

ВЕСТНИК ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА

УПРАВЛЕНИЕ,
ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА
И ИНФОРМАТИКА

TOMSK STATE UNIVERSITY
JOURNAL OF CONTROL AND COMPUTER SCIENCE

Научный журнал

2024

№ 68

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

Учредитель – Национальный исследовательский Томский государственный университет (НИ ТГУ)

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

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

Воробейчиков Сергей Эрикович – д-р физ.-мат. наук, НИ ТГУ
Вишневский Владимир Миронович – проф., д-р техн. наук, Институт проблем управления РАН (г. Москва)
Гордов Евгений Петрович – проф., д-р физ. мат. наук, Институт мониторинга климатических и экологических систем СО РАН (г. Томск)
Кошкин Геннадий Михайлович – проф., д-р физ.-мат. наук, НИ ТГУ
Костюк Юрий Леонидович – проф., д-р техн. наук, НИ ТГУ
Матросова Анжела Юрьевна – проф., д-р техн. наук, НИ ТГУ
Назаров Анатолий Андреевич – проф., д-р техн. наук, НИ ТГУ
Пашинская Татьяна Юрьевна – д-р физ.-мат. наук, НИ ТГУ
Перепелкин Евгений Александрович – проф., д-р техн. наук, Санкт-Петербургский государственный университет аэрокосмического приборостроения (г. Санкт-Петербург)
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
Раймонд Убар (Raimund Ubar) – проф., д-р. (PhD), University of Technology Tallinn, Эстония
Нина Владимировна Евтушенко – проф., д-р. техн. наук, ИСП РАН им. Иванникова В.П., Москва, Россия
Ервант Зорнан (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

Антипина Е.В., Мустафина С.А., Антипин А.Ф. Оптимальное управление процессом протекания каталитической реакции	4
Мансимов К.Б., Керимова А.В. Линеаризованное необходимое условие оптимальности и исследование квазисобых управлений в одной ступенчатой дискретно-непрерывной задаче оптимального управления	14

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

Мелосникова Н.П., Назаров А.А., Фёдорова Е.А. Исследование RQ-системы M/M/1 с катастрофами	28
Rusilko T.V., Salnikov D.A. Asymptotic analysis of a closed G-network with rewards	38
Фунг В.К., Богатырев В.А. Задержки и надежность обслуживания запросов в виртуальном компьютерном кластере	48
Цициашвили Г.Ш., Харченко Ю.Н. Управляемые системы массового обслуживания со стационарным равномерным распределением	59

ОБРАБОТКА ИНФОРМАЦИИ

Игнатьев Н.А., Згуральская Е.Н. Кластерный анализ с применением обучения на основе отношений связности и плотности распределения	66
Канаева И.А., Спицын В.Г. Сегментация дефектов дорожного полотна на основе нейросетевого ансамбля	75
Петровский В.В. Применение нейросетевых алгоритмов для детектирования человека на видеоряде в шахте	86
Смагин В.И., Ким К.С. Идентификация с использованием расширенного экстраполятора Калмана при аномальных ошибках в наблюдениях	92
Шаход Д.М., Агафонов Е.Д. Комбинированная модель локализации акустических источников с применением технологии глубокого обучения	100

ИНФОРМАТИКА И ПРОГРАММИРОВАНИЕ

Чусов А.А., Шкляр А.А., Ефимова Ю.И. Об энергоэффективности и оперативности полиномиального и табличного методов приближения математических функций в системах с плавающей точкой	112
---	-----

ПРОЕКТИРОВАНИЕ И ДИАГНОСТИКА ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

Ефанов Д.В. Синтез самопроверяемых цифровых устройств на основе взвешенных кодов с суммированием, проверочные символы которых описываются самодвойственными булевыми функциями	123
---	-----

CONTENTS

CONTROL OF DYNAMICAL SYSTEMS

Antipina E.V., Mustafina S.A., Antipin A.F. Optimal control of the catalytic reaction process	4
Mansimov K.B., Karimova A.V. Linearized necessary optimality condition and study of quasi-singular controls in one step discrete-continuous optimal control problem	14

MATHEMATICAL MODELING

Meloshnikova N.P., Nazarov A.A., Fedorova E.A. Study of the M/M/1 retrieval queueing system with disasters	28
Rusilko T.V., Salnikov D.A. Asymptotic analysis of a closed G-network with rewards	38
Phung V.Q., Bogatyrev V.A. Delays and reliability of request servicing in a virtual computer cluster	48
Tsitsiashvili G.Sh., Kharchenko Yu.N. Controlled queueing systems with a stationary uniform distribution	59

DATA PROCESSING

Ignatiev N.A., Zguralskaya E.N. Cluster analysis using learning based on connectivity relations and distribution density	66
Kanaeva I.A., Spitsyn V.G. Road damage defects segmentation based on convolutional neural network ensemble	75
Petrovskiy V.V. Application of neural network algorithms for detecting a person in a video sequence in a mine	86
Smagin V.I., Kim K.S. Identification using the extended Kalman extrapolator in case of anomalous errors in observations	92
Shahoud G.M., Agafonov E.D. A combined model for localizing acoustic sources using deep learning technology	100

INFORMATICS AND PROGRAMMING

Chusov A.A., Shklyar A.A., Efimova Yu.I. On energy efficiency and performance of polynomial and LUT-based approximation of elementary functions in floating-point systems	112
---	-----

DESIGNING AND DIAGNOSTICS OF COMPUTER SYSTEMS

Efanov D.V. The self-checking digital devices synthesis based on weight-based sum code, which check symbols are described by self-dual Boolean functions	123
---	-----

УПРАВЛЕНИЕ ДИНАМИЧЕСКИМИ СИСТЕМАМИ

CONTROL OF DYNAMICAL SYSTEMS

Научная статья

УДК 519.6, 004.942

doi: 10.17223/19988605/68/1

Оптимальное управление процессом протекания каталитической реакции

Евгения Викторовна Антипина¹, Светлана Анатольевна Мустафина²,
Андрей Федорович Антипин³^{1,2,3} Уфимский университет науки и технологий, Уфа, Россия¹ stepashinaev@ya.ru² mustafina_sa@mail.ru³ andrejantipin@ya.ru

Аннотация. Статья посвящена разработке алгоритма определения оптимальных режимных параметров каталитической реакции. Сформулированы постановки задач оптимального управления, в которых в качестве параметров управления выступают температура и время протекания реакции. Для их решения приведен генетический алгоритм с вещественным кодированием. Эффективность работы алгоритма подтверждена результатами численных экспериментов для реакции олигомеризации этилена.

Ключевые слова: генетический алгоритм; программный комплекс; каталитическая реакция; оптимальные условия.

Для цитирования: Антипина Е.В., Мустафина С.А., Антипин А.Ф. Оптимальное управление процессом протекания каталитической реакции // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 68. С. 4–13. doi: 10.17223/19988605/68/1

Original article

doi: 10.17223/19988605/68/1

Optimal control of the catalytic reaction process

Evgenia V. Antipina¹, Svetlana A. Mustafina², Andrey F. Antipin³^{1,2,3} Ufa University of Science and Technology, Ufa, Russian Federation¹ stepashinaev@ya.ru² mustafina_sa@mail.ru³ andrejantipin@ya.ru

Abstract. The article is devoted to the development of an algorithm for determining the optimal operating parameters of a catalytic reaction. Statements of optimal control problems are formulated, in which the temperature and time of the reaction act as control parameters. To solve them, a genetic algorithm with real coding is given. The efficiency of the algorithm is confirmed by the results of numerical experiments for the ethylene oligomerization reaction.

Keywords: genetic algorithm; software package; catalytic reaction; optimal conditions.

For citation: Antipina, E.V., Mustafina, S.A., Antipin, A.F. (2024) Optimal control of the catalytic reaction process. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 68. pp. 4–13. doi: 10.17223/19988605/68/1

Введение

В настоящее время оптимальное управление каталитическими процессами может осуществляться с помощью различных методов и технологий, таких как математическое моделирование, оптимизация, автоматизация и использование современного оборудования и программного обеспечения. Одним из подходов, объединяющим перечисленные технологии, является разработка цифровых двойников процессов. Цифровые двойники каталитических процессов – это виртуальные модели этих процессов, созданные на основе данных о состоянии и параметрах процесса. Они могут использоваться для прогнозирования будущих состояний процесса и анализа влияния изменений на параметры процесса. Цифровые двойники позволяют проводить различные сценарные анализы и оптимизацию процессов с использованием математических моделей и алгоритмов оптимизации.

В настоящее время имеются исследования, в которых на основе математических методов определяются оптимальные способы ведения каталитических процессов [1, 2].

Аналитические методы оптимизации процессов химической технологии (принцип максимума Понтрягина, вариационные методы) применяются для решения задач с небольшим количеством независимых переменных [3]. С увеличением количества переменных, а также при наличии ограничений применение аналитических методов становится затруднительным.

Для оптимизации многостадийных процессов химической технологии применяется динамическое программирование, эффективность которого показана в работе [4]. Динамическое программирование хорошо зарекомендовало себя при решении задач, в которых на каждой стадии имеется незначительное число переменных. Его применение для сложных химических реакций, содержащих большое количество промежуточных веществ, затруднительно ввиду высокой размерности задачи и, как следствие, ресурсоемких вычислений.

Одним из развивающихся направлений в области оптимизации и оптимального управления динамическими процессами являются эволюционные вычисления [5, 6]. К методам эволюционного поиска относятся генетические алгоритмы, в основу которых заложены принципы эволюции живой природы [7–9]. Одно из преимуществ генетических алгоритмов – способность преодолевать попадание в локальный экстремум благодаря механизмам кроссовера и мутации. Поскольку поиск решения начинается не с одного начального приближения, а с некоторого множества первоначальных значений, то возможно более полное исследование области допустимых решений. Генетические алгоритмы можно легко модифицировать при увеличении количества переменных, описывающих состояние процесса, что значительно облегчает их практическую реализацию.

В задачах оптимального управления каталитическими реакциями управляющие параметры могут представляться в виде функций (например, температура, давление) или констант (например, длительность процесса, мольное соотношение реагентов). Одновременный поиск оптимальных значений таких параметров при управлении каталитическим процессом – актуальная задача, которая часто возникает на практике. Поэтому целью работы является разработка алгоритма и программного обеспечения для решения задач оптимального управления каталитическими процессами, в которых параметры управления выражаются функциональной зависимостью и / или константой.

1. Постановка задачи

Пусть динамика протекания каталитической реакции описывается системой обыкновенных дифференциальных уравнений

$$\frac{dx}{dt} = f(t, x(t), T(t)) \quad (1)$$

с начальными условиями

$$x(t_0) = x_0, \quad (2)$$

где $x(t) \in R^n$ – вектор концентраций веществ, $t \in [0, \tau]$ – время, $T(t)$ – температура, $f(t, x(t), T(t))$ – непрерывная вместе со своими частными производными вектор-функция.

Фазовыми переменными являются концентрации веществ.

Пусть параметры управления – температура протекания реакции $T(t)$, влияющая на интенсивность химических превращений, и время контакта реагентов τ , определяющее длительность процесса. На значения управляющих параметров наложены физико-химические ограничения

$$T_{\min} \leq T(t) \leq T_{\max}, \quad t \in [0, \tau], \quad (3)$$

$$\tau_{\min} \leq \tau \leq \tau_{\max}, \quad (4)$$

где $T_{\min}$, $T_{\max}$ – минимальное и максимальное допустимые значения температуры, $\tau_{\min}$, $\tau_{\max}$ – минимальное и максимальное допустимые значения продолжительности реакции.

В качестве критерия оптимальности можно рассматривать выход целевых продуктов в конце реакции, содержание побочных веществ, степень превращения реагентов, селективность и другие показатели эффективности химического процесса.

Сформулируем задачи поиска оптимальных режимных параметров каталитической реакции, описываемой системой дифференциальных уравнений (1) с начальными условиями (2).

Задача 1. Определить оптимальный температурный режим $T^*(t) \in [T_{\min}, T_{\max}]$ при заданной продолжительности реакции $\tau \in [\tau_{\min}, \tau_{\max}]$, доставляющий максимум функционалу

$$I(T^*(t)) = g(x(\tau)). \quad (5)$$

Задача 2. Найти оптимальное время контакта веществ $\tau^* \in [\tau_{\min}, \tau_{\max}]$ при заданной постоянной температуре $T(t) = \text{const}$, удовлетворяющей ограничению (3), при котором достигается максимум функционала

$$I(\tau^*) = g(x(\tau^*)). \quad (6)$$

Задача 3. Определить оптимальный температурный режим $T^*(t) \in [T_{\min}, T_{\max}]$ и оптимальное время протекания реакции $\tau^* \in [\tau_{\min}, \tau_{\max}]$, при которых целевой функционал

$$I(T^*(t), \tau^*) = g(x(\tau^*)) \quad (7)$$

принимает максимальное значение.

2. Генетический алгоритм поиска оптимальных режимных параметров каталитической реакции

Для решения сформулированных задач применим генетический алгоритм с вещественным кодированием. Принцип работы генетических алгоритмов основан на имитации процессов эволюции живой природы и включает в себя циклическое применение к популяции особей операций отбора, скрещивания, мутации и обновления.

Будем искать управление $T(t)$ в классе кусочно-постоянных функций $T(t) = T_j$, $t \in [t_j, t_{j+1}]$, $j = \overline{0, m}$, $t_0 < t_1 < t_2 < \dots < t_{m+1}$, $t_0 = 0$, $t_{m+1} = \tau$.

Для каждой задачи оптимального управления каталитической реакцией введем в рассмотрение множество векторов:

$$1) \quad p_i = (T_{i1}, T_{i2}, \dots, T_{im}), \quad i = \overline{1, s};$$

$$2) \quad p_i = (\tau_i), \quad i = \overline{1, s};$$

$$3) \quad p_i = (T_{i1}, T_{i2}, \dots, T_{im}, \tau_i), \quad i = \overline{1, s}.$$

Каждый вектор p_i , представляющий собой возможное решение задачи, будем называть вектором-индивидом (особью), а совокупность всех векторов p_i ($i = \overline{1, s}$) – популяцией. Введем общее обозначение для элементов вектора-индивида: $p_i = (p_{i1}, p_{i2}, \dots, p_{ir}) = (p_{ij})$, $j = \overline{1, r}$, где

$$1) \quad p_{ij} = T_{ij}, \quad r = m \quad \text{– для задачи 1;}$$

2) $p_{ij} = \tau_i$, $r = 1$ – для задачи 2;

3) $p_{ij} = \begin{cases} T_{ij}, & j = \overline{1, m}, \\ \tau_i, & j = m + 1, \end{cases} \quad r = m + 1$ – для задачи 3.

Фитнес-функцией, определяющей, подходит ли вектор-индивид в качестве решения, является целевой функционал (5), (6) или (7) каждой задачи. Так как осуществляется поиск максимального значения целевого функционала, наиболее подходящей (приспособленной) особи соответствует наибольшее значение целевого функционала. Чтобы вычислить значение фитнес-функции вектора-индивида, необходимо найти численное решение системы дифференциальных уравнений (1) с начальными условиями (2).

Работа генетического алгоритма состоит из следующих этапов:

1. Создание начальной популяции векторов-индивидов p_i , $i = \overline{1, S}$. Для каждого индивида вычисляется значение фитнес-функции.

2. Отбор. Из текущей популяции выбираются два индивида p_k, p_l с помощью оператора отбора:

а) панмиксия – случайным образом отбираются два вектора-индивида, при этом вероятность отбора у каждого из них одинакова;

б) турнирный отбор – выполняется два турнира: в первом турнире случайным образом выбирается два различных индивида, из которых случайным образом во втором турнире отбирается один вектор-индивид.

3. Кроссовер. Формируется новый индивид путем применения одного из оператора кроссовера:

а) простейший – создается два потомка $v_1 = (p_{k1}, \dots, p_{kq}, p_{lq+1}, \dots, p_{lr})$, $v_2 = (p_{l1}, \dots, p_{lq}, p_{kq+1}, \dots, p_{kr})$,

где $q \in [1, r - 1]$ – случайное число;

б) арифметический – создается два потомка $v_1 = \lambda p_k + (1 - \lambda) p_l$, $v_2 = \lambda p_l + (1 - \lambda) p_k$, где $\lambda \in (0, 1)$ – случайное число.

4. Мутация. С целью преодоления попадания решения в точку локального экстремума сгенерированный на предыдущем шаге вектор-индивид подвергается действию случайной мутации: случайно выбранная координата каждого из векторов v_1, v_2 заменяется случайным значением из диапазонов, задаваемых неравенствами (3) и / или (4).

5. Обновление популяции. Из текущей популяции выбирается вектор-индивид с наименьшим значением фитнес-функции и заменяется вектором-потомком. Далее осуществляется переход на шаг 2, пока не будет достигнуто условие окончания поиска. В качестве условия завершения вычислений для генетических алгоритмов применяется ограничение на максимальное количество популяций или стабилизация приспособленности популяции на нескольких итерациях метода. Вектор-индивид с наибольшим значением фитнес-функции из последней популяции будет представлять собой приближенное решение задачи оптимального управления каталитической реакцией.

Особенностью алгоритма является его универсальность, так как он позволяет найти приближенное решение задач оптимального управления каталитическим процессом, содержащих управление-функцию (температура реакции $T(t)$) и управление-константу (время реакции τ), а также осуществить их одновременный поиск.

3. Вычислительный эксперимент

Применим генетический алгоритм для поиска оптимальных режимных параметров олигомеризации этилена. Процессы каталитической олигомеризации и полимеризации олефинов имеют важное практическое значение в промышленном производстве. Продуктами олигомеризации этилена являются высшие α -олефины большей молекулярной массы, которые применяются в крупнотоннажных промышленных процессах сополимеризации этилена и пропилена в производстве полиэтилена средней и низкой плотности. Линейные α -олефины используются для получения низкомолекулярных продуктов: высших спиртов, кислот, основ синтетических масел и др. [1].

Одним из перспективных способов синтеза высших α -олефинов (C_4 – C_8) является процесс олигомеризации этилена с использованием металлокомплексных катализаторов. Кинетические закономерности процесса олигомеризации этилена на катализаторе $NiO/B_2O_3-Al_2O_3$ в среде жидкого растворителя гептана рассмотрены в работе [10]. Математическая модель процесса олигомеризации этилена позволяет описать динамику концентраций веществ в проточном реакторе и представляет собой систему обыкновенных дифференциальных уравнений

$$\begin{aligned}\frac{dx_1}{dt} &= -2k_1x_1 - k_2x_1(x_2 + x_3 + x_4 + x_5), \\ \frac{dx_2}{dt} &= k_1x_1 - k_2x_1x_2 - k_3x_2(x_2 + x_3 + x_4), \\ \frac{dx_3}{dt} &= k_2x_1x_2 - k_2x_1x_3 - k_3x_2x_3, \\ \frac{dx_4}{dt} &= k_2x_1x_3 + k_3x_2^2 - k_2x_1x_4 - k_3x_2x_4, \\ \frac{dx_5}{dt} &= k_2x_1x_4 + k_3x_2x_3 - k_2x_1x_5, \\ \frac{dx_6}{dt} &= k_2x_1x_5 + k_3x_2x_4\end{aligned}\quad (8)$$

с начальными условиями

$$x_i(0) = x_i^0, \quad i = \overline{1,6}. \quad (9)$$

Здесь x_i – концентрация углеводорода C_{2i} (моль/л), $t \in [0, \tau]$ – время (с), k_j – константа скорости j -й стадии (1/с для $j = 1$, $dm^3/(mоль \cdot c)$ для $j = 2, 3$), определяемая по уравнению Аррениуса

$$k_j(T) = k_{0j} \exp\left(-\frac{E_j}{RT}\right),$$

где T – температура реакционной смеси (К), k_{0j} – кинетическая константа (1/с для $j = 1$, $dm^3/(mоль \cdot c)$ для $j = 2, 3$), E_j – энергия активации (кДж/моль), R – универсальная газовая постоянная (8,31 Дж/(моль·К)).

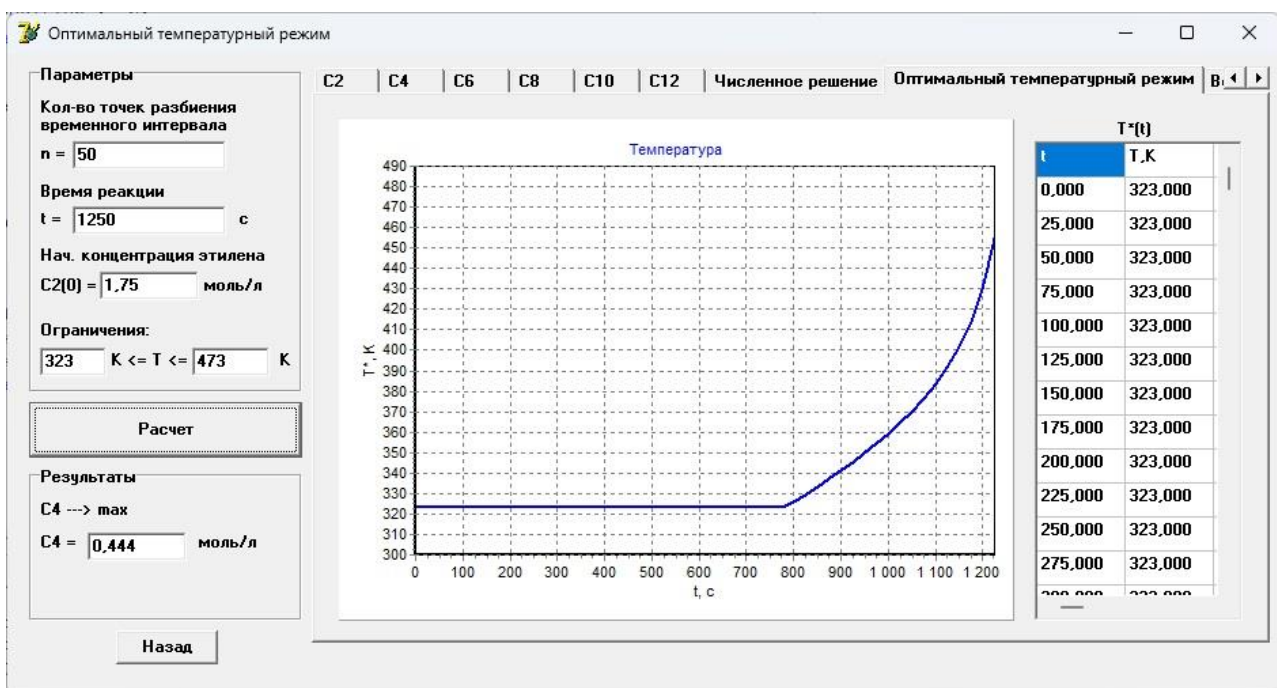


Рис. 1. Окно программного комплекса
Fig. 1. Window of the software package

Пусть на значения управляющих параметров наложены ограничения

$$323 \text{ K} \leq T(t) \leq 473 \text{ K}, \quad t \in [0, \tau], \quad (10)$$

$$60 \text{ с} \leq \tau \leq 1500 \text{ с}. \quad (11)$$

В качестве критерия оптимальности зададим максимальное значение концентрации углеводов фракции C_4 .

Начальные условия (9) заданы значениями (моль/л)

$$x_1(0) = 1,75, \quad x_2(0) = 0, \quad i = \overline{2, 6}. \quad (12)$$

Для решения задач оптимального управления процессом олигомеризации этилена в среде визуального программирования Delphi разработан программный комплекс, реализующий сформулированный генетический алгоритм с вещественным кодированием. Приложение позволяет пользователю выбрать задачу оптимального управления, настроить параметры генетического алгоритма для поиска решения и визуализировать полученные результаты (рис. 1). Для поиска численного решения системы дифференциальных уравнений (8) с начальными условиями (12) в программе реализован предиктор-корректорный метод Адамса второго порядка.

3.1. Определение оптимального температурного режима

Пусть параметром управления является температура реакции. Необходимо найти оптимальный температурный режим $T^*(t)$ процесса олигомеризации этилена, описываемого системой дифференциальных уравнений (8) с начальными условиями (12), удовлетворяющий неравенству (10), при котором целевой функционал

$$I(T^*(t)) = x_2(\tau) \quad (13)$$

принимает максимальное значение (задача 1).

Вычисления проведены для максимально возможной продолжительности реакции $\tau = 1500$ с. Генетический алгоритм применен со следующими параметрами: размер популяции – 70, оператор отбора – турнирный, оператор кроссовера – арифметический, максимальное число популяций – 1000.

На рис. 2 показан рассчитанный оптимальный температурный профиль процесса олигомеризации этилена (задача 1), при котором наибольшее значение концентрации углеводов C_4 составляет 0,444 моль/л.

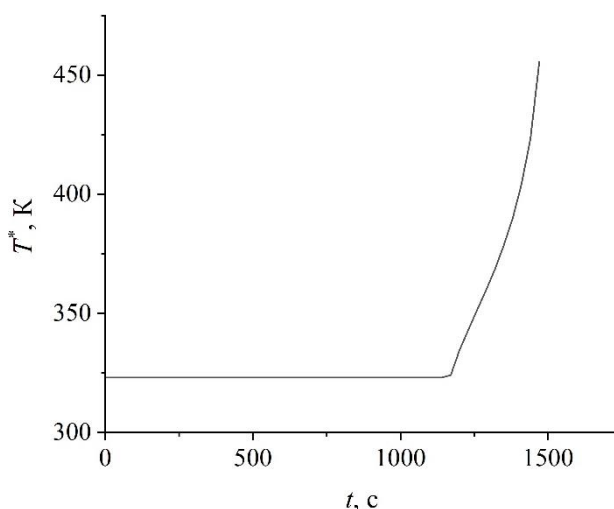


Рис. 2. Оптимальный температурный режим
Fig. 2. Optimal temperature regime

В табл. 1 приведены значения концентрации целевых продуктов реакции, полученных в результате решения системы дифференциальных уравнений (8) с начальными условиями (12) при $\tau = 1500$ с и постоянных значениях температуры, удовлетворяющих неравенству (10).

Значения целевого функционала (13) при различных изотермических режимах

T, K	$x_2(\tau^*), \text{ моль/л}$
323	0,409
348	0,398
373	0,329
398	0,258
423	0,205
448	0,168
473	0,141

Из табл. 1 следует, что при максимально возможной продолжительности реакции олигомеризации этилена в изотермических условиях целевой функционал принимает наибольшее значение при температурном режиме, вычисленном с помощью генетического алгоритма.

3.2. Поиск оптимального времени взаимодействия реагентов

Определим оптимальную продолжительность реакции олигомеризации этилена в изотермических условиях. Пусть требуется для процесса олигомеризации этилена найти время контакта веществ τ^* с учетом ограничений (11), при котором достигается наибольшее значение концентрации углеводородов фракции C_4 (задача 2), т.е.

$$I(\tau^*) = x_2(\tau^*) \rightarrow \max. \quad (14)$$

Решение сформулированной задачи найдено для температуры реакции $T = 423 \text{ K}$ с помощью генетического алгоритма с параметрами: размер популяции – 70, оператор отбора – турнирный, оператор кроссовера – арифметический, максимальное количество популяций – 1 000.

Результаты расчетов показали, что наибольшее значение концентрации целевого продукта реакции, равное 0,386 моль/л, достигается при $\tau^* = 196 \text{ с}$.

На рис. 3 показана динамика концентрации целевых продуктов реакции продолжительностью 1 500 с. Из рисунка видно, что наибольшее значение концентрации x_2 наблюдается в момент времени, равный 196 с, что подтверждает корректную работу генетического алгоритма. Дальнейшее проведение реакции нецелесообразно, поскольку происходит уменьшение содержания в реакционной смеси углеводородов C_4 .

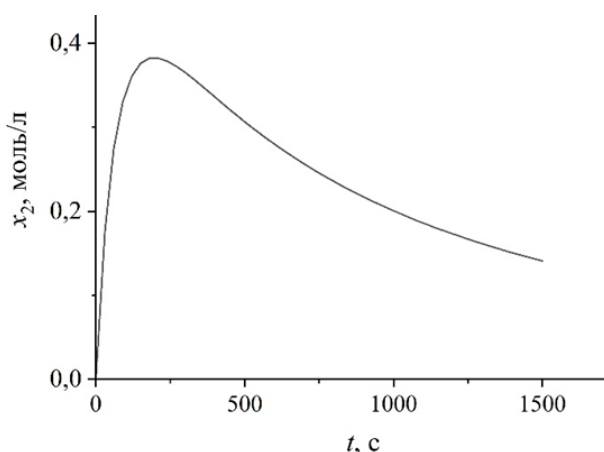


Рис. 3. Изменение концентрации углеводородов C_4
Fig. 3. Changes in concentration of C_4 hydrocarbons

В табл. 2 приведены результаты вычисления оптимального времени взаимодействия веществ для различных изотермических режимов процесса олигомеризации этилена (задача 2).

Таблица 2

Оптимальная продолжительность реакции олигомеризации этилена при постоянных допустимых значениях температуры

T, K	τ^*, c	$x_2(\tau^*), \text{моль/л}$
323	1 500	0,409
348	1 034	0,424
373	690	0,426
398	480	0,423
423	346	0,412
448	257	0,399
473	196	0,383

Из табл. 2 видно, что с повышением температуры оптимальная продолжительность реакции олигомеризации этилена уменьшается. Наибольшее же значение концентрации целевых продуктов реакции, равное 0,426 моль/л, достигается при постоянной температуре реакции, равной 373 К, и продолжительности реакции, равной 690 с.

3.3. Определение оптимального температурного режима и оптимальной продолжительности реакции

Пусть требуется найти оптимальное время контакта веществ τ^* и оптимальный температурный режим $T^*(t)$ процесса олигомеризации этилена, удовлетворяющие условиям (10), (11) и обеспечивающие максимум целевому функционалу (задача 3)

$$I(T^*(t), \tau^*) = x_2(\tau^*). \quad (15)$$

Для решения данной задачи применен генетический алгоритм с параметрами: размер популяции – 100, оператор отбора – турнирный, оператор кроссовера – арифметический, максимальное количество популяций – 2 000.

В результате проведенных вычислений установлено, что концентрация углеводородов C_4 достигает наибольшего значения, равного 0,512 моль/л, при продолжительности реакции $\tau^* = 1\,185$ с и температурном режиме, приведенном на рис. 4.

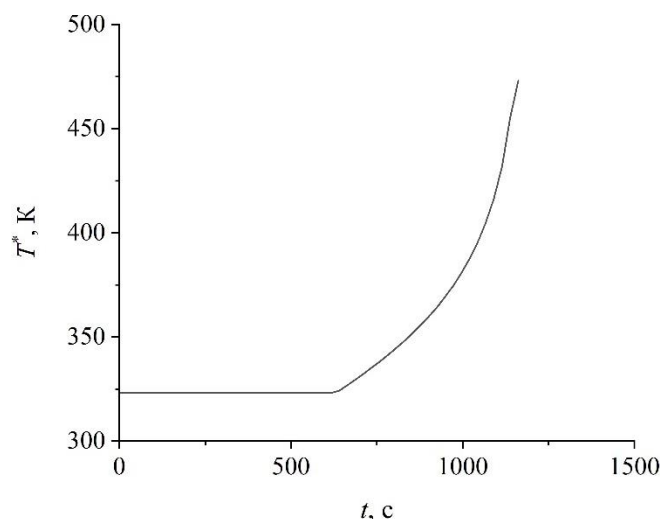


Рис. 4. Оптимальный температурный режим процесса олигомеризации этилена, соответствующий оптимальной продолжительности реакции

Fig. 4. Optimal temperature regime of the ethylene oligomerization process corresponding to optimal reaction duration

Вычисленные значения целевого функционала (15) для различных изотермических режимов (табл. 3) при оптимальной продолжительности реакции не превосходят его значения, рассчитанного для задач 1 и 2.

Значение целевого функционала (15) при постоянных допустимых значениях температуры ($\tau^* = 1\ 185\ c$)

T, K	$x_2(\tau^*), \text{ моль/л}$
323	0,393
348	0,421
373	0,377
398	0,309
423	0,251
448	0,207
473	0,175

Заключение

Таким образом, разработанный генетический алгоритм можно применять для поиска оптимальных режимных параметров ведения каталитических процессов. Алгоритм сочетает в себе возможность поиска управления-функции (температура реакции) и / или управления-константы (время реакции). На основе алгоритма разработан программный комплекс для определения оптимальных условий протекания процесса олигомеризации этилена. Приложение позволяет найти оптимальный температурный режим процесса олигомеризации этилена при заданной продолжительности протекания реакции, определить оптимальное время взаимодействия реагентов в изотермическом режиме, а также вычислить оптимальный температурный профиль и оптимальное время взаимодействия веществ. Разработанный программный комплекс можно применять для исследования закономерностей протекания каталитического процесса олигомеризации этилена, не прибегая к проведению лабораторного эксперимента, и определять стратегию управления процессом.

Список источников

1. Джамбеков А.М., Щербатов И.А. Оптимальное управление процессом каталитического риформинга бензиновых фракций // Вестник Тамбовского государственного технического университета. 2017. Т. 23, № 4. С. 557–571. doi: 10.17277/vestnik.2017.04
2. Антипина Е.В., Мустафина С.А., Антипин А.Ф. Теоретическая оптимизация режимных параметров каталитической реакции с переменным реакционным объемом // Вестник Тверского государственного университета. Сер. Химия. 2022. № 2 (48). С. 67–78. doi: 10.26456/vtchem2022.2.8
3. Cao X., Jia S., Luo Y., Yuan X., Qi Z., Yu K.-T. Multi-objective optimization method for enhancing chemical reaction process // Chemical Engineering Science. 2019. V. 195. P. 494–506. doi: 10.1016/j.ces.2018.09
4. Fu K., Zou Y., Li S. Iterative Unit-based Adaptive Dynamic Programming with Application to Fluid Catalytic Cracker Unit // 2019 Chinese Automation Congress. 2019. P. 5010–5015. doi: 10.1109/CAC48633.2019.8996670
5. Карпенко А.П. Эволюционные операторы популяционных алгоритмов глобальной оптимизации // Математика и математическое моделирование. 2018. № 1. С. 59–89. doi: 10.24108/mathm.0118.0000103
6. Антипина Е.В., Мустафина С.И., Антипин А.Ф., Мустафина С.А. Численный алгоритм решения задачи оптимального управления с терминальными ограничениями для динамических систем // Автометрия. 2020. Т. 56, № 6. С. 132–140. doi: 10.15372/AUT20200615
7. Трокоз Д.А. Метод параметрической оптимизации для широких нейронных сетей с использованием генетических алгоритмов // Известия Самарского научного центра Российской академии наук. 2021. Т. 23, № 2. С. 51–56. doi: 10.37313/1990-5378-2021-23-2-51-56
8. Stastny J., Skorpil V., Balogh Z., Klein R. Job Shop Scheduling Problem Optimization by Means of Graph-Based Algorithm // Applied Sciences. 2021. V. 11, is. 4. Art. 1921. doi: 10.3390/app11041921
9. Gulbaz R., Siddiqui A.B., Anjum N., Alotaibi A.A., Althobaiti T., Ramzan N. Balancer Genetic Algorithm—A Novel Task Scheduling Optimization Approach in Cloud Computing // Applied Sciences. 2021. V. 11, is. 14. Art. 6244. doi: 10.3390/app11146244
10. Волков А.А., Булчевский Е.А., Лавренов А.В. Кинетика олигомеризации этилена на катализаторе $\text{NiO/B}_2\text{O}_3\text{--Al}_2\text{O}_3$ в жидкой фазе // Журнал Сибирского федерального университета. Сер. Химия. 2013. Т. 6, № 4. С. 352–360.

References

1. Dzhambekov, A.M. & Shcherbatov, I.A. (2017) Optimal'noe upravlenie protsessom kataliticheskogo riforminga benzinovykh fraktsiy [Optimal control of the process of catalytic reforming of petrol fractions]. *Vestnik Tambovskogo gosudarstvennogo tekhnicheskogo universiteta*. 23(4). pp. 557–571. DOI: 10.17277/vestnik.2017.04.pp.557-571

2. Antipina, E.V., Mustafina, S.A. & Antipin, A.F. (2022) Teoreticheskaya optimizatsiya rezhimnykh parametrov kataliticheskoy reaktsii s peremennym reaktsionnym ob'emom [Theoretical optimization of regular parameters of a catalytic reaction with a variable reaction volume]. *Vestnik Tverskogo gosudarstvennogo universiteta. Ser. Khimiya*. 2. pp. 67–78. DOI: 10.26456/vtchem2022.2.8
3. Cao, X., Jia, S., Luo, Y., Yuan, X., Qi, Z. & Yu, K.-T. (2019) Multi-objective optimization method for enhancing chemical reaction process. *Chemical Engineering Science*. 195. pp. 494–506. DOI: 10.1016/j.ces.2018.09
4. Fu, K., Zou, Y. & Li, S. (2019) Iterative Unit-based Adaptive Dynamic Programming with Application to Fluid Catalytic Cracker Unit. *2019 Chinese Automation Congress*. pp. 5010–5015. DOI: 10.1109/CAC48633.2019.8996670
5. Karpenko, A.P. (2018) Evolyutsionnye operatory populyatsionnykh algoritmov global'noy optimizatsii [Evolutionary operators for global optimization population-based algorithms]. *Matematika i matematicheskoe modelirovanie*. 1. pp. 59–89. DOI: 10.24108/mathm.0118.0000103
6. Antipina, E.V., Mustafina, S.I., Antipin, A.F. & Mustafina, S.A. (2020) Chislennyy algoritm resheniya zadachi optimal'nogo upravleniya s terminal'nymi ogranicheniyami dlya dinamicheskikh system [A numerical algorithm for solving optimal control problems with terminal constraints for dynamical systems]. *Avtometriya*. 56(6). pp. 132–140. DOI: 10.15372/AUT20200615
7. Trokoz, D.A. (2021) Metod parametricheskoy optimizatsii dlya shirokikh neyronnykh setey s ispol'zovaniem geneticheskikh algoritmov [The parametric optimization method for wide neural networks using genetic algorithms]. *Izvestiya Samarskogo nauchnogo tsentra Rossiyskoy akademii nauk*. 23(2). pp. 51–56. DOI: 10.37313/1990-5378-2021-23-2-51-56
8. Stastny, J., Skorpil, V., Balogh, Z. & Klein, R. (2021) Job Shop Scheduling Problem Optimization by Means of Graph-Based Algorithm. *Applied Sciences*. 11(4). pp. 1921. DOI: 10.3390/app11041921
9. Gulbaz, R., Siddiqui, A.B., Anjum, N., Alotaibi, A.A., Althobaiti, T. & Ramzan, N. (2021) Balancer Genetic Algorithm—A Novel Task Scheduling Optimization Approach in Cloud Computing. *Applied Sciences*. 11(14). pp. 6244. DOI: 10.3390/app11146244
10. Volkov, A.A., Buluchevskiy, E.A. & Lavrenov, A.V. (2013) Kinetika oligomerizatsii etilena na katalizatore NiO/B₂O₃–Al₂O₃ v zhidkoy faze [Kinetics of Ethylene Oligomerization on NiO/B₂O₃–Al₂O₃ in Liquid Phase]. *Zhurnal Sibirskogo federal'nogo universiteta. Ser. Khimiya – Journal of Siberian Federal University. Chemistry*. 6(4). pp. 352–360.

Информация об авторах:

Антипина Евгения Викторовна – кандидат физико-математических наук, старший научный сотрудник управления научных исследований и разработок Уфимского университета науки и технологий (Уфа, Россия). E-mail: stepashinaev@ya.ru

Мустафина Светлана Анатольевна – профессор, доктор физико-математических наук, проректор по цифровой трансформации Уфимского университета науки и технологий (Уфа, Россия). E-mail: mustafina_sa@mail.ru

Антипин Андрей Федорович – доцент, кандидат технических наук, доцент кафедры прикладной информатики и программирования Уфимского университета науки и технологий (Уфа, Россия). E-mail: andrejantipin@ya.ru

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

Information about the authors:

Antipina Evgenia V. (Candidate of Physical and Mathematical Sciences, Senior Researcher, Ufa University of Science and Technology, Ufa, Russian Federation). E-mail: stepashinaev@ya.ru

Mustafina Svetlana A. (Doctor of Physical and Mathematical Sciences, Professor, Vice-Rector for Digital Transformation, Ufa University of Science and Technology, Ufa, Russian Federation). E-mail: mustafina_sa@mail.ru

Antipin Andrey F. (Candidate of Technical Sciences, Associate Professor, Ufa University of Science and Technology, Ufa, Russian Federation). E-mail: andrejantipin@ya.ru

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

Поступила в редакцию 12.03.2024; принята к публикации 03.09.2024

Received 12.03.2024; accepted for publication 03.09.2024

Научная статья
УДК 519.6, 004.942
doi: 10.17223/19988605/68/2

Линеаризованное необходимое условие оптимальности и исследование квазисобых управлений в одной ступенчатой дискретно-непрерывной задаче оптимального управления

Камиль Байрамали оглы Мансимов¹, Айтадж Вагиф кызы Керимова²

^{1, 2} Бакинский государственный университет, Баку, Азербайджан

¹ Институт систем управления Министерства науки и образования Азербайджана, Баку, Азербайджан

¹ kamilbmansimov@gmail.com

² aytac.mansimli@mail.ru

Аннотация. Рассматривается ступенчатая задача оптимального управления, описываемая совокупностью нелинейных разностных и интегро-дифференциальных уравнений типа Вольтерра с начальными условиями Коши и общим нелинейным функционалом типа Больца. Области управления считаются выпуклыми ограниченными множествами. Построена формула приращения второго порядка функционала качества. При определенных предположениях вычислены специальные частные приращения функционала и с их помощью доказан аналог линеаризованного условия максимума. В случае вырождения (квазисобый случай) линеаризованного условия максимума выведен ряд необходимых условий оптимальности квазисобых управлений.

Ключевые слова: разностное уравнение типа Вольтерра; интегро-дифференциальное уравнение Вольтерра; область управления; формула приращения; специальная вариация управления; допустимое управление; условие оптимальности; линеаризованный принцип максимума.

Для цитирования: Мансимов К.Б., Керимова А.В. Линеаризованное необходимое условие оптимальности и исследование квазисобых управлений в одной ступенчатой дискретно-непрерывной задаче оптимального управления // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 68. С. 14–27. doi: 10.17223/19988605/68/2

Original article
doi: 10.17223/19988605/68/2

Linearized necessary optimality condition and study of quasi-singular controls in one step discrete-continuous optimal control problem

Kamil B. Mansimov¹, Aytaj V. Karimova²

^{1, 2} Baku State University, Baku, Azerbaijan,

¹ Institute of Control Systems of the Ministry of Science and Education of Azerbaijan, Baku, Azerbaijan

¹ kamilbmansimov@gmail.com

² aytac.mansimli@mail.ru

Abstract. We consider a stepwise optimal control problem described by a set of nonlinear difference and integro-differential equations of Volterra type with initial Cauchy conditions and a general nonlinear functional of Boltz type. Control areas are considered to be convex bounded sets. A formula for the second-order increment of the quality functional is constructed. Under certain assumptions, special partial increments of the functional are calculated and, with their making use, an analogue of the linearized maximum condition is proved. In the case of degeneration (quasi-special case) of the linearized maximum condition, a number of necessary conditions for the optimality of quasi-special controls are derived.

Keywords: Volterra type difference equation; Volterra type integro-differential equation; control domain; increment formula; special variation of the control; admissible control; optimality condition; linearized maximum principle.

For citation: Mansimov, K.B., Karimova, A.V. (2024) Linearized necessary optimality condition and study of quasi-singular controls in one step discrete-continuous optimal control problem. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 68. pp. 14–27. doi: 10.17223/19988605/68/2

Введение

Как известно, фундаментальный результат неклассического вариационного исчисления (математической теории оптимального управления), принцип максимума Л.С. Понтрягина, первоначально был доказан для процессов, описываемых обыкновенными дифференциальными уравнениями [1]. Затем усилиями многих ученых этот результат при различных предположениях был перенесен на более общие задачи (см., напр.: [1–8]), а также на многоэтапные задачи оптимального управления, описываемые на различных отрезках времени различными обыкновенными дифференциальными уравнениями (см., напр.: [9–14]).

В работе [15] рассмотрена ступенчатая задача оптимального управления, описываемая в различных дискретных узлах прямоугольной сетки дискретными двухпараметрическими системами. Доказан аналог дискретного принципа максимума и исследован случай его вырождения.

Работа [16] посвящена исследованию ступенчатой задачи оптимального управления, описываемой разностными аналогами интегро-дифференциальных уравнений типа Вольтерра. Применяя один вариант метода приращений, авторы доказали аналог принципа максимума Понтрягина, носящий глобальный характер, при предположении выпуклости аналогов множеств допустимых скоростей.

В предлагаемой работе впервые рассматривается одна двухэтапная (ступенчатая) задача оптимального управления, описываемая совокупностью разностных и интегро-дифференциальных уравнений типа Вольтерра, при предположении выпуклости областей управления. Таким образом, рассматривается дискретно-непрерывная ступенчатая задача оптимального управления, имеющая специфические особенности. Доказаны линеаризованные и квадратичные необходимые условия оптимальности первого и второго порядков, носящие явный характер и являющиеся конструктивными.

1. Постановка задачи оптимального управления

Пусть $U_1 \subset R^r$, $U_2 \subset R^q$ – заданные непустые ограниченные и выпуклые множества, $T_1 = \{t_0, t_0 + 1, \dots, t_1 - 1\}$ – заданная конечная последовательность, $T_2 = [t_1, t_2]$ – заданный отрезок.

Предположим, что управляемая дискретно-непрерывная ступенчатая задача оптимального управления описывается следующей системой уравнений:

$$x_1(t+1) = f_1(t, x(t), u(t)) + \sum_{\tau=t_0}^t K_1(t, \tau, x_1(\tau), u_1(\tau)), t \in T_1, \quad (1)$$

$$x_1(t_0) = x_{10}, \quad (2)$$

$$\dot{x}_2(t) = f_2(t, x_2(t), u_2(t)) + \int_{t_1}^t K_2(t, \tau, x_2(\tau), u_2(\tau)) d\tau, t \in T_2, \quad (3)$$

$$x_2(t_1) = G(x_1(t_1)). \quad (4)$$

Здесь $f_1(t, x_1, u_1)$, $K_1(t, \tau, x_1, u_1)$ – заданные дискретные по t и (t, τ) соответственно и непрерывные по (x_1, u_1) вместе с частными производными по (x_1, u_1) до второго порядка включительно n -мерные вектор-функции, $f_2(t, x_2, u_2)$, $K_2(t, \tau, x_2, u_2)$ – заданные непрерывные по совокупности переменных вместе с частными производными по (x_2, u_2) до второго порядка включительно n -мерные вектор-функции, x_{10} – заданный постоянный вектор, $G(x_1)$ – заданная дважды непрерывно дифференцируемая n -мерная вектор-функция, $u_1(t)(u_2(t))$ – $r(q)$ -мерная дискретная (кусочно-непрерывная с конечным числом разрыва первого порядка) вектор-функция управляющих воздействий со значениями из $U_1(U_2)$, т.е.

$$\begin{aligned} u_1(t) &\in U_1 \subset R^r, \quad t \in T_1, \\ u_2(t) &\in U_2 \subset R^q, \quad t \in T_2. \end{aligned} \quad (5)$$

Пару $(u_1(t), u_2(t))$, управляющих функций $u_1(t)$ и $u_2(t)$ с вышеприведенными свойствами назовем допустимым управлением.

Предположим, что каждому допустимому управлению $(u_1(t), u_2(t))$ соответствует, имеет единственное дискретное и ограниченное (кусочно-гладкое) решение $x_1(t)(x_2(t))$ задачи (1), (2) ((3), (4)).

Как видно, уравнение (1) является системой нелинейных разностных уравнений типа Вольтерра, а уравнение (3) – нелинейным интегро-дифференциальным уравнением типа Вольтерра.

Предположим, что $\varphi_i(x_i)$, $i=1,2$ – заданные дважды непрерывно дифференцируемые скалярные функции, $F_1(t, \tau, x_1, u_1), (F_2(t, \tau, x_2, u_2))$ – заданная дискретная по (t, τ) (непрерывная по совокупности переменных) скалярная функция, имеющая вторую непрерывную производную по $(x_1, u_1), ((x_2, u_2))$.

На решениях $(x_1(t), x_2(t))$ задачи (1)–(4), порожденных всевозможными допустимыми управлениями $(u_1(t), u_2(t))$, определим функционал типа Больца вида:

$$J(u_1, u_2) = \varphi_1(x_1(t_1)) + \sum_{t=t_0}^{t_1-1} \left[\sum_{\tau=t_0}^t F_1(t, \tau, x_1(\tau), u_1(\tau)) \right] + \varphi_2(x_2(t_2)) + \int_{t_1}^{t_2} \left[\int_{t_1}^t F_2(t, \tau, x_2(\tau), u_2(\tau)) d\tau \right] dt, \quad (6)$$

и рассмотрим задачу нахождения о минимуме функционала (6) при ограничениях (1)–(5).

Допустимый процесс $(u_1(t), u_2(t), x_1(t), x_2(t))$, доставляющий минимальное значение функционалу (6), назовем оптимальным управлением.

2. Полное приращение функционала (6) и сопряженная система

Считая $(u_1(t), u_2(t), x_1(t), x_2(t))$ и $(\bar{u}_1(t) = u_1(t) + \Delta u_1(t), \bar{u}_2(t) = u_2(t) + \Delta u_2(t), \bar{x}_1(t) = x_1(t) + \Delta x_1(t), \bar{x}_2(t) = x_2(t) + \Delta x_2(t))$ некоторыми допустимыми процессами, запишем приращение функционала (6)

$$\begin{aligned} \Delta J(u_1, u_2) &= \varphi_1(\bar{x}_1(t_1)) - \varphi_1(x_1(t_1)) + \varphi_2(\bar{x}_2(t_2)) - \varphi_2(x_2(t_2)) + \\ &+ \sum_{t=t_0}^{t_1-1} \sum_{\tau=t_0}^t [F_1(t, \tau, \bar{x}_1(\tau), \bar{u}_1(\tau)) - F_1(t, \tau, x_1(\tau), u_1(\tau))] + \int_{t_1}^{t_2} \left[\int_{t_1}^t [F_2(t, \tau, \bar{x}_2(\tau), \bar{u}_2(\tau)) - F_2(t, \tau, x_2(\tau), u_2(\tau))] d\tau \right] dt. \end{aligned} \quad (7)$$

Ясно, что $(x_1(t), x_2(t))$ является решением задачи

$$\begin{aligned} \Delta x_1(t+1) &= [f_1(t, \bar{x}_1(t), \bar{u}_1(t)) - f_1(t, x_1(t), u_1(t))] + \\ &+ \sum_{\tau=t_0}^t [K_1(t, \tau, \bar{x}_1(\tau), \bar{u}_1(\tau)) - K_1(t, \tau, x_1(\tau), u_1(\tau))], \quad t \in T_1, \end{aligned} \quad (8)$$

$$\Delta x(t_0) = 0, \quad (9)$$

$$\begin{aligned} \Delta \dot{x}_2(t) &= [f_2(t, \bar{x}_2(t), \bar{u}_2(t)) - f_2(t, x_2(t), u_2(t))] + \\ &+ \int_{t_1}^t [K_2(t, \tau, \bar{x}_2(\tau), \bar{u}_2(\tau)) - K_2(t, \tau, x_2(\tau), u_2(\tau))] d\tau, \quad t \in T_2, \end{aligned} \quad (10)$$

$$x_2(t_1) = G(\bar{x}_1(t_1)) - G(x_1(t_1)). \quad (11)$$

Пусть $\psi_i = \psi_i(t)$, $i=1,2$ – пока произвольные n -мерные вектор-функции.

