_1} x_2 x_3$, ведущая в 0-концевую вершину, представляет входные наборы 0111, 0110, не обеспечивающие реакцию 101 выходных переменных схемы C .

Простая цепь $y_1 y_2 y_3$, 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.

2. Получение множеств всех тестовых наборов совместно с реакциями исправной и неисправной схем

Рассмотрим схему C в исправном и неисправном состояниях. Речь идет о существенных неисправностях, порождающих другую систему булевых функций и, следовательно, другую входо-выходную функцию схемы.

Исправной схеме соответствуют множества E, I , а неисправной сопоставляются множества E', I' .

Множества E, E' (I, I') обязательно различаются, поскольку в исправной и неисправной схемах реализуются различные системы булевых функций.

Если в E' присутствует входо-выходное состояние, которое не содержится в E , то оно принадлежит множеству I .

Если в E присутствует входо-выходное состояние, которое не содержится в E' , то оно принадлежит множеству I' .

Из вышесказанного следует, что в присутствии неисправности в схеме C могут появиться дополнительные достижимые входо-выходные состояния из I , принадлежащие множеству E' . В то же время из множества E могут исчезнуть некоторые входо-выходные состояния, принадлежащие множеству I' .

Чтобы найти добавленные входо-выходные состояния схемы C в присутствии неисправности, надо выполнить пересечение множества E' с I .

Чтобы найти исчезнувшие из схемы C входо-выходные состояния в присутствии неисправности, надо выполнить пересечение множества E с I' .

Алгоритм получения тестовых наборов с одной из реакций исправной и неисправной схем:

1. Представим входо-выходные функции для исправной и неисправной схем в виде соответствующих однокоренных ROBDD-графов.

2. Для получения тестовых наборов и реакций схемы в исправном состоянии на этих наборах необходимо инвертировать однокоренный граф для неисправной схемы и выполнить его конъюнкцию с однокоренным графом для исправной схемы.

3. Для получения тестовых наборов и реакций схемы в неисправном состоянии на этих наборах необходимо инвертировать однокоренный граф для исправной схемы и выполнить его конъюнкцию с однокоренным графом для неисправной схемы.

Будем иметь в виду, что инвертирование однокоренного графа заключается в переименовании его концевых вершин.

Для каждого элемента множества $E \cap I' (E' \cap I)$ известна только реакция исправной (неисправной) схемы, но не известна реакция неисправной (исправной) схемы.

Алгоритм получения тестовых наборов со второй реакцией:

1. Соответствующие пары реакций можно найти, выделив сначала для каждой реакции из тестовых наборов исправной и неисправной схем обычные ROBDD-графы, представляющие условия обеспечения соответствующих реакций

2. Затем необходимо выполнить пересечения каждого обычного ROBDD-графа исправной (неисправной) схемы с каждым обычным ROBDD-графом для неисправной (исправной) схемы.

Полученные непустые пересечения представляют в виде обычных ROBDD-графов подмножества входных тестовых наборов, каждому из которых сопоставляется одна и та же пара реакций исправной и неисправной схем S .

Проиллюстрируем сказанное примером.

Пусть дана логическая схема, представленная на рис. 2.

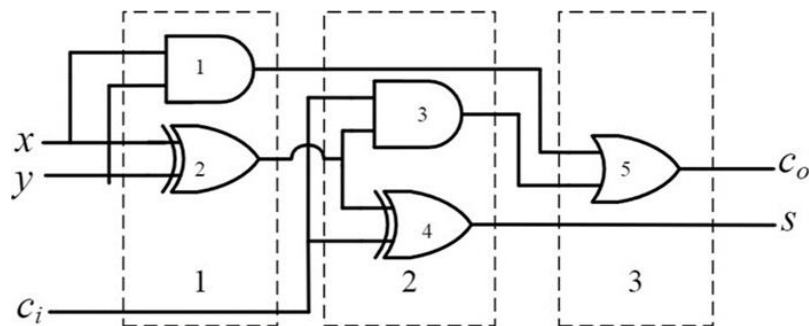


Рис. 2. Логическая схема с двумя выходами и тремя входами
Fig. 2. Logic circuit with two outputs and three inputs

Внесем константную неисправность «1» на выходе второго элемента схемы, при отсутствии неисправности реализующего сумму по модулю два для входных переменных x и y . Построим ROBDD-графы для входо-выходных функций исправной и неисправной схем (рис. 3).