Используя дискретный аналог теоремы Фубини (см. напр.: [17]) и формулу Дирихле (см. напр.: [3]), а также используя начальные условия (9) и (11), из тождеств (8) и (10) получим

$$\begin{aligned} \psi_1'(t_1) \Delta x_1(t_1) + \sum_{t=t_0}^{t_1-1} \psi_1'(t-1) \Delta x_1(t) &= \sum_{\tau=t_0}^{t_1-1} \psi_1'(t) [f_1(t, \bar{x}_1(t), \bar{u}_1(t)) - f_1(t, x_1(t), u_1(t))] + \\ &+ \sum_{t=t_0}^{t_1-1} \left[\sum_{\tau=t}^{t_1-1} \psi_1'(\tau) [K_1(\tau, t, \bar{x}_1(t), \bar{u}_1(t)) - g_1(\tau, t, x_1(t), u_1(t))] \right], \end{aligned} \quad (12)$$

$$\begin{aligned}
 & \psi_2'(t_2)\Delta x_2(t_2) - \psi_2'(t_1)\left(G(\bar{x}_1(t_1)) - G(x_1(t_1))\right) - \int_{t_1}^{t_2} \dot{\psi}_2'(t)\Delta x_2(t)dt = \\
 & = \int_{t_1}^{t_2} \psi_2'(t)\left[f_2(t, \bar{x}_2(t), \bar{u}_2(t)) - f_2(t, x_2(t), u_2(t))\right]dt + \\
 & + \int_{t_1}^{t_2} \int_t^{\tau} \psi_2'(\tau)\left[K_2(\tau, t, \bar{x}_2(t), \bar{u}_2(t)) - K_2(\tau, t, x_2(t), u_2(t))\right]d\tau dt.
 \end{aligned} \tag{13}$$

Здесь и в дальнейшем штрих (') означает операцию транспонирования.

Введем в рассмотрение аналоги Гамильтониана в виде

$$\begin{aligned}
 H_1(t, x_1(t), u_1(t), \psi_1(t)) &= -\sum_{\tau=t}^{t_1-1} F_1(\tau, t, x_1(t), u_1(t)) + \psi_1'(t)f_1(t, x_1(t), u_1(t)) + \\
 &+ \sum_{\tau=t}^{t_1-1} \psi_1'(\tau)K_1(\tau, t, x_1(t), u_1(t)), \\
 H_2(t, x_2(t), u_2(t), \psi_2(t)) &= -\int_t^{t_2} F_2(\tau, t, x_2(t), u_2(t))d\tau + \psi_2'(t)f_2(t, x_2(t), u_2(t)) + \\
 &+ \int_t^{t_2} \psi_2'(\tau)K_2(\tau, t, x_2(t), u_2(t))d\tau, \\
 M(\psi_2(t_1), x_1(t_1)) &= \psi_2'(t_1)G(x_1(t_1)).
 \end{aligned}$$

Учитывая тождества (12) и (13) и выражения функций Гамильтона, приращение (7) функционала (6) представляется в виде:

$$\begin{aligned}
 & \Delta J(u_1, u_2) = \varphi_1(\bar{x}_1(t_1)) - \varphi_1(x_1(t_1)) + \varphi_2(\bar{x}_2(t_2)) - \varphi_2(x_2(t_2)) + \\
 & + \psi_1'(t_1-1)\Delta x_1(t_1) + \sum_{t=t_0}^{t_1-1} \psi_1'(t-1)\Delta x_1(t) - \sum_{t=t_0}^{t_1-1} \left[H_1(t, \bar{x}_1(t), \bar{u}_1(t), \psi_1(t)) - H_1(t, x_1(t), u_1(t), \psi_1(t)) \right] + \\
 & + \psi_2'(t_2)\Delta x_2(t_2) - \left(M(\psi_2(t_1), \bar{x}_1(t_1)) - M(\psi_2(t_1), x_1(t_1)) \right) - \\
 & - \int_{t_1}^{t_2} \dot{\psi}_2'(t)\Delta x_2(t)dt - \int_{t_1}^{t_2} \left[H_2(t, \bar{x}_2(t), \bar{u}_2(t), \psi_2(t)) - H_2(t, x_2(t), u_2(t), \psi_2(t)) \right]dt.
 \end{aligned} \tag{14}$$

В силу условий гладкости, наложенных на данные задачи (1)–(6), можно записать, что

$$\begin{aligned}
 & \varphi_i(\bar{x}_i(t_i)) - \varphi_i(x_i(t_i)) = \frac{\partial \varphi_i'(x_i(t_i))}{\partial x_i} \Delta x_i(t_i) + \\
 & + \frac{1}{2} \Delta x_i'(t_i) \frac{\partial^2 \varphi_i(x_i(t_i))}{\partial x_i^2} \Delta x_i(t_i) + o_i\left(\|\Delta x_i(t_i)\|^2\right), i=1, 2,
 \end{aligned} \tag{15}$$

$$\begin{aligned}
 & M(\psi_2(t_1), \bar{x}_1(t_1)) - M(\psi_2(t_1), x_1(t_1)) = \frac{\partial M'(\psi_2(t_1), x_1(t_1))}{\partial x_1} \Delta x_1(t_1) + \\
 & + \frac{1}{2} \Delta x_1'(t_1) \frac{\partial^2 M(\psi_2(t_1), x_1(t_1))}{\partial x_1^2} \Delta x_1(t_1) + o_3\left(\|\Delta x_1(t_1)\|^2\right),
 \end{aligned} \tag{16}$$

$$\begin{aligned}
 & H_i(t, \bar{x}_i(t), \bar{u}_i(t), \psi_i(t)) - H_i(t, x_i(t), u_i(t), \psi_i(t)) = \\
 & + \frac{\partial H_i'(t, x_i(t), u_i(t), \psi_i(t))}{\partial x_i} \Delta x_i(t) + \frac{\partial H_i'(t, x_i(t), u_i(t), \psi_i(t))}{\partial u_i} \Delta u_i(t) + \\
 & + \frac{1}{2} \left(\Delta x_i'(t) \frac{\partial^2 H_i(t, x_i(t), u_i(t), \psi_i(t))}{\partial x_i^2} \Delta x_i(t) + 2 \Delta u_i'(t) \frac{\partial^2 H_i(t, x_i(t), u_i(t), \psi_i(t))}{\partial u_i \partial x_i} \Delta x_i(t) + \right. \\
 & \left. + \Delta u_i'(t) \frac{\partial^2 H_i(t, x_i(t), u_i(t), \psi_i(t))}{\partial u_i^2} \Delta u_i(t) \right) + o_{i+3}\left(\left[\|\Delta x_i(t)\| + \|\Delta u_i(t)\|\right]^2\right), i=1, 2.
 \end{aligned} \tag{17}$$

Здесь $\|\alpha\|$ – норма вектора $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)'$, определяемая формулой $\|\alpha\| = \sum_{i=1}^n |\alpha_i|$, $o(\alpha^2)$ – величина более высокого порядка малости, чем α^2 , т. е. $o(\alpha^2)/\alpha^2 \rightarrow 0$ при $\alpha \rightarrow 0$.

Теперь предположим, что вектор-функции $\psi_i = \psi_i(t)$, $i=1,2$, являются решениями задач

$$\psi_1(t-1) = \frac{\partial H_1(t, x_1(t), u_1(t), \psi_1(t))}{\partial x_1}, t \in T_1, \quad (18)$$

$$\psi_1(t_1-1) = \frac{\partial M(\psi_2(t_1), x_1(t_1))}{\partial x_1} - \frac{\partial \Phi_1(x_1(t_1))}{\partial x_1}, \quad (19)$$

$$\dot{\psi}_2(t) = -\frac{\partial H_2(t, x_2(t), u_2(t), \psi_2(t))}{\partial x_2}, t \in T_2, \quad (20)$$

$$\psi_2(t_2) = -\frac{\partial \Phi_2(x_2(t_2))}{\partial x_2}. \quad (21)$$

Задачи (18), (19) и (20), (21) назовем сопряженной системой для рассматриваемой задачи управления

Учитывая разложения (15)–(17) и сопряженную систему (18)–(21), формулу приращения (14) функционала можно представить в виде:

$$\begin{aligned} \Delta J(u_1, u_2) = & \frac{1}{2} \Delta x_1'(t_1) \frac{\partial^2 \Phi_1(x_1(t_1))}{\partial x_1^2} \Delta x_1(t_1) + \frac{1}{2} \Delta x_2'(t_2) \frac{\partial^2 \Phi_2(x_2(t_2))}{\partial x_2^2} \Delta x_2(t_2) - \\ & - \frac{1}{2} \Delta x_1'(t_1) \frac{\partial^2 M(\psi_2(t_1), x_1(t_1))}{\partial x_1^2} \Delta x_1(t_1) - \sum_{t=t_0}^{t_1-1} \frac{\partial H_1'(t, x_1(t), u_1(t), \psi_1(t))}{\partial u_1} \Delta u_1(t) - \\ & - \frac{1}{2} \sum_{t=t_0}^{t_1-1} \left[\Delta x_1'(t) \frac{\partial^2 H_1(t, x_1(t), u_1(t), \psi_1(t))}{\partial x_1^2} \Delta x_1(t) + 2 \Delta u_1'(t) \frac{\partial^2 H_1(t, x_1(t), u_1(t), \psi_1(t))}{\partial u_1 \partial x_1} \Delta x_1(t) + \right. \\ & \left. + \Delta u_1'(t) \frac{\partial^2 H_1(t, x_1(t), u_1(t), \psi_1(t))}{\partial u_1^2} \Delta u_1(t) \right] - \int_{t_1}^{t_2} \frac{\partial H_2'(t, x_2(t), u_2(t), \psi_2(t))}{\partial u_2} \Delta u_2(t) dt - \\ & - \frac{1}{2} \int_{t_1}^{t_2} \left[\Delta x_2'(t) \frac{\partial^2 H_2(t, x_2(t), u_2(t), \psi_2(t))}{\partial x_2^2} \Delta x_2(t) + 2 \Delta u_2'(t) \frac{\partial^2 H_2(t, x_2(t), u_2(t), \psi_2(t))}{\partial u_2 \partial x_2} \Delta x_2(t) + \right. \\ & \left. + \Delta u_2'(t) \frac{\partial^2 H_2(t, x_2(t), u_2(t), \psi_2(t))}{\partial u_2^2} \Delta u_2(t) \right] dt + \sum_{i=1}^2 \left[o_i(\|\Delta x_i(t_i)\|^2) - o_3(\|\Delta x_1(t_1)\|^2) - \right. \\ & \left. - \sum_{t=t_0}^{t_1-1} o_4(\|\Delta x_1(t)\| + \|\Delta u_1(t)\|)^2 \right] - \int_{t_1}^{t_2} o_5(\|\Delta x_2(t)\| + \|\Delta u_2(t)\|)^2 dt. \end{aligned} \quad (22)$$

3. Оценка нормы приращений $(x_1(t), x_2(t))$

В дальнейшем нам понадобится оценка нормы приращения траектории.

Задача (8), (9) может быть записана в виде:

$$\begin{aligned} \Delta x_1(t+1) = & \sum_{\tau=t_0}^t \left[\left(f_1(\tau, \bar{x}_1(\tau), \bar{u}_1(\tau)) - f_1(\tau, x_1(\tau), u_1(\tau)) \right) + \right. \\ & \left. + \sum_{s=\tau}^t \left(K_1(s, \tau, \bar{x}_1(\tau), \bar{u}_1(\tau)) - g_1(s, \tau, x_1(\tau), u_1(\tau)) \right) - \Delta x_1(\tau) \right]. \end{aligned} \quad (23)$$

Из тождества (23), переходя к норме, используя условие Липшица, получаем

$$\|\Delta x_1(t+1)\| \leq L_1 \sum_{\tau=t_0}^t [\|\Delta x_1(\tau)\| + \|\Delta u_1(\tau)\|], \quad (24)$$

где $L_1 = \text{const} > 0$ – некоторая постоянная.

Применяя к последнему неравенству дискретный аналог леммы Гронуолла (см. напр.: [16]), приходим к оценке

$$\|\Delta x_2(t)\| \leq L_2 \sum_{\tau=t_0}^t \|\Delta u_1(\tau)\|, \quad (25)$$

$L_2 = \text{const} > 0$.

Далее из задачи (10), (11) получаем, что

$$\begin{aligned} \Delta x_2(t) = & \int_{t_1}^t [f_2(\tau, \bar{x}_2(\tau), \bar{u}_2(\tau)) - f_2(\tau, x_2(\tau), u_2(\tau))] + \\ & + \int_{\tau}^t [K_2(s, \tau, \bar{x}_2(\tau), \bar{u}_2(\tau)) - K_2(s, \tau, x_2(\tau), u_2(\tau))] ds d\tau + (G(\bar{x}_1(t_1)) - G(x_1(t_1))). \end{aligned}$$

Отсюда получаем

$$\|\Delta x_2(t)\| \leq L_3 \left[\|\Delta x_1(t_1)\| + \int_{t_1}^t [\|\Delta x_2(\tau)\| + \|\Delta u_2(\tau)\|] d\tau \right],$$

$L_3 = \text{const} > 0$.

Применяя к последнему неравенству аналог леммы Гронуолла (см. напр., [17]), приходим к оценке

$$\|\Delta x_2(t)\| \leq L_4 \left[\|\Delta x_1(t_1)\| + \int_{t_1}^t \|\Delta u_2(\tau)\| d\tau \right], \quad (26)$$

$L_4 = \text{const} > 0$.

4. Специальное приращение второго порядка функционала

Полагая $\Delta u_2(t) = 0$, приращение $\Delta u_1(t)$ управления $u_1(t)$ определим по формуле

$$\Delta u_1(t; \varepsilon) = \varepsilon (v_1(t) - u_1(t)). \quad (27)$$

Здесь $\varepsilon \in [0, 1]$ – произвольное число, а $v_1(t) \in U_1$, $t \in T_1$, – произвольное допустимое управление.

Через $(\Delta x_1(t; \varepsilon), \Delta x_2(t; \varepsilon))$ определим специальное приращение траектории $(x_1(t), x_2(t))$, отвечающее специальному приращению (27) управляющей функции $u_1(t)$.

Тогда из оценок (25) и (26) следует, что

$$\begin{aligned} \|\Delta x_1(t)\| & \leq L_5 \varepsilon, t \in T_1 \cup t_1, \\ \|\Delta x_2(t)\| & \leq L_6 \varepsilon, t \in T_2, \end{aligned} \quad (28)$$

$L_5 = \text{const} > 0, L_6 = \text{const} > 0$ – некоторые постоянные.

Далее из задач (8), (9) и (10), (11) линеаризацией получаем, что

$$\begin{aligned} \Delta x_1(t+1; \varepsilon) = & \frac{\partial f_1(t, x_1(t), u_1(t))}{\partial x_1} \Delta x_1(t; \varepsilon) + \frac{\partial f_1(t, x_1(t), u_1(t))}{\partial u_1} \Delta u_1(t; \varepsilon) + \\ & + \sum_{\tau=t_0}^t \left(\frac{\partial K_1(t, \tau, x_1(\tau), u_1(\tau))}{\partial x_1} \Delta x_1(\tau; \varepsilon) + \frac{\partial K_1(t, \tau, x_1(\tau), u_1(\tau))}{\partial u_1} \Delta u_1(\tau; \varepsilon) \right) + o_6(\varepsilon), \end{aligned} \quad (29)$$

$$\Delta x_1(t_0; \varepsilon) = 0, \quad (30)$$

$$\Delta \dot{x}_2(t; \varepsilon) = \frac{\partial f_2(t, x_2(t), u_2(t))}{\partial x_2} \Delta x_2(t; \varepsilon) + \int_{t_0}^t \frac{\partial K_2(t, \tau, x_2(\tau), u_2(\tau))}{\partial x_2} d\tau + o_7(\varepsilon), \quad (31)$$

$$\Delta x_2(t_1; \varepsilon) = \frac{\partial G(x_1(t_1))}{\partial x_1} \Delta x_1(t_1; \varepsilon) + o_8(\varepsilon). \quad (32)$$

Имеет место

Лемма 1. Решения $\Delta x_1(t; \varepsilon)$ и $\Delta x_2(t; \varepsilon)$ задач (29), (30) и (31), (32) допускают разложения

$$\begin{aligned} \Delta x_1(t; \varepsilon) &= \varepsilon l_1(t) + o_9(\varepsilon; t), \\ \Delta x_2(t; \varepsilon) &= \varepsilon l_2(t) + o_{10}(\varepsilon; t). \end{aligned} \quad (33)$$

Здесь $l_1(t)$ и $l_2(t)$ являются решениями соответственно задач

$$\begin{aligned} l_1(t+1) &= \frac{\partial f_1(t, x_1(t), u_1(t))}{\partial x_1} l_1(t) + \frac{\partial f_1(t, x_1(t), u_1(t))}{\partial u_1} (v_1(t) - u_1(t)) + \\ &+ \sum_{\tau=t_0}^t \left(\frac{\partial K_1(t, \tau, x_1(\tau), u_1(\tau))}{\partial x_1} l_1(\tau) + \frac{\partial K_1(t, \tau, x_1(\tau), u_1(\tau))}{\partial u_1} (v_1(\tau) - u_1(\tau)) \right), \end{aligned} \quad (34)$$

$$l_1(t_0) = 0, \quad (35)$$

$$\dot{l}_2(t) = \frac{\partial f_2(t, x_2(t), u_2(t))}{\partial x_2} l_2(t) + \int_{t_1}^t \frac{\partial K_2(t, \tau, x_2(\tau), u_2(\tau))}{\partial x_2} l_2(\tau) d\tau, \quad (36)$$

$$l_2(t_1) = \frac{\partial G(x_1(t_1))}{\partial x_1} l_1(t_1). \quad (37)$$

Теперь, предполагая, что $\Delta u_1(t) = 0$, через

$$\Delta u_2(t; \mu) = \mu [v_2(t) - u_2(t)] \quad (38)$$

обозначим специальное приращение допустимой управляющей функции $\Delta u_2(t)$. Здесь $\mu \in [0, 1]$ – произвольное число, а $v_2(t) \in U_2$, $t \in T_2$ – произвольное допустимое управление.

Ясно, что при этом $\Delta x_1(t) \equiv 0$, а специальное приращение $\Delta x_2(t; \mu)$ траектории $x_2(t)$, соответствующее специальному приращению (38) управления $u_2(t)$, будет решением задачи

$$\begin{aligned} \Delta \dot{x}_2(t; \mu) &= [f_2(t, x_2(t) + \Delta x_2(t; \mu), u_2(t) + \Delta u_2(t; \mu)) - f_2(t, x_2(t), u_2(t))] + \\ &+ \int_{t_1}^t [K_2(t, \tau, x_2(\tau) + \Delta x_2(\tau; \mu), u_2(\tau) + \Delta u_2(\tau; \mu)) - K_2(t, \tau, x_2(\tau), u_2(\tau))] d\tau, \end{aligned} \quad (39)$$

$$\Delta x_2(t_1; \mu) = 0. \quad (40)$$

Из задачи (39), (40) получаем, что $\Delta x_2(t; \mu)$ является решением линеаризованной задачи

$$\begin{aligned} \Delta \dot{x}_2(t; \mu) &= \frac{\partial f_2(t, x_2(t), u_2(t))}{\partial x_2} \Delta x_2(t; \mu) + \mu \frac{\partial f_2(t, x_2(t), u_2(t))}{\partial u_2} [v_2(t) - u_2(t)] + \\ &+ o_{11}(\|\Delta x_2(t; \mu)\| + \|\Delta u_2(t; \mu)\|) + \int_{t_1}^t \left[\frac{\partial K_2(t, \tau, x_2(\tau), u_2(\tau))}{\partial x_2} \Delta x_2(\tau; \mu) + \right. \\ &\left. + \mu \frac{\partial K_2(t, \tau, x_2(\tau), u_2(\tau))}{\partial u_2} (v_2(\tau) - u_2(\tau)) + o_{12}(\|\Delta x_2(\tau)\| + \|\Delta u_2(\tau)\|) \right] d\tau, \end{aligned} \quad (41)$$

$$\Delta x_2(t_1; \mu) = 0. \quad (42)$$

Учитывая оценку (см. (26))

$$\|\Delta x_2(t; \mu)\| \leq L_7 \int_{t_1}^t \|\Delta u_2(\tau; \mu)\| d\tau,$$

с помощью линеаризованной задачи (41)–(42) доказывается

Лемма 2. Решение $\Delta x_2(t; \mu)$ линеаризованной задачи (41), (42) допускает разложение

$$\Delta x_2(t; \mu) = \mu l_3(t) + o_{13}(\mu; t), \quad (43)$$

где $l_3(t)$ является решением задачи

$$\begin{aligned} \dot{l}_3(t) = & \frac{\partial f_2(t, x_2(t), u_2(t))}{\partial x_2} l_3(t) + \frac{\partial f_2(t, x_2(t), u_2(t))}{\partial u_2} (v_2(t) - u_2(t)) + \\ & + \int_{t_1}^t \left[\frac{\partial K_2(t, \tau, x_2(\tau), u_2(\tau))}{\partial x_2} l_3(\tau) + \frac{\partial K_2(t, \tau, x_2(\tau), u_2(\tau))}{\partial u_2} (v_2(\tau) - u_2(\tau)) \right] d\tau, \end{aligned} \quad (44)$$

$$l_3(t_1) = 0.$$

Принимая во внимания формулу (28) и разложения (33), специальное частное приращение функционала (6) с учетом (22) записывается в виде:

$$\begin{aligned} J(u_1(t) + \Delta u_1(t; \varepsilon), u_2(t)) - J(u_1(t), u_2(t)) = & \frac{\varepsilon^2}{2} l_1'(t_1) \frac{\partial^2 \varphi_1(x_1(t_1))}{\partial x_1^2} l_1(t_1) + \frac{\varepsilon^2}{2} l_2'(t_2) \frac{\partial^2 \varphi_2(x_2(t_2))}{\partial x_2^2} l_2(t_2) - \\ & - \frac{\varepsilon^2}{2} l_1'(t_1) \frac{\partial^2 M(\psi_2(t_1), x_1(t_1))}{\partial x_1^2} l_1(t_1) - \varepsilon \sum_{t=t_0}^{t_1-1} \frac{\partial H_1'(t, x_1(t), u_1(t), \psi_1(t))}{\partial u_1} (v_1(t) - u_1(t)) - \\ & - \frac{\varepsilon^2}{2} \sum_{t=t_0}^{t_1-1} \left[l_1'(t) \frac{\partial^2 H_1(t, x_1(t), u_1(t), \psi_1(t))}{\partial x_1^2} l_1(t) + 2(v_1(t) - u_1(t))' \frac{\partial^2 H_1(t, x_1(t), u_1(t), \psi_1(t))}{\partial u_1 \partial x_1} l_1(t) + \right. \\ & \left. + (v_1(t) - u_1(t))'(t) \frac{\partial^2 H_1(t, x_1(t), u_1(t), \psi_1(t))}{\partial u_1^2} (v_1(t) - u_1(t)) \right] - \\ & - \frac{\varepsilon^2}{2} \int_{t_1}^{t_2} l_2'(t) \frac{\partial^2 H_2'(t, x_2(t), u_2(t), \psi_2(t))}{\partial x_2^2} l_2(t) dt + o(\varepsilon^2). \end{aligned} \quad (45)$$

Далее из формулы приращения (22) на основании формул (38) и (43) получаем, что

$$\begin{aligned} J(u_1(t), u_2(t) + \Delta u_2(t; \mu)) - J(u_1(t), u_2(t)) = & -\mu \int_{t_1}^{t_2} \frac{\partial H_2'(t, x_2(t), u_2(t), \psi_2(t))}{\partial u_2} [v_2(t) - u_2(t)] dt + \\ & + \frac{\mu^2}{2} l_3'(t_2) \frac{\partial^2 \varphi_2(x_2(t_2))}{\partial x_2^2} l_3(t_2) - \frac{\mu^2}{2} \int_{t_1}^{t_2} \left[l_3'(t) \frac{\partial^2 H_2'(t, x_2(t), u_2(t), \psi_2(t))}{\partial x_2^2} l_3(t) dt + \right. \\ & \left. + 2[v_2(t) - u_2(t)]' \frac{\partial^2 H_2'(t, x_2(t), u_2(t), \psi_2(t))}{\partial u_2 \partial x_2} l_3(t) dt + \right. \\ & \left. + [v_2(t) - u_2(t)]' \frac{\partial^2 H_2'(t, x_2(t), u_2(t), \psi_2(t))}{\partial u_2^2} [v_2(t) - u_2(t)] \right] dt + o(\mu^2). \end{aligned} \quad (46)$$

5. Необходимые условия оптимальности

Установленные частные формулы приращения с учетом независимости друг от друга управляющих функций $u_1(t)$ и $u_2(t)$ позволяют сформулировать необходимые условия оптимальности первого и второго порядков.

Пусть $(u_1(t), u_2(t))$ – оптимальное управление. Тогда из разложений (45) и (46) следует

Теорема 1 (линеаризованное условие оптимальности). Для оптимальности допустимого управления $(u_1(t), u_2(t))$ необходимо, чтобы неравенства

$$-\sum_{t=t_0}^{t_1-1} \frac{\partial H_1'(t, x_1(t), u_1(t), \psi_1(t))}{\partial u_1} (v_1(t) - u_1(t)) \leq 0, \quad (49)$$

$$\int_{t_1}^{t_2} \frac{\partial H_2'(t, x_2(t), u_2(t), \psi_2(t))}{\partial u_2} [v_2(t) - u_2(t)] dt \leq 0 \quad (50)$$

выполнялись для всех $v_1(t) \in U_1$, $t \in T_1$, и $v_2(t) \in U_2$, $t \in T_2$, соответственно.

Неравенства (49) и (50) являются необходимыми условиями оптимальности первого порядка. Однако в некоторых случаях они могут, вырождаясь, становиться неэффективными (см.: [5–7]).

Изучим случай их вырождения.

Определение. Допустимое управление $(u_1(t), u_2(t))$ назовем квазиособым, если для всех $v_1(t) \in U_1$, $t \in T_1$, и $v_2(t) \in U_2$, $t \in T_2$, выполняются соответственно тождества

$$\sum_{t=t_0}^{t_1-1} \frac{\partial H_1'(t, x_1(t), u_1(t), \psi_1(t))}{\partial u_1} (v_1(t) - u_1(t)) = 0, \quad (51)$$

$$\int_{t_1}^{t_2} \frac{\partial H_2'(t, x_2(t), u_2(t), \psi_2(t))}{\partial u_2} [v_2(t) - u_2(t)] dt = 0. \quad (52)$$

Исследование квазиособых управлений в задачах оптимального управления, описываемых обыкновенными дифференциальными уравнениями, началось с работ Р. Габасова и Ф.М. Кирилловой и их учеников (обзор некоторых результатов имеется, например, в [5–8]).

Перейдем к исследованию квазиособых управлений.

Пусть $(u_1(t), u_2(t))$ – квазиособое оптимальное управление.

Тогда из разложений (45) и (46) следует, что

$$\begin{aligned} l_1'(t_1) \frac{\partial^2 \Phi_1(x_1(t_1))}{\partial x_1^2} l_1(t_1) - \sum_{t=t_0}^{t_1-1} \left[l_1'(t) \frac{\partial^2 H_1(t, x_1(t), u_1(t), \psi_1(t))}{\partial x_1^2} l_1(t) + 2(v_1(t) - u_1(t))' \frac{\partial^2 H_1(t, x_1(t), u_1(t), \psi_1(t))}{\partial u_1 \partial x_1} l_1(t) + \right. \\ \left. + (v_1(t) - u_1(t))' (t) \frac{\partial^2 H_1(t, x_1(t), u_1(t), \psi_1(t))}{\partial u_1^2} (v_1(t) - u_1(t)) \right] - l_1'(t_1) \frac{\partial^2 M(\psi_2(t_1), x_1(t_1))}{\partial x_1^2} l_1(t_1) + \\ + l_2'(t_2) \frac{\partial^2 \Phi_2(x_2(t_2))}{\partial x_2^2} l_2(t_2) - \int_{t_1}^{t_2} l_2'(t) \frac{\partial^2 H_2(t, x_2(t), u_2(t), \psi_2(t))}{\partial x_2^2} l_2(t) dt \geq 0, \end{aligned} \quad (53)$$

$$\begin{aligned} l_3'(t_2) \frac{\partial^2 \Phi_2(x_2(t_2))}{\partial x_2^2} l_3(t_2) - \int_{t_1}^{t_2} \left[l_3'(t) \frac{\partial^2 H_2(t, x_2(t), u_2(t), \psi_2(t))}{\partial x_2^2} l_3(t) dt + \right. \\ \left. + 2[v_2(t) - u_2(t)]' \frac{\partial^2 H_2(t, x_2(t), u_2(t), \psi_2(t))}{\partial u_2 \partial x_2} l_3(t) dt + \right. \\ \left. + [v_2(t) - u_2(t)]' \frac{\partial^2 H_2(t, x_2(t), u_2(t), \psi_2(t))}{\partial u_2^2} [v_2(t) - u_2(t)] \right] dt \geq 0. \end{aligned} \quad (54)$$

Неравенства (53) и (54), являясь необходимыми условиями оптимальности квазиособых управлений, носят неявный характер, что делает их использование практически невозможным. Однако из них можно получить довольно общие необходимые условия оптимальности, носящие явный характер.

Пусть $F_i(t, \tau)$, $i = 1, 2, (n \times n)$ – матричные функции, являющиеся решениями задач

$$F_1(t, \tau - 1) = F_1(t, \tau) \frac{\partial f_1(\tau, x_1(\tau), u_1(\tau))}{\partial x_1} + \sum_{s=\tau}^{t-1} F_1(t, s) \frac{\partial K_1(s, \tau, x_1(\tau), u_1(\tau))}{\partial x_1}, \quad F_1(t, t - 1) = E,$$

$$\frac{\partial F_2(t, \tau)}{\partial \tau} = -F_2(t, \tau) \frac{\partial f_2(\tau, x_2(\tau), u_2(\tau))}{\partial x_2} + \int_s^\tau F_2(t, s) \frac{\partial K_2(s, \tau, x_2(\tau), u_2(\tau))}{\partial x_2} ds, \quad F_2(t, t) = E,$$

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

Тогда решения задач (34), (35) и (36), (37) можно представить в виде:

$$l_1(t) = \sum_{\tau=t_0}^{t-1} \left[F_1(t, \tau) \frac{\partial f_1(\tau, x_1(\tau), u_1(\tau))}{\partial u_1} (v_1(\tau) - u_1(\tau)) + \right. \\ \left. + \sum_{s=\tau}^{t-1} F_1(t, s) \frac{\partial K_1(s, \tau, x_1(\tau), u_1(\tau))}{\partial u_1} (v_1(\tau) - u_1(\tau)) \right], \quad (55)$$

$$l_2(t) = F_2(t, t_1) \frac{\partial G(x_1(t_1))}{\partial x_1} l_1(t_1).$$

Отсюда, принимая во внимания представление (55), получим

$$l_2(t) = F_2(t, t_1) \frac{\partial G(x_1(t_1))}{\partial x_1} \left[\sum_{\tau=t_0}^{t-1} \left[F_1(t_1, \tau) \frac{\partial f_1(\tau, x_1(\tau), u_1(\tau))}{\partial u_1} (v_1(\tau) - u_1(\tau)) + \right. \right. \\ \left. \left. + \sum_{s=\tau}^{t_1-1} F_1(t_1, s) \frac{\partial K_1(s, \tau, x_1(\tau), u_1(\tau))}{\partial u_1} (v_1(\tau) - u_1(\tau)) \right] \right]. \quad (56)$$

Введя обозначения

$$Q_1(t, \tau) = F_1(t, \tau) \frac{\partial f_1(\tau, x_1(\tau), u_1(\tau))}{\partial u_1} + \sum_{s=\tau}^{t-1} F_1(t, s) \frac{\partial K_1(s, \tau, x_1(\tau), u_1(\tau))}{\partial u_1},$$

$$Q_2(t, \tau) = F_2(t, t_1) \frac{\partial G(x_1(t_1))}{\partial x_1} \left[F_1(t_1, \tau) \frac{\partial f_1(\tau, x_1(\tau), u_1(\tau))}{\partial u_1} + \sum_{s=\tau}^{t_1-1} F_1(t_1, s) \frac{\partial K_1(s, \tau, x_1(\tau), u_1(\tau))}{\partial u_1} \right],$$

формулы (55) и (56) запишем в виде:

$$l_1(t) = \sum_{\tau=t_0}^{t-1} Q_1(t, \tau) (v_1(\tau) - u_1(\tau)), \quad (57)$$

$$l_2(t) = \sum_{\tau=t_0}^{t_1-1} Q_2(t, \tau) (v_2(\tau) - u_2(\tau)). \quad (58)$$

На основании представлений (57) и (58) доказывается справедливость тождеств

$$l_1'(t_1) \frac{\partial^2 \varphi_1(x_1(t_1))}{\partial x_1^2} l_1(t_1) = \sum_{\tau=t_0}^{t_1-1} \sum_{s=t_0}^{t_1-1} (v_1(\tau) - u_1(\tau))' Q_1(t_1, \tau) \frac{\partial^2 \varphi_1(x_1(t_1))}{\partial x_1^2} Q_1(t_1, s) (v_1(s) - u_1(s)), \quad (59)$$

$$l_1'(t_1) \frac{\partial^2 M(\psi_2(t_1), x_1(t_1))}{\partial x_1^2} l_1(t_1) = \\ = \sum_{\tau=t_0}^{t_1-1} \sum_{s=t_0}^{t_1-1} (v_1(\tau) - u_1(\tau))' Q_1(t_1, \tau) \frac{\partial^2 M(\psi_2(t_1), x_1(t_1))}{\partial x_1^2} Q_1(t_1, s) (v_1(s) - u_1(s)), \quad (60)$$

$$l_2'(t_2) \frac{\partial^2 \varphi_2(x_2(t_2))}{\partial x_2^2} l_2(t_2) = \\ = \sum_{\tau=t_0}^{t_1-1} \sum_{s=t_0}^{t_1-1} (v_2(\tau) - u_2(\tau))' Q_2(t_2, \tau) \frac{\partial^2 \varphi_2(x_2(t_2))}{\partial x_2^2} Q_2(t_2, s) (v_2(s) - u_2(s)), \quad (61)$$

$$\sum_{t=t_0}^{t_1-1} l_1'(t) \frac{\partial^2 H_1(t, x_1(t), u_1(t), \psi_1(t))}{\partial x_1^2} l_1(t) = \\ = \sum_{\tau=t_0}^{t_1-1} \sum_{s=t_0}^{t_1-1} (v_1(\tau) - u_1(\tau))' \left[\sum_{t=\max(\tau, s)+1}^{t_1-1} Q_1(t, \tau) \frac{\partial^2 H_1(t, x_1(t), u_1(t), \psi_1(t))}{\partial x_1^2} Q_1(t, s) \right] (v_1(s) - u_1(s)), \quad (62)$$

$$\sum_{t=t_0}^{t_1-1} (v_1(t) - u_1(t))' \frac{\partial^2 H_1(t, x_1(t), u_1(t), \psi_1(t))}{\partial u_1 \partial x_1} l_1(t) =$$

$$= \sum_{t=t_0}^{t_1-1} \left[\sum_{\tau=t_0}^{t-1} (v_1(t) - u_1(t))' \frac{\partial^2 H_1(t, x_1(t), u_1(t), \psi_1(t))}{\partial u_1 \partial x_1} Q_1(t, \tau) (v_1(\tau) - u_1(\tau)) \right], \quad (63)$$

$$\int_{t_1}^{t_2} l_2'(t) \frac{\partial^2 H_2'(t, x_2(t), u_2(t), \psi_2(t))}{\partial x_2^2} l_2(t) dt =$$

$$= \sum_{\tau=t_0}^{t_1-1} \sum_{s=t_0}^{t_1-1} (v_1(\tau) - u_1(\tau))' \left[\int_{t_1}^{t_2} Q_2'(t, \tau) \frac{\partial^2 H_2'(t, x_2(t), u_2(t), \psi_2(t))}{\partial x_2^2} Q_2(t, s) dt \right] (v_1(s) - u_1(s)). \quad (64)$$

Введя обозначение

$$K_1(\tau, s) = -Q_1'(t_1, \tau) \frac{\partial^2 \varphi_1(x(t_1))}{\partial x_1^2} Q_1(t_1, s) -$$

$$-Q_2'(t_2, \tau) \frac{\partial^2 \varphi_2(x(t_2))}{\partial x_2^2} Q_2(t_2, s) + Q_1(t_1, \tau) \frac{\partial^2 M(\psi_2(t_1), x_1(t_1))}{\partial x_1^2} Q_1(t_1, s) +$$

$$+ \sum_{t=\max(\tau, s)+1}^{t_1-1} Q_1(t, \tau) \frac{\partial^2 H_1(t, x_1(t), u_1(t), \psi_1(t))}{\partial x_1^2} Q_1(t, s) + \int_{t_1}^{t_2} Q_2'(t, \tau) \frac{\partial^2 H_2'(t, x_2(t), u_2(t), \psi_2(t))}{\partial x_2^2} Q_2(t, s) dt$$

и учитывая тождества (59)–(64), неравенство (53) запишем в виде:

$$\sum_{\tau=t_0}^{t_1-1} \sum_{s=t_0}^{t_1-1} (v_1(\tau) - u_1(\tau))' K_1(\tau, s) (v_1(s) - u_1(s)) +$$

$$+ 2 \sum_{t=t_0}^{t_1-1} \left[\sum_{\tau=t_0}^{t-1} (v_1(t) - u_1(t))' \frac{\partial^2 H_1(t, x_1(t), u_1(t), \psi_1(t))}{\partial u_1 \partial x_1} Q_1(t, \tau) (v_1(\tau) - u_1(\tau)) \right] +$$

$$+ \sum_{t=t_0}^{t_1-1} (v_1(t) - u_1(t))' \frac{\partial^2 H_1(t, x_1(t), u_1(t), \psi_1(t))}{\partial u_1^2} (v_1(t) - u_1(t)) \leq 0. \quad (65)$$

Далее решение $l_3(t)$ задачи (64), (65) представляется в виде:

$$l_3(t) = \int_{t_1}^t \left[F_2(t, \tau) \frac{\partial f_2(\tau, x_2(\tau), u_2(\tau))}{\partial u_2} u_2(\tau) + \int_{\tau}^t F_2(t, s) \frac{\partial K_2(s, \tau, x_2(\tau), u_2(\tau))}{\partial u_2} ds \right] (v_2(\tau) - u_2(\tau)) d\tau.$$

Введя обозначение

$$Q_3(t, \tau) = F_2(t, \tau) \frac{\partial f_2(\tau, x_2(\tau), u_2(\tau))}{\partial u_2} u_2(\tau) + \int_{\tau}^t F_2(t, s) \frac{\partial K_2(s, \tau, x_2(\tau), u_2(\tau))}{\partial u_2} ds,$$

из этого представления получаем, что

$$l_3(t) = \int_{t_1}^t Q_3(t, \tau) (v_2(\tau) - u_2(\tau)) d\tau. \quad (66)$$

Поэтому

$$l_3'(t_2) \frac{\partial^2 \varphi_2(x_2(t_2))}{\partial x_2^2} l_3(t_2) = \int_{t_1}^{t_2} \int_{t_1}^{t_2} (v_2(\tau) - u_2(\tau))' Q_3'(t_2, \tau) \frac{\partial^2 \varphi_2(x_2(t_2))}{\partial x_2^2} Q_2(t_2, s) (v_2(s) - u_2(s)) ds d\tau, \quad (67)$$

$$\int_{t_1}^{t_2} [v_2(t) - u_2(t)]' \frac{\partial^2 H_2'(t, x_2(t), u_2(t), \psi_2(t))}{\partial u_2 \partial x_2} l_3(t) dt =$$

$$= \int_{t_1}^{t_2} \left[\int_{t_1}^t [v_2(t) - u_2(t)]' \frac{\partial^2 H_2'(t, x_2(t), u_2(t), \psi_2(t))}{\partial u_2 \partial x_2} Q_3(t, \tau) (v_2(\tau) - u_2(\tau)) d\tau \right] dt, \quad (68)$$

$$\begin{aligned} & \int_{t_1}^{t_2} \int_{t_1}^{t_2} l_3'(t) \frac{\partial^2 H_2'(t, x_2(t), u_2(t), \psi_2(t))}{\partial x_2^2} l_3(t) dt = \\ & = \int_{t_1}^{t_2} \int_{t_1}^{t_2} (v_2(\tau) - u_2(\tau))' \left[\int_{\max(\tau, s)}^{t_1} Q_3'(t, \tau) \frac{\partial^2 H_2'(t, x_2(t), u_2(t), \psi_2(t))}{\partial x_2^2} Q_3(t, s) dt \right] (v_2(s) - u_2(s)) ds d\tau. \end{aligned} \quad (69)$$

Учитывая тождества (67)–(69) и вводя обозначение

$$K_2(\tau, s) = -Q_3'(t_2, \tau) \frac{\partial^2 \varphi_2(x_2(t_2))}{\partial x_2^2} Q_2(t_2, s) + \int_{\max(\tau, s)}^{t_1} Q_3'(t, \tau) \frac{\partial^2 H_2'(t, x_2(t), u_2(t), \psi_2(t))}{\partial x_2^2} Q_3(t, s) dt,$$

представим неравенство (54) в виде:

$$\begin{aligned} & \int_{t_1}^{t_2} \int_{t_1}^{t_2} (v_2(\tau) - u_2(\tau))' K_2(\tau, s) (v_2(s) - u_2(s)) ds d\tau + \\ & + 2 \int_{t_1}^{t_2} \left[\int_{t_1}^t [v_2(t) - u_2(t)]' \frac{\partial^2 H_2'(t, x_2(t), u_2(t), \psi_2(t))}{\partial u_2 \partial x_2} Q_3(t, \tau) (v_2(\tau) - u_2(\tau)) d\tau \right] dt + \\ & + \int_{t_1}^{t_2} [v_2(t) - u_2(t)]' \frac{\partial^2 H_2'(t, x_2(t), u_2(t), \psi_2(t))}{\partial u_2^2} [v_2(t) - u_2(t)] dt \leq 0. \end{aligned} \quad (70)$$

Следовательно, имеет место

Теорема 2 (необходимое условие оптимальности квазиособых управлений). Для оптимальности квазиособого управления $(u_1(t), u_2(t))$ необходимо, чтобы неравенства (65) и (70) выполнялись для всех $v_1(t) \in U_1$, $t \in T_1$, и $v_2(t) \in U_2$, $t \in T_2$, соответственно.

Как видно, необходимые условия оптимальности (65) и (70), являясь довольно общими, носят явный характер. Более того, из них, определяя управляющие функции $v_i(t)$, $i = 1, 2$, специальным образом, можно получить относительно легко проверяемые необходимые условия оптимальности.

Например, имеет место

Теорема 3. Для оптимальности квазиособого управления $(u_1(t), u_2(t))$ необходимо, чтобы неравенства

$$\begin{aligned} & (w_1 - u_1(\theta))' K_1(\theta, \theta) (w_1 - u_1(\theta)) + (w_1 - u_1(\theta)) \frac{\partial^2 H_1(\theta, x_1(\theta), u_1(\theta), \psi_1(\theta))}{\partial u_1^2} (w_1 - u_1(\theta)) \leq 0, \\ & (w_2 - u_2(\theta))' \frac{\partial^2 H_2(\theta, x_2(\theta), u_2(\theta), \psi_2(\theta))}{\partial u_2^2} (w_2 - u_2(\theta)) \leq 0 \end{aligned}$$

выполнялись для всех $w_1 \in U_1$, $\theta \in T_1$, и $w_2 \in U_2$, $\theta \in [t_1, t_2]$, соответственно. Здесь $\theta \in [t_0, t_1]$ – произвольная точка непрерывности управляющей функции $u_2(t)$.

Заключение

В работе ставится и исследуется одна дискретно-непрерывная ступенчатая задача оптимального управления, описываемая системами нелинейных разностных уравнений и нелинейных интегродифференциальных уравнений типа Вольтерра с функционалом качества типа Больца. Области управления являются выпуклыми и ограниченными множествами.

Построена носящий конструктивный характер формула приращения второго порядка на специальных приращениях управляющих функций. Доказан аналог линеаризованного принципа максимума, а затем исследован случай его вырождения (квазиособый случай).

Используя доказанное неявное условие оптимальности установлено общее необходимое условие оптимальности квазиособых управлений, носящее конструктивный характер.

Список источников

1. Понтрягин Л.С., Болтянский В.Г., Гамкрелидзе Р.В., Мищенко Е.Ф. Математическая теория оптимальных процессов. М. : Наука, 1986. 384 с.
2. Габасов Р., Кириллова Ф.М., Альсевич В.В. Методы оптимизации. Минск : Четыре четверти, 2011. 472 с.
3. Алексеев В.М., Тихомиров В.М., Фомин С.В. Оптимальное управление. М. : Физматлит, 2018. 384 с.
4. Альсевич В.В. Оптимизация динамических систем с запаздываниями. Минск : Белорус. гос. ун-т, 2000. 198 с.
5. Габасов Р., Кириллова Ф.М. Особые оптимальные управления. М. : URSS, 2011. 256 с.
6. Габасов Р., Кириллова Ф.М., Мансимов К.Б. Необходимые условия оптимальности второго порядка (обзор). Минск, 1982. 48 с. (Препринт ИМ АН БССР; № 30 (155)).
7. Мансимов К.Б. Особые управления в системах с запаздыванием. Баку : ЭЛМ, 1999. 176 с.
8. Мансимов К.Б. Особые управления в задачах с распределенными параметрами // Современная математика и ее приложения. 2006. Т. 2. С. 39–83.
9. Величенко В.В. Оптимальное управление составными системами // Доклады Академии наук СССР. 1967. Т. 176, № 4. С. 154–156.
10. Захаров Г.К. Оптимизация ступенчатых систем с управляемыми условиями перехода // Автоматика и телемеханика. 1993. № 6. С. 32–36.
11. Агафонова И.А., Гулин Л.Л., Расина И.В. Математическое моделирование и оптимизация процесса метилирования ди-натриевой соли сульфаминоантипина. Иркутск, 1978. 19 с. Деп. в ВИНТИ АН СССР. 10.11.1978. № 3457.
12. Исмаилов Р.Р., Мансимов К.Б. Об условиях оптимальности в одной ступенчатой задаче управления // Журнал вычислительной математики и математической физики. 2006. № 10. С. 1758–1770.
13. Розова В.Н. Оптимальное управление ступенчатыми системами с неинтегральным функционалом // Вестник РУДН. Сер. Прикладная и компьютерная математика. 2002. № 1 (1). С. 131–136.
14. Ащепков Л.Т. Оптимальное управление разрывными системами. Новосибирск : Наука, 1987. 226 с.
15. Мамедова Т.Ф., Мансимов К.Б. Об оптимальности особых управлений в задаче управления ступенчатыми дискретными двухпараметрическими системами // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2018. № 42. С. 12–22.
16. Керимова А.В., Мансимов К.Б. Аналог дискретного принципа максимума Понтрягина в ступенчатой задаче правления для систем, описываемых разностными уравнениями типа Вольтерра // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2023. № 65. С. 4–14.
17. Souyousefain M., Leela S. Stability results for difference equations of Volterra type // Appl. Math. Comput. 1990. V. 36 (1). P. 51–61.

References

1. Pontryagin, L.S., Boltyansky, V.G., Gamkrelidze, R.V. & Mishchenko, E.F. (1986) *Matematicheskaya teoriya optimal'nykh protsessov* [Mathematical Theory of Optimal Processes]. Moscow: Nauka.
2. Gabasov, R., Kirillova, F.M. & Alseovich, V.V. (2011) *Metody optimizatsii* [Optimization Methods]. Minsk: Chetyre chetverti.
3. Alekseev, V.M., Tikhomirov, V.M. & Fomin, S.V. (2018) *Optimal'noe upravlenie* [Optimal Control]. Moscow: Fizmatlit.
4. Alseovich, V.V. (2000) *Optimizatsiya dinamicheskikh sistem s zapazdyvaniyami* [Optimization of Dynamic Systems with Delays]. Minsk: BSU.
5. Gabasov, R. & Kirillova, F.M. (2011) *Osobyie optimal'nye upravleniya* [Special Optimal Controls]. Moscow: URSS.
6. Gabasov, R., Kirillova, F.M. & Mansimov, K.B. (1982) *Neobkhodimye usloviya optimal'nosti vtorogo poryadka (obzor)* [Necessary conditions for second-order optimality (review)]. Minsk: [s.n.].
7. Mansimov, K.B. (1999) *Osobyie upravleniya v sistemakh s zapazdyvaniyem* [Special Controls in Systems with Delay]. Baku: ELM.
8. Mansimov, K.B. (2006) *Osobyie upravleniya v zadachakh s raspredelennymi parametrami* [Special controls in problems with distributed parameters]. *Sovremennaya matematika i ee prilozheniya*. 2. pp. 39–83.
9. Velichenko, V.V. (1967) *Optimal'noe upravlenie sostavnymi sistemami* [Optimal control of composite systems]. *Doklady Akademii nauk SSSR*. 176(4). pp. 154–156.
10. Zakharov, G.K. (1993) *Optimizatsiya stupenchatykh sistem s upravlyаемыми usloviyami perekhoda* [Optimization of step systems with controlled transition conditions]. *Avtomatika i telemekhanika*. 6. pp. 32–36.
11. Agafonova, I.A., Gulina, L.L. & Rasina, I.V. (1978) *Matematicheskoe modelirovanie i optimizatsiya protsessa metilirovaniya dinatrievoy soli sul'famoantiprina* [Mathematical modeling and optimization of the methylation process of sulfaminoantiprine disodium salt]. Irkutsk: [s.n.].
12. Ismailov, R.R. & Mansimov, K.B. (2006) *Ob usloviyakh optimal'nosti v odnoy stupenchatoy zadache upravleniya* [On optimality conditions in one stepwise control problem]. *Zhurnal vychislitel'noy matematiki i matematicheskoy fiziki*. 10. pp. 1758–1770.
13. Rozova, V.N. (2002) *Optimal'noe upravlenie stupenchatymi sistemami s neintegral'nym funktsionalom* [Optimal control of step systems with non-integral functionality]. *Vestnik RUDN. Ser. Prikladnaya i komp'yuternaya matematika*. 1(1). pp. 131–136.
14. Ashchepkov, L.T. (1987) *Optimal'noe upravlenie razryvnymi sistemami* [Optimal control of bursting systems]. Novosibirsk: Nauka.

15. Mammadova, T.F. & Mansimov, K.B. (2018) On optimality of singular controls in control problem of the step discrete two-parametric systems. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 42. pp. 12–22. DOI: 10.17223/19988605/42/2
16. Karimova, A.V. & Mansimov, K.B. (2023) An analogue of Pontryagin's discrete maximum principle in the stepwise control problem for systems described by Volterra type difference equations. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 65. pp. 4–14. DOI: 10.17223/19988605/65/1
17. Souyousefain, M. & Leela, S. (1990) Stability results for difference equations of Volterra type. *Appl. Math. Comput.* 36(1). pp. 51–61.

Информация об авторах:

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

Керимова Айтадж Вагиф кызы – аспирант кафедры «Математическая кибернетика» Бакинского государственного университета (Баку, Азербайджан). E-mail: aytac.mansimli@mail.ru

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

Information about the authors:

Mansimov Kamil B. (Professor, Doctor of Physical and Mathematical Sciences, Institute of Control Systems of the Ministry of Science and Education of Azerbaijan (Baku, Azerbaijan); Baku State University, Baku, Azerbaijan), E-mail: kamilbmansimov@gmail.com

Karimova Aitaj V. (Post-graduate Student, Baku State University, Baku, Azerbaijan). E-mail: aytac.mansimli@mail.ru

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

Поступила в редакцию 28.12.2023; принята к публикации 03.09.2024

Received 28.12.2023; accepted for publication 03.09.2024

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

MATHEMATICAL MODELING

Научная статья

УДК 519.872

doi: 10.17223/19988605/68/3

Исследование RQ-системы M/M/1 с катастрофами

Наталья Павловна Мелосникова¹, Анатолий Андреевич Назаров²,
Екатерина Александровна Фёдорова³^{1, 2, 3} *Национальный исследовательский Томский государственный университет, Томск, Россия*¹ *meloshnikovana@gmail.com*² *nazarov.tsu@gmail.com*³ *ekat_fedorova@mail.ru*

Аннотация. Исследуется однолинейная RQ-система с катастрофами. Катастрофы происходят в моменты наступления событий простейшего потока отрицательных заявок, которые в момент прихода уничтожают все «положительные» заявки, находящиеся в системе, – как на приборе, так и на орбите. Исследуется случайный процесс числа заявок на орбите. Методом производящих функций получена формула для производящей функции стационарного распределения вероятностей числа заявок на орбите. Описаны особенности численной реализации этой функции. Предложена оригинальная аппроксимация модифицированным отрицательным биномиальным распределением. Проведен численный анализ результатов.

Ключевые слова: RQ-системы; отрицательные заявки; орбита; катастрофы; отрицательное биномиальное распределение.

Для цитирования: Мелосникова Н.П., Назаров А.А., Фёдорова Е.А. Исследование RQ-системы M/M/1 с катастрофами // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 68. С. 28–37. doi: 10.17223/19988605/68/3

Original article

doi: 10.17223/19988605/68/3

Study of the M/M/1 retrial queueing system with disasters

Natalya P. Meloshnikova¹, Anatoly A. Nazarov², Ekaterina A. Fedorova³^{1, 2, 3} *National Research Tomsk State University, Tomsk, Russian Federation*¹ *meloshnikovana@gmail.com*² *nazarov.tsu@gmail.com*³ *ekat_fedorova@mail.ru*

Abstract. This paper studies a single-server retrial queueing system with disasters. Disasters occur in moments of negative call arrivals. Each negative call destroys all “positive” call located in the system: on the server and in the orbit. The random process of the number of calls in the orbit is studied. Using the method of generating functions, a formula is for the generating function of the stationary probability distribution of the number of calls in the orbit obtained. The features of this function are described. The approximation by the modified negative binomial distribution is proposed. A numerical analysis of the results was carried out.

Keywords: RQ-systems; negative calls; orbit; disasters; negative binomial distribution.

For citation: Meloshnikova, N.P., Nazarov, A.A., Fedorova, E.A. (2024) Study of the M/M/1 retrial queueing system with disasters. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika* – Tomsk State University Journal of Control and Computer Science. 68. pp. 28–37. doi: 10.17223/19988605/68/3

Введение

Моделирование процессов управления играет фундаментальную роль в обеспечении их работоспособности и оптимизации. Путем создания математических или компьютерных моделей реальных систем можно провести детальный анализ показателей эффективности, предсказать результаты или сценарии поведения систем и разработать стратегии их управления. Это позволяет сократить издержки, увеличить производительность и эффективность реальных систем, а также оптимизировать использование ресурсов.

RQ-системы (системы с повторными вызовами) – это новые математические модели теории массового обслуживания (ТМО), которые часто применяются для анализа, исследования и оптимизации разных технических систем: сотовой связи, сетей передачи информации, call-центров, центров облачных вычислений и др. [1–2]. Особенностью RQ-систем является наличие орбиты для хранения необслуженных заявок, которые могут повторно обратиться к обслуживающему устройству (или блоку приборов) через случайное время. Системы с отрицательными заявками представляют особый интерес, так как для инфокоммуникационных систем характерно наличие негативных воздействий – вирусов, хакерских атак, сбоев и др.

Понятие «отрицательных» заявок введено Е. Gelenbe в работах [3, 4]. В теории массового обслуживания модели с такими заявками упоминаются как G-системы или G-сети. В работе [5] представлен подробный обзор исследований систем массового обслуживания (СМО) с отрицательными заявками. Разнообразное действие отрицательных заявок и методы исследования классических СМО с отрицательными заявками рассмотрены в [6–7].

RQ-системы с отрицательными заявками исследуются в [8–16]. В работах [8–10] рассмотрены случаи, когда отрицательные заявки вызывают поломки приборов (СМО с ненадежным прибором). Одним из вариантов действия отрицательных заявок является опустошение системы; такой эффект также называют «катастрофами». RQ-системы с катастрофами изучаются в работах [11–14]. Стоит отметить, что основным инструментом исследования в указанных работах является матричный алгоритм, не дающий формулы для вычисления распределения вероятностей, но позволяющий получать численные характеристики.

Ранее нами было продемонстрировано влияние различных видов негативных воздействий на характеристики системы [15–16]. Исследованы такие воздействия, как опустошение одного прибора в системе, очищение блока обслуживания и опустошение всей системы (приборов и орбиты). Представлены примеры расчетов разных характеристик системы на основе численных алгоритмов.

В настоящей работе выполнено аналитическое исследование RQ-моделей с катастрофами. Показано, что с вычислительной точки зрения аналитическое решение неудобно для численных расчетов, поэтому предложена оригинальная аппроксимация модифицированным отрицательным биномиальным распределением.

1. Математическая модель и постановка задачи

В работе рассматривается однолинейная RQ-система с катастрофами (рис. 1). На вход системы поступает простейший с параметром λ поток заявок, которые будем называть «положительными». Положительные заявки поступают на обслуживающий прибор. Время обслуживания каждой заявки распределено по экспоненциальному закону с параметром μ . Когда прибор оказывается занятым, входящая заявка идет на орбиту, где осуществляет случайную задержку, продолжительность которой имеет экспоненциальное распределение с параметром σ . После случайной задержки заявка с орбиты

вновь обращается к прибору, и если он свободен, то она занимает его для обслуживания, иначе возвращается на орбиту.

Также на вход системы поступает простейший поток «отрицательных» заявок с параметром γ . Отрицательная заявка не нуждается в обслуживании. При поступлении в систему такая заявка «обнуляет» всю систему, т.е. все находящиеся положительные заявки покидают прибор и орбиту, и система становится пустой. Такую модель называют системой с катастрофами.

Ставится задача нахождения стационарного распределения вероятностей числа заявок на орбите.

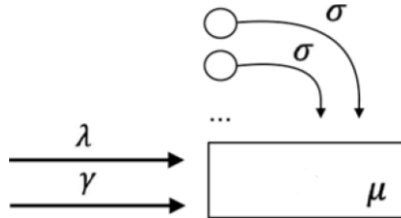


Рис. 1. RQ-система M/M/1 с катастрофами

Fig. 1. RQ-system M/M/1 with disasters

Пусть $i(t)$ – число заявок на орбите, а процесс $k(t)$ определяет состояние прибора следующим образом:

$$k(t) = \begin{cases} 0, & \text{прибор свободен,} \\ 1, & \text{прибор занят.} \end{cases}$$

Очевидно, что процесс $\{k(t), i(t)\}$ изменения состояний данной системы во времени является марковским.

Для распределения вероятностей $P(k, i, t) = P\{k(t) = k, i(t) = i\}$ нетрудно составить систему уравнений Колмогорова, которая в стационарном режиме имеет вид:

$$\begin{cases} -(\lambda + \gamma)P(0, 0) + \mu P(1, 0) + \gamma = 0, \\ -(\lambda + \mu + \gamma)P(1, 0) + \lambda P(0, 0) + \sigma P(0, 1) = 0 \quad (\text{при } i = 0), \\ -(\lambda + i\sigma + \gamma)P(0, i) + \mu P(1, i) = 0, \\ -(\lambda + \mu + \gamma)P(1, i) + \lambda P(0, i) + (i + 1)\sigma P(0, i + 1) + \lambda P(1, i - 1) = 0 \quad (\text{при } i > 0). \end{cases} \quad (1)$$

Для нахождения вероятностей $P(k, i)$ к системе уравнений (1) необходимо также добавить условие нормировки

$$\sum_{i=0}^{\infty} (P(0, i) + P(1, i)) = 1.$$

2. Уравнения для частных производящих функций и их решение

Обозначим частичные производящие функции

$$G(k, z) = \sum_{i=0}^{\infty} P(k, i) z^i.$$

Из системы (1) запишем уравнения для частных производящих функций

$$\begin{cases} -(\lambda + \gamma)G(0, z) + \mu G(1, z) - z\sigma \frac{\partial G(0, z)}{\partial z} = -\gamma, \\ -(\lambda(1 - z) + \mu + \gamma)G(1, z) + \lambda G(0, z) + \sigma \frac{\partial G(0, z)}{\partial z} = 0. \end{cases} \quad (2)$$

Найдем решение $G(0, z)$, $G(1, z)$. Умножим второе уравнение системы (2) на z и сложим с первым уравнением этой системы:

$$[\lambda z - (\gamma + \lambda)]G(0, z) + [\lambda z^2 - z(\lambda + \gamma + \mu) + \mu]G(1, z) + \gamma = 0. \quad (3)$$

Выразим из (3) функцию $G(1, z)$:

$$G(1, z) = \frac{(\lambda + \gamma - \lambda z)G(0, z) - \gamma}{\lambda z^2 - z(\lambda + \gamma + \mu) + \mu}. \quad (4)$$

Подставим (4) в первое уравнение системы (2) и, перегруппировав, получим

$$\frac{\partial G(0, z)}{\partial z} = \frac{1}{\sigma} \frac{(\lambda + \gamma)(\lambda + \gamma + \mu) - \lambda\mu - \lambda z(\lambda + \gamma)}{\lambda z^2 - z(\lambda + \gamma + \mu) + \mu} G(0, z) + \frac{\gamma}{\sigma} \frac{\lambda z - (\lambda + \gamma + \mu)}{\lambda z^2 - z(\lambda + \gamma + \mu) + \mu}. \quad (5)$$

Уравнение (5) сводится к виду линейного неоднородного обыкновенного дифференциального уравнения первого порядка:

$$\frac{\partial G(0, z)}{\partial z} = a(z)G(0, z) + b(z), \quad (6)$$

где коэффициенты $a(z)$ и $b(z)$ имеют вид:

$$a(z) = \frac{1}{\sigma} \frac{(\lambda + \gamma)(\lambda + \gamma + \mu) - \lambda\mu - \lambda z(\lambda + \gamma)}{\lambda z^2 - z(\lambda + \gamma + \mu) + \mu}, \quad (7)$$

$$b(z) = \frac{\gamma}{\sigma} \frac{\lambda z - (\lambda + \gamma + \mu)}{\lambda z^2 - z(\lambda + \gamma + \mu) + \mu}. \quad (8)$$

Решение дифференциального уравнения (6) имеет вид:

$$G(0, z) = e^{-\int_z^1 a(x)dx} \left\{ C - \int_z^1 e^{\int_z^y a(x)dx} b(y)dy \right\}, \quad (9)$$

где $C = const$.

Преобразуем (9). Найдем интеграл от $a(z)$:

$$\int_z^1 a(x)dx = \int_z^1 \frac{1}{\sigma} \frac{(\lambda + \gamma)(\lambda + \gamma + \mu) - \lambda\mu - \lambda x(\lambda + \gamma)}{\lambda x^2 - x(\lambda + \gamma + \mu) + \mu} dx. \quad (10)$$

Найдем корни x_1 и x_2 знаменателя, т.е. квадратного уравнения $\lambda x^2 - x(\lambda + \gamma + \mu) + \mu = 0$:

$$x_1 = \frac{\lambda + \gamma + \mu}{2\lambda} \left(1 - \sqrt{1 - \frac{4\mu\lambda}{(\lambda + \gamma + \mu)^2}} \right), \quad x_2 = \frac{\lambda + \gamma + \mu}{2\lambda} \left(1 + \sqrt{1 - \frac{4\mu\lambda}{(\lambda + \gamma + \mu)^2}} \right).$$

Нетрудно заметить, что для x_1 и x_2 выполняются неравенства $0 < x_1 < 1 < x_2$.

Тогда интеграл (10) можно записать как

$$\int_z^1 \frac{1}{\sigma} \frac{(\lambda + \gamma)(\lambda + \gamma + \mu) - \lambda\mu - \lambda x(\lambda + \gamma)}{\lambda x^2 - x(\lambda + \gamma + \mu) + \mu} dx = \ln \left| \frac{1 - x_1}{z - x_1} \right|^{\frac{A}{\lambda\sigma}} \cdot \left| \frac{1 - x_2}{z - x_2} \right|^{\frac{B}{\lambda\sigma}},$$

где

$$A = \frac{(\lambda + \gamma)(\lambda x_1 - (\lambda + \gamma + \mu)) + \lambda\mu}{x_2 - x_1}, \quad B = -\lambda(\lambda + \gamma) - \frac{(\lambda + \gamma)(\lambda x_1 - (\lambda + \gamma + \mu)) + \lambda\mu}{x_2 - x_1}.$$

С помощью несложных математических преобразований можно получить следующие выражения для постоянных A и B :

$$A = -\frac{(\lambda + \gamma)}{x_2 - x_1} \left[\lambda(1 - x_1) + \frac{\gamma(\gamma + \mu) + \lambda\gamma}{(\lambda + \gamma)} \right], \quad B = -\frac{\mu(\lambda(x_2 - 1) + \gamma)}{x_2(x_2 - x_1)}. \quad (11)$$

Таким образом, A и B принимают отрицательные значения.

Подставив полученное значение интеграла (10) в (9), имеем

$$G(0, z) = \left| \frac{1 - x_1}{z - x_1} \right|^{\frac{A}{\lambda\sigma}} \left| \frac{1 - x_2}{z - x_2} \right|^{\frac{B}{\lambda\sigma}} \left\{ C - \int_z^1 \left| \frac{1 - x_1}{y - x_1} \right|^{\frac{A}{\lambda\sigma}} \cdot \left| \frac{1 - x_2}{y - x_2} \right|^{\frac{B}{\lambda\sigma}} \cdot b(y) dy \right\},$$

или

$$G(0, z) = \frac{\left\{ C - \int_z^1 \left| \frac{1-x_1}{y-x_1} \right|^{\frac{A}{\lambda\sigma}} \left| \frac{1-x_2}{y-x_2} \right|^{\frac{B}{\lambda\sigma}} b(y) dy \right\}}{\left| \frac{1-x_1}{z-x_1} \right|^{\frac{A}{\lambda\sigma}} \left| \frac{1-x_2}{z-x_2} \right|^{\frac{B}{\lambda\sigma}}}. \quad (12)$$

Обратим внимание, что знаменатель производящей функции (12) в точке $z = x_1$ обращается в ноль. Это значение будем называть точкой, подозрительной на разрыв. А так как производящая функция ограничена, то числитель в этой точке также должен обращаться в ноль, поэтому выполняется равенство

$$C = \int_{x_1}^1 \left| \frac{1-x_1}{y-x_1} \right|^{\frac{A}{\lambda\sigma}} \left| \frac{1-x_2}{y-x_2} \right|^{\frac{B}{\lambda\sigma}} b(y) dy.$$

Тогда $G(0, z)$ можно записать в виде:

$$G(0, z) = \frac{\int_{x_1}^z \left| \frac{1-x_1}{y-x_1} \right|^{\frac{A}{\lambda\sigma}} \left| \frac{1-x_2}{y-x_2} \right|^{\frac{B}{\lambda\sigma}} b(y) dy}{\left| \frac{1-x_1}{z-x_1} \right|^{\frac{A}{\lambda\sigma}} \left| \frac{1-x_2}{z-x_2} \right|^{\frac{B}{\lambda\sigma}}}.$$

Аналитическое исследование функции в пределе $z \rightarrow x_1$ с применением правила Лопиталя показывает, что выполняются равенства

$$\lim_{z \rightarrow x_1 + 0} G(0, z) = \lim_{z \rightarrow x_1 - 0} G(0, z) = \frac{\gamma(\lambda x_1 - (\lambda + \gamma + \mu))}{A}.$$

Окончательно получаем, что производящая функция $G(z)$ распределения вероятностей числа заявок на орбите в RQ-системе M/M/1 с катастрофами имеет вид:

$$G(z) = G(0, z) + G(1, z) = \frac{G(0, z)((z-x_1)(z-x_2) + (\lambda + \gamma - \lambda z)) - \gamma}{(z-x_1)(z-x_2)}. \quad (13)$$

При численных расчетах функции (13) возникают проблемы в определении ее значений в окрестности точки x_1 в силу необходимости деления на малые числа. Поэтому в работе предлагается оригинальная аппроксимация этой функции в виде модифицированного отрицательного биномиального распределения (ОБР).

3. Аппроксимация

В работе предлагается оригинальная аппроксимация производящей функции (13) в виде производящей функции трехпараметрического распределения, имеющей вид:

$$G_{\text{app}}(x) = P(0) + (1 - P(0)) \frac{G_p(x) - G_p(0)}{1 - G_p(0)}, \quad (14)$$

где $G_p(x) = \left(\frac{p}{1 - (1-p)x} \right)^r$ – производящая функция двухпараметрического отрицательного биномиального распределения с параметрами r и p .

Выбор данного распределения в качестве аппроксимирующего обусловлен тем, что для классической однолинейной RQ-системы (без отрицательных заявок) было доказано, что распределение числа заявок в системе имеет форму отрицательного биномиального распределения [17]. Однако в силу

наступления «катастроф» вероятность нулевого состояния в представленной модели много выше, чем в классических RQ-системах, поэтому была предложена модификация (14).

3.1. Методы построения аппроксимации

Для построения аппроксимирующего распределения необходимо задать его параметры $P(0)$, p и r . Значение $P(0)$ определим из равенства: $P(0) = G(0)$, которое нетрудно получить из формулы (13). Параметры p и r ОБР можно определить двумя способами:

- 1) приравниванием моментов первого и второго порядка;
- 2) по двум точкам.

Первый способ. Математическое ожидание и дисперсия ОБР вычисляются по формулам

$$M = \frac{r(1-p)}{p}, \quad D = \frac{r(1-p)}{p^2}.$$

Точные моменты исследуемого процесса определяются путем дифференцирования функции (13):

$$G'(z)|_{z=1} = M, \quad G''(z)|_{z=1} = D + M^2 - M.$$

Второй способ. Сложно указать точные значения двух точек, по которым будут определяться значения параметров r и p , однако имеет смысл выбирать точки слева и справа от возникающей «особой» точки $z = \chi_1$. В дальнейших примерах мы выбирали точки $\chi_1 / 2$ и $(\chi_1 + 1) / 2$.

3.2. Результаты аппроксимации

Для анализа области применимости полученных аппроксимаций проведем численное сравнение предлагаемого аналога отрицательно биномиального распределения $P_{app}(i)$ и распределения $P(i) = P(0, i) + P(1, i)$, найденного с помощью численного алгоритма решения системы (1).

В качестве меры расхождения между аппроксимирующим распределением и распределением, полученным в результате применения численного алгоритма, будем использовать расстояние Колмогорова

$$\Delta = \max_n \left| \sum_{i=0}^n (P_{app}(i) - P(i)) \right|.$$

Построим $Ga(z)$ – производящую функцию аппроксимирующего распределения, и $G(z)$ – производящую функцию распределения, полученного выше (рис. 2).

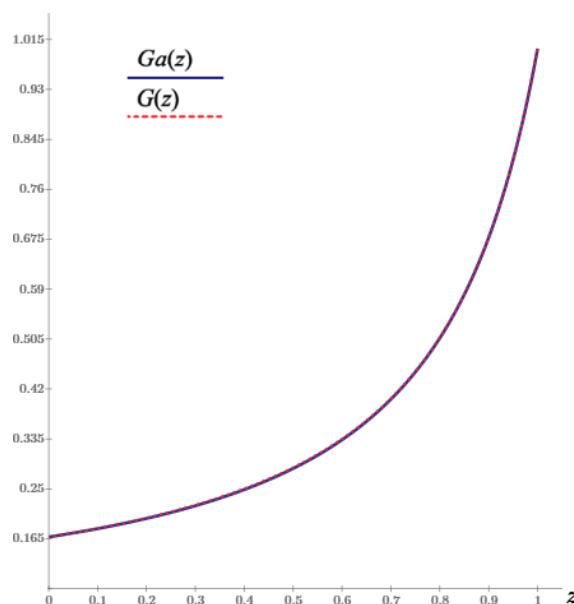


Рис. 2. Сравнение аппроксимации и аналитической производящей функции
Fig. 2. Comparison of approximation and analytical generating functions

Из представленного графика можно сделать вывод, что функция $Ga(z)$ достаточно точно аппроксимирует функцию $G(z)$.

Приведем несколько примеров построения аппроксимации для различных параметров системы (рис. 3, 4).

Зададим параметры системы следующими:

$$\lambda = 0,8; \mu = 1; \gamma = 0,01; \sigma = 1.$$

Применим первый способ построения аппроксимации. Сравнение распределений вероятностей $P_{app}(i)$ и $P(i)$ представлено на рис. 3.

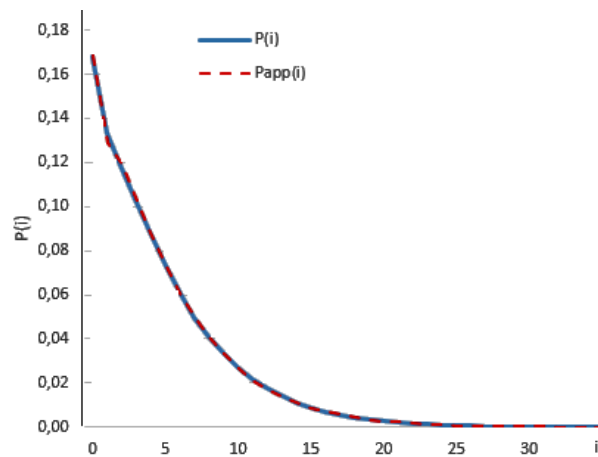


Рис. 3. Сравнение распределений вероятностей $P_{app}(i)$ и $P(i)$
Fig. 3. Comparison of probability distributions $P_{app}(i)$ and $P(i)$

Рассмотрим пример при $\lambda > \mu$. Пусть $\lambda = 2; \mu = 1; \gamma = 0,1; \sigma = 1$. Применим второй способ построения аппроксимации. Сравнение распределений вероятностей $P_{app}(i)$ и $P(i)$ для заданных параметров представлено на рис. 4.

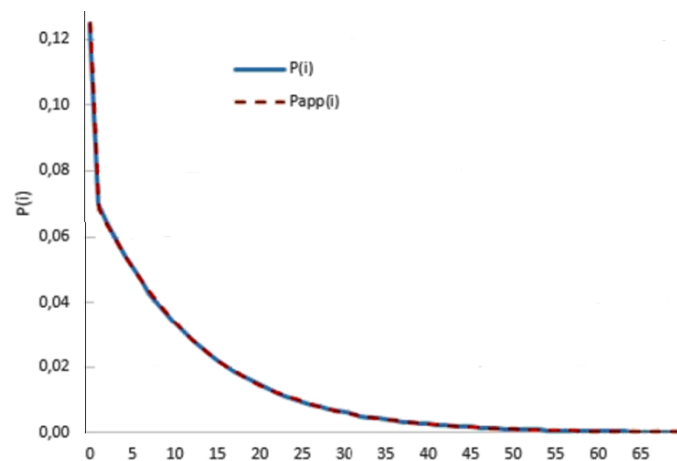


Рис. 4. Сравнение распределений вероятностей $P_{app}(i)$ и $P(i)$
Fig. 4. Comparison of probability distributions $P_{app}(i)$ and $P(i)$

Из рис. 3, 4 видно, что аппроксимирующее распределение достаточно близко к точному распределению. В качестве меры близости выберем расстояние Колмогорова, значения которого для описанных примеров представлены в таблице.

Расстояние Колмогорова между распределениями вероятностей $P_{app}(i)$ и $P(i)$

Пример	Расстояние Колмогорова
Пример 1 (рис. 3)	$\Delta = 0,004$
Пример 2 (рис. 4)	$\Delta = 0,002$

Рассмотрим еще один пример в случае достаточно малой интенсивности наступления катастроф. Пусть $\lambda = 0,7$; $\mu = 1$; $\gamma = 0,01$; $\sigma = 0,1$. Применим второй способ построения аппроксимации. Сравнение распределений вероятностей $P_{\text{app}}(i)$ и $P(i)$ для заданных параметров представлено на рис. 5.

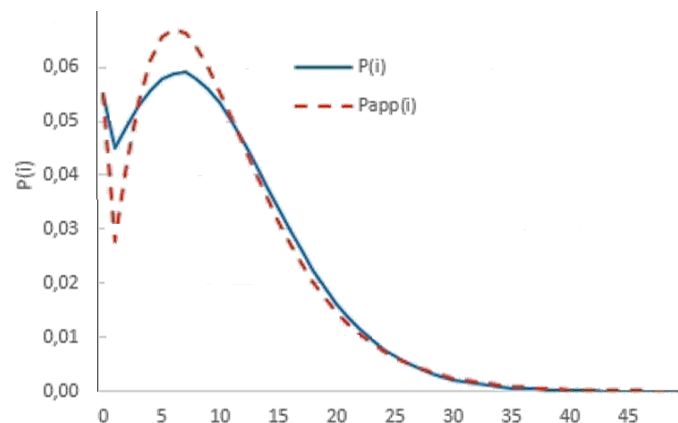


Рис. 5. Сравнение распределений вероятностей $P_{\text{app}}(i)$ и $P(i)$
Fig. 5. Comparison of probability distributions $P_{\text{app}}(i)$ and $P(i)$

Описанный выше пример показывает, что погрешность предлагаемой аппроксимации вырастает с уменьшением параметра γ , что имеет обоснование: система становится близка к модели без катастроф, поэтому требует применения других методов исследования. Однако даже в представленном случае погрешность аппроксимации достаточно мала: $\Delta = 0,022$.

Из представленных примеров можно сделать вывод, что предлагаемое аппроксимирующее распределение является достаточно точным. При этом оно требует меньших вычислительных затрат, чем численный алгоритм, и позволяет осуществлять расчеты без ограничений на значения параметров системы. Таким образом, предложенная аппроксимация может быть использована для определения характеристик реальных систем.

Заключение

В работе рассмотрена однолинейная RQ-система с катастрофами. Проведено исследование системы методом производящих функций, получен вид производящей функции (13) распределения вероятностей числа заявок на орбите. Однако было показано, что данная формула не может быть использована для нахождения числовых характеристик исследуемой системы. В связи этим в работе предлагается также аппроксимация искомого распределения вероятностей модифицированным отрицательным биномиальным распределением. Описан алгоритм нахождения ее параметров двумя способами. На основе проведенного численного сравнения аппроксимирующего распределения и точного, полученного численным алгоритмом, можно сделать вывод о высокой точности предлагаемой аппроксимации.

Список источников

1. Falin G.I., Templeton J.G.C. Retrial queues. London : Chapman and Hall, 1997. 167 p.
2. Artalejo J.R., Gomez-Corral A. Retrial Queueing Systems. Berlin : Springer, 2008. 267 p.
3. Gelenbe E. Product-form queueing networks with negative and positive customers // Journal of Applied Probability. 1991. V. 28 (3). P. 656–663.
4. Gelenbe E., Glynn P., Singman K. Queues with negative arrivals // Journal of Applied Probability. 1991. V. 28. P. 245–250.
5. Do T.V. Bibliography on G-networks, negative customers and applications // Mathematical and Computer Modelling. 2011. V. 53 (1-2). P. 205–212.
6. Pechinkin A.V. Markov queueing system with finite buffer and negative customers affecting the queue end // Autom. Remote Control. 2007. V. 68. P. 1104–1117.
7. Anisimov V.V., Artalejo J.R. Analysis of Markov Multiserver Retrial Queues with Negative Arrivals // Queueing Systems. 2001. V. 39. P. 157–182.

8. Wang J., Zhang P. A discrete-time retrial queue with negative customers and unreliable server // *Computers & Industrial Engineering*. 2009. V. 56 (4). P. 1216–1222.
9. Kirupa K., Udaya C.K. Batch arrival retrial queue with negative customers, multioptional service and feedback // *Communications on Applied Electronics*. 2015. V. 2 (4). P. 14–18.
10. Dimitriou I. A mixed priority retrial queue with negative arrivals, unreliable server and multiple vacations // *Applied Mathematical Modelling*. 2013. V. 37 (3). P. 1295–1309.
11. Shin Y.W. Multi-server retrial queue with negative customers and disasters // *Queueing Systems*. 2007. V. 55 (4). P. 223–337.
12. Shin Y.W. BMAP/G/1 queue with correlated arrivals of customers and disasters // *Oper Res Lett*. 2004. V. 32 (4). P. 364–373.
13. Rani S., Jain M., Dhibar S. Analysis of Markovian Retrial Queue with Double Orbits // *Vacation, Orbital Search, and Disaster Using ANFIS Approach*. 2023. V. 47. P. 1751–1764.
14. Ammar S.I., Rajadurai P. Performance Analysis of Preemptive Priority Retrial Queueing System with Disaster under Working Breakdown Services // *Symmetry*. 2019. V. 11 (3). Art. 419. doi: 10.3390/sym11030419
15. Meloshnikova N.P., Fedorova E.A., Plaksin D.A. Asymptotic Analysis of a Multiserver Retrial Queue with Disasters in the Service Block // *Queueing Theory and Applications. ITMM 2022. Communications in Computer and Information Science*. 2023. P. 55–67.
16. Мелосникова Н.П., Федорова Е.А. Численный анализ влияния отрицательных заявок разного типа действия на многолинейные RQ-системы // *Системы управления, информационные технологии и математическое моделирование : материалы V Всерос. науч.-практ. конф. с междунар. участием. Омск : Омск. гос. техн. ун-т, 2023. С. 245–251.*
17. Фёдорова Е.А. Исследование случайного процесса числа заявок в RQ-системе M/M/1 // *Материалы молодежной научной школы по прикладной теории вероятностей и телекоммуникационным технологиям (АРТСТ–2017). М. : Рос. ун-т дружбы народов, 2017. С. 304–307.*

References

1. Falin, G.I. & Templeton, J.G.C. (1997) *Retrial queues*. London: Chapman and Hall.
2. Artalejo, J.R. & Gomez-Corral, A. (2008) *Retrial Queueing Systems*. Berlin: Springer.
3. Gelenbe, E. (1991) Product-form queueing networks with negative and positive customers. *Journal of Applied Probability*. 28(3). pp. 656–663.
4. Gelenbe, E., Glynn, P. & Singman, K. (1991) Queues with negative arrivals. *Journal of Applied Probability*. 28. pp. 245–250.
5. Do, T.V. (2011) Bibliography on G-networks, negative customers and applications. *Mathematical and Computer Modelling*. 53(1–2). pp. 205–212.
6. Pechinkin, A.V. (2007) Markov queueing system with finite buffer and negative customers affecting the queue end. *Automation and Remote Control*. 68. pp. 1104–1117.
7. Anisimov, V.V. & Artalejo, J.R. (2001) Analysis of Markov Multiserver Retrial Queues with Negative Arrivals. *Queueing Systems*. 39. pp. 157–182.
8. Wang, J. & Zhang, P. (2009) A discrete-time retrial queue with negative customers and unreliable server. *Computers & Industrial Engineering*. 56(4). pp. 1216–1222.
9. Kirupa, K. & Udaya, C.K. (2015) Batch arrival retrial queue with negative customers, multioptional service and feedback. *Communications on Applied Electronics*. 2(4). pp. 14–18.
10. Dimitriou, I. (2013) A mixed priority retrial queue with negative arrivals, unreliable server and multiple vacations. *Applied Mathematical Modelling*. 37(3). pp. 1295–1309.
11. Shin, Y.W. (2007) Multi-server retrial queue with negative customers and disasters. *Queueing Systems*. 55(4). pp. 223–337.
12. Shin, Y.W. (2004) BMAP/G/1 queue with correlated arrivals of customers and disasters. *Operations Research Letters*. 32(4). pp. 364–373.
13. Rani, S., Jain, M. & Dhibar, S. (2023) Analysis of Markovian Retrial Queue with Double Orbits. *Vacation, Orbital Search, and Disaster Using ANFIS Approach*. 47. pp. 1751–1764.
14. Ammar, S.I. & Rajadurai, P. (2019) Performance Analysis of Preemptive Priority Retrial Queueing System with Disaster under Working Breakdown Services. *Symmetry*. 11(3). Art. 419. DOI: 10.3390/sym11030419
15. Meloshnikova, N.P., Fedorova, E.A. & Plaksin D.A. (2023) Asymptotic Analysis of a Multiserver Retrial Queue with Disasters in the Service Block. *Communications in Computer and Information Science. Queueing Theory and Applications. Proc. ITMM 2022*. pp. 55–67.
16. Meloshnikova, N.P. & Fedorova, E.A. (2023) Chislennyy analiz vliyaniya otritsatel'nykh zayavok raznogo tipa deystviya na mno-golineynye RQ-sistemy [Numerical analysis of impact of negative calls of several types on multiserver retrial queues]. *Sistemy upravleniya, informatsionnye tekhnologii i matematicheskoe modelirovanie* [Control Systems, Information Technologies and Mathematical Modeling]. Proc. of the Fifth Conference. Omsk: OSTU. pp. 245–251.
17. Fedorova, E.A. (2017) Issledovanie sluchaynogo protsesssa chisla zayavok v RQ-sisteme M/M/1 [Study of the random process of the number of calls in retrial queueing system M/M/1]. *APTCT–2017. Proceedings of the youth scientific school on applied probability theory and telecommunication technologies*. Moscow: Russian University of Friendship of Peoples. pp. 304–307.

Информация об авторах:

Мелошникова Наталья Павловна – аспирант кафедры теории вероятностей и математической статистики Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: meloshnikovana@gmail.com

Назаров Анатолий Андреевич – профессор, доктор технических наук, профессор кафедры теории вероятностей и математической статистики Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: nazarov.tsu@gmail.com

Фёдорова Екатерина Александровна – кандидат физико-математических наук, доцент кафедры теории вероятностей и математической статистики Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: ekat_fedorova@mail.ru

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

Information about the authors:

Meloshnikova Natalya P. (Post-graduate Student, National Research Tomsk State University, Tomsk, Russia). E-mail: meloshnikovana@gmail.com

Nazarov Anatoly A. (Professor, Doctor of Technical Sciences, Professor, National Research Tomsk State University, Tomsk, Russia). E-mail: nazarov.tsu@gmail.com

Fedorova Ekaterina A. (Candidate of Physical and Mathematical Sciences, Associate Professor, National Research Tomsk State University (Tomsk, Russia). E-mail: ekat_fedorova@mail.ru

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

Поступила в редакцию 27.05.2024; принята к публикации 03.09.2024

Received 27.05.2024; accepted for publication 03.09.2024

Original article

UDC 519.872.5

doi: 10.17223/19988605/68/4

Asymptotic analysis of a closed G-network with rewards**Tatiana V. Rusilko¹, Dmitry A. Salnikov²**^{1,2} Yanka Kupala State University of Grodno, Grodno, Republic of Belarus¹ tatiana.rusilko@gmail.com² dima.saln.gr@gmail.com

Abstract. A closed exponential G-network with positive and negative customers is studied. In addition, the G-network under study generates a sequence of rewards or earnings associated with network transitions from one state to another. The total network reward is a random process governed by the probabilistic relations of the Markov process that determines the number of customers in the network nodes. The purpose of this paper is to asymptotically study the G-network with rewards under the assumption of a large number of customers. The main objective is to calculate the expected total reward of the G-network in the asymptotic case. It is proved that reward density function satisfies the generalized Kolmogorov backward equation. An ordinary differential equation for the expected reward that the G-network will earn in a time t if it starts in a given initial state, is derived.

Keywords: queueing network; G-network; network with rewards; asymptotic analysis method.

Acknowledgments: The research was supported by the State Program of Scientific Research of the Republic of Belarus “Convergence-2025” (sub-program “Mathematical models and methods”, assignment 1.6.01).

For citation: Rusilko, T.V., Salnikov, D.A. (2024) Asymptotic analysis of a closed G-network with rewards. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 68. pp. 38–47. doi: 10.17223/19988605/68/4

Научная статья

doi: 10.17223/19988605/68/4

Асимптотический анализ замкнутой G-сети с вознаграждениями**Татьяна Владимировна Русилко¹, Дмитрий Александрович Сальников²**^{1,2} Гродненский государственный университет им. Янки Купалы, Гродно, Беларусь¹ tatiana.rusilko@gmail.com² dima.saln.gr@gmail.com

Аннотация. Исследуется замкнутая экспоненциальная G-сеть массового обслуживания с положительными и отрицательными заявками. Кроме того, исследуемая G-сеть порождает последовательность вознаграждений или доходов, связанных с переходами сети из одного состояния в другое. Полное вознаграждение такой сети является случайным процессом, управляемым вероятностными связями марковского процесса, определяющего число заявок в узлах сети. Цель данной работы – асимптотический анализ G-сети с вознаграждениями в случае большого числа обслуживаемых заявок и решение задачи прогнозирования полного ожидаемого вознаграждения G-сети. Доказано, что плотность вознаграждения удовлетворяет дифференциальному уравнению в частных производных второго порядка, являющемуся обобщением обратного уравнения Колмогорова. Получено обыкновенное дифференциальное уравнение для ожидаемого вознаграждения G-сети, которое сеть получит за определенное время, исходя из данного начального состояния.

Ключевые слова: сеть массового обслуживания; G-сеть; сеть с вознаграждениями; метод асимптотического анализа.

Благодарности: Работа выполнена в рамках государственной программы научных исследований Республики Беларусь «Конвергенция-2025» (подпрограмма «Математические модели и методы», задание 1.6.01).

Для цитирования: Русилко Т.В., Сальников Д.А. Асимптотический анализ замкнутой G-сети с вознаграждениями // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 68. С. 38–47. doi: 10.17223/19988605/68/4

Introduction

A Markov process is a mathematical model that is useful in the study of complex systems. The basic concepts of a Markov process are «system state» and «state transition». R. Howard proposed that a continuous-time Markov process receives a reward of r_{ij} dollars when the system makes a transition from state i to state j , $i \neq j$; the system earns a reward at the rate of r_i dollars per unit time during all the time that it occupies state i [1. P. 99]. Note that r_{ij} and r_i have different dimensions. According to Howard's concept, it is not necessary that the system earn according to both reward rates and transition rewards [1]. He called r_{ij} the «reward» or «earnings» associated with the process transition from i to j . The set of rewards for the process may be described by a reward matrix $\mathbf{R}$ with elements r_{ij} . The rewards need not be in dollars, they could be energy levels, units of production, or any other physical quantity relevant to the problem. In Howard's interpretation, the Markov process generates a sequence of rewards as it makes transitions from state to state and is called the «Markov process with rewards» [1]. The reward is thus a random process with a probability distribution governed by the probabilistic relations of the Markov process. In [1], the problem of finding the expected total earnings $V_i(t)$ that the system will earn in a time t if it starts in the state i was solved. Note the term «expected» and notation $V_i(t)$ are used in the sense of «prospective» or relating to the future time interval t if the initial state of the process is i .

The state of a closed exponential queueing network is the vector $\mathbf{k}(t)$ with elements representing the number of customers at each node. The process $\mathbf{k}(t)$ is a continuous-time Markov process on the finite state space. The concept of Markov processes with rewards was first used by M. Matalytski as fundamental for defining exponential queueing networks that generate earnings [2, 3]. Queueing networks with earnings were called HM-networks (Howard – Matalytski).

G-networks are generalized queueing networks of queueing nodes with several types of customers: positive customers, negative customers and in some cases triggers. Negative customers and triggers are not serviced, so they are identified as signals. In this article, only negative customers are considered as signals. When a negative customer arrives at a node, one or a group of positive customers is removed or «killed» in a non-empty queue. G-networks were first introduced and studied by E. Gelenbe [4–6]. Their field of application is modelling computing systems and networks, evaluating their performance, modelling biophysical neural networks, pattern recognition tasks and others [7–11].

The purpose of this paper is the asymptotic analysis of a closed exponential G-network with rewards, which implies an approximation method of queueing network study under the critical assumption of a large but limited number of customers [10–11]. We are interested in the expected earnings of the G-network if it operates for a time t with a given initial condition.

1. Formulation of the problem

A closed exponential G-network of a finite number of nodes $S_0, S_1, \dots, S_n$ is the focus of this paper. A fixed number K of customers circulate between the nodes. Customers are homogeneous in terms of their service in the network nodes. The node S_0 is an IS-node (Infinite Server) of K identical exponential servers capable of servicing all customers in the network in parallel. The IS-node plays the role of an «external environment» or a finite source of K customers. The node S_0 generates a Poisson flow of customers with rate $\lambda_0 k_0$, where λ_0 is the flow parameter, k_0 is the number of customers in the node S_0 . This arrival flow is divided into a flow of positive and negative customers. And the probability of a positive customer arriving at short

time interval Δt is $\lambda_0 k_0 p_{0i}^+ \Delta t + o(\Delta t)$, while the probability of a negative customer arriving is $\lambda_0 k_0 p_{0i}^- \Delta t + o(\Delta t)$, $i = \overline{1, n}$, $\sum_{i=1}^n (p_{0i}^+ + p_{0i}^-) = 1$.

Each network node S_i , $i = \overline{1, n}$, is a queueing system with m_i servers and unlimited waiting area for positive customers. The probability of completing the positive customer service at a node S_i during a short time interval Δt is $\mu_i \min(m_i, k_i) \Delta t + o(\Delta t)$, where k_i is the number of customers in the node S_i ; the probability of completing the service two or more customers is $o(\Delta t)$. The completion of the customer service at different network nodes in a short time interval Δt are mutually independent events. Customers are served according to the FIFO rule. When a customer has completed service in the node S_i it is instantly transferred to the node S_j as positive with probability p_{ij}^+ or negative with probability p_{ij}^- , otherwise it is transferred to the IS-node S_0 with probability $p_{i0}^+ = 1 - \sum_{j=1}^n (p_{ij}^+ + p_{ij}^-)$, $i \neq j$, $i, j = \overline{1, n}$. Negative customers arriving at a node are not served by the node servers; therefore, they are considered as signals. A negative customer arriving at a node S_i , $i = \overline{1, n}$, removes one positive customer located at the same node and both of them are immediately transferred to the IS-node S_0 as positive customers.

The state of the G-network under study can be described by a n -dimensional continuous-time Markov process on finite state space $\mathbf{k}(t) = (k_1(t), k_2(t), \dots, k_n(t))$, where $k_i(t)$ represents the number of customers in the node S_i at the moment t , $0 \leq k_i(t) \leq K$, $i = \overline{1, n}$, $t \in [0, +\infty)$. The number of customers in the IS-node S_0 is $k_0(t) = K - \sum_{i=1}^n k_i(t)$. Let us assume that the G-network is in the state $\mathbf{k}$ if at some time t components $k_i(t) = k_i$, $i = \overline{1, n}$, form a vector $\mathbf{k} = (k_1, k_2, \dots, k_n)$.

Let us suppose that the G-network described above earns R_{ij}^+ conventional units (c.u.) when a positive customer makes a transition from the node S_i to the node S_j and it earns R_{ij}^- c.u. when a negative customer makes the same transition, $i \neq j$, $i, j = \overline{0, n}$. We call R_{ij}^+ and R_{ij}^- the «reward» associated with the transition of a positive and a negative customer, respectively, from S_i to S_j ; we call matrices $\mathbf{R}^+ = (r_{ij}^+)_{i,j=\overline{0,n}}$ and $\mathbf{R}^- = (r_{ij}^-)_{i,j=\overline{0,n}}$ the reward matrices. Suppose further that the G-network receives a reward at the rate of $R(\mathbf{k})$ c.u. per unit time during all the time that it occupies the state $\mathbf{k}$. The question of interest is: what will be the expected total earnings of the G-network in a time t , $t \in T$, if the network is now in the state $\mathbf{k}$.

The main objective of the paper is to predict the expected total reward of the G-network. It is obvious that the G-network reward is governed by the Markov process $\mathbf{k}(t)$ and depends on both the initial network state $\mathbf{k}$ and the remaining observation time t , $t \in T$.

2. Asymptotic analysis of a G-network with rewards

General purpose of asymptotic methods in queueing theory is to study servicing processes of queueing systems and networks by finding suitable approximations for them under some critical (limit, asymptotic) assumption. The scientific researches of Tomsk State University are widely known in the field of asymptotic methods [12, 13].

In this paper, a closed exponential G-network with rewards is studied under the asymptotic assumption of a large number of customers K . The passage to the limit from a Markov chain $\mathbf{k}(t)$ to a continuous-state Markov process $\xi(t)$ is used. In contrast to discontinuous processes, continuous processes in any short time interval $\Delta t \rightarrow 0$ have some small change in the state $\Delta x \rightarrow 0$. The mathematical approach used in this paper

is based on a discrete model of a continuous Markov process described in many books on the theory of diffusion Markov processes [14, 15].

Let $V(\mathbf{k}, t)$ be the expected total G-network reward that the network will earn in a time t if it starts in the state $\mathbf{k}$. Notation $v(\mathbf{x}, t)$ is used for reward density in the case of asymptotic approximation, $\mathbf{x}$ is the start state, t is remaining time. The concept of reward density is given in the proof of the following theorem.

Theorem 1. In the asymptotic case of a large number of customers K the reward density function $v(\mathbf{x}, t)$ provided that it is differentiable with respect to t and twice continuously differentiable with respect to x_i , $i = \overline{1, n}$, satisfies up to $o(\varepsilon^2) = o(1/K^2)$, where $\varepsilon = 1/K$, the generalized multidimensional Kolmogorov backward equation:

$$\frac{\partial v(\mathbf{x}, t)}{\partial t} = - \sum_{i=1}^n A_i(\mathbf{x}, t) \frac{\partial v(\mathbf{x}, t)}{\partial x_i} + \frac{\varepsilon}{2} \sum_{i,j=1}^n B_{ij}(\mathbf{x}, t) \frac{\partial^2 v(\mathbf{x}, t)}{\partial x_i \partial x_j} + q(\mathbf{x}), \quad (1)$$

$$\text{where } A_i(\mathbf{x}, t) = \sum_{j=1}^n \mu_j \min(l_j, x_j) (p_{ji}^- - p_{ji}^+ + \delta_{ji}) + \\ + \mu_i \min(l_i, x_i) \sum_{j=1}^n p_{ij}^- (1 - \theta(x_j)) - \lambda \left(1 - \sum_{i=1}^n x_i \right) (p_{0i}^+ - p_{0i}^-) \text{ are drift coefficients,}$$

$$B_{ii}(\mathbf{x}, t) = \sum_{j=1}^n \mu_j \min(l_j, x_j) (p_{ji}^+ + p_{ji}^- + \delta_{ji}) + \\ + \mu_i \min(l_i, x_i) \sum_{j=1}^n p_{ij}^- (1 - \theta(x_j)) + \lambda \left(1 - \sum_{i=1}^n x_i \right) (p_{0i}^+ - p_{0i}^-) \text{ and}$$

$$B_{ij}(\mathbf{x}, t) = \mu_i \min(l_i, x_i) (p_{ij}^- - p_{ij}^+), \quad i \neq j, \text{ are diffusion coefficients;}$$

$$q(\mathbf{x}) = r(\mathbf{x}) + K \left(\sum_{i,j=1}^n \mu_i \min(l_i, x_i) (p_{ij}^+ r_{ij}^+ + p_{ij}^- r_{ij}^- (2 - \theta(x_j))) + \sum_{i=1}^n \mu_i \min(l_i, x_i) p_{i0}^+ r_{i0}^+ + \right. \\ \left. + \sum_{i=1}^n \lambda \left(1 - \sum_{i=1}^n x_i \right) (p_{0i}^+ r_{0i}^+ + p_{0i}^- r_{0i}^-) \right) \text{ is earning rate;}$$

$$\delta_{ji} \text{ is the Kronecker delta.}$$

Proof. Let us introduce the following notations: $\mathbf{I}_i$ is a n -vector of the form $\mathbf{I}_i = \left(\overbrace{0, 0, \dots, 0}^i, 1, 0, 0, \dots, 0 \right)$,

$\theta(x) = \begin{cases} 1, & x > 0, \\ 0, & x \leq 0; \end{cases}$ is the Heaviside step function. Let us assume that at the initial moment of time, Markov

process $\mathbf{k}(t)$, which determines the G-network state, is in the state $\mathbf{k}$. Taking into account the above assumptions, we consider all possible transitions of $\mathbf{k}(t)$ from the given initial state $\mathbf{k}$ at $t = 0$, and the corresponding network earnings in a short time interval Δt . In this way we can relate the expected total reward in a time $t + \Delta t$, $V(\mathbf{k}, t + \Delta t)$, to $V(\mathbf{k}, t)$ by an equation.

– A transition from the state $\mathbf{k}$ to the state $\mathbf{k} + \mathbf{I}_j - \mathbf{I}_i$ with probability

$$\mu_i \min(m_i, k_i) p_{ij}^+ \Delta t + o(\Delta t),$$

which means the customer was served at the node S_i and joined the node S_j as positive, $i, j = \overline{1, n}$. In this case, the network would receive the reward R_{ij}^+ plus the expected total reward $V(\mathbf{k} + \mathbf{I}_j - \mathbf{I}_i, t)$ to be made if it starts in the state $\mathbf{k} + \mathbf{I}_j - \mathbf{I}_i$ with time t remaining, $V(\mathbf{k} + \mathbf{I}_j - \mathbf{I}_i, t)$.

– A transition from the state $\mathbf{k}$ to the state $\mathbf{k} - \mathbf{I}_j - \mathbf{I}_i$ with probability

$$\mu_i \min(m_i, k_i) p_{ij}^- \Delta t + o(\Delta t),$$

which means the customer was served at S_i and transmitted to S_j as negative, $i, j = \overline{1, n}$. The network reward is R_{ij}^- c.u. plus the reward $V(\mathbf{k} - \mathbf{I}_j - \mathbf{I}_i, t)$ that the network would receive for the remaining time t if the initial state was $\mathbf{k} - \mathbf{I}_j - \mathbf{I}_i$.