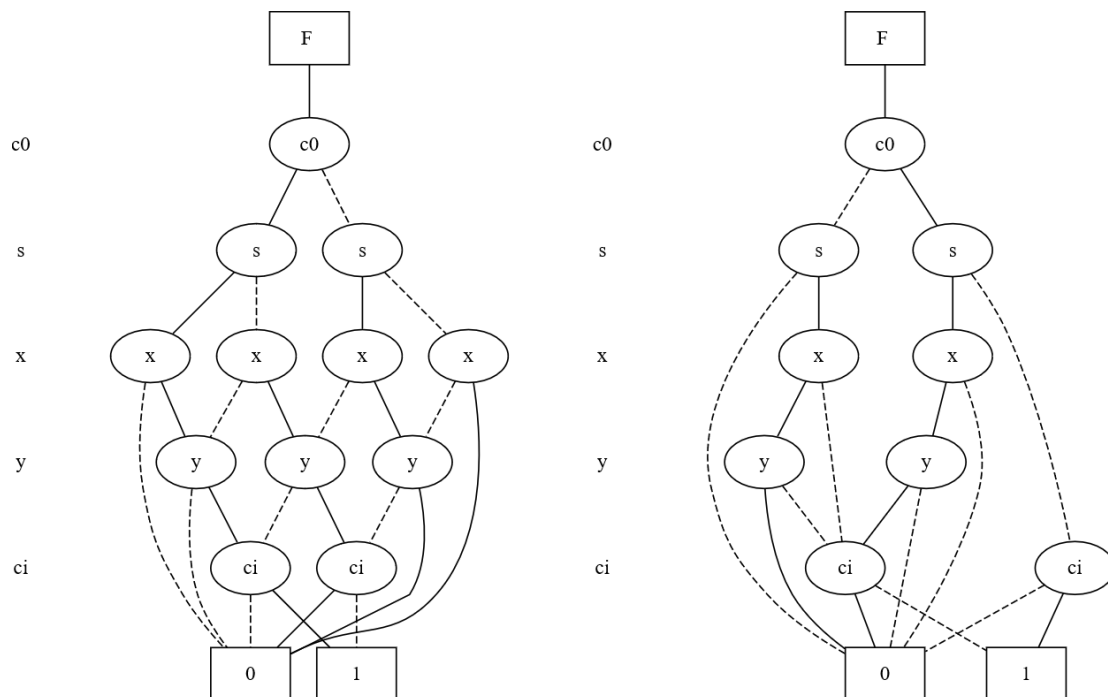


Рис. 3. ROBDD-графы входо-выходной функции исправной и неисправной схем соответственно
Fig. 3. ROBDDs of the input-output functions of fault free and fault circuits, respectively

Обозначим выходные переменные в неисправной схеме как c'_0 и s' . Тогда достижимые состояния исправной и неисправной схем можно привести в таблице истинности (табл. 2).

Таблица 2

Таблица достижимых состояний исправной и неисправной схем

x	y	c_i	c_0	s	c'_0	s'
0	0	0	0	0	0	1
0	0	1	0	1	1	0
0	1	0	0	1	0	1
0	1	1	1	0	1	0
1	0	0	0	1	0	1
1	0	1	1	0	1	0
1	1	0	1	0	1	1
1	1	1	1	1	1	0

Инвертируем ROBDD-граф входо-выходной функции неисправной схемы и выполним его конъюнкцию с ROBDD-графом входо-выходной функции для исправной схемы. В результате получаем граф, содержащий тестовые наборы вместе с реакциями исправной схемы (рис. 4). Инвертируем ROBDD-граф входо-выходной функции исправной схемы и выполним его конъюнкцию с ROBDD-графом входо-выходной функции для неисправной схемы. В результате получаем ROBDD-граф, содержащий тестовые наборы и реакции схемы в неисправном состоянии на этих наборах (см. рис. 4).