– The process transitions from the state $\mathbf{k}$ to the state $\mathbf{k} - \mathbf{I}_i$ in three cases. First, with the probability of

$$\mu_i \min(m_i, k_i) p_{i0}^+ \Delta t + o(\Delta t),$$

which means the customer was served at S_i and transmitted to the IS-node S_0 , $i = \overline{1, n}$. The network earning from this transition is R_{i0}^+ c.u. Secondly, with the probability of

$$\lambda \left(K - \sum_{i=1}^n k_i \right) p_{0i}^- \Delta t + o(\Delta t),$$

which is possible when a negative customer arrives to S_i from the IS-node S_0 , $i = \overline{1, n}$. The network earning in this case is R_{0i}^- c.u. Thirdly, with the probability of

$$\mu_i \min(m_i, k_i) p_{ij}^- (1 - \theta(k_j)) \Delta t + o(\Delta t),$$

when a positive customer is transmitted as a negative from S_i to an empty node S_j , $i, j = \overline{1, n}$. The network earning is R_{ij}^- c.u. In each of the three listed cases, the expected reward $V(\mathbf{k} - \mathbf{I}_i, t)$ is added to the mentioned network earnings.

– A transition from the state $\mathbf{k}$ to the state $\mathbf{k} + \mathbf{I}_i$ with probability

$$\lambda \left(K - \sum_{i=1}^n k_i \right) p_{0i}^+ \Delta t + o(\Delta t),$$

which is possible when a positive customer arrives to S_i from the IS-node S_0 , $i = \overline{1, n}$. The network reward is R_{0i}^+ c.u. plus the expected total reward $V(\mathbf{k} + \mathbf{I}_i, t)$ that the network would earn in the remaining time t if it started in the state $\mathbf{k} + \mathbf{I}_i$.

– During time interval Δt the G-network remains in the state $\mathbf{k}$ with probability

$$1 - \left[\sum_{i,j=1}^n \mu_i \min(m_i, k_i) (1 + p_{ij}^- (1 - \theta(k_j))) + \sum_{i=1}^n \lambda \left(K - \sum_{i=1}^n k_i \right) \right] \Delta t + o(\Delta t),$$

which entails a reward $R(\mathbf{k})\Delta t$ plus the expected reward that the network will earn in the remaining t units of time, $V(\mathbf{k}, t)$. The probabilities of other transitions and rewards are considered to be of order $o(\Delta t)$.

Having regard to the listed above, let us sum up the product of probabilities and rewards over all options. Dividing both sides of the resulting equation by Δt , we take the limit as $\Delta t \rightarrow 0$. As a result, we have the following set of difference-differential equations that completely define $V(\mathbf{k}, t)$:

$$\begin{aligned} \frac{\partial V(\mathbf{k}, t)}{\partial t} = & \sum_{i,j=1}^n \mu_i \min(m_i, k_i) p_{ij}^+ (V(\mathbf{k} + \mathbf{I}_j - \mathbf{I}_i, t) - V(\mathbf{k}, t)) + \\ & + \sum_{i,j=1}^n \mu_i \min(m_i, k_i) p_{ij}^- (V(\mathbf{k} - \mathbf{I}_j - \mathbf{I}_i, t) - V(\mathbf{k}, t)) + \\ & + \sum_{i=1}^n \mu_i \min(m_i, k_i) p_{i0} (V(\mathbf{k} - \mathbf{I}_i, t) - V(\mathbf{k}, t)) + \\ & + \sum_{i=1}^n \lambda \left(K - \sum_{i=1}^n k_i \right) p_{0i}^- (V(\mathbf{k} - \mathbf{I}_i, t) - V(\mathbf{k}, t)) + \\ & + \sum_{i,j=1}^n \mu_i \min(m_i, k_i) p_{ij}^- (1 - \theta(k_j)) (V(\mathbf{k} - \mathbf{I}_i, t) - V(\mathbf{k}, t)) + \\ & + \sum_{i=1}^n \lambda \left(K - \sum_{i=1}^n k_i \right) p_{0i}^+ (V(\mathbf{k} + \mathbf{I}_i, t) - V(\mathbf{k}, t)) + Q(\mathbf{k}), \text{ where} \end{aligned} \quad (3)$$

$$Q(\mathbf{k}) = R(\mathbf{k}) + \sum_{i,j=1}^n \left(\mu_i \min(m_i, k_i) \left(p_{ij}^+ R_{ij}^+ + p_{ij}^- R_{ij}^- (2 - \theta(k_j)) \right) \right) + \\ + \sum_{i=1}^n \mu_i \min(m_i, k_i) p_{i0}^+ R_{i0}^+ + \sum_{i=1}^n \lambda \left(K - \sum_{i=1}^n k_i \right) \left(p_{0i}^- R_{0i}^- + p_{0i}^+ R_{0i}^+ \right).$$

According to R. Howard, let us define a quantity $Q(\mathbf{k})$ as the «earning rate» of the network [1]. This earning rate is composed of combination of reward rates and transition rewards.

The set (3) cannot be solved for large K and n . In this regard, we will use the approximation and study the limiting behavior of the random process $\mathbf{k}(t)$ in the asymptotic case of a large number of customers K , $K \gg 1$. We proceed to the limit from the Markov chain $\mathbf{k}(t)$ to the continuous Markov process

$$\xi(t) = \frac{\mathbf{k}(t)}{K} = \left(\frac{k_1(t)}{K}, \frac{k_2(t)}{K}, \dots, \frac{k_n(t)}{K} \right),$$

as K tends to be large. The phase space of the vector $\xi(t)$ is

$$X = \left\{ \mathbf{x} = (x_1, x_2, \dots, x_n) : x_i \geq 0, i = \overline{1, n}, \sum_{i=1}^n x_i \leq 1 \right\}.$$

The increment of $\xi_i(t)$ in the short time $\Delta t \rightarrow 0$ is $\Delta x_i = \varepsilon = 1/K$, and $\Delta x_i \rightarrow 0$ as $K \rightarrow \infty$. Therefore, the process $\xi_i(t)$ tends to be continuous as $K \rightarrow \infty$ ($\varepsilon \rightarrow 0$), and the vector $\xi(t)$ is a continuous-state Markov process on X . In the considered asymptotic case, the total reward of the G-network is a continuously changing process depending on the initial state $\mathbf{x}$, $\mathbf{x} \in X$, and the upcoming time t .

In physics, the mass density (volumetric mass density or specific mass) is a substance's mass per unit of volume. Mathematically, density is defined as mass divided by volume:

$$\rho = \frac{\text{mass}}{\text{volume}} = \lim_{\varepsilon \rightarrow 0} \frac{m(x_1 \leq \xi_1 < x_1 + \varepsilon, x_2 \leq \xi_2 < x_2 + \varepsilon, \dots, x_n \leq \xi_n < x_n + \varepsilon)}{\varepsilon^n}.$$

In probability theory, probability density is the probability per unit volume. By analogy, let us introduce into consideration the concept of «reward density», meaning the reward that the network earns per unit of state space in time t based on the initial state at the point $\mathbf{x}$:

$$v(\mathbf{x}, t) = \frac{\text{reward}}{\text{state space}} = \lim_{\varepsilon \rightarrow 0} \frac{V(x_1 \leq \xi_1 < x_1 + \varepsilon, x_2 \leq \xi_2 < x_2 + \varepsilon, \dots, x_n \leq \xi_n < x_n + \varepsilon, t)}{\varepsilon^n}.$$

Realizing the passage to the limit as $K \rightarrow \infty$, a n -dimensional lattice with vertices at discrete points $\frac{\mathbf{k}}{K} = \left(\frac{k_1}{K}, \frac{k_2}{K}, \dots, \frac{k_n}{K} \right)$ transforms into set of points $\mathbf{x} = (x_1, x_2, \dots, x_n) \in X$, the «point density» increases as K increases. It is required to take into account the continuous change in reward on X . Let us use the approximation $V(\mathbf{x}K, t) = v(\mathbf{x}, t)\varepsilon^n$ for reward when $x_i \leq \xi_i < x_i + \varepsilon$, $i = \overline{1, n}$. Similarly, based on reward rates and transition rewards, $R(\mathbf{x})$ and R_{ij} , we introduce parameters of earning rate per unit of state space X , $r(\mathbf{x})$ and r_{ij} , into consideration. Therefore, $r(\mathbf{x})\varepsilon^n$, $r_{ij}\varepsilon^n$ are earning parameters in case of small change in $\frac{\mathbf{k}}{K}$.

Let $\mathbf{e}_i = \mathbf{I}_i \varepsilon$.

Thus, in the asymptotic case under study, equation (3) can be represented as the following partial differential equation:

$$\begin{aligned} \frac{\partial v(\mathbf{x}, t)}{\partial t} = & K \sum_{i,j=1}^n \mu_i \min(\varepsilon m_i, x_i) p_{ij}^+ (v(\mathbf{x} + \mathbf{e}_j - \mathbf{e}_i, t) - v(\mathbf{x}, t)) + \\ & + K \sum_{i,j=1}^n \mu_i \min(\varepsilon m_i, x_i) p_{ij}^- (v(\mathbf{x} - \mathbf{e}_j - \mathbf{e}_i, t) - v(\mathbf{x}, t)) + \\ & + K \sum_{i=1}^n \mu_i \min(\varepsilon m_i, x_i) p_{i0} (v(\mathbf{x} - \mathbf{e}_i, t) - v(\mathbf{x}, t)) + \end{aligned} \quad (4)$$

$$\begin{aligned}
 & +K \sum_{i=1}^n \lambda \left(1 - \sum_{i=1}^n x_i\right) p_{0i}^- (v(\mathbf{x} - \mathbf{e}_i, t) - v(\mathbf{x}, t)) + \\
 & +K \sum_{i,j=1}^n \mu_i \min(\varepsilon m_i, x_i) p_{ij}^- (1 - \theta(x_j)) (v(\mathbf{x} - \mathbf{e}_i, t) - v(\mathbf{x}, t)) + \\
 & +K \sum_{i=1}^n \lambda \left(1 - \sum_{i=1}^n x_i\right) p_{0i}^+ (v(\mathbf{x} + \mathbf{e}_i, t) - v(\mathbf{x}, t)) + q(\mathbf{x}),
 \end{aligned}$$

where $q(\mathbf{x})$ is earning rate on state space X and it is defined by (2).

If $v(\mathbf{x}, t)$ is a twice continuously differentiable function with respect to x_i , then we can use the Taylor series to second order about the point x for functions $v(\mathbf{x} + \mathbf{e}_j - \mathbf{e}_i, t)$, $v(\mathbf{x} - \mathbf{e}_i, t)$, $v(\mathbf{x} + \mathbf{e}_i, t)$, $i, j = \overline{1, n}$. Equation (4) becomes:

$$\begin{aligned}
 \frac{\partial v(\mathbf{x}, t)}{\partial t} = & \sum_{i,j=1}^n \mu_i \min(\varepsilon m_i, x_i) p_{ij}^+ \left(\left(\frac{\partial v(\mathbf{x}, t)}{\partial x_j} - \frac{\partial v(\mathbf{x}, t)}{\partial x_i} \right) + \frac{\varepsilon}{2} \left(\frac{\partial^2 v(\mathbf{x}, t)}{\partial x_j^2} - 2 \frac{\partial^2 v(\mathbf{x}, t)}{\partial x_j \partial x_i} + \frac{\partial^2 v(\mathbf{x}, t)}{\partial x_i^2} \right) \right) + \\
 & + \sum_{i,j=1}^n \mu_i \min(\varepsilon m_i, x_i) p_{ij}^- \left(- \left(\frac{\partial v(\mathbf{x}, t)}{\partial x_j} + \frac{\partial v(\mathbf{x}, t)}{\partial x_i} \right) + \frac{\varepsilon}{2} \left(\frac{\partial^2 v(\mathbf{x}, t)}{\partial x_j^2} + 2 \frac{\partial^2 v(\mathbf{x}, t)}{\partial x_j \partial x_i} + \frac{\partial^2 v(\mathbf{x}, t)}{\partial x_i^2} \right) \right) + \\
 & + \sum_{i=1}^n \mu_i \min(\varepsilon m_i, x_i) p_{i0}^+ \left(- \frac{\partial v(\mathbf{x}, t)}{\partial x_i} + \frac{\varepsilon}{2} \frac{\partial^2 v(\mathbf{x}, t)}{\partial x_i^2} \right) + \sum_{i=1}^n \lambda \left(1 - \sum_{i=1}^n x_i\right) p_{0i}^- \left(- \frac{\partial v(\mathbf{x}, t)}{\partial x_i} + \frac{\varepsilon}{2} \frac{\partial^2 v(\mathbf{x}, t)}{\partial x_i^2} \right) + \\
 & + \sum_{i,j=1}^n \mu_i \min(\varepsilon m_i, x_i) p_{ij}^- (1 - \theta(x_j)) \left(- \frac{\partial v(\mathbf{x}, t)}{\partial x_i} + \frac{\varepsilon}{2} \frac{\partial^2 v(\mathbf{x}, t)}{\partial x_i^2} \right) + \\
 & + \sum_{i=1}^n \lambda \left(1 - \sum_{i=1}^n x_i\right) p_{0i}^+ \left(\frac{\partial v(\mathbf{x}, t)}{\partial x_i} + \frac{\varepsilon}{2} \frac{\partial^2 v(\mathbf{x}, t)}{\partial x_i^2} \right) + q(\mathbf{x}) + o(\varepsilon^2).
 \end{aligned}$$

Having grouped first-order and second-order partial derivatives of a function $v(\mathbf{x}, t)$ in the resulting equation, we conclude that the compact mathematical expression (1) is valid up to $o(\varepsilon^2) = o(1/K^2)$. The equation (1) differs from the well-known multidimensional Kolmogorov backward equation only by the earning component $q(\mathbf{x})$. The theorem is proved.

3. Mathematical model for the mean expected reward of the G-network

It is obvious that the diffusion coefficients $B_{ij}(\mathbf{x}, t)$ of the equation (1) are of order ε , then the term

$\frac{\varepsilon}{2} \sum_{i,j=1}^n B_{ij}(\mathbf{x}, t) \frac{\partial^2 v(\mathbf{x}, t)}{\partial x_i \partial x_j}$ on the right side of (1) is $O(\varepsilon^2)$. In this regard, up to terms of order $O(\varepsilon^2)$, the re-

ward density is given by the following equation:

$$\frac{\partial v(\mathbf{x}, t)}{\partial t} = - \sum_{i=1}^n A_i(\mathbf{x}, t) \frac{\partial v(\mathbf{x}, t)}{\partial x_i} + q(\mathbf{x}). \quad (5)$$

Integrating the density $v(\mathbf{x}, t)$ within a n -dimensional region D , $D \subseteq X$, we get the expected total reward that the network will earn in time t if it starts in state $\mathbf{x} \in D$:

$$V_D(t) = \int \int \dots \int_D v(\mathbf{x}, t) d\mathbf{x}.$$

Applying this transformation to both parts of (5), using the rules of integration, Leibniz integral rule, the linearity of coefficients $A_i(\mathbf{x}, t)$ in $\mathbf{x}$ and the boundary condition $A(\mathbf{x}, t)v(\mathbf{x}, t) = 0$, $\mathbf{x} \in \Gamma(D)$, where $\Gamma(D)$ is the reflecting boundary of the region D [15], we obtain a first order ordinary linear differential equation for the expected reward $V_D(t)$:

$$\frac{d}{dt}V_D(t) = \sum_{i=1}^n \frac{\partial A_i(\mathbf{x}, t)}{\partial x_i} \cdot V_D(t) + \iint \dots \int_D q(\mathbf{x}) d\mathbf{x}. \quad (6)$$

The differential equation (6) is a mathematical model of the expected total reward of the G-network that the network will earn in time t if its initial state belongs to set D . Now we have a linear differential equation that completely define $V_D(t)$ when $V_D(0)$ is known.

By analogy with a center of gravity of a material body D in physics, we can find the equilibrium point of the expected reward, when $\mathbf{x} \in D$:

$$E_i^D(t) = \frac{1}{V_D(t)} \iint \dots \int_D x_i v(\mathbf{x}, t) d\mathbf{x}, \quad i = \overline{1, n}.$$

Example. Consider a closed exponential G-network, including four nodes S_i , $i = \overline{1, 4}$, and IS-node S_0 . Customers are transmitted between network nodes with the following non-zero probabilities: $p_{01}^+ = 0,98$, $p_{01}^- = 0,02$, $p_{12}^+ = 0,6$, $p_{12}^- = 0,02$, $p_{10}^+ = 0,38$, $p_{23}^+ = 0,75$, $p_{23}^- = 0,1$, $p_{20}^+ = 0,15$, $p_{34}^+ = 0,9$, $p_{34}^- = 0,1$, $p_{41}^+ = 0,1$, $p_{40}^+ = 0,9$. The structure of this network is presented graphically in Fig. 1. The node S_0 generates the Poisson flow of customers routed to the node S_1 , the flow parameter is $\lambda = 0,1$ customers per unit of time. The number of customers circulating in the network is $K = 10\,000$.

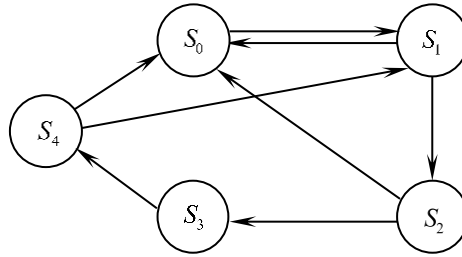


Fig. 1. G-network structure

We assume that each node S_i has four identical servers $m_i = 4$, $i = \overline{1, 4}$, the service rates are $\mu_1 = 0,017$, $\mu_2 = 0,001$, $\mu_3 = 0,005$, $\mu_4 = 0,002$ respectively. The network reward is determined by the following non-zero parameters: $r_{01}^+ = 4$, $r_{01}^- = -0,5$, $r_{10}^+ = 1$, $r_{12}^+ = 5$, $r_{12}^- = -0,5$, $r_{20}^+ = 2$, $r_{23}^+ = 3$, $r_{23}^- = -0,5$, $r_{34}^+ = 6$, $r_{34}^- = -0,5$, $r_{40}^+ = 4$, $r_{41}^+ = 0,2$, $r_{41}^- = -0,2$. A reward rate is a linear function $r(\mathbf{x}) = 250 + 2x_1 + 3x_2 - 7x_3 - 5x_4$; the given coefficients determine the contribution of each queue to the reward rate. We are interested in the expected G-network reward if the network starts in the state $\mathbf{x} \in D$ with time t remaining, where $D = \{\mathbf{x} = (x_1, x_2, x_3, x_4): 50\epsilon < x_i \leq 450\epsilon, i = \overline{1, 4}\}$. The numerical solution of (6) when $V_D(0) = 0$ is presented graphically in Fig. 2.

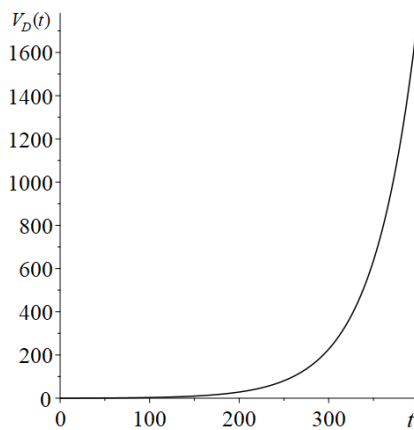


Fig. 2. Earnings of G-network in monetary units with remaining time

The resulting model allows us to predict the expected G-network reward $V_D(t)$ in monetary units as a function of remaining time t , when the initial state of the network is known, $\mathbf{x} \in D$. There is a tendency that the expected network reward exponentially increases or decreases, which depends on the ratio between earnings and losses from transmissions of a large number of customers between network nodes. In addition, calculations allowed us to conclude that the expected network reward is sensitive to the total number of network customers K , remaining time t , the volume of region D and nodal service parameters.

Conclusions

In this paper, a closed exponential G-network with rewards is studied. The earning sequence created by the network occurs through customer transmissions between nodes. The purpose of the study is to develop the asymptotic method for analyzing G-network reward in the case of a large number of customers. The paper presents the results of mathematical modeling that were used to predict the expected total reward of the G-network as a function of remaining time t , when the initial state of the network is known. The numerical example is presented.

Queueing networks are widely used in modeling complex systems and processes [16]. G-networks with rewards allow, along with the analysis of the probabilistic-time characteristics of modeled objects, to calculate their earnings associated with transmissions positive and negative customers. Real systems typically handle a large number of customers. In this regard, the results obtained in the asymptotic case of a large number of customers will be relevant with high accuracy.

It is planned to continue the study of queueing networks with rewards and to develop a mathematical model for the earned network reward by the time t , if its state at the moment t is known. The research carried out allows us to conclude that the generalized Fokker – Planck – Kolmogorov forward equation will serve as a mathematical model for the density of such reward.

References

1. Howard, R.A. (1960) *Dynamic Programming and Markov Processes*. Cambridge: Massachusetts Institute of Technology.
2. Matalytsky, M.A. & Pankov, A.V. (2004) Probabilistic analysis of income in banking networks. *Vestnik BSU. Series 1. Physics, Mathematics, Computer Science*. 9. pp. 79–92.
3. Matalytsky, M. & Pankov, A. (2003) Analysis of the stochastic model of changing of incomes in the open banking network. *Computer Science*. 3(5). pp. 19–29.
4. Gelenbe, E. (1991) Product form queueing networks with negative and positive customers. *Journal of Applied Probability*. 28. pp. 656–663. DOI: 10.2307/3214499
5. Gelenbe, E. (1993) G-networks with triggered customer movement. *Journal of Applied Probability*. 30. pp. 742–748. DOI:10.2307/3214781
6. Gelenbe, E. (1993) G-networks with signals and batch removal. *Probability in the Engineering and Informational Sciences*. 7(3). pp. 335–342. DOI: 10.1017/S0269964800002953
7. Caglayan, M.U. (2017) G-networks and their applications to machine learning, energy packet networks and routing: introduction to the special issue. *Probability in the Engineering and Informational Sciences*. 31(4). pp. 381–395. DOI: 10.1017/S0269964817000171
8. Zhang, Y. (2020) G-Networks and the performance of ICT with renewable energy. *SN Computer Science*. 1. Art. 56. DOI: 10.1007/s42979-019-0056-2
9. Zhang, Y. (2021) Optimal energy distribution with energy packet networks. *Probability in the Engineering and Informational Sciences*. 35(1). pp. 75–91. DOI: 10.1017/S0269964818000566
10. Rusilko, T. (2022) Asymptotic analysis of a closed G-network of unreliable nodes. *Journal of Applied Mathematics and Computational Mechanics*. 21(2). pp. 91–102. DOI: 10.17512/jamcm.2022.2.08
11. Rusilko, T.V. (2023) The G-network as a stochastic data network model. *Journal of the Belarusian State University. Mathematics and Informatics*. 2. pp. 45–54. DOI: 10.33581/2520-6508-2023-2-45-54
12. Nazarov, A.A. & Pavlova, E.A. (2022) Study of SMO type MMPP|M|N with feedback by the method of asymptotic analysis. *Vestnik Tomskogo gosudarstvennogo universiteta Upravlenie vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 58. pp. 47–57. DOI: 10.17223/19988605/58/5
13. Moiseeva, S.P., Bushkova, T.V. & Pankratova, E.V. (2022) Asymptotic analysis of resource heterogeneous QS under equivalently increasing service time. *Automation and Remote Control*. 83(8). pp. 1213–1227. DOI: 10.1134/S0005117922080057
14. Tikhonov, V.A. & Mironov, M.A. (1977) *Markovskie protsessy* [Markov processes]. Moscow: Sovetskoe radio.

15. Gardiner, K.V. (1986) *Stokhasticheskie metody v estestvennykh naukakh* [Stochastic Methods in Natural Sciences]. Translated from English. Moscow: Mir.
16. Rusilko, T.V. & Pankov, A.V. (2024) Queueing network model of a call center with customer retrials and impatient customers. *Izvestiya of Saratov University. Mathematics. Mechanics. Informatics*. 24(2). pp. 287–297. DOI: 10.18500/1816-9791-2024-24-2-287–297.

Information about the authors:

Rusilko Tatiana V. (Candidate of Physical and Mathematical Sciences, Associate Professor, Associate Professor of Fundamental and Applied Mathematics Department, Yanka Kupala State University of Grodno, Grodno, Republic of Belarus). E-mail: tatiana.rusilko@gmail.com

Salnikov Dmitry A. (Master's student of the Department of Mathematical and Information Support of Economic Systems, Yanka Kupala State University of Grodno, Grodno, Republic of Belarus). E-mail: dima.saln.gr@gmail.com

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

Информация об авторах:

Русилко Татьяна Владимировна – доцент, кандидат физико-математических наук, доцент кафедры фундаментальной и прикладной математики Гродненского государственного университета им. Янки Купалы (Гродно, Беларусь). E-mail: tatiana.rusilko@gmail.com

Сальников Дмитрий Александрович – магистрант кафедры математического и информационного обеспечения экономических систем Гродненского государственного университета им. Янки Купалы (Гродно, Беларусь). E-mail: dima.saln.gr@gmail.com

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

Received 22.05.2024; accepted for publication 03.09.2024

Поступила в редакцию 22.05.2024; принята к публикации 03.09.2024

Научная статья

УДК 004.942

doi: 10.17223/19988605/68/5

**Задержки и надежность обслуживания запросов
в виртуальном компьютерном кластере****Фунг Ван Кю¹, Владимир Анатольевич Богатырев²**^{1,2} Университет ИТМО, Санкт-Петербург, Россия² Санкт-Петербургский государственный университет аэрокосмического приборостроения,
Санкт-Петербург, Россия¹ phungvanquy97@gmail.com² vladimir.bogatyrev@gmail.com

Аннотация. Проведено исследование возможностей повышения качества обслуживания запросов в двух-уровневом виртуальном кластере с целью обоснования выбора числа мест ожидания в очередях виртуальных машин второго уровня и числа разворачиваемых в них контейнеров. Особенностью предлагаемой модели виртуального кластера является учет возможного снижения интенсивности обслуживания в каналах (контейнерах) из-за разделения между ними общих ограниченных вычислительных ресурсов системы.

Ключевые слова: виртуальный кластер; виртуальная машина; контейнер; система массового обслуживания.

Для цитирования: Фунг В.К., Богатырев В.А. Задержки и надежность обслуживания запросов в виртуальном компьютерном кластере // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 68. С. 48–58. doi: 10.17223/19988605/68/5

Original article

doi: 10.17223/19988605/68/5

Delays and reliability of request servicing in a virtual computer cluster**Van Quy Phung¹, Vladimir A. Bogatyrev²**^{1,2} National Research University of Information Technologies, Mechanics and Optics (ITMO University),
Saint Petersburg, Russian Federation² Saint Petersburg State University of Aerospace Instrumentation, Saint Petersburg, Russian Federation¹ phungvanquy97@gmail.com² vladimir.bogatyrev@gmail.com

Abstract. This research investigates the possibilities of enhancing service quality in a two-tier virtual cluster with the aim of justifying the selection of the number of waiting slots in the queues of second-tier virtual machines and the number of containers deployed within them. The proposed virtual cluster model accounts for the potential decrease in service rate within channels (containers) due to the division of shared limited computational resources among them.

Keywords: virtual cluster; virtual machine; container; queuing system.

For citation: Phung, V.Q., Bogatyrev, V.A. (2024) Delays and reliability of request servicing in a virtual computer cluster. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 68. pp. 48–58. doi: 10.17223/19988605/68/5

Введение

В настоящее время одной из ключевых задач является повышение производительности и надежности систем распределенной обработки данных, а также сокращение задержек в компьютерных системах с кластерной архитектурой [1–5], включая задержки доступа к ним через сеть [6–8]. Решение этих задач связано с необходимостью развития методологии оценки и обеспечения надежности [9–10] и производительности инфокоммуникационных систем и сетей, включая оптимизацию совместного использования ограниченных ресурсов системы [11–13].

Актуальность развития методологии оценки и обоснования выбора проектных решений для обеспечения надежности и минимизации задержек обслуживания запросов в распределенных системах постоянно возрастает в связи с непрерывным появлением новых технологических решений. Это включает в себя интенсивно развивающиеся технологии виртуализации, широко применяемые при проектировании компьютерных систем с кластерной архитектурой.

Выделяется две основные технологии виртуализации: виртуальные машины (ВМ) и контейнеры [14]. Виртуальные машины [15] включают полный стек аппаратного обеспечения и операционной системы, что обеспечивает их более надежную изоляцию, но требует большего объема ресурсов. Контейнеры [16], используя общие ресурсы физического узла (хоста), разделяют общее ядро операционной системы, что делает их более удобными в управлении, однако менее изолированными по сравнению с ВМ.

Интерес к технологиям виртуализации стимулирует множество исследований, направленных на повышение производительности систем. Так, в работе [18] представлен инновационный подход к масштабированию на различных уровнях виртуализации. При этом в [18] показана эффективность виртуализации на популярных облачных платформах. Для полного использования ресурсов при соблюдении требуемого уровня обслуживания в [19] предложен подход, автоматически генерирующий правила, позволяющие определить максимально допустимое количество объектов в облачной системе. Применение методов машинного обучения при формировании стратегий масштабирования, например ElasticDocker, приводит к повышению качества обслуживания для конечных пользователей при оптимизации затрат как для клиентов контейнеров, так и для провайдеров облачных сервисов.

Известные модели компьютерных кластеров, представленные системами массового обслуживания [20–22], не учитывают влияния разделения ограниченных ресурсов компьютерной системы между виртуальными машинами на снижение интенсивности обработки запросов.

Развитие методологии моделирования распределенных систем, особенно с учетом предоставления сервиса облачных вычислений на основе виртуализации, требует исследования влияния мониторинга и распределения общих вычислительных ресурсов на производительность и задержки распределенных вычислений [23–29].

Для виртуальной машины с развертыванием в ней контейнеров исследование влияния их числа на снижение интенсивности обслуживания в каждом из контейнеров из-за разделения между ними общих вычислительных ресурсов экспериментально установлено в статье [17]. При исследовании задержек обслуживания запросов виртуальная машина представляется в виде многоканальной системы массового обслуживания (СМО) с бесконечной очередью. При этом каждый канал обслуживания представляет отдельный контейнер. Особенностью предлагаемой в [17] аналитической модели СМО является учет установленного экспериментально влияния на интенсивность обслуживания в каждом контейнере их общего числа и числа контейнеров, выполняющих обслуживание поступавших в СМО запросов.

Для исследования возможностей виртуализации облачных и кластерных компьютерных систем в предлагаемой статье строится модель виртуального двухуровневого кластера. На первом уровне виртуального кластера с помощью виртуальной машины с бесконечной очередью проводится диспетчеризация запросов, распределяемых на обслуживание в виртуальные машины второго уровня. В каждой из виртуальных машин второго уровня развернуто несколько контейнеров и организована очередь с ограниченным числом мест ожидания.

Цель статьи – исследование возможностей повышения качества обслуживания запросов в двух-уровневом виртуальном кластере при обосновании выбора числа мест ожидания в очередях виртуальных машин второго уровня и числа разворачиваемых в них контейнеров с учетом возможного снижения производительности контейнеров из-за разделения между ними общих ограниченных вычислительных ресурсов системы.

1. Организация виртуального компьютерного кластера

Рассмотрим виртуальный компьютерный кластер, организация которого представлена на рис. 1. Термин «виртуальный» подразумевает, что в отличие от традиционных кластеров, состоящих из физических серверов, данный кластер формируется путем объединения нескольких изолированных виртуальных машин с выделением для каждой из них части общих вычислительных ресурсов физического сервера. Одна из сформированных виртуальных машин выступает в роли мастера (ведущего), обеспечивая равномерное распределение нагрузки на все ведомые (slave) виртуальные машины.

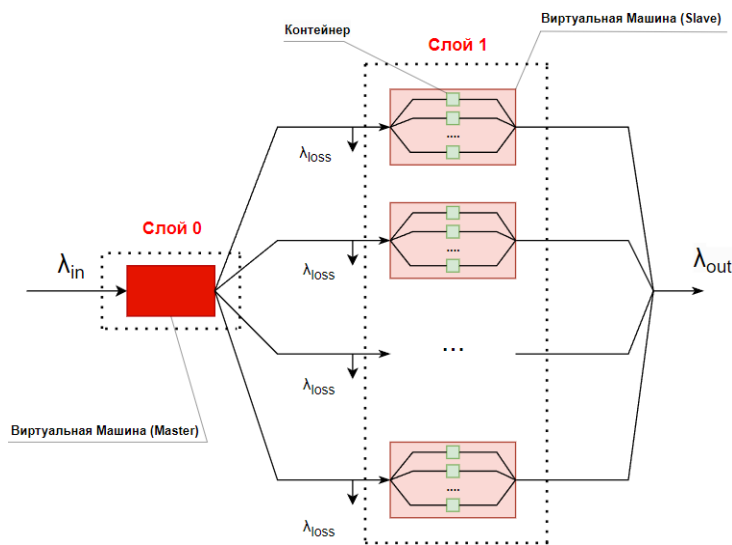


Рис. 1. Организация виртуального компьютерного кластера
Fig. 1. Organization of a virtual computer cluster

В рассматриваемом виртуальном кластере можно выделить два уровня (слоя). Первый слой функционирует как машина-мастер (диспетчер) и выполняет роль высокоскоростного балансировщика запросов для всего кластера. Второй слой состоит из множества виртуальных машин, на каждой из которых развернут один и тот же сервис в виде нескольких контейнеров.

Виртуальная машина-мастер функционирует как одноканальная система массового обслуживания (М/М/1) с неограниченной очередью [22].

Подчиненные (слейв)-виртуальные машины представляются как многоканальные системы массового обслуживания с ограниченной очередью. Особенностью предлагаемой модели многоканальной системы обслуживания является учет того, что интенсивность обслуживания запросов в каждом канале обслуживания (контейнере) меняется в зависимости от количества сформированных в системе контейнеров и количества активных контейнеров. Под активными подразумеваются контейнеры, задействованные в обслуживании поступающих в систему запросов. Такой подход отличается от традиционного тем, что позволяет учесть разделение ограниченных вычислительных ресурсов физического компьютера между активными контейнерами.

При построении модели кластерной системы предположим, что виртуальная машина, выполняющая роль мастера-балансировщика нагрузки, работает с высокой интенсивностью, в результате чего ее влияние на увеличение задержек в кластере будет незначительным. Параметры реализации второго уровня виртуального кластера зададим как $\{N, C, n\}$, при этом C – число мест ожидания в

очереди каждой из N виртуальных машин второго уровня, в каждой из которых развернуто n контейнеров.

Определим среднее время пребывания запроса в кластере T как сумму средних задержек на его первом и втором уровнях:

$$T = T_1 + T_2,$$

$$T_1 = \frac{1/\mu_1}{1 - (\lambda_{in}/\mu_1)},$$

где λ_{in} – интенсивность поступления запросов в кластер; μ_1 – интенсивность обслуживания балансировщиком.

Предположим, что во втором слое все N виртуальных машин идентичны и работают независимо друг от друга с выделением для каждой из них равной доли общих вычислительных ресурсов системы. В этом случае для оценки среднего времени пребывания запроса во втором слое достаточно рассмотреть работу одной виртуальной машины с интенсивностью поступления запросов $\lambda = \lambda_{in}/N$.

Интенсивность поступающего на обслуживание потока, так же как для классической многоканальной СМО с ограниченной очередью, задается как

$$\lambda_k = \begin{cases} \lambda, & \text{при } k < C + n, \\ 0, & \text{при } k \geq C + n. \end{cases}$$

При поступлении запросов на одну виртуальную машину второго слоя с интенсивностью λ , если количество запросов в системе k не превышает числа сформированных контейнеров n , количество активных контейнеров m будет равно числу запросов в системе ($m = k$). Интенсивность обслуживания в этом случае равна $k\mu(n, k)$, где $\mu(n, m)$ – интенсивность обслуживания в одном контейнере, зависящая от общего числа сформированных контейнеров и числа активных контейнеров в рассматриваемой ВМ. В случае если $k > n$, интенсивность обслуживания не зависит от числа запросов в очереди r (где $r = k - n$), равна $n\mu(n, n)$. При этом количество активных контейнеров m равно общему числу контейнеров в системе n ($m = n$). Таким образом, интенсивность обслуживания виртуальной машиной определяется как

$$\mu_k = \begin{cases} k\mu(n, k), & \text{при } k \leq n, \\ n\mu(n, n), & \text{при } n < k \leq n + C. \end{cases}$$

Предлагаемый подход к построению модели СМО виртуального кластера предполагает определение зависимости $\mu(n, m)$, которая при $0 \leq m \leq n$ установлена экспериментально [17].

После определения $\mu(n, m)$ вероятность незанятости системы вычисляется как

$$p_0 = \begin{cases} \left(1 + \sum_{i=1}^n \frac{\lambda^i}{i! \prod_{j=1}^i \mu(n, j)} + \frac{\lambda^n \cdot \rho}{n! \prod_{k=1}^n \mu(n, k)} \cdot \frac{1 - \rho^C}{1 - \rho} \right)^{-1}, & \text{при } \rho \neq 1, \\ \left(1 + \sum_{i=1}^n \frac{\lambda^i}{i! \prod_{j=1}^i \mu(n, j)} + \frac{\lambda^n \cdot C}{n! \prod_{k=1}^n \mu(n, k)} \right)^{-1}, & \text{при } \rho = 1, \end{cases}$$

где $\rho = \frac{\lambda}{n\mu(n, n)}$.

Вероятность нахождения в системе k запросов

$$p_k = \begin{cases} \frac{\lambda^k}{k! \prod_{i=1}^k \mu(n, i)} \cdot p_0, & \text{при } k \leq n, \\ \frac{\rho^{k-n} \cdot \lambda^n}{n! \prod_{i=1}^n \mu(n, i)} \cdot p_0, & \text{при } k > n. \end{cases}$$

Среднее количество запросов, находящихся на обслуживании, вычисляется как

$$L_{об} = \sum_{k=1}^n p_k \cdot k + \sum_{k=n+1}^{n+C} p_k \cdot n =$$

$$= \begin{cases} \sum_{k=1}^n \frac{\lambda^k p_0}{(k-1)! \prod_{i=1}^k \mu(n, i)} + \frac{\lambda^n p_0}{(n-1)! \prod_{i=1}^n \mu(n, i)} \cdot \frac{1-\rho^C}{1-\rho}, & \text{при } \rho \neq 1, \\ \sum_{k=1}^n \frac{\lambda^k p_0}{(k-1)! \prod_{i=1}^k \mu(n, i)} + \frac{\lambda^n \cdot p_0 \cdot C}{(n-1)! \prod_{i=1}^n \mu(n, i)}, & \text{при } \rho = 1. \end{cases}$$

Среднее количество запросов в очереди:

$$L_{оч} = \sum_{r=1}^C p_{n+r} \cdot r =$$

$$= \begin{cases} \frac{p_0 \lambda^{n+1}}{n \cdot \mu(n, n) \cdot n! \prod_{i=1}^n \mu(n, i)} \cdot \frac{1-\rho^C (C+1-C \cdot \rho)}{(1-\rho)^2}, & \text{при } \rho \neq 1, \\ \frac{p_0 \lambda^{n+1}}{n \cdot \mu(n, n) \cdot n! \prod_{i=1}^n \mu(n, i)} \cdot \frac{C}{2} (1+C), & \text{при } \rho = 1. \end{cases}$$

Вероятность отказа в обслуживании запроса определяется как вероятность того, что все C мест ожидания в очереди заняты:

$$P_{отк} = p_{n+C} = \frac{\rho^C \cdot \lambda^n}{n! \prod_{i=1}^n \mu(n, i)} \cdot p_0.$$

По формуле Литтла

$$T_2 = \frac{L_{об} + L_{оч}}{Q \cdot \lambda} = N \cdot \frac{L_{об} + L_{оч}}{Q \cdot \lambda_{in}}, \text{ где } Q = (1 - P_{отк}).$$

Следует заметить противоречие между рассматриваемыми временными и вероятностными показателями. Уменьшение мест ожидания в очереди приводит к росту вероятности потери запросов, однако снижает средние задержки их обработки в кластере. Для поиска разрешения данного технического противоречия определения оптимального числа мест ожидания и числа формируемых контейнеров предлагается использовать следующий обобщенный показатель эффективности обслуживания запросов в системе:

$$\Omega = (1 - P_{отк}) \cdot (T_0 - T) = Q \cdot (T_0 - T),$$

где T_0 – установленное значение среднего времени пребывания запросов в кластере. Таким образом Ω представляет собой средний запас времени относительно установленного значения для обслуженных запросов (запросов, не получивших отказ в обслуживании из-за занятой очереди).

2. Результаты моделирования и экспериментов

Эксперименты на объекте проводились с целью установления функциональной зависимости $\mu(n, m)$ снижения интенсивности обслуживания из-за разделения общих вычислительных ресурсов между сформированными (m) и активными контейнерами (n). При проведении натурных экспериментов задействованы средства: компьютер $4 \times \text{Intel(R) Core(TM) i5-4570 @ 3,20 ГГц}$, версия ядра Linux 6.2.16-3-рве. Виртуализация и управление серверами осуществлялись средствами Proxmox. Использована конфигурация виртуального сервера, включающая одно виртуальное ядро, один виртуальный сокет, 4 ГБ оперативной памяти.

Экспериментальные исследования проводились при использовании однопоточного веб-сервера, упакованного в контейнер и развернутого в нескольких репликах на одном узле (ВМ) кластера Kubernetes k3s.

В результате натурных испытаний установлено влияние разделения общих ресурсов системы на интенсивность обслуживания в контейнерах в зависимости от числа активных контейнеров и их общего числа. Полученная экспериментально зависимость представлена на рис. 2.

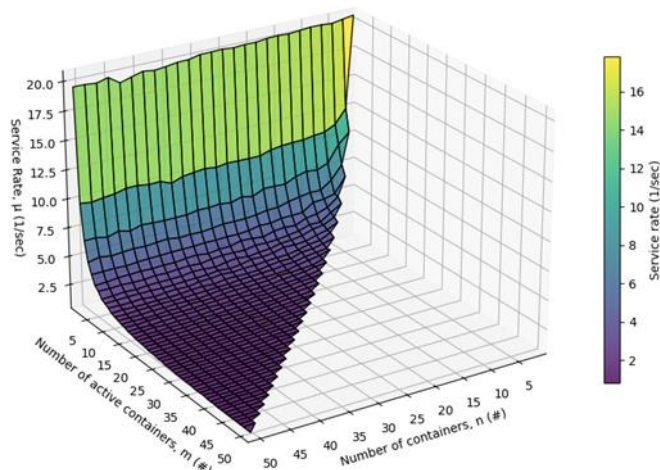


Рис. 2. Функциональная зависимость $\mu(n, m)$ снижения интенсивности обслуживания при разделении общих вычислительных ресурсов между сформированными (n) и активными контейнерами (m)

Fig. 2. Functional dependency $\mu(n, m)$ of service rate reduction when sharing common computational resources between formed (n) and active containers (m)

Исследование возможностей виртуализации кластерных систем проводится в сочетании аналитического моделирования, натурных экспериментов на объекте и имитационного моделирования. В данной работе имитационные модели строятся с использованием инструмента Simpy [30] – библиотеки Python. Этот инструмент моделирования обеспечивает возможность настройки сложных параметров системы в соответствии с требованиями исследователя.

Следует отметить, что для исследований данной статьи наиболее интересен анализ влияния числа контейнеров на эффективность обслуживания запросов.

Искомые зависимости среднего времени пребывания запросов и вероятность их потери от количества контейнеров в каждой виртуальной машине при фиксированных значениях числа виртуальных машин ($N = 3$) и размера буфера ($C = 32$), представлены на рис. 3 и 4. Из представленных результатов видно, что средняя задержка запросов в кластере с увеличением числа контейнеров уменьшается только до определенного значения, после чего начинает увеличиваться. В то же время вероятность отказа обслуживания запросов при увеличении количества контейнеров до определенной границы уменьшается, после чего изменяется незначительно.

На рис. 5 приведены результаты моделирования по определению обобщенного показателя эффективности системы при $T_0 = 3$ (с). В зависимости от интенсивности запросов при увеличении числа контейнеров до некоторого значения обобщенный показатель эффективности возрастает, однако после достижения определенного порога начинает снижаться.

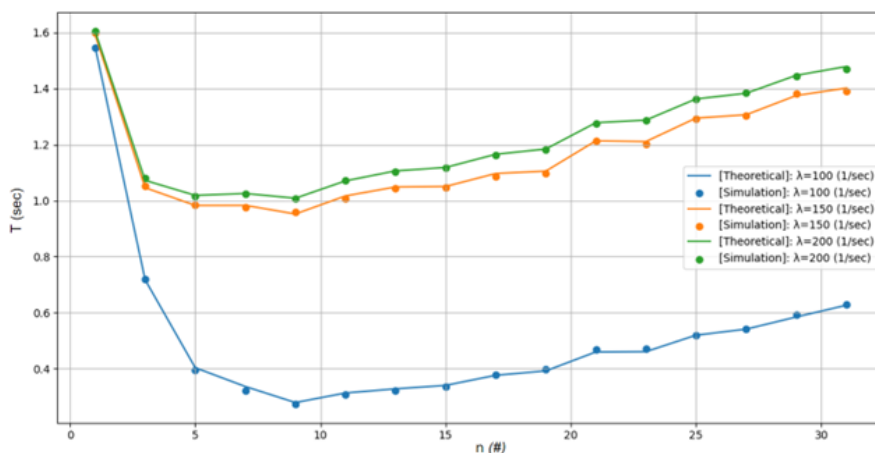


Рис. 3. Среднее время пребывания запроса в системе в зависимости от количества контейнеров

Fig. 3. Average Request Processing Time as a Function of the Number of Containers

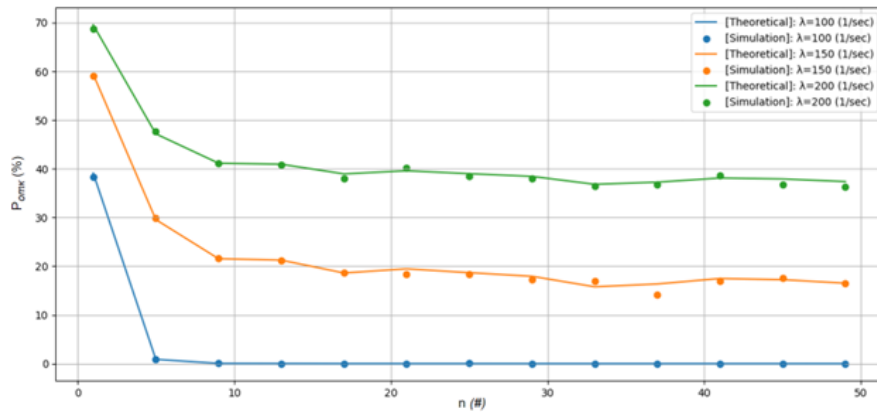


Рис. 4. Вероятность отказа в обслуживании запросов системы в зависимости от количества контейнеров
Fig. 4. Service request rejection probability as a function of the number of containers

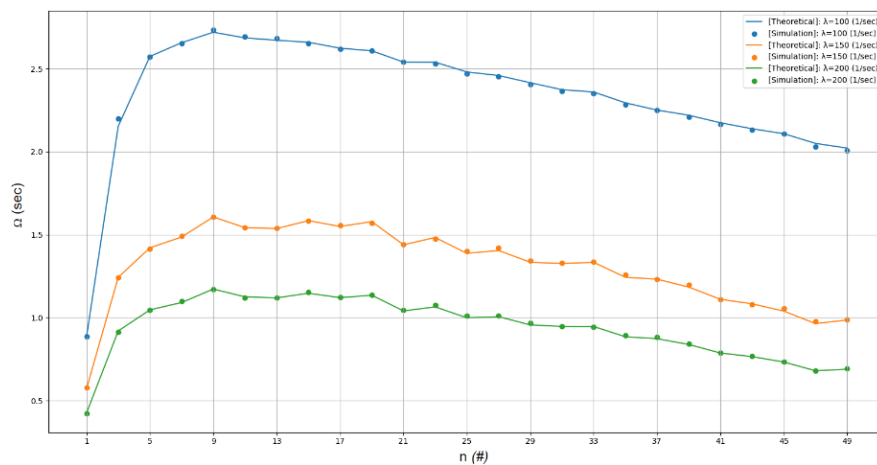


Рис. 5. Зависимость обобщенного показателя эффективности системы Ω от количества контейнеров
Fig. 5. Dependency of the generalized system efficiency indicator Ω on the number of containers

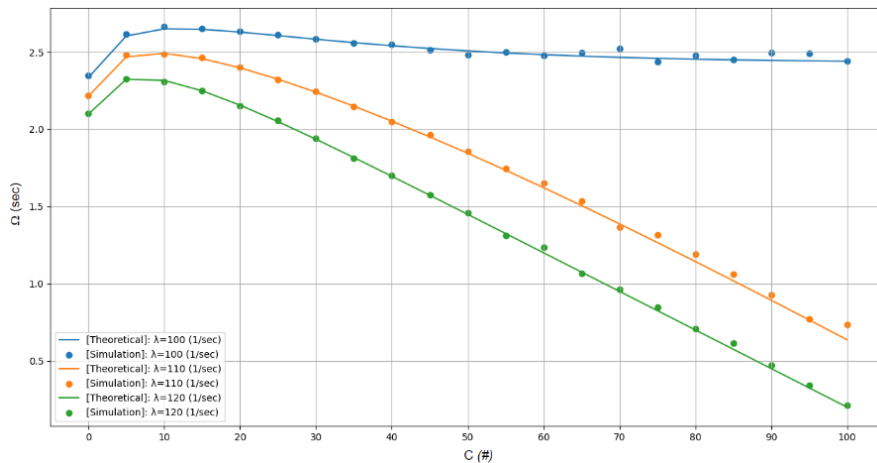


Рис. 6. Зависимость обобщенного показателя эффективности системы Ω от размеров буфера ВМ
Fig. 6. Dependency of the generalized system efficiency indicator Ω on VM Buffer Sizes

На рис. 6 показана зависимость обобщенного показателя эффективности Ω обслуживания запросов от размера буфера при $T_0 = 3$ (с). С ростом числа мест ожидания эффективность обслуживания запросов растет до определенного уровня, а затем начинает падать, причем такое падение более резко при высокой интенсивности запросов. Это обусловлено тем, что при низкой интенсивности запросов вероятность отказа незначительна и среднее время ожидания практически не меняется при увеличении размера буфера.

Заключение

Проведено исследование возможностей повышения качества обслуживания запросов в двух-уровневом виртуальном кластере. На первом уровне кластера посредством виртуальной машины с неограниченной очередью реализуется диспетчеризация общего потока запросов, выполнение которых реализуется вторым уровнем виртуального кластера. Второй уровень объединяет множество виртуальных машин с ограниченной очередью, в каждой из которых сформировано несколько функционально одинаковых контейнеров.

Предлагаемые исследования выполнены при сочетании аналитического и имитационного моделирования, а также натурных экспериментов на объекте.

Предложена аналитическая модель виртуального кластера. Виртуальная машина первого уровня кластера, осуществляющая диспетчеризацию запросов, представляется в виде одноканальной СМО с бесконечной очередью. Виртуальные машины, размещаемые на втором уровне кластера, представляются в виде многоканальных систем массового обслуживания с ограниченной очередью, при этом каждый канал обслуживания соответствует контейнеру. Особенностью предлагаемой модели виртуального кластера является учет возможного снижения интенсивности обслуживания в каналах (контейнерах) из-за разделения между ними общих ограниченных вычислительных ресурсов системы. Зависимость снижения интенсивности обслуживания запросов в каждом контейнере от числа активных и пассивных контейнеров установлена в результате натурных экспериментов.

При построении имитационной модели написана программа на языке программирования Python с использованием библиотеки SimPy.

Исследование противоречивого влияния мест ожидания в очереди, числа развертываемых виртуальных машин и формируемых в них контейнеров на эффективность обслуживания запросов в виртуальном кластере проведено по обобщенному показателю, выражающему для обслуженных запросов средний запас времени их пребывания относительно установленного граничного значения.

Список источников

1. Bennis M., Debbah M., Poor V. Ultra-reliable and low-latency wireless communication: Tail, risk and scale // Proc. of the IEEE. 2018. V. 106. P. 1834–1853. doi: 10.1109/JPROC.2018.2867029
2. Feng D., Lai L., Luo J. et al. Ultra-reliable and low-latency communications: applications, opportunities and challenges // Sci. China Inf. Sci. 2021. V. 64. P. 1–12. doi: 10.1007/s11432-020-2852-1
3. Ge X. Ultra-reliable low-latency communications in autonomous vehicular networks // IEEE Trans. Veh. Technol. 2019. V. 68 (2). P. 5005–5016. doi: 10.1109/TVT.2019.2903793
4. Philip P.C., Abdel-Hafez M. A Review on Ultra Reliable and Low Latency Communications (PHY and MAC Layer Perspectives) // Journal of Computer Science. 2020. V. 16 (8). P. 1100–1114.
5. Astakhova T., Shamin A., Verzun N., Kolbanev M. A model for estimating energy consumption seen when nodes of ubiquitous sensor networks communicate information to each other // CEUR Workshop Proc. 2019. № 2. URL: <https://ceur-ws.org/Vol-2344/paper5.pdf>
6. Korobeynikov A.G., Shukalov A.V., Zharinov I.O., Zharinov O.O. Cyber-physical and human control of technological process and equipment sensitive to failures // IOP Conference Series: Earth and Environmental Science. 2021. V. 839 (4). Art. 040281. doi: 10.1088/1755-1315/839/4/040281
7. Татарникова Т.М., Архипцев Е.Д., Кармановский Н.С. Определение размера кластера и числа реплик высоконагруженных информационных систем // Известия высших учебных заведений. Приборостроение. 2023. Т. 66, № 8. С. 646–651.
8. Gurjanov A.V., Korobeynikov A.G., Zharinov I.O., Zharinov O.O. Edge, fog and cloud computing in the cyber-physical systems networks // III International Workshop on Modeling, Information Processing and Computing (MIP: Computing-2021). Krasnoyarsk, 2021. P. 103–108.
9. Bogatyrev V.A., Bogatyrev A.V., Bogatyrev S.V. Multipath transmission of heterogeneous traffic in acceptable delays with packet replication and destruction of expired replicas in the nodes that make up the path // Communications in Computer and Information Science. 2023. V. 1748. P. 104–121. doi: 10.1007/978-3-031-30648-8_9
10. Богатырев В.А., Богатырев С.В., Богатырев А.В. Оценка готовности компьютерной системы к своевременному обслуживанию запросов при его совмещении с информационным восстановлением памяти после отказов // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 3. С. 608–617.
11. Bogatyrev A.V., Bogatyrev V.A., Bogatyrev S.V. The probability of timeliness of a fully connected exchange in a redundant real-time communication system // WECONF 2020. 2020. P. 1–4. doi: 10.1109/WECONF48837.2020.9131517

12. Андреев С.Д., Самуйлов К.Е., Тюрликов А.М. Управление потоками в гетерогенных мобильных сетях радиодоступа с соединениями устройство-устройство // Вестник Российского университета дружбы народов. Сер. Математика, информатика, физика. 2018. Т. 26. № 4. С. 357–371.
13. Bogatyrev V.A. Increasing the fault tolerance of a multi-trunk channel by means of inter-trunk packet forwarding // Automatic Control and Computer Sciences. 1999. V. 33 (2). P. 70–76.
14. Bhardwaj A., Krishna C.R. Virtualization in Cloud Computing: Moving from Hypervisor to Containerization - A Survey // Arab. J. Sci. Eng. 2021. V. 46. P. 8585–8601. doi: 10.1007/s13369-021-05553-3
15. Bermejo B., Juiz C. Virtual machine consolidation: a systematic review of its overhead influencing factors // J. Supercomput. 2020. V. 76. P. 324–361. doi: 10.1007/s11227-019-03025-y
16. Bentaleb O., Belloum A.S.Z., Sebaa A. et al. Containerization technologies: taxonomies, applications and challenges // J Supercomput. 2022. V. 78 (1). P. 1144–1181. doi: 10.1007/s11227-021-03914-1
17. Фунг В.К., Богатырев В.А., Кармановский Н.С., Лэ В.Х. Оценка вероятностно-временных характеристик компьютерной системы с контейнерной виртуализацией // Научно-технический вестник информационных технологий, механики и оптики. 2024. Т. 24, № 2. С. 249–255. doi: 10.17586/2226-1494-2024-24-2-249-255
18. Podolskiy V., Jindal A., Gerndt M. Multilayered autoscaling performance evaluation: can virtual machines and containers co-scale // Int. J. Appl. Math. Comput. Sci. 2019. V. 29 (2). P. 227–244. doi: 10.2478/amcs-2019-0017
19. Abbaspour M., Khendek F., Toeroe M. A model-based approach for design time elasticity rules generation // Proc. of the 23rd International Conference on Engineering of Complex Computer Systems, IEEE. 2018. P. 93–103. doi: 10.1109/ICECCS2018.2018.00018
20. Vilaplana J., Solsona F., Teixidó I. et al. A queuing theory model for cloud computing // J Supercomput. 2014. V. 69. P. 492–507. doi: 10.1007/s11227-014-1177-y
21. Chandy K.M., Sauer C.H. Approximate methods for analyzing queueing network models of computing systems // ACM Comput. Surv. 1978. V. 10 (3). P. 281–317. doi: 10.1145/356733.356737
22. Клейнрок Л. Теория массового обслуживания : учеб. пособие. М. : Машиностроение, 1979. 432 с.
23. Бочков А.П., Хомоненко А.Д., Барановский А.М. Модель формирования кластеров информативных узлов интегрированной и распределенной обработки данных в вычислительной сети // Научные технологии в космических исследованиях Земли. 2021. Т. 13, № 1. С. 44–57.
24. Bogatyrev V.A., Bogatyrev S.V., Bogatyrev A.V. Control of multipath transmissions in the nodes of switching segments of reserved paths // Proc. of the 2022 International Conference on Information, Control, and Communication Technologies (ICCT). 2022. P. 1–5. doi: 10.1109/ICCT56057.2022.9976839
25. Khalil M.M., Khomonenko A.D., Matushko M.D. Measuring the effect of monitoring on a cloud computing system by estimating the delay time of requests // Journal of King Saud University. Computer and Information Sciences. 2022. V. 34 (7). P. 3968–3972. doi: 10.1016/j.jksuci.2021.02.001
26. Ермаков С.Г., Халил М.М., Хомоненко А.Д., Бухарова К.А. Марковская модель оценивания оперативности туманных вычислений в интернете вещей // T-Comm: Телекоммуникации и транспорт. 2022. Т. 16, № 12. С. 46–52.
27. Khalil M.M., Khomonenko A.D., Gindin S.I. Load Balancing Cloud Computing with Web-Interface Using Multi-channel Queuing Systems with Warming up and Cooling // Proc. Intelligent Distributed Computing XIII. 2020. P. 385–393. doi: 10.1007/978-3-030-32258-8_45
28. Nazarov A.N. Processing streams in a monitoring cloud cluster // Russian Technological Journal. 2019. V. 7 (6). P. 56–67. doi: 10.32362/2500-316X-2019-7-6-56-67
29. Altahtat M.A., Agarwal A., Goel N., Zaman M. Neural network based regression model for virtual machines migration method selection // Proc. IEEE International Conference on Communications Workshops, ICC Workshops (2021). 2021. P. 1–6. doi: 10.1109/ICCWorkshops50388.2021.9473749
30. Matloff N. Introduction to discrete-event simulation and the simpy language. Davis, CA, 2009. P. 1–33. Dep. of Computer Science: University of California at Davis.

References

1. Bennis, M., Debbah, M. & Poor, V. (2018) Ultra-reliable and low-latency wireless communication: Tail, risk and scale. *Proceedings of the IEEE*. 106. pp.1834–1853. DOI: 10.1109/JPROC.2018.2867029
2. Feng, D., Lai, L., Luo, J. et al. (2021) Ultra-reliable and low-latency communications: applications, opportunities and challenges. *Sci. China Inf. Sci.* 64. pp. 1–12. DOI: 10.1007/s11432-020-2852-1
3. Ge, X. (2019) Ultra-reliable low-latency communications in autonomous vehicular networks. *IEEE Trans. Veh. Technol.* 68(2). pp. 5005–5016. DOI: 10.1109/TVT.2019.2903793
4. Prinu Chacko, P. & Abdel-Hafez, M. (2020) A Review on Ultra Reliable and Low Latency Communications (PHY and MAC Layer Perspectives). *Journal of Computer Science*. 16(8). pp. 1100–1114.
5. Astakhova, T., Shamin, A., Verzun, N. & Kolbanev, M. (2019) A model for estimating energy consumption seen when nodes of ubiquitous sensor networks communicate information to each other. *CEUR Workshop Proceedings*. 2344. no. 2.
6. Korobeynikov, A.G., Shukalov, A.V., Zharinov, I.O. & Zharinov, O.O. (2021) Cyber-physical and human control of technological process and equipment sensitive to failures. *IOP Conference Series: Earth and Environmental Science*. 839(4).

7. Tatarnikova, T.M., Arkhitektsev, E.D. & Karmanovskiy, N.S. (2023) Opredelenie razmera klastera i chisla replik vysokonagruzhennykh informatsionnykh sistem [Determination of the cluster size and the number of replicas of highly loaded information systems]. *Izvestiya vysshikh uchebnykh zavedeniy. Priborostroenie*. 66(8). pp. 646–651.
8. Gurjanov, A.V., Korobeynikov, A.G., Zharinov, I.O. & Zharinov, O.O. (2021) Edge, fog and cloud computing in the cyber-physical systems networks. *III International Workshop on Modeling, Information Processing and Computing (MIP: Computing-2021)*. pp. 103–108.
9. Bogatyrev, V.A., Bogatyrev, A.V. & Bogatyrev, S.V. (2023) Multipath transmission of heterogeneous traffic in acceptable delays with packet replication and destruction of expired replicas in the nodes that make up the path. *Communications in Computer and Information Science*. 1748. pp. 104–121. DOI: 10.1007/978-3-031-30648-8_9
10. Bogatyrev, V.A., Bogatyrev, S.V. & Bogatyrev, A.V. (2023) Otsenka gotovnosti komp'yuternoy sistemy k svoevremennomu obsluzhivaniyu zaprosov pri ego sovmeshchenii s informatsionnym vosstanovleniem pamyati posle otkazov [Assessing the readiness of a computer system for timely request servicing when combined with information memory recovery after failures]. *Nauchno-tekhnicheskii vestnik informatsionnykh tekhnologiy, mekhaniki i optiki*. 23(3). pp. 608–617.
11. Bogatyrev, A.V., Bogatyrev, V.A. & Bogatyrev, S.V. (2020) The probability of timeliness of a fully connected exchange in a redundant real-time communication system. *WECONF*. pp. 1–4. DOI: 10.1109/WECONF48837.2020.9131517
12. Andreev, S.D., Samuylov, K.E. & Tyurlikov, A.M. (2018) Upravlenie potokami v geterogennykh mobil'nykh setyakh radio-dostupa s soedineniyami ustroystvo-ustroystvo [Flow Management in Heterogeneous Mobile Radio Access Networks with Device-to-Device Connections]. *Vestnik Rossiyskogo universiteta druzhby narodov. Ser. Matematika, informatika, fizika*. 26(4). pp. 357–371.
13. Bogatyrev, V.A. (1999) Increasing the fault tolerance of a multi-trunk channel by means of inter-trunk packet forwarding. *Automatic Control and Computer Sciences*. 33(2). pp. 70–76.
14. Bhardwaj, A. & Krishna, C.R. (2021) Virtualization in Cloud Computing: Moving from Hypervisor to Containerization – A Survey. *Arabian Journal for Science and Engineering*. 46. pp. 8585–8601. DOI: 10.1007/s13369-021-05553-3
15. Bermejo, B. & Juiz, C. (2020) Virtual machine consolidation: a systematic review of its overhead influencing factors. *Journal of Supercomputing*. 76. pp. 324–361. DOI: 10.1007/s11227-019-03025-y
16. Bentaleb, O., Belloum, A.S.Z., Sebaa, A. et al. (2022) Containerization technologies: taxonomies, applications and challenges. *Journal of Supercomputing*. 78(1). pp. 1144–1181. DOI: 10.1007/s11227-021-03914-1
17. Phung, V.Q., Bogatyrev, V.F., Karmanovskiy, N.S. & Le, V.H. (2024) Otsenka veroyatnostno-vremennykh kharakteristik komp'yuternoy sistemy s konteynernoy virtualizatsiei [Estimation of probabilistic-temporal characteristics of a computer system with container virtualization]. *Nauchno-tekhnicheskii vestnik informatsionnykh tekhnologiy, mekhaniki i optiki*. 24(2). pp. 249–255. DOI: 10.17586/2226-1494-2024-24-2-249-255
18. Podolskiy, V., Jindal, A. & Gerndt, M. (2019) Multilayered autoscaling performance evaluation: can virtual machines and containers co-scale. *International Journal of Applied Mathematics and Computer Science*. 29(2). pp. 227–244. DOI: 10.2478/amcs-2019-0017
19. Abbaspour, M., Khendek, F. & Toeroe, M. (2018) A model-based approach for design time elasticity rules generation. *Proceedings of the 23rd International Conference on Engineering of Complex Computer Systems, IEEE*. pp. 93–103. DOI: 10.1109/ICECCS2018.2018.00018
20. Vilaplana, J., Solsona, F., Teixidó, I. et al. (2014) A queuing theory model for cloud computing. *Journal of Supercomputing*. 69. pp. 492–507. DOI: 10.1007/s11227-014-1177-y
21. Chandy, K.M. & Sauer, C.H. (1978) Approximate methods for analyzing queueing network models of computing systems. *ACM Comput. Surv.* 10(3). pp. 281–317. DOI: 10.1145/356733.356737
22. Kleinrock, L. (1979) *Teoriya massovogo obsluzhivaniya* [Theory of Queuing Systems]. Moscow: Mashinostroenie.
23. Bochkov, A.P., Khomonenko, A.D. & Baranovskiy, A.M. (2021) Model' formirovaniya klasterov informativnykh uzlov integrirovannoy i raspredelennoy obrabotki dannykh v vychislitel'noy seti [Model of forming clusters of informative nodes for integrated and distributed data processing in a computer network]. *Naukoemkie tekhnologii v kosmicheskikh issledovaniyakh Zemli*. 13(1). pp. 44–57.
24. Bogatyrev, V.A., Bogatyrev, S.V. & Bogatyrev, A.V. (2022) Control of multipath transmissions in the nodes of switching segments of reserved paths. *Proceedings of the 2022 International Conference on Information, Control, and Communication Technologies (ICCT)*. pp. 1–5. DOI: 10.1109/ICCT56057.2022.9976839
25. Khalil, M.M., Khomonenko, A.D. & Matushko, M.D. (2022) Measuring the effect of monitoring on a cloud computing system by estimating the delay time of requests. *Journal of King Saud University. Computer and Information Sciences*. 34(7). pp. 3968–3972. DOI: 10.1016/j.jksuci.2021.02.001
26. Ermakov, S.G., Khalil, M.M., Khomonenko, A.D. & Bukharova K.A. (2022) Markovskaya model' otsenivaniya operativnosti tumannykh vychisleniy v internete veshchey [Markov model for assessing the efficiency of fuzzy computations in the Internet of Things]. *T-Comm: Telekommunikatsii i transport*. 16(12). pp. 46–52.
27. Khalil, M.M., Khomonenko, A.D. & Gindin, S.I. (2020) Load Balancing Cloud Computing with Web-Interface Using Multi-channel Queuing Systems with Warming up and Cooling. *Proceedings Intelligent Distributed Computing XIII*. pp. 385–393. DOI: 10.1007/978-3-030-32258-8_45
28. Nazarov, A.N. (2019) Processing streams in a monitoring cloud cluster. *Russian Technological Journal*. 7(6). pp. 56–67. DOI: 10.32362/2500-316X-2019-7-6-56-67

29. Altahat, M.A., Agarwal, A., Goel, N. & Zaman, M. (2021) Neural network based regression model for virtual machines migration method selection. *Proceedings IEEE Int. Conference on Communications Workshops*. ICC Workshops 2021 – Proceedings. pp. 1–6. DOI: 10.1109/ICCWorkshops50388.2021.9473749
30. Matloff, N. (2009) *Introduction to discrete-event simulation and the SimPy language*. Dep. of Computer Science: University of California at Davis. pp. 1–33.

Информация об авторах:

Фунг Ван Кю – аспирант факультета программной инженерии и компьютерной техники Университета ИТМО (Санкт-Петербург, Россия). E-mail: phungvanquy97@gmail.com

Богатырев Владимир Анатольевич – доктор технических наук, профессор факультета программной инженерии и компьютерной техники Университета ИТМО (Санкт-Петербург, Россия); профессор Санкт-Петербургского государственного университета аэрокосмического приборостроения (Санкт-Петербург, Россия). E-mail: vladimir.bogatyrev@gmail.com

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

Information about the authors:

Phung Van Quy (Post-graduate Student, National Research University of Information Technologies, Mechanics and Optics (ITMO University), Saint Petersburg, Russia Federation). E-mail: phungvanquy97@gmail.com

Bogatyrev Vladimir A. (Doctor of Technical Sciences, Professor, National Research University of Information Technologies, Mechanics and Optics (ITMO University), Saint Petersburg, Russia Federation; Professor in Saint Petersburg State University of Aerospace Instrumentation, Saint Petersburg, Russia Federation). E-mail: vladimir.bogatyrev@gmail.com

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

Поступила в редакцию 03.04.2024; принята к публикации 03.09.2024

Received 03.04.2024; accepted for publication 03.09.2024

Научная статья

УДК 519.217.2

doi: 10.17223/19988605/68/6

**Управляемые системы массового обслуживания
со стационарным равномерным распределением****Гурами Шалвович Цициашвили¹, Юрий Николаевич Харченко²**^{1,2} *Институт прикладной математики ДВО РАН, Владивосток, Россия*¹ *guram@iam.dvo.ru*² *har@iam.dvo.ru*

Аннотация. Рассматривается способ введения в систему массового обслуживания дополнительных переходных интенсивностей, создающих равномерное стационарное распределение на множестве состояний. В его основе лежит выделение состояний, в которых сумма входных интенсивностей не равна сумме выходных интенсивностей. Между данными состояниями вводятся новые переходные интенсивности, которые преобразуют эти неравенства в равенства. Такое преобразование, с одной стороны, связано с процедурой очистки системы обслуживания от заявок. С другой стороны, эта процедура в какой-то степени моделирует процесс передачи информации от кратковременной памяти в долговременную. В работе процедура введения в граф переходных интенсивностей новых элементов понимается как управление в системе обслуживания. Примененные в работе алгоритмы введения новых переходных интенсивностей тесно связаны с процедурой определения максимального потока в сети и с методом северо-западного угла в линейном программировании.

Ключевые слова: стационарное распределение; переходные интенсивности; эргодичность; потоки в сетях.

Благодарности: Исследование выполнено в рамках государственного задания Института прикладной математики ДВО РАН (№ 075-00459-24-00).

Для цитирования: Цициашвили Г.Ш., Харченко Ю.А. Управляемые системы массового обслуживания со стационарным равномерным распределением // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 68. С. 59–65. doi: 10.17223/19988605/68/6

Original article

doi: 10.17223/19988605/68/6

Controlled queuing systems with a stationary uniform distribution**Gurami Sh. Tsitsiashvili¹, Yury N. Kharchenko²**^{1,2} *Institute for Applied Mathematics, FEB RAS, Vladivostok, Russian Federation*¹ *guram@iam.dvo.ru*² *har@iam.dvo.ru*

Abstract. The paper considers a method for introducing additional transient intensities into the queuing system, creating a uniform stationary distribution over a set of states. It is basing on the selection of states in which the sum of the input intensities is not equal to the sum of the output intensities. New transitional intensities are introducing between these states, which transform these inequalities into equalities. On the one hand, such a transformation is connecting with the procedure for clearing the service system of customers. On the other hand, this procedure to some extent simulates the process of transferring information from short-term memory to long-term memory. In this paper, the procedure for introducing new elements into the graph of transient intensities is understood as management in the queuing system. The algorithms used in the work for introducing new transient intensities are closely relating to the procedure for determining the maximum flow in the network and to the northwest angle method in linear programming.

Keywords: stationary distribution; transient intensities; ergodicity; flows in networks.

Acknowledgments: The research was carried out within the framework of the state assignment of the Institute of Applied Mathematics Far Eastern Branch of RAS (No. 075-00459-24-00).

For citation: Tsitsiashvili, G.Sh., Kharchenko, Yu.N. (2024) Controlled queuing systems with a stationary uniform distribution. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika* – Tomsk State University Journal of Control and Computer Science. 68. pp. 59–65. doi: 10.17223/19988605/68/6

Введение

В настоящей работе вводится и анализируется способ управления одноканальной системой массового обслуживания с отказами с целью придания ей свойств, отображающих интересные для приложений закономерности реальных процессов. Эта система массового обслуживания является частным случаем систем стохастической очистки, которые принимают и накапливают входные данные случайных величин в течение случайных интервалов времени до тех пор, пока не будут выполнены определенные заранее критерии. Затем происходит мгновенная очистка некоторых или всех этих входных данных (см., напр.: [1–4]). В то же время имеется целая серия анатомио-психологических исследований по превращению кратковременной памяти в долговременную с участием гиппокампа [5–8].

Обращает на себя внимание, что для того, чтобы эти системы были приемлемы для многократного использования, они должны обладать достаточно сглаженным распределением на множестве состояний. Для этого здесь по аналогии со статистической механикой [9–11] конструируется система массового обслуживания с равномерным распределением на множестве состояний. Внимание именно к равномерному распределению объясняется тем, что для него характерно отсутствие каких-либо приоритетов для отдельных состояний. Причем равномерное стационарное распределение позволяет уменьшить вероятность попадания в состояние, характеризующееся уходом заявок из системы или переходом информации, описываемой числом заявок в системе, из кратковременной / оперативной памяти в долговременную.

На языке математического моделирования такие системы могут быть описаны [12, 13] с помощью стационарных уравнений Колмогорова–Чепмена, приводящих для заданного равномерного распределения к транспортной задаче линейного программирования (см., напр.: [14]). В настоящей работе рассматривается задача определения переходных интенсивностей в системе обслуживания, строящейся на основе системы с отказами и конечным числом мест ожидания с возможностью очистки и восстановления системы от заявок в случае заполнения всех мест ожидания. В работе процедура введения в граф переходных интенсивностей, описывающих систему обслуживания, новых элементов понимается как управление.

1. Система массового обслуживания с отказами

Рассмотрим модель одноканальной системы массового обслуживания с отказами $M | M | 1 | m$ при превышении числом заявок верхней границы m . Опишем эту систему массового обслуживания дискретным марковским процессом $x(t)$ с состояниями $0, 1, \dots, m$, характеризующими число заявок в системе. Предположим, что пуассоновский входной поток в эту систему обслуживания имеет интенсивность λ , а интенсивность обслуживания равна μ . Тогда переходные интенсивности между состояниями системы имеют вид:

$$\gamma(i, i+1) = \lambda, \gamma(i+1, i) = \mu \quad (i = 0, 1, \dots, m-1), \quad (1)$$

и представлены на рис. 1.

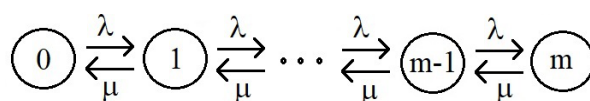


Рис. 1. Переходные интенсивности в системе с отказами $M | M | 1 | m$

Fig.1. Transient intensities in a system with failures

Обозначим $\rho = \lambda/\mu$ коэффициент загрузки системы, тогда при $\rho < 1$ стационарная вероятность $p(i)$ нахождения системы в состоянии $i, i = 0, 1, \dots, m$, удовлетворяет равенству

$$p_i = \frac{(1-\rho)\rho^i}{1-\rho^{m+1}}, \rho \neq 1; p_i = \frac{1}{m+1}, \rho = 1. \quad (2)$$

2. Система массового обслуживания с равномерным распределением стационарных вероятностей

Рассмотрим теперь задачу преобразования системы с отказами $M|M|1|m$ в систему с равномерным распределением стационарных вероятностей. Для этого введем в систему $M|M|1|m$ дополнительные интенсивности перехода, характеризующие очистку системы от заявок и восполнение системы удаленными заявками

$$\gamma(0, m) = \lambda, \gamma(m, 0) = \mu. \quad (3)$$

Переходные интенсивности в системе с очисткой от заявок представлены на рис. 2.

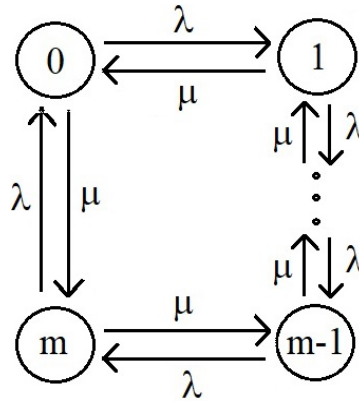


Рис. 2. Переходные интенсивности в системе с очисткой от заявок
Fig. 2. Transient intensities in a system with clearance of customers

Элементарный расчет показывает, что состояния $0, 1, \dots, m$ образуют цикл, в котором сумма интенсивностей, входящих в состояние i , равна сумме интенсивностей, выходящих из состояния i (и равна $\lambda + \mu$). Отсюда следует, что стационарные вероятности нахождения в состояниях расширенной системы удовлетворяют равенству

$$p(i) = \frac{1}{m+1}, i = 0, 1, \dots, m, \quad (4)$$

а само стационарное распределение является равномерным. Тем самым введением только двух новых переходных интенсивностей, приведенных в формуле (3), стационарное распределение (2) преобразуется в равномерное стационарное распределение (4). Особенно сильно такое изменение сказывается для состояния 0 , в которое дискретный марковский процесс $x(t)$, характеризующий число заявок в системе в момент t , переходит из состояния m .

Замечание. Следует отметить, что введенные на рис. 2 переходные интенсивности $\gamma(0, m) = \mu$, $\gamma(m, 0) = \lambda$ могут быть и другими. Действительно, в вершинах $1, \dots, m-1$ графа, изображающего состояния системы $M|M|1|m$ и ее переходные интенсивности (см. рис. 1), разность между суммой выходных интенсивностей и суммой входных интенсивностей равна нулю. В вершине 0 эта разность равна $\lambda - \mu$, а в вершине m она равна $\mu - \lambda$. Поэтому, вводя между вершинами $0, m$ переходные интенсивности $\gamma(0, m) = \mu', \gamma(m, 0) = \lambda', \lambda + \mu' = \lambda' + \mu$, можно получить равенство нулю соответствующих разностей. Причем при $\lambda > \mu$ можно выбрать $\lambda' = \lambda - \mu + \varepsilon, \mu' = \varepsilon, \varepsilon > 0$. Аналогично, если $\mu > \lambda$, то, вводя между вершинами $0, m$ переходные интенсивности $\gamma(0, m) = \lambda', \gamma(m, 0) = \mu', \lambda + \mu' = \lambda' + \mu$, мож-

но при $\mu' = \mu - \lambda + \varepsilon$, $\lambda' = \varepsilon$, $\varepsilon > 0$, получить равенство нулю соответствующих разностей входящих и выходящих интенсивностей.

3. Объединение систем со стационарным равномерным распределением

Рассмотрим граф состояний и переходных интенсивностей для двух систем типа $M | M | 1 | m$, по которым перемещаются заявки разного типа (например, звуковые и текстовые сообщения). Пусть нулевые состояния в этих системах описывают хранилище долговременной памяти. Такую систему можно изобразить с помощью рис. 3. Тогда новые переходные интенсивности между крайними состояниями этой системы можно определять отдельно для каждой из изображаемых подсистем типа $M | M | 1 | m$. Такая постановка вопроса позволяет перейти к более сложным системам, изображаемым, например, ациклическим ориентированным графом. Особенностью такой системы является тот факт, что стационарная вероятность каждого состояния равна $1/(m + m' + 2) < 1/(m + 1)$. Примером такого объединения двух систем, подобной системе, представленной на рис. 2, является система, изображенная на рис. 3. Присоединяя к состоянию 0 не одну подобную рис. 2 систему, можно достичь существенного уменьшения стационарной вероятности нахождения в состояниях объединенной системы. Особенно это важно отнести к состоянию 0, являющемуся общим для объединяемых подсистем.

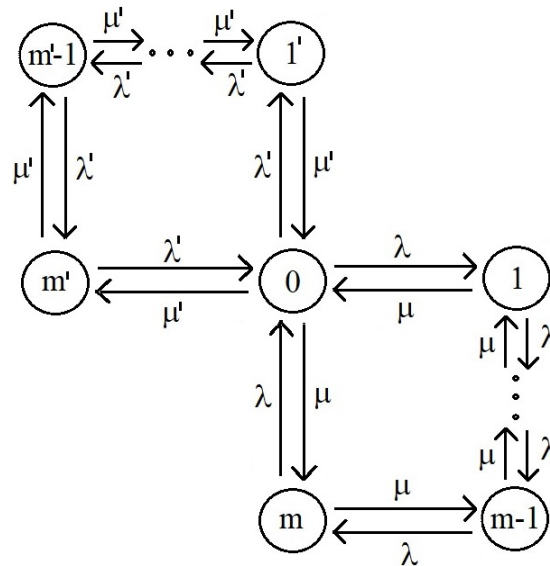


Рис. 3. Переходные интенсивности в объединенной системе
Fig. 3. Transient intensities in the combined system

4. Теорема об эргодичности стационарного распределения

В заключение следует отметить, что все рассмотренные в работе утверждения основаны на следующей теореме.

Теорема. Предположим, что дискретный марковский процесс $x(t)$ с конечным множеством состояний I и с переходными интенсивностями $\gamma(i, j)$, $i, j \in I$, $i \neq j$, удовлетворяет следующему условию связности состояний для любых $i, j \in I$, $i \neq j$. Пусть все состояния этого процесса являются сообщающимися, т.е. для $\forall i, j \in I \exists i_1, \dots, i_r \in I$, $r \geq 0$, такими, что

$$\gamma(i, i_1) > 0, \gamma(i_1, i_2) > 0, \dots, \gamma(i_r, j) > 0, \quad (5)$$

и для любых $i \in I$ выполняются равенства

$$\sum_{j \in I, j \neq i} \gamma(i, j) = \sum_{j \in I, j \neq i} \gamma(j, i). \quad (6)$$

Тогда марковский процесс является эргодическим, и его стационарное распределение $p(i)$ является равномерным на множестве состояний I .

Действительно, из соотношений (5), (6) следует эргодичность (а значит, и существование единственного стационарного распределения) у дискретного марковского процесса $x(t)$ [13, 14]. Причем из равенств (6) следует, что равномерное распределение $p(i), i \in I$, удовлетворяет стационарному уравнению Колмогорова–Чепмена. Теорема позволяет строить различные обобщения переходных интенсивностей, приводящих к стационарному равномерному распределению соответствующих им дискретных марковских процессов.

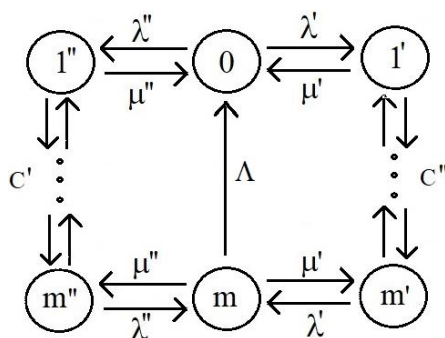


Рис. 4. Максимальная интенсивность перехода из выходного состояния m во входное состояние 0 , обеспечивающая стационарное равномерное распределение сети

Fig. 4. The maximum intensity of the transition from the output state m to the input state 0 , which ensures a stationary uniform distribution of the network

В качестве примера рассмотрим объединение двух линейных цепочек состояний с общим начальным состоянием 0 и с длинами $m'+2, m''+2$, соединенных в общем конечном состоянии m (рис. 4). Пусть максимально допустимая разность интенсивностей (пропускная способность) $\lambda' - \mu'$ у ребер первой цепочки равна $c' > 0$, а у ребер второй цепочки пропускная способность $\lambda'' - \mu''$ равна $c'' > 0$. Тогда сумма выходных интенсивностей из состояния 0 равна сумме входных интенсивностей в состояние m и равна $\Lambda = c' + c''$. Эту величину можно трактовать как максимальный поток [15] в сети без ребра $(m, 0)$ с заданными пропускными способностями ребер. Поэтому переходная интенсивность Λ из состояния m в состояние 0 обеспечивает равенство сумм входных и выходных интенсивностей для этих состояний и является максимальной.

Заключение

В работе предпринята попытка введения в граф переходных интенсивностей, описывающих функционирование системы массового обслуживания, новых «взвешенных» ребер, обеспечивающих выполнение стационарных уравнений Колмогорова–Чепмена во всех узлах графа. Если исходный граф переходных интенсивностей представляет поток в терминах теоремы Форда–Фалкерсона, то имеются возможности введением нового ребра из выходной вершины во входную добиться выполнения этих уравнений во всех вершинах графа. Причем возникает возможность оптимизационной постановки для этого графа с помощью теоремы о максимальном потоке в детерминированной сети.

Список источников

1. Ghosh S., Hassin R. Inefficiency in stochastic queueing systems with strategic customers // European J. Oper. Res. 2021. V. 295 (1). P. 1–11.
2. He Q.M., Bookbinder J.H., Cai Q. Optimal policies for stochastic clearing systems with time-dependent delay penalties // Naval Res. Logist. 2020. V. 67 (7). P. 487–502.
3. Kim B.K., Lee D.H. The M/G/1 queue with disasters and working breakdowns // Appl. Math. Model. 2014. V. 38 (5-6). P. 1788–1798.

4. Missbauer H., Stolz R. Schneckenreither M. Order release optimisation for time-dependent and stochastic manufacturing systems // *Int. J. Prod. Res.* 2024. V. 62 (7). P. 2415–2434.
5. Вартанов А.В., Козловский С.А. и др. Память человека и анатомические особенности гиппокампа // *Вестник Московского университета. Сер. 14. Психология.* 2009. № 4. С. 3–16.
6. Lye T.C., Grayson D.A. et al. Predicting memory performance in normal ageing using different measures of hippocampal size // *Neuroradiology.* 2006. V. 48 (2). P. 90–99.
7. Lupien S.J., Evans A. et al. Hippocampal volume is as variable in young as in older adults: Implications for the notion of hippocampal atrophy in humans // *Neuroimage.* 2007. V. 34 (2). P. 479–485.
8. Squire L.R. The legacy of patient H.M. for neuroscience // *Neuron.* 2009. V. 61 (1). P. 6–9.
9. Сивухин Д.В. *Общий курс физики. М. : Наука, 1990. Т. 2: Термодинамика и молекулярная физика.* 591 с.
10. Pathria R.K., Beale P.D. *Statistical Mechanics.* Elsevier, 2011. 540 p.
11. Gyenis B. Maxwell and the normal distribution: a coloured story of probability, independence, and tendency toward equilibrium // *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics.* 2017. V. 57. P. 53–65.
12. Ивченко Г.И., Каштанов В.А., Коваленко И.Н. *Теория массового обслуживания. М. : Высш. школа, 1982.* 256 с.
13. Коваленко И.Н., Кузнецов Н.Ю., Шуренков В.М. *Случайные процессы. Киев : Наукова думка, 1983.* 366 с.
14. Галютдинов Р.Р. Транспортная задача: метод Северо-Западного угла // Сайт преподавателя экономики. 2013. URL: <https://galyautdinov.ru/post/metod-severo-zapadnogo-ugla> (дата обращения: 22.02.2024).
15. Кормен Т.Х. и др. *Алгоритмы: построение и анализ. М. : Вильямс, 2006.* 1296 с.

References

1. Ghosh, S. & Hassin, R. (2021) Inefficiency in stochastic queueing systems with strategic customers. *European Journal of Operational Research.* 295(1). pp. 1–11.
2. He, Q.M., Bookbinder, J.H. & Cai, Q. (2020) Optimal policies for stochastic clearing systems with time-dependent delay penalties. *Naval Research Logistics.* 67(7). pp. 487–502.
3. Kim, B.K. & Lee, D.H. (2014) The M/G/1 queue with disasters and working breakdowns. *Applied Mathematical Modelling.* 38(5-6). pp. 1788–1798.
4. Missbauer, H., Stolz, R. & Schneckenreither, M. (2024) Order release optimisation for time-dependent and stochastic manufacturing systems. *International Journal of Production Research.* 62(7). pp. 2415–2434
5. Vartanov, A.V. & Kozlovsky, S.A. et al. (2009) Human memory and anatomical features of the hippocampus. *Vestnik Moskovskogo universiteta. Ser. 14. Psikhologiya.* 14(4). pp. 3–16.
6. Lye, T.C. & Grayson, D.A. et al. (2006) Predicting memory performance in normal ageing using different measures of hippocampal size. *Neuroradiology.* 48(2). pp. 90–99.
7. Lupien, S.J. & Evans, A. et al. (2007) Hippocampal volume is as variable in young as in older adults: Implications for the notion of hippocampal atrophy in humans. *Neuroimage.* 34(2). pp. 479–485.
8. Squire, L.R. (2009) The legacy of patient H.M. for neuroscience. *Neuron.* 61(1). pp. 6–9.
9. Sivukhin, D.V. (1990) *Obshchiy kurs fiziki* [General Course of Physics]. Vol. 2. Moscow: Nauka.
10. Pathria, R.K. & Beale, P.D. (2011) *Statistical Mechanics.* Elsevier.
11. Gyenis, B. (2017) Maxwell and the normal distribution: A coloured story of probability, independence, and tendency toward equilibrium. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics.* 57. pp. 53–65.
12. Ivchenko, G.I., Kashtanov, V.A. & Kovalenko, I.N. (1982) *Teoriya massovogo obsluzhivaniya* [Theory of Queuing]. Moscow: Vysshaya shkola.
13. Kovalenko, I.N., Kuznetsov, N.Yu. & Shurenkov, V.M. (1983) *Sluchaynye protsessy* [Random Processes]. Kyiv: Naukova dumka.
14. Galyautdinov, R.R. (2013) *Transportnaya zadacha: metod Severo-Zapadnogo ugla* [Transport task: The method of the North-Western corner]. [Online] Available from: <https://galyautdinov.ru/post/metod-severo-zapadnogo-ugla> (Accessed: 22nd February 2024).
15. Kormen, T.H. et al. (2006) *Algoritmy: postroenie i analiz* [Algorithms: Construction and Analysis]. Translated from English. Moscow: Vil'yams.

Информация об авторах:

Цицишвили Гурами Шалвович – профессор, доктор физико-математических наук, главный научный сотрудник Института прикладной математики ДВО РАН (Владивосток, Россия). E-mail: guram@iam.dvo.ru.

Харченко Юрий Николаевич – кандидат физико-математических наук, научный сотрудник Института прикладной математики ДВО РАН (Владивосток, Россия). E-mail: har@iam.dvo.ru

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

Information about the authors:

Tsitsiashvili Gurami Sh. (Doctor of Physical and Mathematical Sciences, Professor, Main Researcher of the Institute of Applied Mathematics, the Far Eastern Branch of the Russian Academy of Sciences, Vladivostok, Russian Federation).

E-mail: guram@iam.dvo.ru

Kharchenko Yury N. (Candidate of Physical and Mathematical Sciences, Researcher of the Institute of Applied Mathematics, the Far Eastern Branch of the Russian Academy of Sciences, Vladivostok, Russian Federation). E-mail: har@iam.dvo.ru

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

Поступила в редакцию 12.04.2024; принята к публикации 03.09.2024

Received 12.04.2024; accepted for publication 03.09.2024

ОБРАБОТКА ИНФОРМАЦИИ

DATA PROCESSING

Научная статья

УДК 519.95

doi: 10.17223/19988605/68/7

**Кластерный анализ с применением обучения на основе отношений связанности
и плотности распределения****Николай Александрович Игнатьев¹, Екатерина Николаевна Згуральская²**¹ *Национальный университет Узбекистана им. Мирзо Улугбека, Ташкент, Узбекистан*² *Ульяновский государственный технический университет, Ульяновск, Россия*¹ *n_ignatev@rambler.ru*² *iatu@inbox.ru*

Аннотация. Рассматривается разбиение объектов на группы на основе отношений их связанности и плотности распределения. Плотность по локальной области в форме гипершара вычисляется относительно его центра в объекте выборки. Принадлежность объектов к непересекающимся классам определяется по объединению локальных областей с заданными ограничениями на значения плотности распределения. Определено условие выделения шумовых объектов классов из множества граничных. Проверка условия необходима для корректировки состава классов и группировки объектов по отношению их связанности по системе пересекающихся гипершаров. Обоснованы выделение эталонов минимального покрытия классов и их применение при оценке качества группировки. Для оценки формы конфигурации группы предложен показатель, значением которого является длина минимального кратчайшего пути между эталонами.

Ключевые слова: отношение связанности; оценки качества кластеризации; плотность распределения; шумовые объекты; эталоны минимального покрытия.

Для цитирования: Игнатьев Н.А., Згуральская Е.Н. Кластерный анализ с применением обучения на основе отношений связанности и плотности распределения // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 68. С. 66–74. doi: 10.17223/19988605/68/7

Original article

doi: 10.17223/19988605/68/7

Cluster analysis using learning based on connectivity relations and distribution density**Nikolay A. Ignatiev¹, Ekaterina N. Zguralskaya²**¹ *National University of Uzbekistan named after Mirzo Ulugbek, Tashkent, Republic of Uzbekistan*² *Ulyanovsk State Technical University, Ulyanovsk, Russian Federation*¹ *n_ignatev@rambler.ru*² *iatu@inbox.ru*

Abstract. The division of objects into groups based on the relationships of their connectivity and distribution density is considered. Density over a local area in the shape of a hyperball is calculated relative to its center in the sample object. The belonging of objects to disjoint classes is determined by combining local areas with specified restrictions on the distribution density values. The condition for selecting noise objects of classes from the set of boundary ones is determined. Checking the condition is necessary to adjust the composition of classes and grouping of objects in relation to their connectivity along a system of intersecting hyperballs. The identification of standards for minimum class coverage and their use in assessing the quality of the grouping are justified. To assess the shape

of the group configuration, an indicator is proposed, the value of which is the length of the minimum shortest path between the standards.

Keywords: connectedness relation; assessing the quality of clustering; distribution density; noise objects; minimum coverage standards.

For citation: Ignatiev, N.A., Zguralskaya, E.N. (2024) Cluster analysis using learning based on connectivity relations and distribution density. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 68. pp. 66–74. doi: 10.17223/19988605/68/7

Введение

В задачах группировки с частичным обучением считается, что для части объектов выборки классификация уже известна. Классификация может быть задана экспертами либо введена по результатам разведочного анализа данных. Оценки плотности распределения в локальных областях признакового пространства рассматриваются в качестве примера такого анализа.

Для количественной оценки качества кластеризации используют внешние и внутренние меры [1–3]. Внешние меры основаны на сравнении результата кластеризации с априори известным разделением на группы. Эти меры используют дополнительные знания о кластеризуемом множестве: распределение по кластерам, количество кластеров и т.д. Внутренние меры отображают качество кластеризации только по информации в данных.

Существующие меры оценки качества кластерного анализа не распространяются на результаты ряда алгоритмов. Если конфигурация кластеров не является сферической, то идея использования методов вычисления их центров является абсурдной. Центр может находиться за пределами конфигурации кластера. Например, вероятность такого события существует для результатов алгоритма кластеризации наборов многомерных данных на основе плотности распределения DBSCAN [2]. Алгоритм может обнаруживать кластеры любой формы и автоматически определять выбросы.

Имеются исследования с описанием методов для понижения чувствительности результатов от выбора параметров ε и k для DBSCAN, в которых ε представляет максимальный радиус гипершара с центром в объекте наблюдения, а k – ограничение на минимальное количество объектов выборки, содержащихся в таком гипершаре [4]. Несмотря на все попытки внесения модификаций не было предложено ни одного критерия оценки качества кластеризации по алгоритму DBSCAN.

Для устойчивого разбиения на кластеры в [5] предложен метод локального осреднения для вычисления плотности распределения данных как точек в метрическом пространстве. При выделении областей сгущения пространства применялось осреднение Гауссовского типа $P(x, x_i) = \exp\left(-\pi\left(\frac{\rho(x, x_i)}{R}\right)^2\right)$ по

радиусу R , которое инвариантно относительно поворотов (отображений, не меняющих взаимные расстояния), и суммарная вероятность вхождения точки x_i в пространство равна 1. Обоснованы условия выбора значений локальных радиусов осреднения и вхождения точек в шар единичного объема. Локальные радиусы могут различаться для разных кластеров. Перечислены причины, по которым алгоритм DBSCAN может как разъединить кластер, так и объединить в один несколько различных кластеров.

Существует потребность в разработке и использовании новых мер оценки качества кластеризации с учетом топологических свойств и плотности распределения объектов. В [6] проводились исследования топологических свойств по отношению связанности объектов классов по системе пересекающихся гипершаров. Условием связанности было наличие в пересечении граничных объектов классов. Шумовые объекты при машинном обучении рассматривались как подмножество граничных. Удаление шумовых объектов использовалось как средство цензурирования обучающих выборок с целью повышения обобщающей способности алгоритмов распознавания. Косвенной оценкой качества распознавания в порядковой шкале измерений служила мера компактности, значение которой зависело от числа эталонов минимального покрытия выборки, шумовых объектов и их составов.

Разработка новых мер оценки качества кластерного анализа расширяет возможности обнаружения скрытых закономерностей в рамках информационных моделей, основанных на знаниях. В качестве нового знания экспертам из предметных областей становится доступной информация о структуре кластеров со сложной конфигурацией.

1. Предмет исследования

При отсутствии каких-либо предположений о природе среды для анализа данных в локальных областях признакового пространства используют непараметрические методы оценки плотности распределения. В методах « k ближайших соседей» и «парzenовское окно» локальные области для вычисления плотности имеют сферическую форму. Геометрическое разнообразие форм конфигурации локальных областей с высокой плотностью распределения свойственно для результатов группировки по алгоритму DBSCAN [4]. Потребность в анализе разнообразия форм конфигураций послужила причиной разработки новых критериев (индексов) качества кластеризации на основе топологических свойств объектов классов.

Для анализа качества кластеризации с учетом топологии признакового пространства необходимо иметь данные о принадлежности хотя бы части объектов выборки к непересекающимся классам. Включение таких данных об объектах при кластерном анализе рассматривается как задание частичного обучения. Процесс выделения множества шумовых объектов классов интерпретируется в новом смысле. Шумовой объект – это кандидат на смену значения целевого признака класса или перехода из состава одной группы в другую. Представление о форме конфигурации группы формируется по результатам анализа структуры отношений между эталонами минимального покрытия классов.

Введение классификации на множестве не разделенных по категориям объектов рассматривается как предобработка данных с использованием эвристик. Выбор эвристик может быть связан с формой представления локальной области и вычислением значений плотности распределения в ней. Например, с представлением в виде гипершаров с фиксированным радиусом и центрами в каждом объекте выборки, вычисление значений плотности распределения в которых зависит как от числа объектов, так и от их расстояний от центра.

Классификация объектов на основе плотности их распределения связана с использованием метрик, методов нормирования данных, выбором размеров локальной области, правилами деления множества объектов на непересекающиеся подмножества. Такие термины и понятия, как граничный объект, меры компактности, эталоны минимального покрытия, информативные признаки, шумовые объекты, гипотеза о компактности, используемые при классификации, становятся востребованными и при оценке качества кластерного анализа.

Следующая последовательность операций над данными из табличного набора (выборки объектов) служит примером предобработки для введения классификации:

- вычисление плотности распределения объектов выборки;
- формирование упорядоченной по значениям плотности последовательности;
- разбиение последовательности на непересекающиеся подпоследовательности (классы).

Наличие классификации объектов позволяет:

- выделять шумовые объекты классов по множеству граничных;
- корректировать составы классов за счет шумовых объектов;
- производить разбиение объектов классов на непересекающиеся группы по отношению их связанности;
- определять эталоны минимального покрытия;
- вычислять принадлежность произвольного допустимого объекта к группе по мере его близости к эталонам минимального покрытия.

Условия принадлежности граничного объекта к множеству шумовых при кластерном анализе оставались неисследованными. Разъясним смысл, придаваемый понятиям шумового и граничного объектов при классификации, и покажем, как их можно адаптировать для кластерного анализа.

Разбиение объектов на непересекающиеся группы [6] с учетом топологических свойств признакового пространства производится в каждом классе по системе пересекающихся гипершаров. Количество групп вычисляется по отношению связанности, и их состав зависит от множества граничных объектов. Частичное представление о форме конфигурации группы дают эталоны минимального покрытия. Структура расположения эталонов зависит от множества граничных по заданной метрике объектов классов. Изменение состава эталонов и структуры отношений между ними связано с выделением шумовых объектов из множества граничных. Перенос шумовых объектов из одного класса в другой является причиной изменения числа и состава эталонов на выборке данных.

При включении граничного объекта [6] в состав шумовых использовались отношения расстояний до ближайших объектов из своего класса и его дополнения. Для получения оптимального с точки зрения точности распознавания решения вводились ограничения (пороги) на значения оценок отношений. Удаление шумовых объектов как части граничных в обучающей выборке приводило к увеличению отступа между классами и рассматривалось в качестве средства повышения обобщающей способности алгоритмов распознавания. Отступ является относительной величиной, которая влияет на вычисление значений параметров локальных метрик и состав эталонов минимального покрытия выборки.

В отличие от задач классификации при кластерном анализе шумовой объект не удаляется из выборки. Условия на отступы между граничными объектами классов определяют переход из одной группы в другую.

Разработка новых внутренних критериев для кластерного анализа и получения дополнительных сведений о форме конфигурации групп связана с условиями вычисления отступов между классами. Предлагается перечень дополнительных сведений (знаний) о конфигурации группы, которые можно получить по результатам кластерного анализа:

1. Число и описание эталонов группы.
2. Длина кратчайшего незамкнутого пути (КНП) между эталонами.
3. Число эталонов, являющихся граничными объектами группы.
4. Среднее число объектов группы, притягиваемое одним эталоном минимального покрытия.