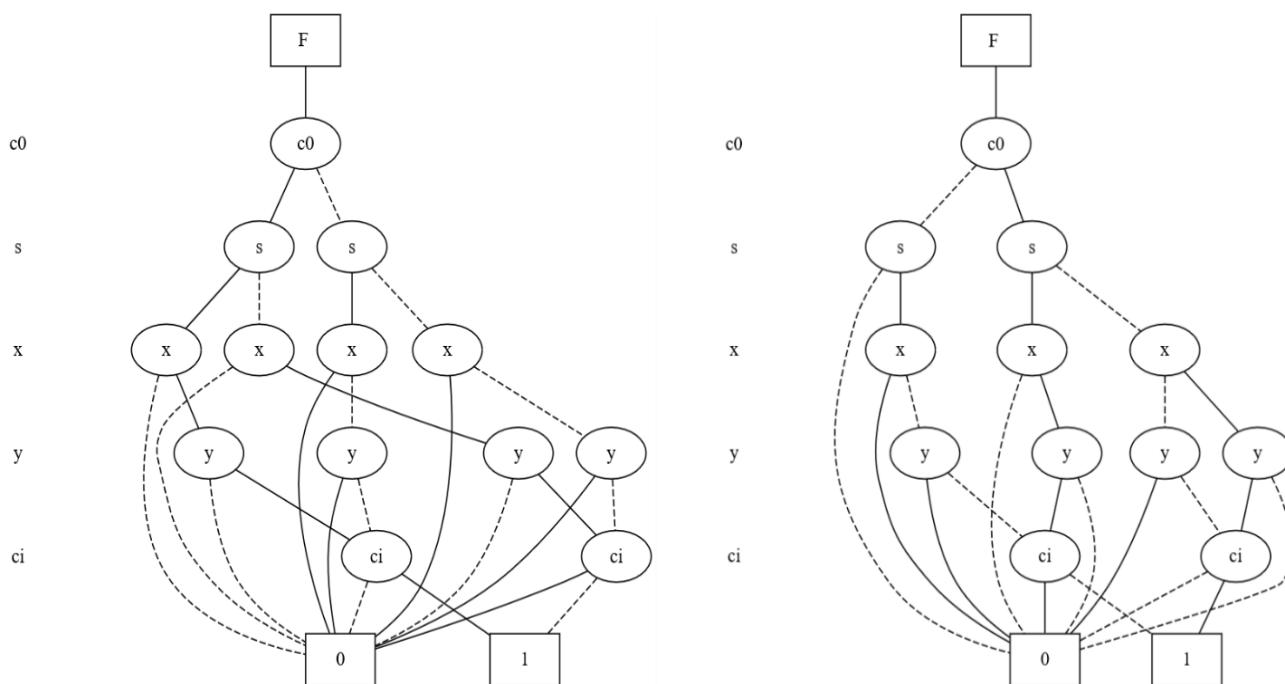


Рис. 4. ROBDD-графы, содержащие тестовые наборы совместно с реакциями исправной и неисправной схем соответственно
Fig. 4. ROBDDs, representing test patterns with the corresponding reactions of fault free and fault circuits, respectively

Выпишем значения переменных в графах различий для исправной и неисправной схем, которые соответствуют путям, ведущим в 1-терминальную вершину, причем обеспечивают различные реакции этих схем на одном и том же входном наборе. Полученные входные тестовые наборы и их реакции представлены в табл. 3.