2. Постановка задачи. Классификация объектов на основе плотности их распределения

Для описания множества (выборки) объектов $E_0 = \{S_1, \dots, S_m\}$ применяется набор разнотипных признаков $X(n) = (x_1, \dots, x_n)$. Определены процедуры для:

– вычисления плотности распределения $\beta_i = \varphi(S_i, \varepsilon)$ объекта $S_i \in E_0$ в локальной области в форме гипершара радиуса ε по метрике $\rho(x, y)$;

– деления объектов E_0 по множеству $\{\beta_i\}$ на непересекающиеся классы K_1, K_2 .

Объекты $S_u, S_v \in K_i, i = 1, 2$, считаются связанными между собой (обозначается $S_u \leftrightarrow S_v$), если

$$\{S \in B(E_0, \rho) | \rho(S, S_u) < r_u \text{ and } \rho(S, S_v) < r_v\} \neq \emptyset, \quad (1)$$

где r_u (r_v) – расстояние до ближайшего от S_u (S_v) объекта из K_{3-i} по метрике $\rho(x, y)$, $B(E_0, \rho) =$

$= \left\{ S \in E_0 \mid \rho(S_u, S) = \min_{S_u \in K_q, S_v \in K_{3-q}} \rho(S_u, S_v), u=1, \dots, m \right\}$ – множество граничных объектов классов, определяемое на E_0 .

Множество $B(E_0, \rho)$ используется для описания объектов E_0 в пространстве из бинарных признаков по таблице $T = \{t_{ij}\}_{m,d}, d = |B(E_0, \rho)|$. Значение элемента таблицы $t_{ij} = 1$, если выполняется условие (1) и $t_{ij} = 0$ в противном случае. По таблице T определяется минимальное число групп из связанных и несвязанных объектов классов. Составы групп не пересекаются и представлены объектами класса K_1 или K_2 . Для любых объектов $S_b, S_a \in K_i$, принадлежащих одной группе, всегда существует множество пар пересекающихся по $B(E_0, \rho)$ гипершаров для формирования цепочки $S_a \leftrightarrow S_t \leftrightarrow \dots \leftrightarrow S_b$.

Требуется:

– произвести разбиения объектов E_0 в каждом классе $K_t, t = 1, 2$, по отношению связанности (1) на непересекающиеся группы $J_{t1}, \dots, J_{t(l)}$ и определить объекты минимального покрытия эталонами;

– оценить качество разбиения на группы по классам и выборке в целом.

Пусть размер локальной области задан по значению радиуса гипершара ε . Значение плотности распределения по гипершару с центром в объекте $S \in E_0$ определим как

$$\varphi(S, \varepsilon) = \sum_{\rho(S, S_i) < \varepsilon} \left(1 - \frac{\rho(S, S_i)}{\varepsilon} \right). \quad (2)$$

Тогда значением математического ожидания по (2) будет

$$M(\varepsilon) = \sum_{i=1}^m \varphi(S_i, \varepsilon) / m. \quad (3)$$

По (2), (3) определим исходный (без корректировки) состав классов:

$$K_1 = \{S_j \mid \rho(S_k, S_j) < \varepsilon \text{ and } \varphi(S_k, \varepsilon) > M(\varepsilon)\} \text{ и } K_2 = E_0 \setminus K_1. \quad (4)$$

Корректировка состава классов по (4) связана с отбором шумовых объектов. Число групп по отношению связанности определяется алгоритмическим путем.

3. Отбор шумовых объектов из множества граничных

Решение о включении (не включении) граничного объекта в множество шумовых принимается по отношению двух ближайших от него расстояний до объектов из своего класса и его дополнения. Для анализа требуется определить условия, на основе которых принимается такое решение.

На множестве граничных объектов $B = B(E_0, \rho)$ сформируем множество пар $BG = \{(S_i, S_j)\}$, $S_i \in K_t \cap B$, $t = 1, 2$, $S_j \in CK_t \cap B$, $\rho(S_i, S_j) = \min_{S_v \in B \cap CK_t} \rho(S_i, S_v)$. Для $(S_i, S_j) \in BG$ введем обозначения

$r(S_i) = \rho(S_i, S_j)$, $d(S_i) = \rho(S_i, S_v)$, где $S_v = \arg \min_{S_a \in E_0 \cap K_t \setminus \{S_i\}} \rho(S_j, S_a)$. Аналогично для $S_j \in CK_t \cap B$ определим $r(S_j) = \rho(S_k, S_j) = \min_{S_v \in B \cap K_t} \rho(S_v, S_j)$, $d(S_j) = \rho(S_j, S_\mu)$, где $S_\mu = \arg \min_{S_a \in E_0 \cap CK_t \setminus \{S_j\}} \rho(S_k, S_a)$. Отноше-

ние $\frac{r(S_i)}{d(S_i)} < 1$ рассматривается как необходимое условие отнесения объекта $S_i \in K_t \cap B$ к множеству

шумовых. Достаточным условием являются неравенства

$$\frac{r(S_i)}{d(S_i)} < 1 \quad \text{и} \quad \frac{r(S_j)}{d(S_j)} \geq 1. \quad (5)$$

Иллюстрация определения принадлежности граничного объекта $S_i \in K_1$ к множеству шумовых на выборке $E_0 = K_1 \cup K_2$ по отношениям расстояний $r(S_i)$, $d(S_j)$, $d(S_i)$ показана на рис. 1.

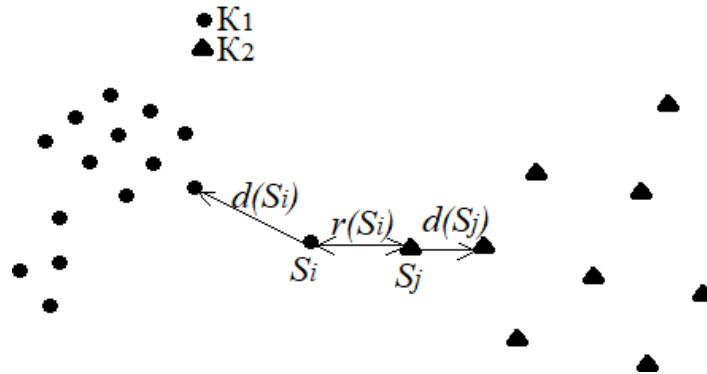


Рис. 1. Отношение расстояний $r(S_i)$, $d(S_j)$, $d(S_i)$ для определения принадлежности граничного объекта $S_i \in K_1$ к множеству шумовых

Fig. 1. Ratio of distances $r(S_i)$, $d(S_j)$, $d(S_i)$ to determine whether the boundary object $S_i \in K_1$ belongs to the set of noise

Формирование многообразия структур отношений объектов зависит от используемых мер расстояния, способов нормирования, составов наборов признаков и т.д. Как правило, выводы на основе оценок качества кластерного анализа делаются по результатам вычислительных экспериментов.

С изменением структуры отношений внутри групп и между группами связан процесс выделения шумовых объектов и перехода из состава одной группы в другую. В рамках поставленной в работе задачи результаты отнесения объекта к множеству шумовых зависят от значения параметра ϵ для вычисления плотности распределения по (2) и классификации по (4).

Выделение шумовых объектов по (5) используется для коррекции состава множества граничных объектов классов. Коррекция граничных объектов приводит к изменению отношений связанности объектов при разбиении их на группы. Меняются число и состав эталонов минимального покрытия. Исследовать структуру отношений внутри группы можно через КНП между эталонами минимального покрытия и расстояния объектов до них. Анализ расстояний до эталонов и их интерпретация для оценки качества кластерного анализа в данной работе не рассматриваются. Значения КНП вычисляются при числе эталонов больше или равном 2. Источниками информации для обнаружения скрытых закономерностей по конфигурации группы являются:

- пересечение множества эталонов и граничных объектов;
- соотношение между числом эталонов и числом граничных объектов;
- среднее число объектов группы, притягиваемое одним эталоном минимального покрытия или компактность группы.

По результатам анализа эталонов и граничных объектов из групп можно получить представление об их конфигурациях. Например, как это показано на рис. 2.

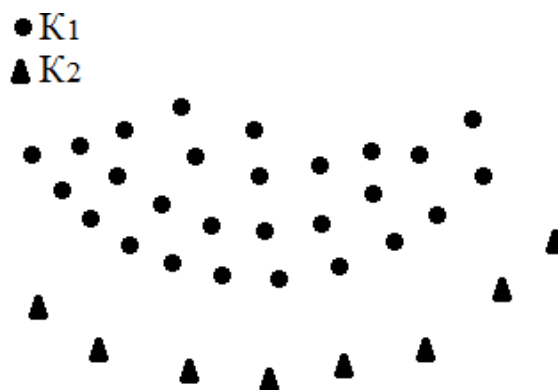


Рис. 2. Визуальное представление о конфигурациях групп
Fig. 2. Visual representation of group configurations

Визуальное представление на рис. 2 группы из класса K_2 можно интерпретировать следующим образом:

- все эталоны являются граничными объектами;
- среднее число объектов группы, притягиваемое одним эталоном минимального покрытия, меньше, чем у группы из K_1 .

4. Внешняя мера оценки качества кластеризации

Пусть $J_{11}, \dots, J_{1u}$ ($J_{11} \cup \dots \cup J_{1u} = K_1$) и $J_{21}, \dots, J_{2v}$ ($J_{21} \cup \dots \cup J_{2v} = K_2$) – множества непересекающихся групп, полученных по отношению связанности объектов (1) по классам K_1 и K_2 . Каждая группа определяется по непересекающемуся подмножеству граничных объектов $B(E_0, \rho)$, принадлежащих одному из классов: K_1 или K_2 . Мера компактности кластеров по классам и выборки в целом [7] используется для оценки качества кластеризации и определяется как

$$\Omega(K_1) = \frac{|J_{11}|^2 + \dots + |J_{1u}|^2}{|K_1|^2}, \quad (6)$$

$$\Omega(K_2) = \frac{|J_{21}|^2 + \dots + |J_{2v}|^2}{|K_2|^2}, \quad (7)$$

$$\Omega(E_0) = (|K_1|\Omega(K_1) + |K_2|\Omega(K_2))/m. \quad (8)$$

Множество допустимых значений (8) в $[\kappa; 1]$, $\kappa > 0$, позволяет интерпретировать результаты группировки с помощью лингвистических переменных.

5. Вычислительный эксперимент

В качестве материала для исследования использовалась выборка данных из [8] с общим числом объектов 351, описываемых 34 количественными признаками. Размер локальной области определялся по радиусу гипершара, равному среднему расстоянию до k ближайших соседей от объектов выборки данных $\varepsilon = \varepsilon(k)$. Результаты разбиения на группы при разных значениях $\varepsilon(k)$ в исходном признаковом пространстве приводятся в табл. 1. В скобках указано число шумовых объектов, включенных в состав класса.

Таблица 1

Мера компактности классов по метрике Евклида в исходном признаковом пространстве

Радиус $\varepsilon(k)$	Класс K_1		Класс K_2	
	Число объектов	Компактность по (6)	Число объектов	Компактность по (7)
$\varepsilon(2)$	143 (1)	1,0000	208 (2)	0,6380
$\varepsilon(3)$	157 (4)	0,9749	194 (5)	0,5912
$\varepsilon(7)$	183 (0)	1,0000	168 (1)	0,5192
$\varepsilon(10)$	189 (0)	1,0000	162 (2)	0,2639

Анализ структуры по двум самым большим по числу объектов группам из классов K_1 и K_2 приводится в табл. 2. В скобках указано число эталонов минимального покрытия.

Таблица 2

Анализ структуры кластеров по метрике Евклида в исходном признаковом пространстве

Радиус $\varepsilon(k)$	Класс K_1		Класс K_2	
	Число объектов	Длина КНП	Число объектов	Длина КНП
$\varepsilon(2)$	165 (2)	7,1717	144 (5)	7,5729
$\varepsilon(3)$	156 (5)	9,9618	148 (2)	7,3027
$\varepsilon(7)$	184 (3)	4,6596	120 (5)	24,0466
$\varepsilon(10)$	191 (3)	5,1324	81 (2)	7,5759

По аналогии с табл. 1, 2 результаты в нормированном в $[0; 1]$ признаковом пространстве по метрике Хэмминга представлены в табл. 3, 4.

Таблица 3

Мера компактности классов по метрике Хемминга в нормированном признаковом пространстве

Радиус $\varepsilon(k)$	Класс K_1		Класс K_2	
	Число объектов	Компактность по (6)	Число объектов	Компактность по (7)
$\varepsilon(2)$	161 (0)	1,0000	190 (1)	0,8677
$\varepsilon(3)$	171 (0)	1,0000	180 (1)	0,5437
$\varepsilon(7)$	189 (0)	1,0000	162 (1)	0,7467
$\varepsilon(10)$	186 (0)	1,0000	165 (1)	0,8709

Таблица 4

Анализ структуры кластеров по метрике Хемминга в нормированном признаковом пространстве

Радиус $\varepsilon(k)$	Класс K_1		Класс K_2	
	Число объектов	Длина КНП	Число объектов	Длина КНП
$\varepsilon(2)$	162 (10)	35,7786	176 (10)	107,6708
$\varepsilon(3)$	172 (7)	32,8450	131 (5)	54,8325
$\varepsilon(7)$	190 (6)	28,8066	139 (7)	79,4127
$\varepsilon(10)$	187 (5)	16,6060	153 (8)	85,1761

Выводы, следующие из анализа результатов, представленных в табл. 1–4:

1. Для групп объектов с высокой плотностью распределения (класс K_1) свойственны более высокие значения меры компактности, чем для групп с низкой плотностью (класс K_2).
2. Нормирование данных (см. табл. 4) привело к резкому различию конфигураций групп по длине КНП с высокой (класс K_1) и низкой (класс K_2) плотностями распределений.

Заключение

В статье предложена новая методика оценки качества кластеризации при интеллектуальном анализе данных групп объектов со сложной конфигурацией. Методика ориентирована на исследование топологических свойств объектов через их классификацию в локальных областях многомерных данных с различными плотностями распределений. Численная оценка топологических свойств выражается через значения меры компактности групп по отношению связанности объектов классов. Отношение связанности применялось для поиска минимального покрытия выборки эталонами. Длина КНП между эталонами группы служит информацией для формирования представления о ее (группе) конфигурации.

Список источников

1. Kleinberg J. An Impossibility Theorem for Clustering. URL: <https://www.cs.cornell.edu/home/kleinber/nips15.pdf> (accessed: 02.06.2024).
2. Сивоголовко Е.В. Методы оценки качества четкой кластеризации // Компьютерные инструменты в образовании. 2011. № 4. С. 14–31.
3. Шмидт Ю.Д., Лободина О.Н., Грищенко В.В., Нинбо Ц. Об оценке уровня инновационного развития регионов России // Региональная экономика: теория и практика. 2017. Т. 15, № 2. С. 342–354.
4. Shili H. Clustering in big data analytics: a systematic review and comparative analysis (review article) // Scientific and Technical Journal of Information Technologies, Mechanics and Optics. 2023. V. 23 (5). P. 967–979. doi: 10.17586/2226-1494-2023-23-5-967-979
5. Айдагулов Р.Р., Главатский С.Т., Михалёв А.В. Методы осреднения в задачах кластеризации больших данных // Интеллектуальные системы. Теория и приложения. 2021. № 25 (4). С. 12–18.
6. Игнатьев Н.А., Турсунмуротов Д.Х. Цензурирование обучающих выборок с использованием регуляризации отношений связанности объектов классов // Научно-технический вестник информационных технологий, механики и оптики. 2024. Т. 24 (2). С. 322–329. doi: 10.17586/2226-1494-2024-24-2-322-329
7. Ignatyev N.A. Structure Choice for Relations between Objects in Metric Classification Algorithms // Pattern Recognition and Image Analysis. 2018. № 4. P. 590–597.
8. UCI repository of machine learning databases. Ionosphere. URL: <https://archive.ics.uci.edu/dataset/52/ionosphere> (accessed: 02.06.2024).

References

1. Kleinberg, J. (n.d.) *An Impossibility Theorem for Clustering*. [Online] Available from: <https://www.cs.cornell.edu/home/kleinber/nips15.pdf> (Accessed: 2nd June 2024).
2. Sivogolovko, E.V. (2011) *Metody otsenki kachestva chetkoy klasterizatsii* [Methods for assessing the quality of clear clustering]. *Komp'yuternye instrumenty v obrazovanii*. 4. pp. 14–31.
3. Shmidt, Yu.D., Lobodina, O.N., Grishchenko, V.V. & Ninbo, C. (2017) *Ob otsenke urovnya innovatsionnogo razvitiya regionov Rossii* [On assessing the level of innovative development of Russian regions]. *Regional'naya ekonomika: teoriya i praktika*. 15(2). pp. 342–354.
4. Shili, H. (2023) Clustering in big data analytics: a systematic review and comparative analysis (review article). *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*. 23(5). pp. 967–979. DOI: 10.17586/2226-1494-2023-23-5-967-979
5. Aydagulov, R.R., Glavatskiy, S.T. & Mikhalev, A.V. (2021) *Metody osredneniya v zadachakh klasterizatsii bol'shikh dannykh* [Averaging methods in big data clustering problems]. *Intellektual'nye sistemy. Teoriya i prilozheniya*. 25(4). pp. 12–18.
6. Ignatiev, N.A. & Tursunmurotov, D.Kh. (2024) *Tsenzurovanie obuchayushchikh vyborok s ispol'zovaniem regulyarizatsii otoshe-niy svyazannosti ob'ektov klassov* [Censoring training samples using regularization of relatedness relations of class objects]. *Nauchno-tekhnicheskiy vestnik informatsionnykh tekhnologiy, mekhaniki i optiki*. 24(2). pp. 322–329. DOI: 10.17586/2226-1494-2024-24-2-322-329
7. Ignatyev, N.A. (2018) Structure Choice for Relations between Objects in Metric Classification Algorithms. *Pattern Recognition and Image Analysis*. 4. pp. 590–597.

8. *UCI repository of machine learning databases. Ionosphere.* [Online] Available from: <https://archive.ics.uci.edu/dataset/52/ionosphere> (Accessed: 2nd June 2024).

Информация об авторах:

Игнатьев Николай Александрович – профессор, доктор физико-математических наук, профессор кафедры «Искусственный интеллект» Национального университета Узбекистана им. М. Улугбека (Ташкент, Узбекистан). E-mail: n_ignatev@rambler.ru

Згуральская Екатерина Николаевна – кандидат технических наук, доцент кафедры «Информационные технологии и общенаучные дисциплины» Ульяновского государственного технического университета (Ульяновск, Россия). E-mail: iatu@inbox.ru

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

Information about the authors:

Ignatev Nikolay A. (Professor, Doctor of Physical and Mathematical Sciences, Professor of the Department of Artificial Intelligence at the National University of Uzbekistan named after M. Ulugbek, Tashkent, Uzbekistan). E-mail: n_ignatev@rambler.ru

Zguralskaya Ekaterina N. (Candidate of Technical Sciences, Associate Professor of the Department of Information Technologies and General Scientific Disciplines of Ulyanovsk State Technical University (Ulyanovsk, Russian Federation). E-mail: iatu@inbox.ru

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

Поступила в редакцию 19.06.2024; принята к публикации 03.09.2024

Received 19.06.2024; accepted for publication 03.09.2024

Научная статья

УДК 519.7

doi: 10.17223/19988605/68/8

Сегментация дефектов дорожного полотна на основе нейросетевого ансамбля**Ирина Александровна Канаева¹, Владимир Григорьевич Спицын²**^{1,2} *Национальный исследовательский Томский политехнический университет, Томск, Россия*² *Национальный исследовательский Томский государственный университет, Томск, Россия*¹ *iap15@tpu.ru*² *spvg@tpu.ru*

Аннотация. Своевременная диагностика качества дорожного покрытия и обнаружение дефектов являются немаловажным этапом обеспечения безопасности на дороге. Цель данной работы – разработка набора данных и нейросетевой модели для сегментации разрушений на изображениях автомобильных дорог. Собирается и аннотирован на уровне пикселей набор данных SegmRDD, содержащий 4 420 изображений с дефектами трех классов: «трещины», «сетка трещин», «выбоины». Набор данных сбалансирован и охватывает дорожную обстановку пяти стран, включая Россию. Разработан ансамбль на основе трех независимо обученных нейросетевых моделей YOLOv8, U-Net, Mask R-CNN и оригинального алгоритма объединения результатов, позволяющий получить агрегированный показатель полноты и точности F_1 , равный 70%.

Ключевые слова: семантическая сегментация; обнаружение дефектов покрытия; сверточные нейронные сети; нейросетевой ансамбль.

Для цитирования: Канаева И.А., Спицын В.Г. Сегментация дефектов дорожного полотна на основе нейросетевого ансамбля // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 68. С. 75–85. doi: 10.17223/19988605/68/8

Original article

doi: 10.17223/19988605/68/8

Road damage defects segmentation based on convolutional neural network ensemble**Irina A. Kanaeva¹, Vladimir G. Spitsyn²**^{1,2} *National Research Tomsk Polytechnic University, Tomsk, Russian Federation*² *National Research Tomsk State University, Tomsk, Russian Federation*¹ *iap15@tpu.ru*² *spvg@tpu.ru*

Abstract. Early monitoring road conditions and defect detection are an important step in ensuring road safety. The work presents a new road damage segmentation dataset SegmRDD. It contains 4420 images with defects of three classes "cracks", "alligator crack", "potholes" well annotated at the pixel level. The dataset is balanced and covers the roads of five countries, including Russia. Developed ensemble model based on three parallel-trained neural network models YOLOv8, U-Net, Mask R-CNN with combining results, and achieved an F1-score of 70% for all defects.

Keywords: semantic segmentation; road damage detection; convolutional neural networks; neural network ensemble.

For citation: Kanaeva, I.A., Spitsyn, V.G. (2024) Road damage defects segmentation based on convolutional neural network ensemble. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 68. pp. 75–85. doi: 10.17223/19988605/68/8

Введение

На начало 2024 г. протяженность дорог в России составляет более 1,6 млн км. Высокая интенсивность трафика и сложные погодные условия часто приводят к деформации и частичному разрушению асфальтобетонных покрытий. Для своевременного мониторинга состояния дорог и оценки качества дорожного покрытия используются передвижные дорожные лаборатории, оснащенные системой видеокамер, гиринерциальной системой позиционирования, приемником GPS, системой измерения ровности и расстояния, а также другими датчиками. С помощью системы видеокамер осуществляется видеосъемка на всей протяженности обследуемого участка. В настоящий момент дальнейшая обработка данных, полученных в ходе диагностики, практически полностью не автоматизирована. Специалист вручную отмечает на видео все необходимые для анализа объекты, такие как линии разметки, кромка дорожного полотна, дефекты дорожного покрытия. Трудоемкость процесса выделения дефектов, а также бурное развитие современных методов компьютерного зрения позволяют судить об актуальности автоматизации процесса обнаружения дефектов.

За последнее десятилетие учеными из разных стран предпринят ряд попыток разработать метод анализа качества дорожного покрытия по изображению или видеозаписи камеры на основе технологий обработки изображений для более эффективной диагностики дорожного покрытия. Исторически вначале исследовались изображения, полученные непосредственно сверху, над дорогой, и использовались методы фильтрации для отделения трещин от основной текстуры покрытия [1, 2].

Обширные исследования проводились в области построения различных классификаторов для определения того, содержит ли фрагмент покрытия дороги определенный тип разрушения. Для этого применялись различные методы машинного обучения: нейронные сети [3], машины опорных векторов [4], градиентный бустинг [5], деревья решений [6], вероятностные генеративные модели [7].

Успехи в разработке эффективных архитектур глубоких сверточных нейронных сетей (СНС) в 2016 г. привлекли внимание ученых, и была продемонстрирована возможность их успешного применения к данной задаче [8, 9]. В 2018 г. на соревновании Road Damage Detection Challenge был предоставлен в открытом доступе набор данных для детектирования дефектов покрытия RDD2018, который описан в работе японского ученого Н. Maeda и соавт. [10] и используется для обучения глубоких СНС. Данное событие подняло интерес к детектированию различных дефектов по общему изображению, снятому из машины. Лучшее решение задачи детекции 4-х классов на расширенном наборе данных RDD2022 достигло оценки $F_1 = 77\%$ и основано на ансамбле из моделей YOLOv5, YOLOv7, Faster R-CNN и SWIN трансформер [11].

Переходом от задачи выделения дефектов ограничивающей рамкой к выделению пиксельной маски разрушения является работа исследователей из Новой Зеландии [12], которые вручную аннотировали по 143 изображения из четырех наборов данных, включая RDD2018, для сегментации выбоин на дороге. В результате авторы обучили по полученным данным модель Mask R-CNN и получили высокую точность на тестовом наборе.

На данный момент задача детектирования дефектов по изображениям из разных стран с различными ракурсами съемки является наиболее актуальной. Однако постепенно происходит переход от детекции к сегментации дефектов. Так, в конце 2023 г. китайскими исследователями был продемонстрирован первый бенчмарк в области интеллектуальной дорожной инспекции UDTIRI [13]. Бенчмарк представляет собой хорошо аннотированный на уровне пикселей набор данных из 1 000 изображений выбоин, который можно использовать для обучения и оценки сетей обнаружения объектов, семантической сегментации и сегментации экземпляров. Еще одним перспективным направлением в сфере дорожной инспекции является обработка 3D-изображений, полученных с помощью лазерного сканирования или съемки стереопары [14].

Анализ имеющихся публично открытых наборов данных показал, что сегодня не существует единой выборки для решения задачи сегментации основных типов дефектов покрытия (трещин и выбоин). Цель данной работы – разработка набора данных и нейросетевой модели для сегментации дефектов дорожного полотна на изображениях автомобильных дорог. Получены следующие результаты:

представлен набор данных SegmRDD, содержащий 4 420 изображений автомобильных дорог с дефектами трех классов: трещинами, сеткой трещин и выбоинами. Набор данных сбалансирован и охватывает дорожную обстановку Бразилии, Канады, Чехии, Японии и России. Разработанный ансамбль на основе трех независимо обученных нейросетевых моделей YOLOv8, U-Net, Mask R-CNN и оригинального алгоритма объединения результатов позволил получить точность сегментации по агрегированному критерию качества F_1 , равную 70%.

1. Накопленный опыт мониторинга состояния дорожного покрытия

1.1. Статистический анализ дорожно-транспортных происшествий в России

Выбоина – это местное разрушение дорожного покрытия, имеющее вид углубления с резко очерченными краями. Одной из причин образования выбоин является развитие одиночных трещин или сетки трещин. Выбоины на дороге способны привести к авариям и повреждениям автомобилей, так как они могут вызывать внезапные рывки в управлении автомобилем, особенно на высоких скоростях. По статистике Госавтоинспекции, в дорожно-транспортных происшествиях (ДТП) в России за 2022 г. погибли свыше 14 тыс. человек и пострадали более 160 тыс. Более 75% всех ДТП происходит в городах, при этом основной или сопутствующей причиной каждого третьего происшествия является неудовлетворительное состояние дорог и улиц. В результате наличия данногоотягчающего фактора только в 2022 г. погибли свыше 4 тыс. человек (рис. 1).

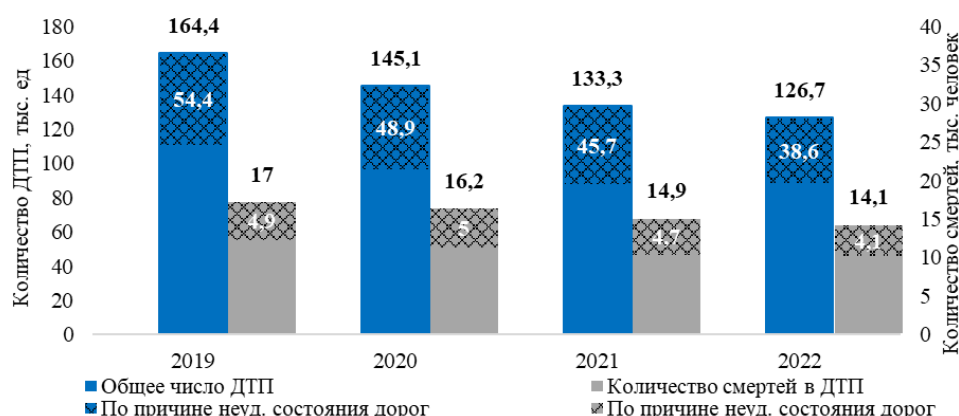


Рис. 1. Статистика ДТП в России по причине неудовлетворительного состояния улиц и дорог

Fig. 1. Russian road accident statistics due to poor road quality surfaces

С 2019 г. в Российской Федерации реализуется национальный проект «Безопасные качественные дороги», цель которого – повышение качества дорог и улучшение доступности инфраструктуры. Реализация данного проекта позволила значительно сократить общую смертность на дорогах: –17% в 2022 г. относительно 2019 г., а показатель смертности на 100 тыс. населения составил 9,8. Данная статистика еще раз подчеркивает важность своевременной диагностики и выявления дефектов дорожного покрытия.

Трещины покрытия дороги не отображены в ГОСТ Р 50597–2017 как вид дефекта, так как сами по себе не несут опасности для участников движения. Однако важность обнаружения именно трещин на ранних стадиях подтверждается тем, что они служат «первоосновой» более сложных дефектов и требуют меньших затрат на ремонт.

1.2. Классификация существующих методов детектирования дефектов

Наиболее распространенным способом выявления дефектов является подход на основе анализа изображений автомобильных дорог благодаря невысоким затратам по получению таких данных и широкому спектру различных технологий компьютерного зрения. Обновленная классификация

накопленного опыта в области методов обнаружения дефектов покрытия по изображению представлена в табл. 1 [15].

Таблица 1

Классификация накопленного опыта

Класс выделяемых повреждений	Тип ракурса	Тип решаемой задачи	Сложность применяемой модели
1. Трещины (одиночные, сетка трещин). 2. Разрушения (ямы, выбоины, проломы). 3. Износ дорожной разметки. 4. Выпотевание вяжущего	1. Вид сверху. 2. Вид водителя. 3. Общий вид. 4. Вид камеры заднего вида	1. Классификация. 2. Детектирование. 3. Сегментация	1. Обработка изображений. 2. Машинное обучение. 3. Глубокое обучение

По типу ракурса исторически наиболее развито направление обработки изображений, полученных непосредственно сверху над поверхностью автомобильной дороги. Однако сам процесс получения таких изображений сопряжен с трудностями размещения камер для полного охвата проезжей части, освещения поверхности и высокого разрешения картинки, часто это сопряжено с невысокой скоростью езды и загрязнением камеры. Расположение камеры спереди транспортного средства позволяет получать полную информацию о дорожной ситуации, однако дорожное полотно удалено и часто перекрыто капотом в нижней части изображения.

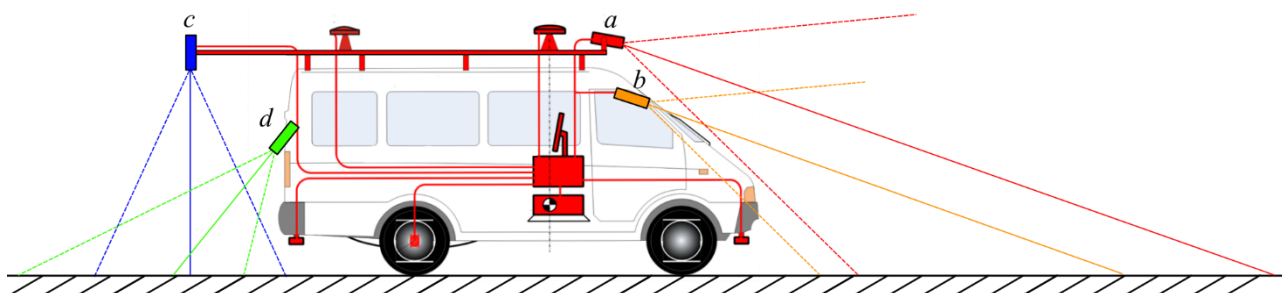


Рис. 2. Способы расположения камеры на дорожной лаборатории для ракурсов вида:
a – общий, b – водителя, c – сверху, d – камеры заднего вида

Fig. 2. Representation of the camera placement used by vehicle platform for different views:
a) street view, b) wide view, c) top-down view, d) rear camera view

На рис. 2 схематично представлены используемые способы расположения камер для регистрации дорожного покрытия.

1.3 Существующие наборы данных

Несмотря на многообразие представленных методов детектирования дефектов автомобильных дорог, в настоящее время количество общедоступных наборов данных для сегментации дефектов невелико. Так как исторически вначале развивались подходы, основанные на применении фильтров для выделения трещин на покрытии, исследователи создавали наборы изображений, снятые непосредственно над покрытием, сверху. Во многих публикациях используются собственные закрытые наборы данных, что затрудняет воспроизводимость и сопоставимость показанных в работах результатов. В табл. 2 представлено сравнение наиболее популярных и открытых наборов данных для решения задач выделения дефектов дорожного полотна.

Таблица 2

Свободно-доступные наборы данных для обнаружения дефектов дорожного покрытия

Наименование и место сбора	Количество, штук	Размер, пиксель	Тип ракурса	Классы дефектов
Классификация				
FHWA/LTPP [9] (США и Канада)	1 056	2 048 × 3 072	Вид сверху	2 (неповрежденная дорога, трещина)

Наименование и место сбора	Количество, штук	Размер, пиксель	Тип ракурса	Классы дефектов
<i>Детектирование</i>				
GAPs v2 [16] (Германия)	2 468	1 920 × 1 080	Вид сверху	6 (неповрежденная дорога, трещина, заплатка, заплатка с пропиткой, выбоина, открытый шов покрытия)
Pavement Image Datasets (PID) [17] (США)	7 237	640 × 640	Вид водителя и полученный вид сверху	10 (8 типов трещин, выбоины, выкрашивания)
CQU-BPDD [18] (Китай)	60 059	1 200 × 900	Вид сверху	6 (неповрежденная дорога, 5 типов трещин)
RDD2018 [10] (Япония)	9 053	600 × 600	Вид водителя	8 (5 типов трещин, выбоины, износ разметки пешеходного перехода, износ разметки линий)
RDD2019 [19] (Япония)	13 135	600 × 600	Вид водителя	9 (те же в RDD2018 + ревизионный люк)
RDD2020 [20] (Япония, Индия, Чехия)	26 620	600 × 600, 720 × 720	Вид водителя	4 (продольная трещина, поперечная трещина, сетка трещин, выбоина)
RDD2022 [21] (Япония, Индия, Чехия, Норвегия, США, Китай)	47 420	512 × 512, 600 × 600, 720 × 720, 3 650 × 2 044	Вид водителя, вид сверху, общий вид	4 (продольная трещина, поперечная трещина, сетка трещин, выбоина)
SVRDD [22] (Китай)	8 000 (6 000–1 000–1 000)	1 024 × 1 024	Общий вид	7 (продольная трещина, поперечная трещина, сетка трещин, выбоина, продольная заплатка, поперечная заплатка и ревизионный люк)
<i>Сегментация</i>				
CRACKTree200 [1] (Китай)	206	800 × 600	Вид сверху	2 (неповрежденная дорога, трещина)
CFD [23] (Китай)	118	480 × 320	Вид сверху	2 (неповрежденная дорога, трещина)
EdmCrack600 [24] (Канада)	600	1 920 × 1 080	Вид камеры заднего вида	2 (фон, трещина)
DNIT [25] (Бразилия)	2 235	1 024 × 640, 1 024 × 630	Вид водителя	4 (трещина, выбоина, полоса движения, фон)
Farzad Road Damage [26] (Интернет)	780 (720–60)	640 × 640	Различный	2 (фон, выбоина)
UDTIRI [13] (Интернет)	1 000 (600–100–300)	Различный	Различный	2 (фон, выбоина)

Ввиду различных топологий дефектов, в частности диагональных трещин, прямоугольная область выделения может захватывать множество лишней информации, что вносит трудности в определение площади поврежденного покрытия. Для целей оценки качества автомобильной дороги наилучшим вариантом является выделение дефектов с помощью пиксельной маски, однако создание такого набора данных является трудозатратным.

2. Создание набора данных для задачи сегментации дефектов

Итоговый набор данных SegmRDD для обнаружения дефектов основан на разметке четырех свободно распространяемых наборов данных и двух новых, собранных авторами:

1. DNIT – исходная аннотация на уровне пикселей была переведена в полигональные сегментационные маски, класс Crack был разделен на классы «трещины» и «сетка трещин»; применен фильтр Гаусса к полосе встречного движения для маскировки неразмеченных дефектов.

2. EdmCrack600 – разметка трещин переведена в полигональные сегментационные маски, размечены дополнительно встречающиеся выбоины.

3. UDTIRI – отобранные изображения, содержащие только класс «выбоины».

4. RDD2020 – отобраны изображения из коллекций Japan и Czech, произведена разметка на уровне пикселей для трех классов дефектов.
5. RusRDD – размечены изображения российских дорог.
6. SynRDD – синтетические изображения для классов «трещины» и «выбоины», генерация которых описана в работе [15].

На рис. 3 представлены распределения классов в выборках, описанных выше.

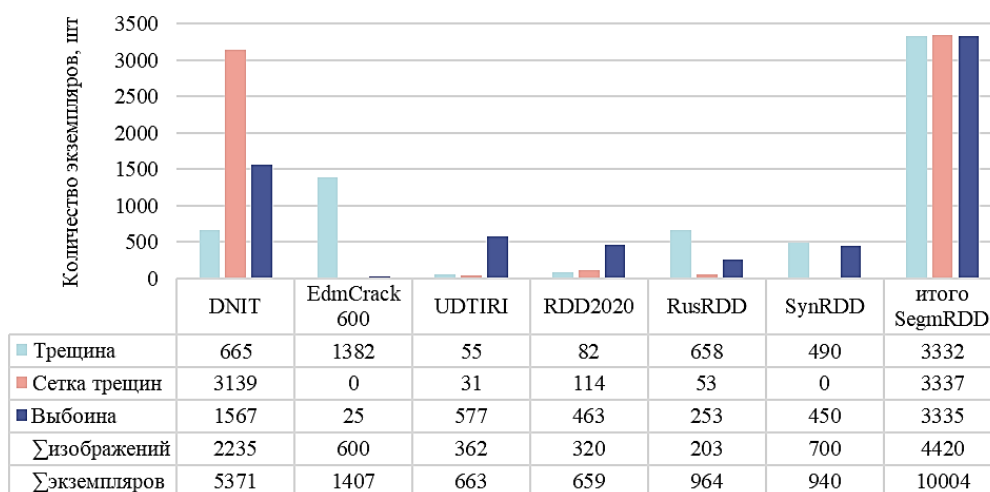


Рис. 3. Распределение классов в разработанном наборе данных для сегментации дефектов SegmRDD
Fig. 3. Statistic of the number of damage instances included in the presented segmentation dataset SegmRDD

Полученный набор данных состоит из 4 420 изображений дорог из пяти стран: Бразилии, Канады, Чехии, Японии, России. Коллекция RDD2020 включает в себя также дорожную обстановку Индии, которая ввиду значительных отличий по сравнению с российскими дорогами не рассматривалась. В работе авторов положительно оценивается применимость японской модели RDD2020 для других стран только в случае использования модели как базовой и обязательного дополнения изображениями из самой страны применения. Авторы рекомендуют смешивать коллекции дефектов покрытия разных стран для увеличения количества данных, предотвращения переобучения нейросетевых моделей и увеличения их обобщаемости.

Благодаря включению изображений SynRDD со сгенерированными трещинами и выбоинами полученный набор данных сбалансирован и имеет примерно по 3 330 экземпляров дефектов каждого класса.

3. Исследование и разработка нейросетевых моделей обнаружения дефектов

3.1. Модель детектирования дефектов

В 2024 г. китайские ученые представили открытый набор данных общего вида SVRDD для детектирования таких видов разрушений, как продольная трещина, поперечная трещина, сетка трещин, выбоина, продольная заплатка, поперечная заплатка и ревизионный люк. Набор данных содержит 8 000 изображений $1\,024 \times 1\,024$ пикселей.

Модель YOLO является одноступенчатой архитектурой, которая обрабатывает изображение в исходном масштабе за один проход, рассматривая задачу детектирования как задачу регрессии ограничивающих рамок и классов объектов. Архитектура YOLO подразделяется на CHC для вычисления признаков изображения, называемую backbone, подсеть для обработки признаков neck и head – объединение частей, отвечающих за предсказание ограничивающей рамки. В [27] показано, что модель YOLOv8 позволяет достигать самых современных результатов в различных тестах обнаружения объектов и семантической сегментации (YOLOv8-seg) при сохранении высокой скорости и эффективности. Авторы фреймворка представили пять конфигураций модели YOLOv8, по-разному опти-

мизированных для улучшения скорости и точности и предварительно обученных на наборе данных COCO.

Для обучения нейросетевой модели YOLOv8 все изображения были масштабированы до размера входа сети 640×640 пикселей. Для ускорения обучения и улучшения точности модели часто применяют передачу обучения (transfer learning), заключающуюся в адаптации существующей обученной нейронной сети в разрабатываемую архитектуру. Для повышения производительности модели во время обучения YOLOv8 использует мозаичную аугментацию: объединяет четыре изображения в одно для изучения контекстов объектов в разных местах и на разном фоне. Обучение нейросети производилось в течение 90 эпох с помощью метода AdamW с начальной скоростью обучения 0,0001 и конечной 0,01, размер пакета равен 16. Оценка результатов лучшей модели YOLOv8s на проверочной и тестовой выборках по 1 000 изображений представлена в табл. 3.

Таблица 3

Точность детектирования YOLOv8s на выборке SVRDD

Класс	Валидационная выборка, (%)				Тестовая выборка, (%)			
	<i>P</i>	<i>R</i>	<i>mAP</i> _{0,5}	<i>mAP</i> _{0,5:0,95}	<i>P</i>	<i>R</i>	<i>mAP</i> _{0,5}	<i>mAP</i> _{0,5:0,95}
Все	0,755	0,621	0,693	0,435	0,760	0,617	0,690	0,433
Продольная трещина	0,717	0,638	0,686	0,444	0,697	0,630	0,664	0,438
Поперечная трещина	0,714	0,541	0,628	0,354	0,710	0,573	0,627	0,358
Сетка трещин	0,744	0,716	0,737	0,479	0,745	0,725	0,763	0,485
Выбоина	0,667	0,461	0,553	0,322	0,745	0,452	0,552	0,308
Ревизионный люк	0,853	0,715	0,796	0,498	0,849	0,657	0,766	0,481
Продольная заплатка	0,810	0,716	0,774	0,532	0,804	0,706	0,766	0,539
Поперечная заплатка	0,782	0,559	0,677	0,418	0,768	0,576	0,689	0,420

Полученные результаты $F_1 = 0,681$ сопоставимы с полученными лучшими результатами в [22] путем обучения нейросетей YOLOv5 ($F_1 = 0,709$; $mAP_{0,5:0,95} = 0,417$) и YOLOX ($F_1 = 0,691$; $mAP_{0,5:0,95} = 0,420$) и превосходят их по метрике $mAP_{0,5:0,95}$, обеспечивающей более полную оценку, так как данная метрика учитывает более широкий диапазон пороговых значений *IoU*, фиксируя как высокое, так и низкое перекрытие между предсказанными и реальными объектами.

Применение обученной модели к данным российских дорог показало точность 33%, что подтверждает необходимость разработки комплексного решения на основе нейросетевого ансамбля.

3.2. Модель сегментации дефектов

Анализ существующих датасетов показал, что использование накопленного массива данных для детектирования дефектов позволит уменьшить дисперсию модели сегментации. Обучение разных моделей на различных наборах данных и агрегация результатов часто работают лучше в сравнении с отдельной нейросетевой моделью [28]. Для построения нейросетевого ансамбля параллельным методом первым шагом является выбор сетей-компонентов. Для этого решено было использовать три современные нейросети: YOLOv8 для детектирования дефектов, U-Net для семантической сегментации и Mask R-CNN для инстанс-сегментации дефектов трех типов.

В состав ансамбля входит ранее описанная модель детекции YOLOv8, обученная на наборе данных SVRDD для выделения ограничивающих рамок семи классов, модель U-Net и Mask R-CNN для сегментации классов трещин, сетки трещин и выбоин, обученные на SegmRDD.

Модель U-Net соответствует представленной в работе [15], в качестве кодирующей части сети использовались начальные слои ResNet50, предобученной на наборе данных ImageNet.

Для обучения CHC Mask R-CNN использовалась технология переноса обучения с помощью предварительно обученной модели ResNet101 на наборе данных MS-COCO. На вход CHC подаются RGB-изображения размера $1\,024 \times 1\,024$, а в качестве масштабов анкоров используются следующие значения: 0,33; 0,5; 1; 2; 3. Обучение производилось на протяжении 40 эпох по 400 итераций с применением мини-масок размером 56×56 пикселей для оптимизации используемой памяти компьютера.

Вторым этапом создания ансамбля является разработка алгоритма объединения результатов предсказаний моделей, основные моменты которого включают:

1. Независимое получение предсказаний моделей с пороговой точностью более 20%.
 2. Отсеивание дефектов класса «канализационный люк» или «заплата».
 3. Применение морфологической операции открытие над результатом U-Net, что позволяет удалить маленькие объекты.
 4. Перебор всех экземпляров класса «трещины» Mask R-CNN и объединение с сегментационными масками U-Net в соответствующем классе.
 5. Объединение результатов в классах «выбоины» и «сетка трещин».
- В результате работы нейросетевой ансамбль повышает точность сегментации (табл. 4).

Таблица 4

Точность нейросетевого ансамбля на тестовой выборке SegmRDD

Класс	YOLOv8		Mask R-CNN		U-Net		Ансамбль	
	F_1 (box)	$mAP_{0,5}$ (box)	F_1 (mask)	$mAP_{0,5}$ (mask)	F_1 (mask)	$mAP_{0,5}$ (mask)	F_1 (mask)	$mAP_{0,5}$ (mask)
Все	0,329	0,254	0,559	0,515	0,631	0,607	0,701	0,668
Трещина	0,255	0,163	0,377	0,292	0,632	0,609	0,652	0,612
Сетка трещин	0,328	0,264	0,611	0,600	0,691	0,663	0,701	0,685
Выбоина	0,403	0,335	0,688	0,652	0,570	0,550	0,750	0,707

Таким образом, использование трех независимо обученных нейросетевых моделей и оригинального алгоритма объединения результатов моделей позволило создать ансамбль с точностью сегментации по агрегированному критерию качества F_1 , равному 70%. Примеры результатов сегментации на тестовых изображениях из набора данных дефектов SegmRDD приведены на рис. 4.

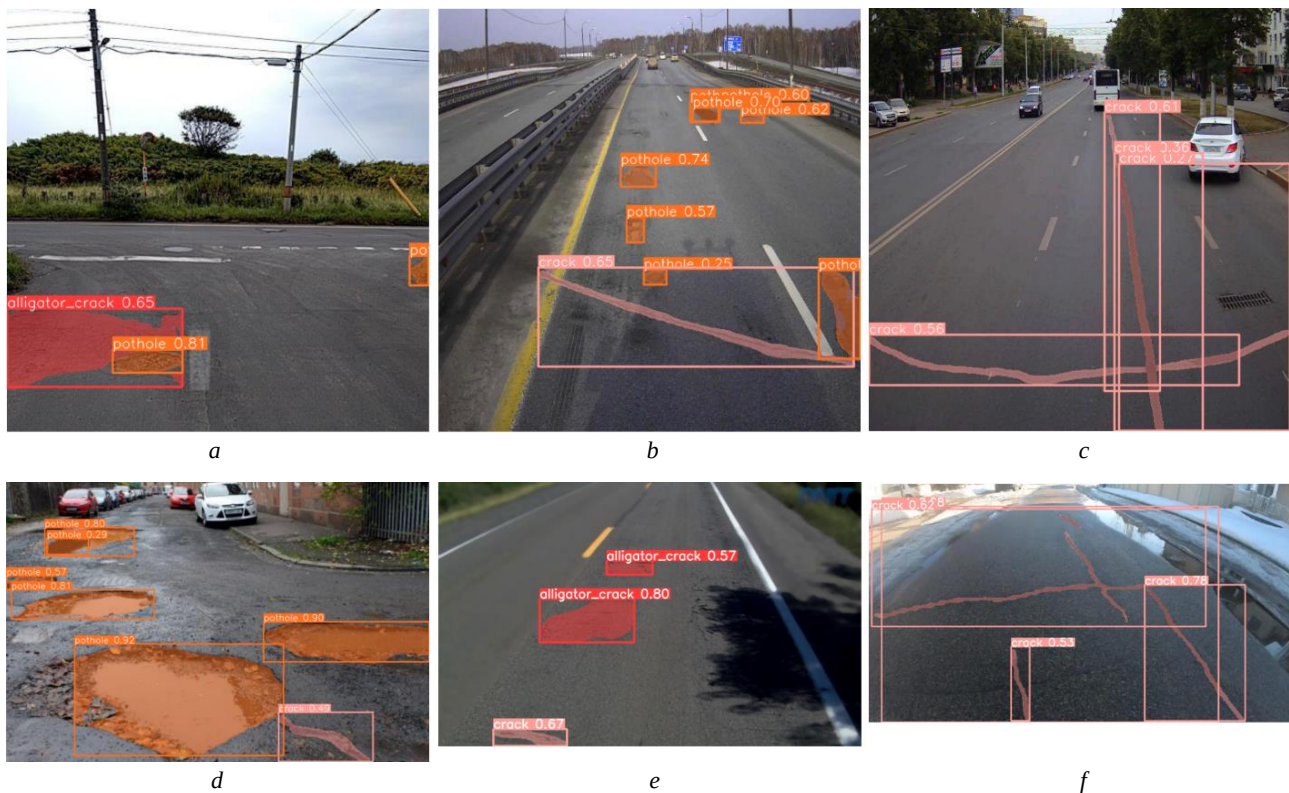


Рис. 4. Примеры результатов сегментации на тестовых изображениях из набора данных дефектов SegmRDD по выборкам: *a* – RDD2020, *b* и *c* – RusRDD, *d* – UDTIRI, *e* – DNIT, *f* – EdmCrack600

Fig. 4. Examples of segmentation results on test images from the SegmRDD dataset by sets: *a*) RDD2020, *b-c*) RusRDD, *d*) UDTIRI, *e*) DNIT, *f*) EdmCrack600

Ввиду наличия в наборе данных SegmRDD образцов с разными типами ракурса к дорожному покрытию (вид водителя, вид сверху и т.д.), случаются ложные срабатывания в областях, не принадлежащих дороге. Также наблюдаются трудности в сегментации выбоин с нетипичным мусором. Так как в обучающем наборе данных присутствуют экземпляры выбоин с лужами и гравием, то в целом модель хорошо справляется с погодными проявлениями на дорожном покрытии.

Заключение

В работе проведен анализ актуальности темы автоматизации процесса дефектовки дорожного покрытия, приведена классификация накопленного опыта в обнаружении дефектов, выполнен обзор существующих открытых наборов данных, рассмотрены современные архитектуры СНС.

Представлен созданный сбалансированный мультинациональный набор данных SegmRDD, содержащий 4 420 изображений автомобильных дорог с дефектами трех классов: «трещины», «сетка трещин», «выбоины».

Разработан новый алгоритм для решения задач обнаружения, выделения и сегментации дефектов дорожного покрытия на основе ансамбля глубоких СНС, обеспечивающий высокие показатели полноты и точности ($F_1 = 70\%$).

Список источников

1. Zou C., Cao Y., Li Q., Mao Q., Wang S. CrackTree: Automatic crack detection from pavement images // Pattern Recognition Letters. 2012. № 3 (33). P. 227–238. doi: 10.1016/j.patrec.2011.11.004
2. Salman M., Mathavan S., Kamal K., Rahman M. Pavement crack detection using the Gabor filter // Proc. IEEE Conference on Intelligent Transportation Systems (ITSC 2013). 6–9 October. Hague, Netherlands. 2013. P. 2039–2044. doi: 10.1109/ITSC.2013.6728529
3. Shatnawi N. Automatic pavement cracks detection using image processing techniques and neural network // International Journal of Advanced Computer Science and Applications. 2018. № 9 (9). P. 399–402. doi: 10.14569/IJACSA.2018.090950
4. Varadharajan S., Jose S., Sharma K., Wander L., Mertz C. Vision for road inspection // Proc. IEEE Winter Conference on Applications of Computer Vision. 2014. P. 115–122. doi: 10.1109/WACV.2014.6836111
5. Chatterjee S., Saeedfar P., Tofangchi S., Kolbe L. Intelligent Road maintenance: a machine learning approach for surface defect detection // Proc. Twenty-Sixth European Conference on Information Systems (ECIS2018). Portsmouth, UK. 2018. P. 1–16.
6. Cubero-Fernandez A., Rodriguez-Lozano F.J., Villatoro R., Olivares J., Palomares J.M. Efficient pavement crack detection and classification // EURASIP Journal on Image and Video Processing. 2017. № 39 (2017). P. 1–11. doi: 10.1186/s13640-017-0187-0
7. Ai D., Jiang G., Siew-Kei L., Li C. Automatic pixel-level pavement crack detection using information of multi-scale neighborhoods // IEEE Access. 2018. № 6. P. 24452–24463. doi: 10.1109/ACCESS.2018.2829347
8. Zhang L., Yang F., Zhang Y.D., Zhu Y.J. Road crack detection using deep convolutional neural network // Proc. IEEE International Conference on Image Processing (ICIP 2016). 25–28 September. Phoenix, USA. 2016. P. 3708–3712. doi: 10.1109/ICIP.2016.7533052.
9. Gopalakrishnan K., Khaitan S.K., Choudhary A., Agrawal A. Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection // Construction and Building Materials. 2017. № 157 (September). P. 322–330. doi: 10.1016/j.conbuildmat.2017.09.110
10. Maeda H., Sekimoto Y., Seto T., Kashiyama T., Omata H. Road damage detection and classification using deep neural networks with smartphone images // Computer-Aided Civil and Infrastructure Engineering. 2018. № 12 (33). P. 1127–1141. doi: 10.1111/mice.12387
11. Arya D., Maeda H., Sekimoto Y. From global challenges to local solutions: A review of cross-country collaborations and winning strategies in road damage detection // Advanced Engineering Informatics. 2024. № 60. Art. 102388. doi: 10.1016/j.aei.2024.102388
12. Dhiman A., Klette R. Pothole detection using computer vision and learning // IEEE Transactions on Intelligent Transportation Systems. 2020. № 21 (8). P. 3536–3550. doi: 10.1109/TITS.2019.2931297
13. Guo S., Li L., Feng Y., Zhou D., Zhang D., Chen C., Su S., Zhu X., Chen Q., Fan R. UDTIRI: An online open-source intelligent road inspection benchmark suite // IEEE Transactions on Intelligent Transportation Systems. 2024. № 1. P. 1–12. doi: 10.1109/TITS.2024.3351209
14. Fan R., Ozgunalp U., Hosking B., Liu M., Pitas I. Pothole detection based on disparity transformation and road surface modeling // IEEE Transactions on Image Processing. 2020. № 29. P. 897–908. doi: 10.1109/TIP.2019.2933750
15. Канаева И.А., Иванова Ю.А., Спицын В.Г. Сегментация дефектов дорожного покрытия на основе формирования синтетических выборок с помощью глубоких генеративно-состязательных сверточных сетей // Компьютерная оптика. 2021. Т. 45, № 6. С. 907–916. doi: 10.18287/2412-6179-CO-844

16. Stricker R., Eisenbach M., Sesselmann M., Debes K., Gross H.-M. Improving visual road condition assessment by extensive experiments on the extended GAPs dataset // Proc. International Joint Conference on Neural Networks (IJCNN 2019). 14–19 July. Budapest, Hungary. IEEE, 2019. P. 1–8. doi: 10.1109/IJCNN.2019.8852257
17. Majidifard H., Jin P., Adu-Gyamfi Y., Buttlar W.G. Pavement Image Datasets: A new benchmark dataset to classify and densify pavement distresses // Transportation Research Record: Journal of the Transportation Research Board. 2020. № 2 (2674). P. 328–339. doi: 10.1177/0361198120907283
18. Tang W., Huang S., Zhao Q., Li R., Huangfu L. An iteratively optimized patch label inference network for automatic pavement distress detection // IEEE Transactions on Intelligent Transportation Systems. 2022. V. 23 (7). P. 8652–8661. doi: 10.1109/TITS.2021.3084809.
19. Maeda H., Kashiya T., Sekimoto Y., Seto Y., Omata H. Generative adversarial network for road damage detection // Computer-Aided Civil and Infrastructure Engineering. 2021. № 36. P. 47–60. doi: 10.1111/mice.12561
20. Arya D., Maeda H., Ghosh S.K., Toshniwal D., Sekimoto Y. RDD2022: A multi-national image dataset for automatic Road Damage Detection Background and Summary. URL: <https://arxiv.org/abs/2209.08538> (accessed: 22.06.2024).
21. Arya D., Maeda H., Ghosh S.K., Toshniwal D., Mraz A., Kashiya T., Sekimoto Y. Transfer learning-based Road Damage Detection for multiple countries. URL: <http://arxiv.org/abs/2008.13101> (accessed: 22.06.2024).
22. Ren M., Zhang X., Zhi X., Wei Y., Feng Z. An annotated street view image dataset for automated road damage detection // Scientific Data. 2024. № 1 (11). Art. 407. doi: 10.1038/s41597-024-03263-7
23. Shi Y., Cui L., Qi Z., Meng F., Chen Z. Automatic road crack detection using random structured forests // IEEE Transactions on Intelligent Transportation Systems. 2016. № 12 (17). P. 3434–3445. doi: 10.1109/TITS.2016.2552248
24. Mei Q., Gul M. A cost effective solution for pavement crack inspection using cameras and deep neural networks // Construction and Building Materials. 2020. № 256 (November). Art. 119397. doi: 10.1016/j.conbuildmat.2020.119397
25. Passos B.T., Cassaniga M.J., Fernandes A.M.R., Medeiros K.B., Comunello E. Cracks and potholes in road images. URL: <https://biankatpas.github.io/Cracks-and-Potholes-in-Road-Images-Dataset> (accessed: 20.05.2024).
26. Farzad. Pothole Segmentation YOLOv8 Dataset. URL: https://universe.roboflow.com/farzad/pothole_segmentation_yolov8 (accessed: 25.03.2024)
27. Jocher G., Chaurasia A., Qiu J. Ultralytics YOLO (Version 8.0.0). URL: <https://github.com/ultralytics/ultralytics> (accessed: 22.06.2024).
28. Харченко И.К., Боровской И.Г., Шельмина Е.А. Использование ансамбля сверточных нейронных сетей для распознавания дорожных знаков // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2022. № 61. С. 88–96. doi: 10.17223/19988605/61/9

References

1. Zou, C., Cao, Y., Li, Q., Mao, Q. & Wang, S. (2012) CrackTree: Automatic crack detection from pavement images. *Pattern Recognition Letters*. 33(3). pp. 227–238. DOI: 10.1016/j.patrec.2011.11.004
2. Salman, M., Mathavan, S., Kamal, K. & Rahman, M. (2013) Pavement crack detection using the Gabor filter. *Proceedings IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*. October 6–9. Hague, Netherlands. pp. 2039–2044. DOI: 10.1109/ITSC.2013.6728529
3. Shatnawi, N. (2018) Automatic Pavement Cracks Detection using Image Processing Techniques and Neural Network. *International Journal of Advanced Computer Science and Applications*. 9(9). pp. 399–402. DOI: 10.14569/IJACSA.2018.090950
4. Varadharajan, S., Jose, S., Sharma, K., Wander, L. & Mertz, C. (2014) Vision for road inspection. *Proceedings IEEE Winter Conference on Applications of Computer Vision*. pp. 115–122. DOI: 10.1109/WACV.2014.6836111
5. Chatterjee, S., Saeedfar, P., Tofangchi, S. & Kolbe, L. (2018) Intelligent Road maintenance: a machine learning approach for surface defect detection. *Proceedings Twenty-Sixth European Conference on Information Systems (ECIS2018)*. Portsmouth, UK. pp. 1–16.
6. Cubero-Fernandez, A., Rodriguez-Lozano, F.J., Villatoro, R., Olivares, J. & Palomares, J.M. (2017) Efficient pavement crack detection and classification. *EURASIP Journal on Image and Video Processing*. 39. pp. 1–11. DOI: 10.1186/s13640-017-0187-0
7. Ai, D., Jiang, G., Siew-Kei, L. & Li, C. (2018) Automatic pixel-level pavement crack detection using information of multi-scale neighborhoods. *IEEE Access*. 6. pp. 24452–24463. DOI: 10.1109/ACCESS.2018.2829347
8. Zhang, L., Yang, F., Zhang, Y.D. & Zhu, Y.J. (2016) Road crack detection using deep convolutional neural network. *Proceedings IEEE International Conference on Image Processing (ICIP 2016)*. September 25–28. Phoenix, USA. pp. 3708–3712. DOI: 10.1109/ICIP.2016.7533052
9. Gopalakrishnan, K., Khaitan, S.K., Choudhary, A. & Agrawal, A. (2017) Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection. *Construction and Building Materials*. 157. pp. 322–330. DOI: 10.1016/j.conbuildmat.2017.09.110
10. Maeda, H., Sekimoto, Y., Seto, T., Kashiya, T. & Omata, H. (2018) Road damage detection and classification using deep neural networks with smartphone images. *Computer-Aided Civil and Infrastructure Engineering*. 33(12). pp. 1127–1141. DOI: 10.1111/mice.12387
11. Arya, D., Maeda, H. & Sekimoto, Y. (2024) From global challenges to local solutions: A review of cross-country collaborations and winning strategies in road damage detection. *Advanced Engineering Informatics*. 60. Art. 102388. DOI: 10.1016/j.aei.2024.102388
12. Dhiman, A. & Klette, R. (2020) Pothole detection using computer vision and learning. *IEEE Transactions on Intelligent Transportation Systems*. 8(21). pp. 3536–3550. DOI: 10.1109/TITS.2019.2931297

13. Guo, S., Li, L., Feng, Y., Zhou, D., Zhang, D., Chen, C., Su, S., Zhu, X., Chen, Q. & Fan, R. (2024) UDTIRI: An online open-source intelligent road inspection benchmark suite. *IEEE Transactions on Intelligent Transportation Systems*. 1. pp. 1–12. DOI: 10.1109/TITS.2024.3351209
14. Fan, R., Ozgunalp, U., Hosking, B., Liu, M. & Pitas, I. (2020) Pothole detection based on disparity transformation and road surface modeling. *IEEE Transactions on Image Processing*. 29. pp. 897–908. DOI: 10.1109/TIP.2019.2933750
15. Канаева, И.А., Иванова, Ю.А. & Спицын, В.Г. (2021) Segmentatsiya defektov dorozhnogo pokrytiya na osnove formirovaniya sinteticheskikh vyborok s pomoshch'yu glubokikh generativno-sostyazatel'nykh svertochnykh setey [Deep convolutional generative adversarial network-based synthesis of datasets for road pavement distress segmentation]. *Komp'yuternaya optika*. 45(6). pp. 907–916. DOI: 10.18287/2412-6179-CO-844
16. Stricker, R., Eisenbach, M., Sesselmann, M., Debes, K. & Gross, H.M. (2019) Improving visual road condition assessment by extensive experiments on the extended GAPs dataset. *Proceedings International Joint Conference on Neural Networks (IJCNN 2019)*. July 14–19. Budapest. Hungary. pp. 1–8. DOI: 10.1109/IJCNN.2019.8852257
17. Majidifard, H., Jin, P., Adu-Gyamfi, Y. & Buttlar, W.G. (2020) Pavement Image Datasets: A new benchmark dataset to classify and densify pavement distresses. *Transportation Research Record: Journal of the Transportation Research Board*. 2674(2). pp. 328–339. DOI: 10.1177/0361198120907283
18. Tang, W., Huang, S., Zhao, Q., Li, R. & Huangfu, L. (2022) An iteratively optimized patch label inference network for automatic pavement distress detection. *IEEE Transactions on Intelligent Transportation Systems*. 23(7). pp. 8652–8661. DOI: 10.1109/TITS.2021.3084809
19. Maeda, H., Kashiya, T., Sekimoto, Y., Seto, Y. & Omata, H. (2021) Generative adversarial network for road damage detection. *Computer-Aided Civil and Infrastructure Engineering*. 36. pp. 47–60. DOI: 10.1111/mice.12561
20. Arya, D., Maeda, H., Ghosh, S.K., Toshniwal, D. & Sekimoto, Y. (n.d.) RDD2022: A multi-national image dataset for automatic Road Damage Detection Background and Summary. URL: <https://arxiv.org/abs/2209.08538> (Accessed: 22nd June 2024).
21. Arya, D., Maeda, H., Ghosh, S.K., Toshniwal, D., Mraz, A., Kashiya, T. & Sekimoto, Y. (n.d.) Transfer learning-based Road Damage Detection for multiple countries. [Online] Available from: <http://arxiv.org/abs/2008.13101> (Accessed: 22nd June 2024).
22. Ren, M., Zhang, X., Zhi, X., Wei, Y. & Feng, Z. (2024) An annotated street view image dataset for automated road damage detection. *Scientific Data*. 11(1). Art. 407. DOI: 10.1038/s41597-024-03263-7
23. Shi, Y., Cui, L., Qi, Z., Meng, F. & Chen, Z. (2016) Automatic road crack detection using random structured forests. *IEEE Transactions on Intelligent Transportation Systems*. 17(12). pp. 3434–3445. DOI: 10.1109/TITS.2016.2552248
24. Mei, Q. & Gul, M. (2020) A cost effective solution for pavement crack inspection using cameras and deep neural networks. *Construction and Building Materials*. 256. pp. 119397–119397. DOI: 10.1016/j.conbuildmat.2020.119397
25. Passos, B.T., Cassaniga, M.J., Fernandes, A.M.R., Medeiros, K.B. & Comunello, E. (n.d.) Cracks and potholes in road images. [Online] Available from: <https://biankatpas.github.io/Cracks-and-Potholes-in-Road-Images-Dataset>. (Accessed: 20th May 2024).
26. Farzad. (n.d.) Pothole Segmentation YOLOv8 Dataset. [Online] Available from: https://universe.roboflow.com/farzad/pothole_segmentation_yolov8 (Accessed: 25th March 2024).
27. Jocher, G., Chaurasia, A. & Qiu, J. (2023) Ultralytics YOLO (Version 8.0.0). [Online] Available from: <https://github.com/ultralytics/ultralytics> (Accessed: 22nd June 2024).
28. Kharchenko, I.K., Borovskoy, I.G. & Shelmina, E.A. (2022) Usage of convolutional neural network ensemble for traffic sign recognition. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 61. pp. 88–96. DOI: 10.17223/19988605/61/9

Информация об авторах:

Канаева Ирина Александровна – аспирант отделения информационных технологий Инженерной школы информационных технологий и робототехники Национального исследовательского Томского политехнического университета (Томск, Россия). E-mail: iap15@tpu.ru

Спицын Владимир Григорьевич – доктор технических наук, профессор отделения информационных технологий Инженерной школы информационных технологий и робототехники Национального исследовательского Томского политехнического университета (Томск, Россия); профессор Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: spvg@tpu.ru

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

Information about the authors:

Kanaeva Irina A. (Post-graduate Student, National Research Tomsk Polytechnic University, Tomsk, Russian Federation). E-mail: iap15@tpu.ru

Spitsyn Vladimir G. (Doctor of Technical Sciences, Professor, National Research Tomsk Polytechnic University, Tomsk, Russian Federation; National Research Tomsk State University, Tomsk, Russian Federation). E-mail: spvg@tpu.ru

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

Поступила в редакцию 01.06.2024; принята к публикации 03.09.2024

Received 01.06.2024; accepted for publication 03.09.2024

Научная статья

УДК 004.896

doi: 10.17223/19988605/68/9

**Применение нейросетевых алгоритмов для детектирования человека
на видеоряде в шахте****Владислав Васильевич Петровский***Национальный исследовательский Томский политехнический университет, Томск, Россия,
petrovskiy.vlad.98@gmail.com*

Аннотация. Рассматривается задача применения нейросетевых алгоритмов для детектирования человека на видеоряде в шахте. Производится анализ сверточных нейронных сетей Faster R-CNN, YOLOv5 и YOLOv8 со сборками n, m, x (Nano, Medium и Extra Large) и SSG для детектирования объектов на видеоряде с классами «шахтер», «лицо», «голова с каской», «каска».

Ключевые слова: детектирование; CNN, Faster R-CNN; YOLO, SSG.

Для цитирования: Петровский В.В. Применение нейросетевых алгоритмов для детектирования человека на видеоряде в шахте // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 68. С. 86–91. doi: 10.17223/19988605/68/9

Original article

doi: 10.17223/19988605/68/9

**Application of neural network algorithms for detecting a person
in a video sequence in a mine****Vladislav V. Petrovskiy***National Research Tomsk Polytechnic University, Tomsk, Russian Federation,
petrovskiy.vlad.98@gmail.com*

Abstract. The problem of using neural network algorithms to detect a person in a video sequence in a mine is considered. Convolutional neural networks are analyzed: Faster R-CNN, YOLOv5 and YOLOv8 with n, m, x (Nano, Medium and Extra Large) and SSG assemblies for detecting objects in video with classes: miner, face, head with a helmet, helmet.

Keywords: detection; CNN, Faster R-CNN, YOLO, SSG.

For citation: Petrovskiy, V.V. (2024) Application of neural network algorithms for detecting a person in a video sequence in a mine. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 68. pp. 86–91. doi: 10.17223/19988605/68/9

Введение

Существует несколько типов шахт в зависимости от способа добычи: открытые и глубокие. Если в открытых шахтах добыча происходит на поверхности, то в глубоких – подземно на значительной глубине. Для этого используются вертикальные или наклонные шахты, а также горизонтальные разработки. Помимо этого, для транспортировки материалов и рабочих в шахтах используются подземные железные дороги, конвейеры, дизельные вагонетки и т.д.

Общий процесс работы в шахте связан с высокими рисками для здоровья человека. Поэтому в большинстве случаев шахты оборудуются:

- источником пресной воды; она может использоваться как для употребления человеком, так и для очистки горной породы и предотвращения скопления пыли;
- источниками света; обычно используется несколько источников: светодиодные лампы на стенах шахты и шахтерские каски с фонариком;
- вентиляцией по всей длине шахты;
- обогревательными трубами.

При работе горнодобывающего оборудования нередко происходят несчастные случаи попадания человека под транспорт или подвижный прибор. Решение подобной проблемы было найдено в применении искусственного интеллекта в мониторинге безопасности. Используя камеры видеонаблюдения и сверточные нейронные сети (СНН) для детектирования объектов, можно минимизировать риски аварийных и опасных ситуаций. Камера будет транслировать видео на нейронную сеть, которая покадрово обрабатывает изображение и детектирует человека в области действия оборудования. Информация о возникновении человека в опасной зоне передается как на экран рубки к наблюдателю, так и в «триггер» через микроконтроллер в оборудование / транспорт для немедленного отключения.

Одними из наиболее популярных СНН [1], которые обладают высокой эффективностью в распознавании объектов на видеоряде за миллисекунды, являются: Faster R-CNN (Region-Based Convolutional Neural Network) [2], YOLOv5 и v8 (You Only Look Once version 5 and 8) [3, 4] и SSD (Single Shot Detector) [5].

В [6] сравниваются производительность и точность архитектуры Faster R-CNN с предыдущими моделями R-CNN, Fast R-CNN, что показано в табл. 1.

Таблица 1

Производительность и точность архитектур R-CNN, Fast R-CNN и Faster R-CNN

	R-CNN	Fast R-CNN	Faster R-CNN
Test time per image	50 seconds	2 seconds	0,2 seconds
Speed-up	1x	25x	250x
mAP (VOC 2007)	66,0%	66,9%	66,9%

Для обучения моделей использовалось 2 000 размеченных изображений из датасета PASCAL VOC 2007 [7] разного разрешения.

Из результатов видно, что версия Faster R-CNN имеет высокие показатели по скорости и точности распознавания объекта по метрике mAP (оценки среднего нахождения области под кривой точности -отзыва) на изображении по сравнению с предыдущими версиями.

Авторы [8] сравнивают зависимости количества параметров (parameters; рис. 1, а) и производительности графического процессора NVIDIA Latency A100 TensorRT FP16 (рис. 1, б) от точности распознавания объектов (COCO mAP 50–95%) на следующих версиях этой архитектуры: YOLO v5, v6, v7 и v8 в соответствии с их сборкой: Nano (n), Small (s), Medium (m), Large (l) и Extra Large (x).

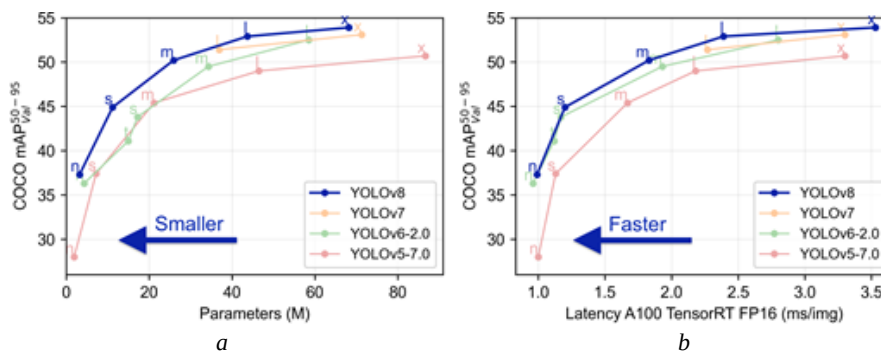


Рис. 1. Влияние параметров и производительности на точность у разных версий YOLO
Fig. 1. Impact of parameters and performance on accuracy for different versions of YOLO

При обучении архитектур использовался датасет COCO [9] из 2 000 размеченных изображений в разрешении 640×640 пикселей.

Все варианты YOLO-v8 обеспечивают лучшую пропускную способность при одинаковом количестве параметров, что указывает на аппаратно-эффективные архитектурные реформы.

В [10] сравнивается производительность тестовых данных в Faster R-CNN и SSD на метриках: точность (precision), полнота (Recall), показатель f1 (f1 score) и средний IoU (mean IoU) выходного двоичного изображения модели, что показано в табл. 2.

Таблица 2

Сравнение производительности на тестовых данных в Faster R-CNN и SSD

Model	Precision	Recall	f1 score	Mean IoU	Inference Time of 1 152 × 1 152 Image in Seconds
Faster R-CNN	0,65	0,68	0,66	0,85	0,23
SSD	0,66	0,68	0,67	0,84	0,21

При обучении было использовано 450 размеченных изображений в разрешении $1\,152 \times 1\,152$ пикселя из датасета, сделанных вручную из изображений полей, снятых на камеру дрона.

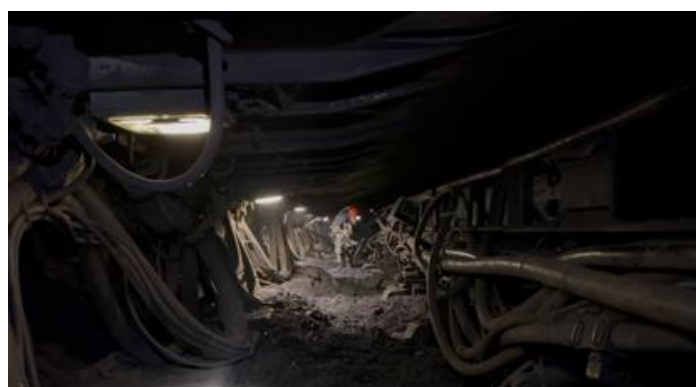
Как видно из опыта, точность, полнота, показатель f1 и среднее значение IoU обеих моделей были одинаковыми, но модель SSD была немного быстрее в исполнении, чем Faster R-CNN.

Faster R-CNN, YOLOv5 и YOLOv8 со сборками n, m, x (Nano, Medium и Extra Large) и SSG + VGG16 для детектирования человека на видеоряде в шахте

Использование таких датасетов, как COCO и PASCAL VOC 2007, не приведет к качественным результатам, так как из-за малой освещенности и разной высоты от пола до потолка шахты человек может находиться в разных положениях, – это сказывается на дальности и качестве детектирования.

Исходные данные

Для данной работы был собран датасет из видеорядов и снимков рабочей зоны угольной шахты. Протяженность рядов в среднем занимает от 10 до 15 с. Была проведена раскадровка, в результате чего в общей сумме получено 500 снимков (рис 2, а). Затем для каждого снимка с помощью программы Label Studio были размечены рамки классов формата txt (рис. 2, b).



а

0a0d393d-196.txt – Блокнот
 Файл Правка Формат Вид Справка
 1 0.5038461538461538 0.34551282051282056 0.4230769230769229 0.10641025641025643
 1 0.3384615384615385 0.4544871794871795 0.10769230769230774 0.15
 1 0.676923076923077 0.45897435897435895 0.11282051282051278 0.13589743589743591
 0 0.5025641025641027 0.43141025641025643 0.4923076923076924 0.2705128205128205

b

Рис. 2. Пример данных из датасета

Fig. 2. Example of data from a dataset

Всего было выбрано 4 класса:

- Face (лицо) – RGB: 255, 60, 60;
- Helmet (голова с каской) – RGB: 255, 160, 150;
- Miner (шахтер) – RGB: 255, 110, 30;
- Helmes (каска) – RGB: 255, 150, 60.

Из 350 снимков и их разметок в датасете 450 идут на обучение, 50 – на валидацию. Для теста принимается видеоряд процесса работы шахтера в узком пространстве.

Обучение и валидация

Для обучения и тестирования нейронных сетей был использован графический процессор с поддержкой CUDA Toolkit 11.8. Результаты обучения и валидации представлены в табл. 3.

Таблица 3

Обучение СНН на 100 циклах

	Precision	Recall	Train box loss	Train class loss	Val. box loss	Val. class loss	mAP 50	mAP 50-95	Studying time, sec
Faster R-CNN	0,901	0,836	0,747	0,523	1,545	0,786	0,893	0,531	10 251
SSD + VGG16	0,854	0,747	0,711	0,482	1,412	0,872	0,885	0,502	8 432
YOLOv5n	0,857	0,722	0,908	0,589	1,721	0,911	0,814	0,404	343
YOLOv5m	0,828	0,741	0,667	0,407	1,753	0,933	0,829	0,414	852
YOLOv5x	0,875	0,774	0,555	0,364	1,718	0,901	0,844	0,428	29 547
YOLOv8n	0,807	0,792	0,829	0,521	1,804	0,927	0,805	0,395	355
YOLOv8m	0,813	0,763	0,570	0,354	1,851	0,975	0,803	0,393	952
YOLOv8x	0,830	0,793	0,382	0,287	1,704	0,875	0,829	0,431	31 733

Результаты проведенных экспериментов показывают высокую точность и надежность обучения Faster R-CNN, на что потребовалось большое время обучения.

Детектор SSD с моделью VVG16 показал средние результаты и имеет похожее время обучения с Faster R-CNN. Модели YOLO версий 5 и 8 со сборками Nano и Medium имеют неплохие показатели точности и полноты обучения, а их скорость обучения намного выше, чем у других архитектур. YOLOv5 и YOLOv8 со сборкой Extreme Large показывают лучшие результаты по сравнению с другими сборками моделей YOLO, однако их время обучения слишком большое: 29 547 и 31 733 с соответственно.

Тесты

Для отрисовки рамок на видеоряде используется библиотека OpenCV python (рис. 3).



Рис. 3. Пример выделения объектов при тестировании моделей СНН с применением библиотеки OpenCV python
Fig. 3. Example of object selection when testing SNN models using the OpenCV python library

Результаты тестов средней производительности (Average FPS) и среднего времени предсказания (Average time) при изменении доверительного порога представлены в табл. 4.

Таблица 4

Влияние доверительного порога на производительность и время предсказания за кадр

	Precision > 0,001	Precision > 0,50	Precision > 0,75	Precision > 0,90
Faster R-CNN	17,929 FPS, 52 ms	18,293 FPS, 47 ms	21,274 FPS, 43 ms	21,221 FPS, 41 ms
SSD + VGG16	25,562 FPS, 52 ms	28,252 FPS, 47 ms	32,503 FPS, 43 ms	33,015 FPS, 41 ms
YOLOv5n	30,926 FPS, 36 ms	57,583 FPS, 21 ms	62,289 FPS, 19 ms	61,121 FPS, 19 ms
YOLOv5m	34,108 FPS, 31 ms	43,409 FPS, 26 ms	52,989 FPS, 21 ms	54,270 FPS, 20 ms
YOLOv5x	37,256 FPS, 34 ms	45,346 FPS, 25 ms	53,232 FPS, 21 ms	53,643 FPS, 21 ms
YOLOv8n	28,942 FPS, 38 ms	39,105 FPS, 30 ms	58,965 FPS, 21 ms	60,228 FPS, 20 ms
YOLOv8m	32,531 FPS, 33 ms	45,131 FPS, 25 ms	51,297 FPS, 22 ms	53,262 FPS, 20 ms
YOLOv8x	35,352 FPS, 34 ms	52,634 FPS, 23 ms	53,297 FPS, 21 ms	55,262 FPS, 19 ms

Тесты показывают, что при доверительном пороге в 0,001 ($> 0,1\%$) производительность наименьшая на всех моделях. Это связано с увеличенной нагрузкой на процессор при отрисовке рамок классов, которые в большинстве случаев несут некорректную информацию. При пороге выше 0,90 ($> 90\%$) время на предсказание и отрисовку рамок в кадр лучше остальных, однако детектирование объектов происходит только на близком расстоянии от камеры слежения.

В ходе тестирования доверительными порогами детектирования выбраны значения 0,5 и 0,75 ($> 50\%$ и $> 75\%$). При данных порогах модели достаточно уверены в своих предсказаниях, но не слишком строги, что позволяет улавливать большинство верных рамок объектов и не нагружать процессор. Лучшие результаты на этих доверительных порогах показала архитектура YOLOv5n.

Заключение

Проведен анализ эффективности сверточных нейронных сетей для распознавания человека в шахте. По результатам обучения и тестирования моделей можно сделать следующие выводы:

- для обеспечения малых вычислительных затрат и высокой скорости обработки входящей информации с видеоряда предпочтительнее выбирать архитектуры YOLO со сборкой Nano;
- если важна точность распознавания человека или объекта, стоит воспользоваться архитектурами Faster R-CNN, SSD + CNN VGG16, YOLOv5, v8 со сборкой Extreme Large;
- во избежание ложных предсказаний (или отсутствия самих предсказаний) при работе с CNN рекомендуется использовать доверительный порог в интервале от 50 до 90%.

Список источников

1. Snehal J. Top Object Detection Models in 2024. URL: <https://www.hitechbpo.com/blog/top-object-detection-models.php> (accessed: 27.05.2024).
2. Ren S., He K., Girshick R., Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks // arXiv. 2016. 6 Jan. doi: 1506.01497v3. URL: <https://arxiv.org/pdf/1506.01497> (accessed: 27.05.2024).
3. Исчерпывающее руководство по Ultralytics YOLOv5 // Ultralytics YOLO Docs. URL: <https://ultralytics.com/ru/yolov5/> (дата обращения: 27.05.2024).
4. YOLOv8 // Ultralytics YOLO Docs. URL: <https://docs.ultralytics.com/ru/models/yolov8/> (дата обращения: 27.05.2024).
5. Liu W., Anguelov D., Dumitru E., Szegedy C., Reed S., Fu C.-Y., Berg A. SSD: Single Shot MultiBox Detector // arXiv. 2016. 29 Dec. doi: 1512.02325. URL: <https://arxiv.org/pdf/1512.02325> (accessed: 27.05.2024).
6. Umam A. 1. Introduction to How Faster R-CNN, Fast R-CNN and R-CNN Works. 2017. 31 March. URL: <https://www.youtube.com/watch?v=v5bFVbQvFRk> (accessed: 28.05.2024).
7. PASCAL VOC 2007 : dataset // Paperswithcode.com. URL: <https://paperswithcode.com/dataset/pascal-voc-2007> (accessed: 30.05.2024).
8. Ultralytics YOLOv8 // Github. URL: <https://github.com/ultralytics/ultralytics> (accessed: 28.05.2024).
9. MS COCO (Microsoft Common Objects in Context) : dataset // Paperswithcode.com. URL: <https://paperswithcode.com/dataset/coco> (accessed: 30.05.2024).

10. Sivakumar A.N.V., Li J., Scott S., Psota E., Jhala A.J. Comparison of object detection and patch-based classification deep learning models on mid to late-season weed detection in UAV imagery // J. Biological Systems Engineering. V. 3. P. 2–22. doi: 10.3390/rs12132136

References

1. Snehal, J. (2024) *Top Object Detection Models in 2024*. [Online] Available from: <https://www.hitechbpo.com/blog/top-object-detection-models.php> (Accessed: 27th May 2024).
2. Ren, S., He, K., Girshick, R. & Sun, J. (2016) *Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks*. arXiv. 6th January. DOI: 1506.01497v3
3. Ultralytics.com. (n.d.) *Ischerpyvayushchee rukovodstvo po Ultralytics YOLOv5* [A comprehensive guide to Ultralytics YOLOv5]. [Online] Available from: <https://docs.ultralytics.com/ru/yolov5/> (Accessed: 27th May 2024).
4. Ultralytics.com. (n.d.) *YOLOv8. Ultralytics YOLO Docs*. [Online] Available from: <https://docs.ultralytics.com/ru/models/yolov8/> (Accessed: 27th May 2024).
5. Liu, W., Anguelov, D., Dumitru, E., Szegedy, C., Reed, S., Fu, C-Y. & Berg, A. (2016) *SSD: Single Shot MultiBox Detector*. arXiv. 29th December. DOI: 1512.02325
6. Umam, A. (2017) *Introduction to How Faster R-CNN, Fast R-CNN and R-CNN Works*. 31st May. [Online] Available from: <https://www.youtube.com/watch?v=v5bFVbQvFRk> (Accessed: 27th May 2024).
7. Paperswithcode.com. (2007) *PASCAL VOC 2007: dataset*. [Online] Available from: <https://paperswithcode.com/dataset/pascal-voc-2007> (Accessed: 29th May 2024).
8. Github.com. (n.d.) *Ultralytics YOLOv8*. [Online] Available from: <https://github.com/ultralytics/ultralytics> (Accessed: 29th May 2024).
9. Paperswithcode.com. (n.d.) *MS COCO (Microsoft Common Objects in Context). Dataset*. [Online] Available from: <https://paperswithcode.com/dataset/coco> (Accessed: 29th May 2024).
10. Sivakumar, A.N.V., Li, J., Scott, S., Psota, E. & Jhala, A.J. (2020) Comparison of object detection and patch-based classification deep learning models on mid to late-season weed detection in UAV imagery. *Journal of Biological Systems Engineering*. 3. pp. 2–22. DOI: 10.3390/rs12132136

Информация об авторе:

Петровский Владислав Васильевич – аспирант отделения информационных технологий Инженерной школы информационных технологий и робототехники Национального исследовательского Томского политехнического университета (Томск, Россия). E-mail: petrovskiy.vlad.98@gmail.com

Автор заявляет об отсутствии конфликта интересов.

Information about the author:

Petrovskiy Vladislav V. (Post-graduate Student, Information Technology Department, School of Information Technology and Robotics Engineering, National Research Tomsk Polytechnic University, Tomsk, Russian Federation). E-mail: petrovskiy.vlad.98@gmail.com

The author declares no conflicts of interests.

Поступила в редакцию 06.06.2024; принята к публикации 03.09.2024

Received 06.06.2024; accepted for publication 03.09.2024

Научная статья

УДК 681.5

doi: 10.17223/19988605/68/10

**Идентификация с использованием расширенного экстраполятора Калмана
при аномальных ошибках в наблюдениях****Валерий Иванович Смагин¹, Константин Станиславович Ким²**^{1,2} *Национальный исследовательский Томский государственный университет, Томск, Россия*¹ *vsm@mail.tsu.ru*² *kks93@rambler.ru*

Аннотация. Рассматривается задача идентификации для дискретной стохастической системы при аномальных ошибках в наблюдениях. Предложено для решения задачи идентификации использовать алгоритмы одновременного оценивания состояния и аномальных ошибок (неизвестного входа в наблюдениях). Для повышения точности идентификации оценки неизвестного входа предложено вычислять с использованием алгоритмов дополнительного сглаживания. Представлены результаты моделирования.

Ключевые слова: идентификация; аномальные ошибки; сглаживание.

Для цитирования: Смагин В.И., Ким К.С. Идентификация с использованием расширенного экстраполятора Калмана при аномальных ошибках в наблюдениях // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 68. С. 92–99. doi: 10.17223/19988605/68/10

Original article

doi: 10.17223/19988605/68/10

**Identification using the extended Kalman extrapolator
in case of anomalous errors in observations****Valery I. Smagin¹, Konstantin S. Kim²**^{1,2} *National Research Tomsk State University, Tomsk, Russian Federation*¹ *vsm@mail.tsu.ru*² *kks93@rambler.ru*

Abstract. The identification problem for a discrete stochastic system with anomalous errors in observations is considered. It is proposed to use algorithms for simultaneous assessment of the state and anomalous errors (unknown input in observations) to solve the identification problem. To increase the accuracy of identification, estimates of an unknown input are proposed to be calculated using additional smoothing algorithms. The simulation results are presented.

Keywords: identification; anomalous errors; smoothing.

For citation: Smagin, V.I., Kim, K.S. (2024) Identification using the extended Kalman extrapolator in case of anomalous errors in observations. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 68. pp. 92–99. doi: 10.17223/19988605/68/10

Введение

Проблема оценки параметров стохастических линейных динамических систем была исследована в значительной степени из-за ее важности для построения моделей [1]. Хорошо известно, что общий случай задачи идентификации стохастических систем, описываемых дифференциальными и разностными уравнениями, приводит к задаче нелинейного оценивания. Было предложено несколько методов,

большинство из которых основано на расширенном фильтре Калмана (ЕКФ) [1, 2]. В настоящее время при решении различных задач идентификации, требующих применения аппарата нелинейной фильтрации, часто используются сигма-точечные алгоритмы нелинейной фильтрации Unscented Kalman filter (UKF) [3, 4] и кубатурные алгоритмы Cubature Kalman filter (CKF) [5, 6]. Алгоритмы рекуррентного оценивания при аномальных ошибках в канале наблюдений рассматривались в работах [7, 8]. В работах [9, 10] исследовались задачи, в которых строились оценки вектора состояния и неизвестного входа в канале наблюдений. В [11, 12] предложено в алгоритмах оценивания неизвестного входа с целью повышения точности дополнительно использовать процедуры непараметрического сглаживания.

В настоящей работе рассмотрена задача идентификации с использованием расширенного экстраполятора Калмана при аномальных неизвестных входах в канале наблюдений. На примере оценки параметров модели нелинейной дискретной системы показано, что применение методов рекуррентного оценивания и дополнительного сглаживания в системах с аномальными ошибками измерения может повысить точность идентификации. При моделировании выполнено сравнение точности при использовании различных методов сглаживания.

1. Постановка задачи

Пусть модель объекта и канала наблюдений имеет вид:

$$x(k+1) = f(\theta, x(k), k) + q(k), \quad x(0) = x_0, \quad (1)$$

$$y(k) = S(k)x(k) + H\eta(k) + v(k), \quad (2)$$

где $x(k) \in \mathbb{R}^n$ – вектор состояния, x_0 – случайный вектор (предполагаются известными дисперсионная матрица $N_0 = M\{(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T\}$ и математическое ожидание $\bar{x}_0 = M\{x_0\}$), $\theta \in \mathbb{R}^p$ – неизвестный постоянный параметр, вектор $f(\theta, x(k))$ – заданная функция, $q(k) \in \mathbb{R}^n$ и $v(k) \in \mathbb{R}^r$ – последовательности нормально распределенных случайных векторов со следующими характеристиками:

$$M\{q(k)\} = 0, \quad M\{q(k)q^T(j)\} = Q(k)\delta_{kj},$$

$$M\{v(k)\} = 0, \quad M\{v(k)v^T(j)\} = V(k)\delta_{kj}. \quad (3)$$

Здесь δ_{kj} – символ Кронекера. В (2) $y(k) \in \mathbb{R}^l$ – вектор наблюдений; $S(k)$ – известная матрица канала наблюдений; $\eta(k) \in \mathbb{R}^s$ – неизвестный вектор (неизвестный вход в канале наблюдений, являющийся моделью аномальной ошибки в канале наблюдений); H – матрица влияния неизвестного возмущения. Предполагается, что последовательности $q(k)$, $v(k)$ и x_0 независимы между собой, матрица $Q(k)$ – неотрицательно определена, а матрица $V(k)$ – положительно определена.

Важной задачей, часто возникающей в различных областях науки и техники, является разработка алгоритмов параметрической идентификации для математической модели (1), (2) с учетом высказанных выше предположений. Задача заключается в нахождении оценок неизвестных параметров $\theta \in \mathbb{R}^p$ по данным наблюдений $y(k)$ при аномальных ошибках в наблюдениях.

Итак, по информации, поступившей в момент $k \in [0; T]$, требуется найти оценку $\hat{\theta}(k)$ на основе минимизации следующего критерия:

$$J[0; T] = M \left\{ \sum_{k=0}^T (\theta - \hat{\theta}(k))^T (\theta - \hat{\theta}(k)) \right\}. \quad (4)$$

2. Применение расширенного экстраполятора Калмана

Для решения задачи идентификации будем использовать расширенный экстраполятор Калмана, вычисляющий оценку экстраполяции на один такт вперед. Построим оценки $\hat{x}(k)$ и $\hat{\theta}(k)$ при условии,

что θ является неизвестным постоянным вектором. Дополнительно в этом разделе будем предполагать, что вектор $\eta(k)$ известен точно, оценку этого вектора опишем в разд. 3.

В качестве динамической модели вектора θ можно использовать разностное уравнение

$$\theta(k+1) = \theta(k), \quad \theta(0) = \theta_0, \quad (5)$$

где θ_0 – случайная величина с математическим ожиданием $\bar{\theta}_0$ и дисперсионной матрицей P_{θ_0} .

Введем в рассмотрение расширенный вектор

$$X(k) = \begin{pmatrix} x(k) \\ \theta(k) \end{pmatrix}. \quad (6)$$

Уравнения динамики модели (1) с учетом (5) можно записать в виде:

$$X(k+1) = F(X(k), k) + \bar{q}(k), \quad (7)$$

где векторная функция $F(X(k), k) = \begin{pmatrix} f(x(k), \theta(k), k) \\ \theta(k) \end{pmatrix}$, $\bar{q}(k)$ – последовательность нормально распределенных случайных векторов со следующими характеристиками:

$$M\{\bar{q}(k)\} = 0, \quad M\{\bar{q}(k)\bar{q}^T(j)\} = \bar{Q}(k)\delta_{kj}, \quad \bar{Q}(k) = \begin{pmatrix} Q(k) & 0 \\ 0 & 0 \end{pmatrix}.$$

Канал наблюдений примет вид:

$$y(k) = \bar{S}X(k) + H\eta(k) + v(k), \quad (8)$$

где $\bar{S} = [S \ 0]$.

Оценку вектора найдем с помощью рекуррентного алгоритма (экстраполятора):

$$\begin{aligned} \hat{X}(k+1) &= \begin{pmatrix} \hat{x}(k+1) \\ \hat{\theta}(k+1) \end{pmatrix} = F(\hat{X}(k)) + K(k)[y(k) - \bar{S}\hat{X}(k) - H\eta(k)] = \\ &= \begin{pmatrix} f(\hat{x}(k), \hat{\theta}(k)) \\ \hat{\theta}(k) \end{pmatrix} - K(k)[y(k) - S\hat{x}(k) - H\eta(k)]. \end{aligned} \quad (9)$$

Здесь

$$\begin{aligned} \hat{X}(k) &= \begin{pmatrix} \hat{x}(k) \\ \hat{\theta}(k) \end{pmatrix}, \\ K(k) &= \bar{A}(k)P(k)\bar{S}^T[\bar{S}P(k)\bar{S}^T + V(k)]^{-1}, \end{aligned} \quad (10)$$

где

$$\bar{A}(k) = \frac{\partial F}{\partial X} \Big|_{X=\hat{X}(k)} = \begin{pmatrix} \frac{\partial f}{\partial x} \Big|_{x=\hat{x}(k), \hat{\theta}(k)} & \frac{\partial f}{\partial \theta} \Big|_{x=\hat{x}(k), \hat{\theta}(k)} \\ 0 & E \end{pmatrix}, \quad (11)$$

E – единичная матрица размерности $p \times p$.

Дисперсионная матрица $P(k)$ определится из уравнения

$$P(k+1) = (\bar{A}(k) - K(k)\bar{S})P(k)(\bar{A}(k) - K(k)\bar{S})^T + K(k)V(k)K(k)^T + \bar{Q}(k), \quad P(0) = P_{\theta_0}. \quad (12)$$

3. Построение оценок неизвестного входа

В качестве алгоритма оценивания неизвестного входа $\eta(k)$ будем использовать алгоритмы МНК, в этом случае оценку можно построить на основе минимизации дополнительного критерия [7]:

$$J(\eta(k)) = \sum_{t=1}^k \left\{ \|y(t) - S\hat{x}(t)\|_{W_1}^2 + \|\eta(t)\|_{W_2}^2 \right\}, \quad (13)$$

где $W_1 > 0$, $W_2 > 0$ – весовые матрицы.

Оценка $\hat{\eta}(k)$ определяется из условий минимума критерия (13). В результате получим

$$\hat{\eta}(k) = (H^T W_1 H + W_2)^{-1} H^T W_1 [y(k) - S\hat{x}(k)]. \quad (14)$$

Отметим, что оценку вектора состояния системы определим по формуле (9) с заменой $\eta(k)$ на ее оценку:

$$\begin{aligned} \hat{X}(k+1) &= \begin{pmatrix} \hat{x}(k+1) \\ \hat{\theta}(k+1) \end{pmatrix} = F(\hat{X}(k)) + K(k)[y(k) - \bar{S}\hat{X}(k) - H\hat{\eta}(k)] = \\ &= \begin{pmatrix} f(\hat{x}(k), \hat{\theta}(k)) \\ \hat{\theta}(k) \end{pmatrix} - K(k)[y(k) - S\hat{x}(k) - H\hat{\eta}(k)]. \end{aligned} \quad (15)$$

Оценку (14) можно модифицировать, выполнив дополнительное сглаживание [11–13].

Для вычисления оценки сглаживания можно использовать метод взвешенного скользящего среднего [11]:

$$\hat{\eta}(k, \vec{g}) = \sum_{i=0}^{T_s-1} g_{T_s-i} \hat{\eta}(k-i), \quad (16)$$

где $g_i \geq 0$ – весовые коэффициенты, для которых

$$\sum_{i=0}^{T_s-1} g_{T_s-i} = 1. \quad (17)$$

Здесь в (16) вычисления реализуются от $k = T_s$ (T_s – целое число, определяющее величину скользящего интервала), значения g_i можно найти, решив оптимизационную задачу

$$J(\vec{g}) = \sum_{i=0}^{T_s-1} (\hat{\eta}(j) - \hat{\eta}(j, \vec{g}))^2 \Rightarrow \min_{g_i} \quad (18)$$

$$g_i \geq 0, \quad \sum_{i=1}^{T_s} g_i = 1, \quad i = \overline{1, T_s}. \quad (19)$$

при ограничениях

Для сглаживания можно также использовать алгоритм экспоненциального сглаживания [11].

Модификацию оценки можно выполнить на основе непараметрического сглаживания [12, 13], тогда оценка определится по формуле

$$\hat{\eta}^{(NP)}(k) = (H^T W_1 H + W_2)^{-1} H^T W \hat{\Omega}(k), \quad (20)$$

где j -я компонента вектора $\hat{\Omega}(k)$ имеет вид:

$$\hat{\Omega}_j(k) = \frac{\sum_{i=1}^k [y(i) - S\hat{x}(i)]_j G\left(\frac{k-i+1}{\mu_j}\right)}{\sum_{i=1}^k G\left(\frac{k-i+1}{\mu_j}\right)}. \quad (21)$$

В соотношении (21) $G(\cdot)$ является ядерной функцией, μ_j – коэффициенты размытости.

Для непараметрического сглаживания можно также использовать оценку (20), вычисленную на скользящем интервале T_s . В этом случае $\hat{\Omega}_j(k)$ вычисляется по формуле

$$\hat{\Omega}_j(k) = \frac{\sum_{i=k-T_s+1}^k [y(i) - S\hat{x}(i)]_j G\left(\frac{k-i+1}{\mu_j}\right)}{\sum_{i=k-T_s+1}^k G\left(\frac{k-i+1}{\mu_j}\right)}. \quad (22)$$

4. Результаты моделирования

Перейдем к рассмотрению результатов моделирования алгоритма идентификации в задаче передачи скрытой цифровой информации [14, 15] при аномальных ошибках в наблюдениях на основе

оценок неизвестного входа с дополнительным сглаживанием процесса $\hat{\eta}(k)$ при использовании в качестве детектора алгоритма идентификации (15).

Задача передачи скрытой информации реализуется с использованием модели аттрактора с неизвестным параметром при наличии в канале передачи информации дополнительных ненаблюдаемых аномальных помех. Модель аттрактора использовалась следующая [14]:

$$x(k+1) = f(\theta, x(k)) = \theta - 1 - \theta|x(k)|, \quad x(0) = x_0, \quad (23)$$

где θ – неизвестный параметр, $x(k)$ – переменная состояния. Канал передачи информации имеет вид:

$$y(k) = x(k) + \eta(k) + v(k), \quad (24)$$

где $\eta(k)$ и $v(k)$ – скалярные процессы ($\eta(k)$ – ненаблюдаемые аномальные помехи, $v(k)$ – последовательности нормально распределенных случайных величин с нулевым средним и математическим ожиданием: $M\{v(k)v(j)\} = V\delta_{kj}$).

Значения параметров модели задавались: $\theta = 1,55$ (соответствует передаваемому цифровому сигналу «0») и $\theta = 1,95$ (соответствует передаваемому цифровому сигналу «1»). Параметры модели принимались следующими:

$$x(0) = 0,1; \quad \hat{x}(0) = 0,12; \quad \hat{\theta}(0) = 1,75; \quad T_s = 3; \quad Q = 0; \quad P(0) = 10; \quad V = 0,01. \quad (25)$$

Для диагностики значения передаваемого сигнала можно использовать следующее правило:

Если значение оценки $\hat{\theta}(k)$ больше 1,75 (среднее арифметическое возможных значений θ), то считается, что передается сигнал верхнего уровня «1», если значение оценки $\hat{\theta}(k)$ меньше 1,75, то считается, что передается сигнал нижнего уровня «0».