Следует иметь в виду, что сложность однокоренных графов многовыходных схем, в частности число их внутренних вершин, помеченных символами выходных переменных, экспоненциально растет с ростом числа выходов схемы. Предложенный подход позволяет получить все тестовые наборы входных переменных с соответствующими реакциями исправной и неисправной схем и может быть применен для схем с числом выходов порядка 10 и менее.

Таблица тестовых наборов

x	y	c_i	c_0	s	c'_0	s'
0	0	0	0	0	0	1
0	0	1	0	1	1	0
1	1	0	1	0	1	1
1	1	1	1	1	1	0

Часто в реальных схемах присутствие неисправности искажает значения на некотором подмножестве выходов, связанных между собой подмножеством входных переменных исправной и неисправной схем. В этих ситуациях рассматриваемый подход также может быть применен при наличии подходящих вычислительных ресурсов.

Будем иметь в виду, что обычно нет необходимости в получении всех тестовых наборов и их реакций. В то же время знание возможных реакций схемы позволяет выбирать входные тестовые наборы с полезными для нас реакциями. Учитывая это обстоятельство, можно поступить следующим образом.

Воспользуемся алгоритмом поиска достижимых реакций схемы без получения условий их обеспечения [2]. Представим реакции графом от выходных переменных схемы S . Эти графы существенно проще однокоренных графов многовыходных схем, рассматриваемых нами ранее. Так, для однокоренного графа граф, представляющий только достижимые реакции схемы, показан на рис. 5.

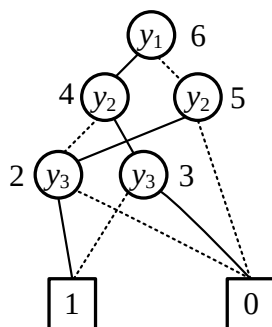


Рис. 5. ROBDD-граф достижимых реакций схемы
Fig. 5. ROBDD of achievable reactions of the scheme

В условиях использования графов реакций предлагается выполнить следующие шаги:

1. Выбираем из графов достижимых реакций подмножество пар реакций исправной и неисправной схем, так чтобы эти реакции по возможности характеризовались большим расстоянием по Хеммингу.
2. Для каждой пары выбранных реакций строим соответствующие пары обычных ROBDD-графов от входных переменных, обеспечивающих эти реакции.
3. Выполняем пересечение этих графов. Непустые ROBDD-графы, представляющие результаты пересечений, задают множество тестовых наборов с одной и той же парой реакций на них.
4. Из подмножества графов, полученных в п. 3 выбираем, например, по одному тестовому набору и путем двоичного моделирования на наборах этого множества находим подмножества элементов, изменивших свои значения в исправной схеме, пытаясь определить элементы, на которых последствия неисправностей впервые проявились при движении от входов к выходам исправной схемы, т.е. для дебаггинга.

Проведенные нами предварительные эксперименты [4–8] над моделями реальных исправных и неисправных схем показали:

- 1) различные тестовые наборы для одной и той же пары реакций могут порождать различные множества элементов, изменяющих значения своих выходов в присутствии неисправности;

2) чем больше расстояние по Хеммингу между реакциями этих схем для одного и того же тестового набора, тем больше элементов исправной схемы изменяет свои значения на соответствующем тестовом наборе.

Заключение

Предложен точный алгоритм поиска всех тестовых наборов с соответствующими им реакциями исправной и неисправной схем, основанный на выполнении операций над ROBDD-графами, зависящими только от входных переменных, только от выходных переменных и от графов, зависящих от обоих множеств переменных. Для каждой пары реакций исправной и неисправной схем строится ROBDD-граф от входных переменных, компактно представляющий соответствующее множество тестовых наборов. Обсуждаются некоторые возможности применения тестовых наборов совместно с их реакциями с целью дебаггинга.

Список источников

1. 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
2. Провкин В.А. Построение множества всех достижимых реакций комбинационных схем // Математическое и программное обеспечение информационных, технических и экономических систем : материалы IX Междунар. науч. конф., Томск, 26–28 мая 2022 г. Томск : Изд-во Том. ун-та, 2022. С. 99–108.
3. Bryant R.E. Graph-Based Algorithms for Boolean Function Manipulation // IEEE Transactions on Computers. 1986. V. C-35, is. 8. P. 677–691.
4. ISCAS'89 Benchmarks // Digital Design & Dependability Research Group. URL: <https://ddd.fit.cvut.cz/www/prj/Benchmarks> (accessed: 25.05.2024).
5. Simple Blifparser. URL: <https://github.com/mario33881/blifparser> (Дата обращения 25.05.2024).
6. dd package // Github.com. URL: <https://github.com/tulip-control/dd> (accessed: 25.05.2024).
7. The CUDD package, BDD, ADD Tutorial and examples // David Kebo Houngninou. URL: <https://www.davidkebo.com/cudd#cudd1> (accessed: 25.05.2024)
8. SIS // Matteo Iervasi : embedded systems engineer. URL: <https://jackhack96.github.io/logic-synthesis/sis.html#overview> (accessed: 25.05.2024).

References

1. Provkina, 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
2. Provkina, V.A. (2022) Postroenie mnozhestva vseh dostizhimykh reaktsiy kombinatsionnykh skhem [Construction of the set of all achievable reactions of combinational circuits]. *Matematicheskoe i programnoye obespechenie informatsionnykh, tekhnicheskikh i ekonomicheskikh sistem* [Mathematical and software support for information, technical and economic systems]. Proc. of the 9th International Conference. May 26–28, 2022. Tomsk. pp. 99–108.
3. Bryant, R.E. (1986) Graph-Based Algorithms for Boolean Function Manipulation. *IEEE Transactions on Computers*. C-35(8). pp. 677–691.
4. Czechia. (1989) *ISCAS'89 Benchmarks*. [Online] Available from: <https://ddd.fit.cvut.cz/www/prj/Benchmarks>. (Accessed: 25th May 2024).
5. *Simple Blifparser*. [Online] Available from: <https://github.com/mario33881/blifparser>. (Accessed: 25th May 2024).
6. *dd package*. [Online] Available from: <https://github.com/tulip-control/dd>. (Accessed: 25th May 2024).
7. *The CUDD package, BDD, ADD Tutorial and examples*. [Online] Available from: <https://www.davidkebo.com/cudd#cudd1>. (Accessed: 25th May 2024).
8. SIS. [Online] Available from: <https://jackhack96.github.io/logic-synthesis/sis.html#overview>. (Accessed: 25th May 2024).

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

Провкин Виктор Алексеевич – старший преподаватель кафедры компьютерной безопасности Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: prowkan@mail.ru

Сухорученко Ксения Алексеевна – магистрант кафедры компьютерной безопасности Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: ksenia_s.a@bk.ru

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

Information about the authors:

Matrosova Anzhela Y. (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

Provkin Viktor A. (Institute of Applied Mathematics and Computer Science, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: prowkan@mail.ru

Sukhoruchenko Ksenia A. (Master's Student, Institute of Applied Mathematics and Computer Science, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: ksenia_s.a@bk.ru

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

Received 16.08.2024; accepted for publication 02.12.2024

Поступила в редакцию 16.08.2024; принята к публикации 02.12.2024

Научный журнал

**ВЕСТНИК
ТОМСКОГО
ГОСУДАРСТВЕННОГО
УНИВЕРСИТЕТА**

**УПРАВЛЕНИЕ,
ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА
И ИНФОРМАТИКА**

**TOMSK STATE UNIVERSITY
JOURNAL OF CONTROL AND COMPUTER SCIENCE**

2024. № 69

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 27.12.2024 г. Формат 60x84^{1/8}.
Гарнитура Times. Усл. печ. л. 16,7.
Тираж 250 экз. Заказ № 6151. Цена свободная.

Дата выхода в свет 15.01.2025 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