Решение задачи идентификации выполнялось для пяти вариантов алгоритма идентификации:

- без учета оценок неизвестного входа;
- с учетом оценки неизвестного входа (14) (без сглаживания);
- с учетом оценки неизвестного входа с использованием экспоненциального сглаживания;
- с учетом оценки неизвестного входа с использованием непараметрического сглаживания (со скользящим интервалом);
- с учетом оценки неизвестного входа с использованием взвешенного скользящего сглаживания.

Моделирование выполнялось на временном интервале $[0, N]$ ($N = 50$) для следующих видов неизвестного входа (ненаблюдаемых аномальных ошибок) в канале наблюдений:

- неизвестный постоянный вход,
- неизвестный колебательный вход ($\eta(k) = \cos(2k)$),
- неизвестный ступенчатообразный вход вида:

$$\eta(k) = \begin{cases} -4, & 0 \leq k < 10, \\ -2, & 10 \leq k < 20, \\ 2, & 20 \leq k < 30, \\ 4, & 30 \leq k < 40, \\ 8, & 40 \leq k < 50. \end{cases} \quad (26)$$

Расчет среднеквадратических ошибок оценивания параметра θ выполнялся по формуле

$$\sigma = \sqrt{\frac{\sum_{k=1}^N (\theta - \hat{\theta}(k))^2}{N-1}}. \quad (27)$$

В таблице приведены среднеквадратические ошибки оценок идентификации σ для различных методов оценивания неизвестного входа в канале наблюдений. Усреднение осуществлялось по 100 реализациям.

При непараметрическом сглаживании использовалась ядерная функция гауссовского вида:

$$G(u) = \frac{\exp\left(-\frac{u^2}{2}\right)}{\sqrt{2\pi}}.$$

Среднеквадратические ошибки оценок идентификации 6 параметра θ

Неизвестный вход	Метод				
	Без учета оценок неизвестного входа	С учетом оценки неизвестного входа без сглаживания (15)	С учетом оценки неизвестного входа с использованием экспоненциального сглаживания	С учетом оценки неизвестного входа с использованием непараметрического сглаживания	С учетом оценки неизвестного входа с использованием взвешенного скользящего сглаживания
Постоянный	1,644	0,222	0,085	0,077	0,071
Колебательный	1,277	0,211	0,152	0,116	0,084
Ступенчатообразный	1,636	0,202	0,361	0,087	0,067

На рис. 1, 2 приведены результаты моделирования для случая неизвестного входа в виде ступенчатой функции при передаче сигнала верхнего уровня «1» (см. рис. 1) и при передаче сигнала нижнего уровня «0» (см. рис. 2). Результаты моделирования приведены для двух случаев:

- оценки неизвестного входа вычислены без дополнительного сглаживания (см. рис 1, а, 2, а);
- оценки неизвестного входа вычислены с использованием дополнительного сглаживания по алгоритму взвешенного скользящего среднего (см. рис 1, б, 2, б).

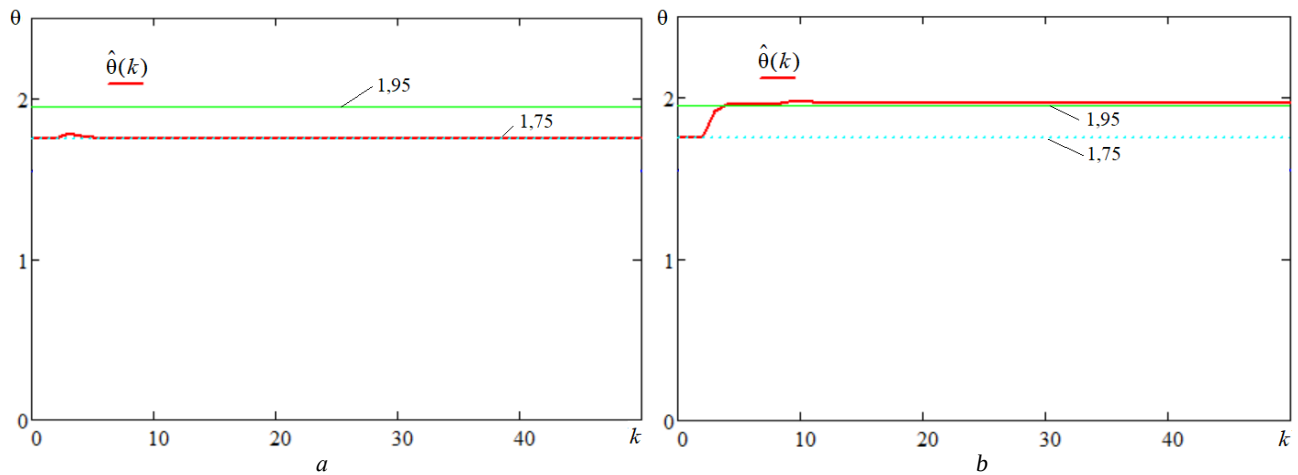


Рис. 1. Оценка параметра $\hat{\theta}(k)$ при неизвестном входе в виде ступенчатой функции без использования дополнительного сглаживания (а), с использованием дополнительного сглаживания (алгоритм (16)) (б) (передается сигнал верхнего уровня «1»)

Рис. 1. Estimation of the $\hat{\theta}(k)$ with an unknown input in the form of a step function without using additional smoothing (а), using additional smoothing (algorithm (18)) (б) (an upper-level signal “1” is transmitted).

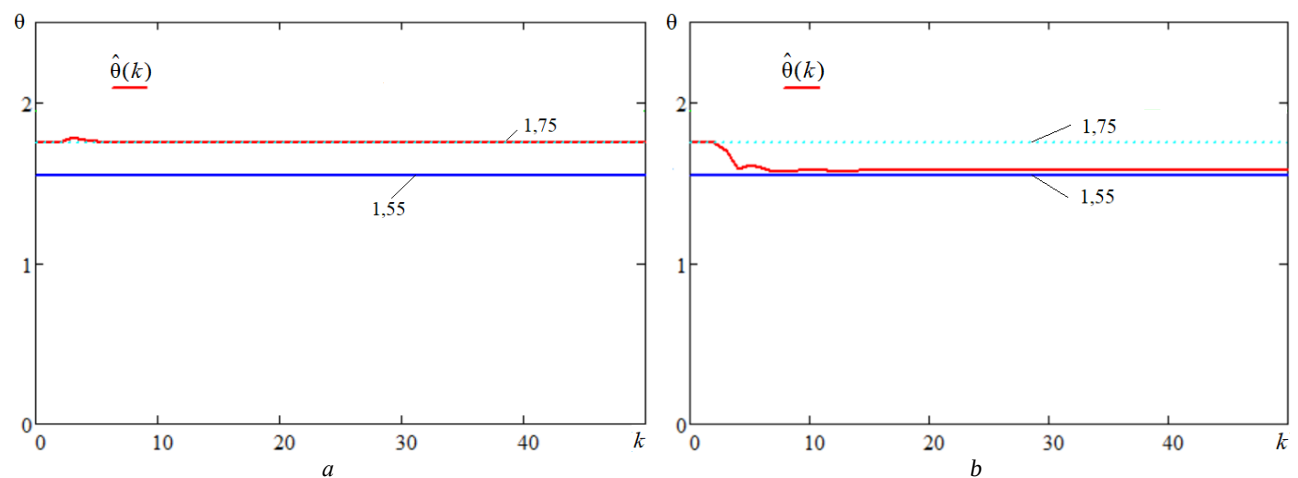


Рис. 2. Оценка параметра $\hat{\theta}(k)$ при неизвестном входе в виде ступенчатой функции без использования дополнительного сглаживания (а), с использованием скользящего сглаживания (алгоритм (16)) (б) (передается сигнал нижнего уровня «0»).

Рис. 2. Estimation of the $\hat{\theta}(k)$ with an unknown input in the form of a step function without using additional smoothing (а), using sliding smoothing (algorithm (18)) (б) (a signal of the lower level “0” is transmitted)

Из рисунков и таблицы видно, что при наличии ненаблюдаемых помех в канале наблюдений применение дополнительного сглаживания в алгоритме идентификации позволяет существенно повысить точность оценки $\hat{\theta}(k)$ и, как следствие, точность определения передаваемого цифрового двоичного сигнала. Из таблицы видно, что лучшая точность оценивания параметра θ достигается для алгоритмов с дополнительным использованием методов взвешенного скользящего сглаживания и непараметрического сглаживания.

Заключение

Рассматривается задача идентификации для дискретной стохастической системы с неизвестными аномальными ошибками в наблюдениях. Предложено для решения задачи идентификации использовать алгоритмы одновременного рекуррентного оценивания вектора состояния и неизвестного входа с использованием алгоритмов дополнительного сглаживания. На примере решения задачи передачи скрытой цифровой информации показано, что лучшая точность оценивания неизвестного параметра θ достигается при дополнительном использовании методов взвешенного скользящего сглаживания и непараметрического сглаживания оценок неизвестного входа в канале наблюдений.

Список источников

1. Astrom K.J. Introduction to stochastic control theory. New York : Academic Press, 1970. 322 с.
2. Sage A.P., Melsa J.L. Estimation Theory with Applications to Communications and Control. McGraw-Hill, 1971. 529 p.
3. Julier S.J., Uhlmann J.K., Durrant-Whyte H. A new approach for filtering nonlinear systems // Proc. of the American Control Conference. Seattle, WA, 1995. P. 1628–1632.
4. Särkkä S. On unscented Kalman filter for state estimation of continuous-time nonlinear systems. // IEEE Trans. Automat. Control. 2007. V. 52 (9). P. 1631–1641. doi: 10.1109/TAC.2007.904453
5. Arasaratnam I., Haykin S. Cubature Kalman filters // IEEE Trans. Automat. Control. 2009. V. 54 (6). P. 1254–1269. doi: 10.1109/TAC.2009.2019800
6. Цыганова Ю.В., Куликова М.В. Об эффективных методах параметрической идентификации линейных дискретных стохастических систем // Автоматика и телемеханика. 2012. № 6. С. 34–51.
7. Janczak D., Grishin Yu. State estimation of linear dynamic system with unknown input and uncertain observation using dynamic programming // Control and Cybernetics. 2006. V. 4. P. 851–862.
8. Дёмин Н.С., Рожкова О.В., Рожкова С.В. Фильтрация в динамических системах по наблюдениям с памятью при наличии аномальных помех // Известия Томского политехнического университета. 2009. Т. 314, № 5. С. 16–20.
9. Xiao M., Zhang Y., Fu H. Three-stage unscented Kalman filter for state and fault estimation of nonlinear system with unknown input // Journal of the Franklin Institute. 2017. V. 354. P. 8421–8443. doi: 10.1016/j.jfranklin.2017.09.031
10. Hsieh C.-S. Unbiased minimum-variance input and state estimation for systems with unknown inputs: a system reformation approach // Automatica. 2017. V. 84. P. 236–240. doi: 10.1016/j.automatica.2017.06.037
11. Булашев С.В. Статистика для трейдеров. М. : Спутник+, 2003. 245 с.
12. Smagin V.I., Koshkin G.M. Kalman filtering and control algorithms for systems with unknown disturbances and parameters using nonparametric technique // Proc. 20th Int. Conference on Methods and Models in Automation and Robotics (MMAR 2015). 24–27 August. 2015. Miedzyzdroje, Poland, 2015. P. 247–251.
13. Koshkin G., Smagin V. Kalman filtering and forecasting algorithms with use of nonparametric functional estimators // Springer Proc. in Mathematics and Statistics. Nonparametric Statistics. 2016. V. 175. P. 75–84.
14. Ruan H., Zhai T., Yaz E. A Chaotic Secure Communication Scheme with Extended Kalman Filter Based Parameter Estimation // Proc. IEEE Int. Conference on Control Applications. Istanbul, Turkey, 2003. P. 404–408.
15. Sadeghian H., Salarieh H., Alasty A., Meghdari A. On the fractional-order extended Kalman filter and its application to chaotic cryptography in noisy environment // Applied Mathematical Modelling. 2014. V. 38, is. 3. P. 961–973.

References

1. Astrom, K.J. (1970) *Introduction to Stochastic Control Theory*. New York: Academic Press.
2. Sage, A.P. & Melsa, J.L. (1971) *Estimation Theory with Applications to Communications and Control*. McGraw-Hill.
3. Julier, S.J., Uhlmann, J.K. & Durrant-Whyte, H. (1995) A new approach for filtering nonlinear systems. *Proceedings of the American Control Conference*. Seattle. WA. pp. 1628–1632.
4. Särkkä, S. (2007) On unscented Kalman filter for state estimation of continuous-time nonlinear systems. *IEEE Trans. Automat. Control*. 52(9). pp. 1631–1641. DOI: 10.1109/TAC.2007.904453

5. Arasaratnam, I. & Haykin, S. (2009) Cubature Kalman filters. *IEEE Trans. Automat. Control*. 54(6). pp. 1254–1269. DOI: 10.1109/TAC.2009.2019800
6. Tsyganova, Yu.V. & Kulikova, M.V. (2012) Ob effektivnykh metodakh parametricheskoy identifikatsii lineynykh diskretnykh stokhasticheskikh system [On effective methods of parametric identification of linear discrete stochastic systems]. *Avtomatika i telemekhanika*. 6. pp. 34–51.
7. Janczak, D. & Grishin, Yu. (2006) State estimation of linear dynamic system with unknown input and uncertain observation using dynamic programming. *Control and Cybernetics*. 4. pp. 851–862.
8. Demin, N.S., Rozhkova, O.V. & Rozhkova, S.V. (2009) Fil'tratsiya v dinamicheskikh sistemakh po nablyudeniyam s pamyat'yu pri nalichii anomal'nykh pomekh [Filtering in dynamic systems based on observations with memory in the presence of abnormal interference]. *Izvestiya Tomskogo politekhnicheskogo universiteta*. 314(5). pp.16–20.
9. Xiao, M., Zhang, Y. & Fu, H. (2017) Three-stage unscented Kalman filter for state and fault estimation of nonlinear system with unknown input. *Journal of the Franklin Institute*. 354. pp. 8421–8443. DOI: 10.1016/j.jfranklin.2017.09.031
10. Hsieh, C.-S. (2017) Unbiased minimum-variance input and state estimation for systems with unknown inputs: A system reformation approach. *Automatica*. 84. pp. 236–240. DOI: 10.1016/j.automatica.2017.06.037
11. Bulashev, S.V. (2003) *Statistika dlya treyderov* [Statistics for traders]. Moscow: Sputnik+.
12. Smagin, V.I. & Koshkin, G.M. (2015) Kalman filtering and control algorithms for systems with unknown disturbances and parameters using nonparametric technique. *Proceedings 20th Int. Conference on Methods and Models in Automation and Robotics (MMAR 2015)*. August 24–27. Miedzyzdroje. Poland. pp. 247–251.
13. Koshkin, G. & Smagin, V. (2016) Kalman filtering and forecasting algorithms with use of nonparametric functional estimators. *Springer Proceedings in Mathematics and Statistics. Nonparametric Statistics*. 175. pp. 75–84.
14. Ruan, H., Zhai, T. & Yaz, E. (2003) A Chaotic Secure Communication Scheme with Extended Kalman Filter Based Parameter Estimation. *Proceedings IEEE Int. Conference on Control Applications*. Istanbul, Turkey. pp. 404–408.
15. Sadeghian, H., Salarieh, H., Alasty, A. & Meghdari, A. (2014) On the fractional-order extended Kalman filter and its application to chaotic cryptography in noisy environment. *Applied Mathematical Modelling*. 38(3). pp. 961–973.

Информация об авторах:

Смагин Валерий Иванович – профессор, доктор технических наук, профессор кафедры прикладной математики Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: vsm@mail.tsu.ru

Ким Константин Станиславович – кандидат физико-математических наук, доцент Высшей IT школы Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: kks93@rambler.ru

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

Information about the authors:

Smagin Valery I. (Doctor of Technical Science, Professor, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: vsm@mail.tsu.ru

Kim Konstantin S. (Candidate of Physical and Mathematical Sciences, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: kks93@rambler.ru

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

Поступила в редакцию 24.04.2024; принята к публикации 03.09.2024

Received 24.04.2024; accepted for publication 03.09.2024

Научная статья

УДК 004.032.26, 004.048

doi: 10.17223/19988605/68/11

**Комбинированная модель локализации акустических источников
с применением технологии глубокого обучения****Джиах Михаил Шаход¹, Евгений Дмитриевич Агафонов²**^{1,2} *Сибирский федеральный университет, Красноярск, Россия*² *Сибирский государственный университет науки и технологий им. акад. М.Ф. Решетнева, Красноярск, Россия*¹ *ghiathlovealaa@gmail.com*² *evgeny.agafonov@mail.ru*

Аннотация. Представлена модель глубокого обучения для локализации акустического источника, которая рассматривает локализацию как проблему классификации направлений источников. Предложен подход, основанный на интеграции признаков акустической интенсивности и признаков GCC-PHAT (Generalized cross-correlation – phase transform) в качестве входных данных для сверточных нейронных сетей. С учетом условий моделирования были созданы наборы данных с целью обучения, валидации и тестирования модели с пространственным разрешением 10° и 2°. Результаты моделирования продемонстрировали эффективность предложенной модели при локализации источника с высокой точностью в замкнутой среде и при наличии реверберации. Предложенная модель с разрешением 10° превзошла модель, принимавшую только признаки акустической интенсивности в качестве входных данных, достигнув улучшения точности на 6,57% и точности прогнозирования на 2,86%, в то время как модель с разрешением 2° достигла улучшения точности на 15,57% и точности прогнозирования на 2,04%.

Ключевые слова: глубокое обучение; локализация акустического источника; классификация; акустическая интенсивность; сверточная нейронная сеть.

Для цитирования: Шаход Д.М., Агафонов Е.Д. Комбинированная модель локализации акустических источников с применением технологии глубокого обучения // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 68. С. 100–111. doi: 10.17223/19988605/68/11

Original article

doi: 10.17223/19988605/68/11

A combined model for localizing acoustic sources using deep learning technology**Ghiath M. Shahoud¹, Evgeny D. Agafonov²**^{1,2} *Siberian Federal University, Krasnoyarsk, Russian Federation*² *Reshetnev Siberian State University of Science and Technology, Krasnoyarsk, Russian Federation*¹ *ghiathlovealaa@gmail.com*² *evgeny.agafonov@mail.ru*

Abstract. This paper presents a deep learning model for sound source localization, which considers localization as a source direction classification problem. An approach is proposed based on the integration of sound intensity features and GCC-PHAT (Generalized cross-correlation - phase transform) features as input data for convolutional neural networks. Taking into account the modeling conditions, datasets were created for the purpose of training, validating and testing the model with spatial resolutions of 10° and 2°. Simulation results demonstrated the effectiveness of the proposed model in localizing the source with high accuracy in a closed environment and in the presence of reverberation. The proposed model with a resolution of 10° outperformed the model fed with only sound intensity features as input features, achieving improvement in accuracy by 6,57% and in prediction accuracy by 2,86%, while the model with a resolution of 2° achieved an improvement in accuracy by 15,57% and in prediction accuracy by 2,04%.

Keywords: deep learning; sound source localization; classification; sound intensity; convolutional neural network.

For citation: Shahoud, G.M., Agafonov, E.D. (2024) A combined model for localizing acoustic sources using deep learning technology. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 68. pp. 100–111. doi: 10.17223/19988605/68/11

Введение

Задачу локализации акустических источников (Sound Source Localization; SSL) можно определить в одном из ее видов как нахождение направления акустических источников или объектов, отражающих акустические сигналы, что может применяться в различных сложных средах. Необходимость решения этой задачи возникает во многих областях и приложениях техники и технологии. В частности, SSL широко применяется при автоматическом слежении за камерой для телеконференций, взаимодействии человека с роботом, распознавании речи на расстоянии, громкой связи. Также локализация имеет исключительную важность в геофизике и неразрушающем контроле материалов и изделий.

Для регистрации акустических сигналов с дальнейшим анализом их характеристик применяются микрофонные решетки, которые состоят из набора микрофонов, расположенных в пространстве определенным образом для получения пространственной информации об акустическом источнике. Пространственно-временная информация, полученная с микрофонной решетки, может использоваться для оценки различных параметров источника или извлечения предполагаемого исходного сигнала.

Первоначально задача локализации акустических источников решалась с помощью традиционных методов обработки сигналов, таких как Time Difference of Arrival (TDOA) [1], Delay-And-Sum beamformer (DAS) [2], Multiple Signal Classification (MUSIC) [3] и Generalized cross-correlation – phase transform (GCC-PHAT) [4]. Однако эти методы имеют недостатки, связанные со сложностью акустических характеристик окружающей среды. В последние годы, с появлением и развитием методов глубокого обучения (Deep Learning, DL), таких как Convolutional Neural Networks (CNN) [5], Recurrent Neural Network (RNN) [6], Convolutional Recurrent Neural Network (CRNN) [7] и Residual Neural Networks (ResNet) [8], и их широким использованием в области акустических приложений, был наметен новый вектор развития направления локализации источников.

Основным преимуществом SSL на основе методов DL является включение в процесс обучения информации об акустических свойствах, которые входят в число признаков, в то время как традиционные методы основаны только на пространственной информации [4]. В результате методы, основанные на данных, такие как DL, могут превосходить традиционные методы за счет работы с большим объемом данных, реальных или смоделированных. Эффективность модели DL определяется ее способностью обобщать различные аспекты конфигурации (например, расстояние между источником и микрофонной решеткой, уровни шума, время реверберации и т.д.), а именно правильно классифицировать новые данные с признаками, отличающимися от полученных в ходе обучения.

Процедуры представления и подготовки данных, эффективное извлечение признаков и способность этих признаков к высокоточной оценке местоположения источников являются методологией, предложенной в данной статье.

Использование вектора акустической интенсивности (Sound Intensity, SI) для локализации акустического источника на основе DL было впервые предложено в [9], где была показана превосходящая эффективность по сравнению с традиционными методами. SI как один из входных признаков для методов DL доказала свою способность работать в условиях шума и реверберации при использовании микрофонных решеток небольшого размера [10], однако не удалось удовлетворительно обобщить глубокую модель при изменении некоторых условий моделирования, особенно расстояния между источником и центром решетки. С другой стороны, признаки GCC-PHAT доказали свою способность к обобщению [11], поскольку они зависят от пространственной информации, тогда как акустическая интенсивность зависит от акустических характеристик.

Цель, поставленная в работе, – выработка универсального и эффективного подхода к решению проблемы локализации путем интеграции признаков SI и признаков GCC-PHAT в качестве входных данных для методов DL для достижения преимуществ обеих разновидностей признаков.

1. Общий принцип создания систем локализации акустического источника на основе методов глубокого обучения

Общий принцип создания систем SSL на основе методов DL схематично представлен на рис. 1. Признаки, представляющие входные данные глубоких нейронных сетей (Deep Neural Network; DNN), извлекаются из многоканальных акустических сигналов, записанных с помощью микрофонной решетки. Входные признаки передаются в DNN, которая дает оценку местоположения акустического источника.

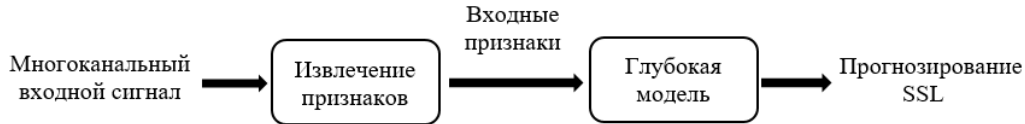


Рис. 1. Схема локализации акустического источника на основе глубокого обучения

Fig. 1. Deep learning-based sound source localization scheme

Многоканальные сигналы содержат информацию о местоположении источника. Фактически микрофоны в решетке всегда расположены существенно ближе друг к другу, чем расстояние от решетки до источника. При этом сигналы, полученные с микрофонов, демонстрируют различия с точки зрения задержки и амплитуды. Такие различия между сигналами обусловлены различными путями распространения звука от источника к различным микрофонам решетки. Это имеет место как для прямого пути (прямая видимость между источником и микрофоном), так и для многочисленных отражений, которые составляют эхо в помещении (реверберация). Другими словами, в случае допустимости линейного описания процессов распространения звука в среде сигнал источника $s_j(t)$ сворачивается с импульсной характеристикой помещения $a_{i,j}(t)$ (Room Impulse Response; RIR), которая, в свою очередь, зависит от местоположения источника, местоположения и направления микрофона.

Когда источник возбуждает идеальный импульс, все отражения, достигающие микрофонов, вместе составляют RIR, которая в идеале представляет собой последовательность импульсов, каждый из которых соответствует одному отражению [12]. Время, в течение которого уровень RIR падает на 60 дБ, определяется как время реверберации (RT60) [13].

Сигнал микрофона формулируется следующими соотношениями [14]:

$$x_{i,j}(t) = a_{i,j}(t) * s_j(t) + n_i(t), \quad (1)$$

$$x_{i,j}(t) = \sum_{\tau=0}^{T-1} a_{i,j}(\tau) s_j(t - \tau) + n_i(t), \quad (2)$$

где $x_{i,j}(t)$ – полезный сигнал, воспринятый микрофоном i от источника j , $n_i(t)$ – фоновый шум и, возможно, шум микрофона, $*$ обозначает свертку. Сигналы являются цифровыми, t и τ – дискретные временные индексы, а T – эффективная длина RIR.

Основная предпосылка разработки систем SSL на основе DNN заключается в том, что в условиях, когда взаимосвязь между информацией в многоканальном сигнале и местоположением источника в целом сложна (особенно в многоисточниковой реверберирующей и шумной конфигурации), DNN представляет собой эффективную модель, способную автоматически идентифицировать и использовать эту взаимосвязь, если она обучена на большом количестве репрезентативных примеров.

2. Процедуры представления и подготовки данных для систем глубокого обучения

Предполагается, что плоская ортогональная микрофонная решетка состоит из четырех направленных микрофонов, расположенных, как показано на рис. 2. Микрофоны M_1 и M_3 расположены вдоль оси x ортогонально микрофонам M_2 и M_4 вдоль оси y , d представляет размер решетки микрофонов, а O – центр решетки. Используется модель распространения в дальнем поле [15], где направлением акустического источника является угол θ , определенный по отношению к положительной оси x , т.е. $\theta \in [-180^\circ, 180^\circ]$.

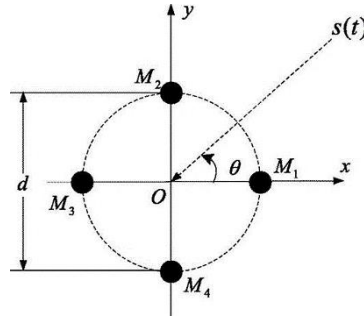


Рис. 2. Конфигурация микрофонной решетки
Fig. 2. Microphone array configuration

Можно определить акустическую интенсивность как энергию, переносимую акустическими волнами на единицу площади в направлении r , перпендикулярном этой площади, и, следовательно, эта величина содержит информацию о местонахождении источника. Уравнение (3) связывает скорость частиц $v_r(t)$ в направлении r и акустическое давление $p(t)$ [10]:

$$v_r(t) = -\frac{1}{\rho} \int_{-\infty}^t \frac{\partial p(\tau)}{\partial r} d\tau, \quad (3)$$

где ρ – плотность воздуха. Градиент давления можно аппроксимировать конечной разностью между двумя близкими точками на расстоянии Δr по направлению r .

$$v_r(t) \approx -\frac{1}{\Delta r \rho} \int_{-\infty}^t [p_{r2}(t) - p_{r1}(t)] dt. \quad (4)$$

Таким образом, заменив две точки в предыдущей аппроксимации расположением каждой пары микрофонов (M_1 – M_3) и (M_2 – M_4), можно выразить (4) в частотно-временной области как [16]

$$V_x(\omega, t) = j \frac{[P_3(\omega, t) - P_1(\omega, t)]}{\omega \rho d}, \quad (5)$$

$$V_y(\omega, t) = j \frac{[P_4(\omega, t) - P_2(\omega, t)]}{\omega \rho d}, \quad (6)$$

где $P_i(\omega, t)$ – кратковременное преобразование Фурье (Short-time Fourier transform; STFT) для сигнала на микрофоне M_i , j – мнимая единица. Акустическое давление в центре декартовых координат O можно оценить путем усреднения сигналов, полученных от всех микрофонов в частотно-временной области, оно формулируется следующим образом [16]:

$$P_O(\omega, t) = \frac{P_1(\omega, t) + P_2(\omega, t) + P_3(\omega, t) + P_4(\omega, t)}{4}. \quad (7)$$

Можно выразить SI в обоих направлениях x и y как [17]

$$I_{Ox}(\omega, t) = P_O(\omega, t) V_x^*(\omega, t), \quad (8)$$

$$I_{Oy}(\omega, t) = P_O(\omega, t) V_y^*(\omega, t), \quad (9)$$

где $*$ указывает на комплексное сопряжение.

Информация о местоположении встроена в действительную часть SI [17], поэтому извлеченные признаки SI для направлений x и y могут быть выражены как

$$I_x(\omega, t) = \Re[I_{Ox}(\omega, t)], \quad (10)$$

$$I_y(\omega, t) = \Re[I_{Oy}(\omega, t)]. \quad (11)$$

В [10] было предложено использовать отбеливающий вес (whitening weight), который выражается в уравнении (12), для решения как проблемы реверберации помещения, так и проблемы аддитивного шума:

$$W(\omega, t) = \sqrt{|P_O(\omega, t)|^2 + \beta(|V_x(\omega, t)|^2 + |V_y(\omega, t)|^2)}. \quad (12)$$

Также было доказано, что этот вес обеспечивает наилучшую эффективность локализации при $\beta = 10^6$. Следовательно, можно переписать полученные признаки для направлений и как

$$I_x^{(W)}(\omega, t) = \frac{I_x(\omega, t)}{W(\omega, t)}, \quad (13)$$

$$I_y^{(W)}(\omega, t) = \frac{I_y(\omega, t)}{W(\omega, t)}. \quad (14)$$

Подмассивы представляют собой группы из трех соседних микрофонов ($M_4-M_1-M_2$), ($M_1-M_2-M_3$), ($M_2-M_3-M_4$) и ($M_3-M_4-M_1$). Признаки извлекаются для ($M_4-M_1-M_2$) следующим образом. Скорость частиц в направлениях (M_2-M_1) и (M_4-M_1) определяется выражениями

$$V_{M_2, M_1}(\omega, t) = j\sqrt{2} \frac{[P_1(\omega, t) - P_2(\omega, t)]}{\omega p d}, \quad (15)$$

$$V_{M_4, M_1}(\omega, t) = j\sqrt{2} \frac{[P_1(\omega, t) - P_4(\omega, t)]}{\omega p d}, \quad (16)$$

где M_1 – общий микрофон. Таким образом, признаки SI можно записать как

$$I_{M_2, M_1}^{(W)}(\omega, t) = \frac{1}{W_{M_1}(\omega, t)} \Re [P_{M_1}(\omega, t) V_{M_2, M_1}^*(\omega, t)], \quad (17)$$

$$I_{M_4, M_1}^{(W)}(\omega, t) = \frac{1}{W_{M_1}(\omega, t)} \Re [P_{M_1}(\omega, t) V_{M_4, M_1}^*(\omega, t)], \quad (18)$$

где $W_{M_1}(\omega, t)$ и $P_{M_1}(\omega, t)$ выражаются следующим образом:

$$W_{M_1} = \sqrt{|P_{M_1}(\omega, t)|^2 + \beta \left(|V_{M_2, M_1}(\omega, t)|^2 + |V_{M_4, M_1}(\omega, t)|^2 \right)}, \quad (19)$$

$$P_{M_1}(\omega, t) = \frac{P_4(\omega, t) + P_1(\omega, t) + P_2(\omega, t)}{3}. \quad (20)$$

Аналогично извлекаются признаки для остальных соседних микрофонов в решетке. Можно определить тензор δ размером $M \times N \times C$, чтобы собрать все признаки вместе. M представляет количество кадров преобразования STFT, N – элементы разрешения по частоте, а $C = 10$, поскольку основной массив плюс четыре подмассива генерируют признаки SI в двух направлениях.

$$\delta(:, :, i) = I_{(K)}^{(W)}, \quad (21)$$

$$K = [x, y, (M_2, M_1), (M_4, M_1), (M_1, M_2), (M_3, M_2), (M_2, M_3), (M_4, M_3), (M_3, M_4), (M_1, M_4)]. \quad (22)$$

GCC-PHAT рассчитывается как обратное преобразование Фурье перекрестного спектра мощности (Cross-Power Spectrum; CPS) между сигналами двух микрофонов.

Пусть x_i и x_j будут двумя сигналами, CPS определяется как [18]

$$CPS = \frac{X_i(f) X_j^*(f)}{|X_i(f) X_j^*(f)|}, \quad (23)$$

где $X_i(f)$ и $X_j(f)$ – N -точечные дискретные преобразования Фурье двух сигналов.

Тогда GCC-PHAT определяется как

$$\hat{R}_{PHAT}(\tau) = \sum_{f=0}^{F-1} \frac{X_i(f) X_j^*(f)}{|X_i(f) X_j^*(f)|} e^{j2\pi \left(\frac{f\tau}{N} \right)}, \quad (24)$$

где τ – задержка между сигналами каждой пары микрофонов в решетке.

3. Предлагаемая модель

Спектрограммы акустических сигналов, как известно, имеют локальные взаимосвязи во временной и частотной областях, а способность CNNs эффективно моделировать эти взаимосвязи че-

рез локальные соединения сделала их широко используемыми для локализации акустических источников.

На рис. 3 показана структура предлагаемой модели, представленная объединением двух CNN с использованием полносвязных слоев (Fully Connected; FC) с целью классификации.

Каждая CNN имеет одинаковую структуру, представленную входным слоем (признаки SI с размерностью $14 \times 511 \times 10$ представляют собой входной слой для одной, а признаки GCC-PHAT с размерностью $16 \times 96 \times 6$ – входной слой для другой) и двумя двумерными сверточными слоями Conv2D (каждый сверточный слой использует 64 сверточных ядра размера 3×3 с размером шага 1×1 для обучения локальных взаимосвязей между извлеченными признаками), после каждого сверточного слоя используется слой пакетной нормализации (Batch Normalization) для улучшения стабильности сети и ускорения сходимости моделей.

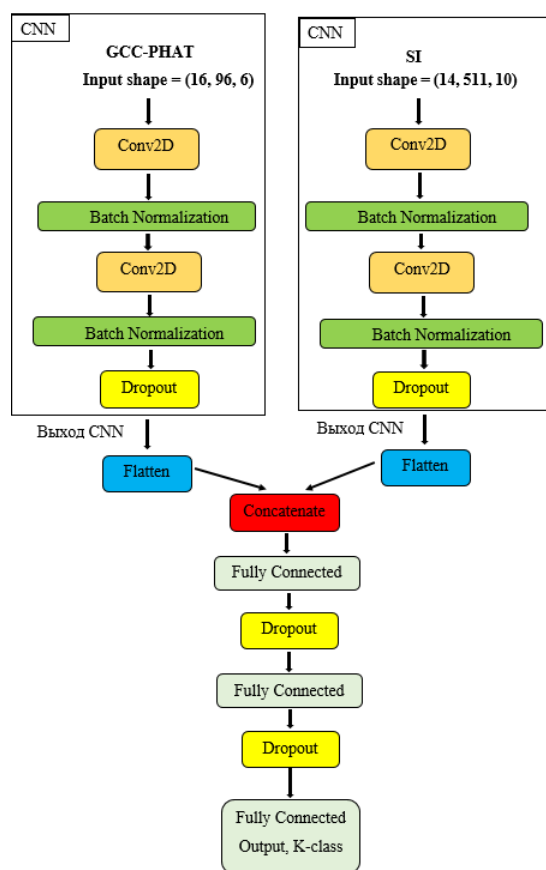


Рис. 3. Структура предлагаемой модели

Fig. 3. Structure of the proposed model

К выходным данным каждого блока CNN был применен сглаживающий слой (Flatten). Для объединения выходных данных каждой CNN использовался слой конкатенации (Concatenate). На выходе слоя Concatenate использовалось два полносвязных слоя, и после каждого полносвязного слоя производилась процедура регуляризации Dropout, чтобы избежать переобучения [19]. Функция ReLU использовалась в качестве функции активации для сверточных слоев и полносвязных слоев. В предлагаемой архитектуре выходной слой представляет собой полносвязный слой (классификационный слой) и состоит из K нейронов (K -class), в этом слое использовалась Softmax в качестве функции активации.

Было обучено две модели: одна с $K = 36$, поскольку она классифицирует направление источника с пространственным разрешением 10° , а другая с $K = 180$, поскольку она классифицирует направление с разрешением 2° .

В качестве оптимизатора (алгоритма обучения) был использован Adam с инициализированной скоростью обучения, равной 10^{-5} . Перекрестная энтропия (Cross-Entropy) использовалась как функ-

ция потерь. Максимальное количество эпох обучения выбрано равным 200, при этом используется ранняя остановка (Early Stopping) с терпением (patience) 25 эпох (количество эпох, в течение которых не происходит улучшения точности валидации и после которых обучение будет остановлено), чтобы избежать переобучения [20].

4. Моделирование

Будем считать, что задача локализации решается в замкнутой среде, представленной помещением с размерами $15 \times 9 \times 4$ м³. Центр ортогональной микрофонной решетки расположен в позиции (7,5; 4,5; 1,5) м с размером решетки $d = 0,2$ м. Расстояние между источником и центром решетки установлено равным 2 м, а акустический источник находится на той же высоте, что и решетка (1,5 м).

В ходе моделирования рассматривается два случая с пространственным разрешением направления источника: 10° и 2° . Когда пространственное разрешение установлено равным 10° , общее количество потенциальных направлений источников составляет $K = 36$ и равномерно распределено от -180 до 170° . Для разрешения 2° общее количество потенциальных направлений источников составляет $K = 180$ и равномерно распределено от -180 до 178° .

Сигналы акустических источников представляют собой аудиофайлы из базы данных TIMIT [21] и выбираются случайным образом с длительностью 500 мс (с целью уменьшения времени обработки данных) и частотой дискретизации 16 кГц.

RIRs между источниками и микрофонами генерируются с помощью программного обеспечения RIR GENERATOR [22], которое основано на методе источника изображения [23]. Принимая длину RIR равной 4 096 отсчетов и $RT60 = 0,36$ с, максимальный порядок отражений звука автоматически задается программой обеспечения RIR GENERATOR.

Случайным образом и с равномерным распределением генерируется 6 000 и 1 000 направлений для обучения и валидации модели соответственно, чтобы все потенциальные направления были охвачены равномерно. Для синтеза обучающих данных из обучающего набора базы данных TIMIT случайным образом выбирается 300 предложений, каждое предложение соответствует 20 различным сгенерированным направлениям. Предложения, связанные с направлениями, представляют собой акустические источники, которые свернуты с соответствующими RIRs для формирования микрофонных сигналов (исходные данные).

Метки, соответствующие каждому направлению, определяются для обеих моделей путем нахождения класса, соответствующего наименьшей разнице между значением направления (взятым из 6 000 направлений) и значениями направлений в классах, соответствующих моделям.

Еще 100 предложений используется в качестве данных валидации, и каждое предложение соответствует 10 различным направлениям.

Для тестирования модели с разрешением 10° из тестовых данных базы данных TIMIT случайным образом выбирается 10 предложений, каждое предложение соответствует 36 различным направлениям.

Для тестирования модели с разрешением 2° из тестовых данных базы данных TIMIT случайным образом выбирается 10 предложений, каждое предложение соответствует 180 различным направлениям.

Итоговый объем обучающей выборки составил 6 000 образцов для каждой из двух моделей, объем валидационного набора данных – 1 000 образцов, объем тестовых данных – 360 образцов для модели с разрешением 10° и 1 800 образцов для модели с разрешением 2° .

Для извлечения признаков SI к исходным данными (сигналы четырех микрофонов в каждом образце) было применено преобразование STFT с окном Ханнинга шириной 1 024 точек, что соответствует временной длине в 64 мс и 511 элементам разрешения по частоте. Для увеличения размера выборки была проведена аугментация данных, заключающаяся в перекрытии временных сигналов на 50%. Полученные спектрограммы сигналов микрофона имеют размерность 14×511 . На основании формулы (21) было зафиксировано два других измерения тензора δ : $M = 14$ и $N = 511$. В итоге полученные

признаки SI для каждого образца имеют размерность $14 \times 511 \times 10$, что соответствует размерности первого входного слоя модели.

Признаки GCC-PHAT извлекаются для каждой пары микрофонных сигналов в каждом образце, т.е. количество пар в каждом образце равно 6. К этим сигналам было применено STFT с окном Ханнинга шириной 2 048 точек с перекрытием 25%.

В предложенной модели наибольшая временная задержка между сигналами микрофонов составляет примерно 0,58 мс, что соответствует примерно 9 отсчетам. То есть чтобы использовать алгоритм GCC-PHAT для поиска временных задержек между сигналами, достаточно представить GCC-PHAT на 18 отсчетов из-за его симметрии. Чтобы использовать признаки GCC-PHAT в качестве входных данных для модели, их следует представить в несколько большем количестве отсчетов, чтобы получить больше пользы от дифференциации этих признаков между микрофонами.

Признаки GCC-PHAT, извлеченные для каждой пары микрофонных сигналов, представлены в виде матрицы 16×96 (16 – количество кадров, а 96 – количество отсчетов, на которых представлена функция GCC-PHAT), поскольку каждый кадр соответствует 96 отсчетам; значения этих отсчетов представляют собой амплитуды. В итоге полученные признаки GCC-PHAT для каждого образца имеют размерность $16 \times 96 \times 6$, что соответствует размерности второго входного слоя модели.

5. Метрики оценки

5.1. Точность (Accuracy)

Точность относится к количеству направлений источника, которые модель правильно прогнозирует, из общего количества направлений источника, включенных в тестовый набор данных, и определяется в процентах следующей формуле:

$$Accuracy(\%) = \frac{N_c}{N_s} \times 100, \quad (25)$$

где N_s представляет количество оцениваемых направлений источника, а N_c указывает количество правильно спрогнозированных направлений источника.

4.2. Точность прогнозирования (Prediction Accuracy; PA)

Точность прогнозирования определяется следующим образом:

$$PA(\%) = \frac{N_p}{N_s} \times 100, \quad (26)$$

где N_s – количество оцениваемых направлений источника, а N_p – количество правильно распознанных направлений источника. При этом направление источника считается правильно распознанным, если прогнозируемое направление находится в пределах пространственного разрешения модели, т.е. отклонение прогнозируемого направления от фактического находится в пределах $\pm\theta_0$ для пространственного разрешения θ_0 [11].

6. Оценка эффективности моделей

График точности моделей в процессе обучения на обучающей и валидационной выборках представлен на рис. 4. Модель с разрешением 10° достигла точности 99,3% на валидационной выборке и 99,97% на обучающей выборке в процессе 55 эпох обучения, а модель с разрешением 2° достигла точности 95,8% на валидационной выборке и 99,85% на обучающей выборке в процессе 200 эпох обучения. Стоит отметить, что в начале обучения точность валидации выше точности обучения. Это можно объяснить использованием слоев Dropout в архитектуре модели, где точность обучения рассчитывается с учетом действующего процесса исключения, в этом случае некоторая информация

о каждой выборке теряется (не все признаки используются), и, таким образом, это может в некоторой степени снизить точность обучения. Однако в процессе валидации Dropout неактивно и все блоки модели доступны, поэтому модель имеет полную вычислительную мощность и, следовательно, может работать лучше, чем при обучении. Но по мере увеличения эпох обучения модель становится более стабильной, эффективно обучается, а точность обучения становится незначительно выше точности валидации.

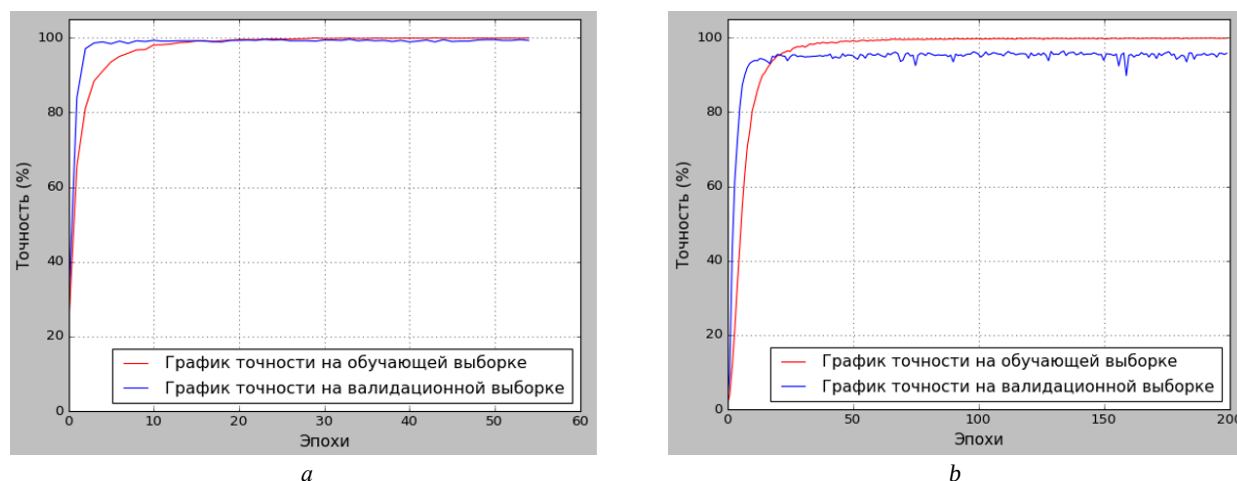


Рис. 4. Оценка эффективности моделей: *a* – график точности модели с разрешением 10° в процессе обучения на обучающей и валидационной выборках; *b* – график точности модели с разрешением 2° в процессе обучения на обучающей и валидационной выборках

Fig. 4. Assessment of models efficiency: *a*) graph of model accuracy with a resolution of 10° during the learning process on training and validation samples; *b*) graph of model accuracy with a resolution of 2° during the training process on the training and validation samples

Эффективность предложенной модели сравнивается с моделью SI-CNN, предложенной в [10], для обоих пространственных разрешений: 10° и 2° .

В модели SI-CNN использовались только признаки SI в качестве входных данных для CNN. Модель SI-CNN обучалась и тестировалась в тех же условиях моделирования и на том же наборе данных обучения и тестирования, которые учитывались при обучении и тестировании предложенной модели. В таблице представлены значения метрик оценки эффективности обученных моделей, которые рассчитываются в ходе тестирования моделей.

Значения метрик оценки эффективности моделей в ходе тестирования

Модель	Accuracy, %	PA, %
SI-CNN с разрешением 10°	93,06	97,22
Предлагаемая с разрешением 10°	99,17	100
SI-CNN с разрешением 2°	82,39	97,83
Предлагаемая с разрешением 2°	95,22	99,83

Значения метрик, представленные в таблице, показывают, что предлагаемая модель превосходит модель SI-CNN при обоих пространственных разрешениях: 10° и 2° .

Для разрешения 10° предложенная модель достигла улучшения показателя точности на 6,57% и точности прогнозирования на 2,86%.

Для разрешения 2° предложенная модель достигла улучшения показателя точности на 15,57% и точности прогнозирования на 2,04%.

Заключение

В работе предложена новая классификационная модель локализации акустического источника, основанная на интеграции признаков акустической интенсивности и признаков GCC-PHAT. Структура

глубокой модели представлена объединением двух CNN с использованием полносвязных слоев, где в одну из CNN подавались признаки акустической интенсивности, а в другую – признаки GCC-RNAT.

Было рассмотрено две модели, одна из которых предназначена для оценки направления акустического источника с пространственным разрешением 10° , а другая – с разрешением 2° . В процессе настройки и валидации модели данные были взяты из базы данных TIMIT с учетом условий моделирования, представленных размером помещения, расположением микрофонной решетки и расстоянием между источником и центром микрофонной решетки. Признаки были извлечены из исходных сигналов, и созданы наборы данных с целью обучения моделей, проверки их производительности в процессе обучения, а затем и их тестирования. Модели были построены с применением методов регуляризации.

Предложенная модель с разрешением 10° достигла точности 99,17% и точности прогнозирования 100% в рамках нахождения источников в пределах пространственного разрешения модели, превосходя глубокую модель SI-CNN, основанную на использовании только признаков акустической интенсивности в качестве входных данных, с улучшением точности на 6,57% и точности прогнозирования на 2,86%; модель с разрешением 2° достигла точности 95,22% и точности прогнозирования 99,83%, превосходя модель SI-CNN с улучшением точности на 15,57% и точности прогнозирования на 2,04%.

Список источников

1. Zhu N., Reza, T. A modified cross-correlation algorithm to achieve the time difference of arrival in sound source localization // *Measurement and Control*. 2019. V. 52 (3-4). P. 212–221. doi: 10.1177/0020294019827977
2. Chiariotti P., Martarelli M., Castellini P. Acoustic beamforming for noise source localization – Reviews, methodology and applications // *Mechanical Systems and Signal Processing*. 2019. V. 120. P. 422–448.
3. Zhong Y., Xiang J., Chen X., Jiang Y., Pang J. Multiple Signal Classification-Based Impact Localization in Composite Structures Using Optimized Ensemble Empirical Mode Decomposition // *Applied Sciences*. 2018. V. 8 (9). Art. 1447. doi: 10.3390/app8091447
4. Desai D., Mehendale N. A Review on Sound Source Localization Systems // *Archives of Computational Methods in Engineering*. 2022. V. 29 (7). P. 4631–4642. doi: 0.1007/s11831-022-09747-2
5. Nguyen T.N.T., Gan W.S., Ranjan R., Jones D.L. Robust source counting and DOA estimation using spatial pseudo-spectrum and convolutional neural network // *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. 2020. V. 28. P. 2626–2637. doi: 10.1109/TASLP.2020.3019646
6. Nguyen T.N.T., Nguyen N.K., Phan H., Pham L., Ooi K., Jones D.L., Gan W.S. A general network architecture for sound event localization and detection using transfer learning and recurrent neural network // *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Proc. (ICASSP)*. 2021. P. 935–939. doi: 10.1109/ICASSP39728.2021.9414602
7. Adavanne S., Politis A., Nikunen J., Virtanen T. Sound event localization and detection of overlapping sources using convolutional recurrent neural networks // *IEEE Journal of Selected Topics in Signal Proc.* 2018. V. 13 (1). P. 34–48.
8. He W., Motlicek P., Odobez J.M. Adaptation of multiple sound source localization neural networks with weak supervision and domain-adversarial training // *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Proc. (ICASSP)*. 2019. P. 770–774.
9. Perotin L., Serizel R., Vincent E., Guérin A. CRNN-based joint azimuth and elevation localization with the Ambisonics intensity vector // *2018 16th International Workshop on Acoustic Signal Enhancement (IWAENC)*. 2018. P. 241–245. doi: 10.1109/IWAENC.2018.8521403
10. Liu N., Chen H., Songgong K., Li Y. Deep learning assisted sound source localization using two orthogonal first-order differential microphone arrays // *The Journal of the Acoustical Society of America*. 2021. V. 149 (2). P. 1069–1084. doi: 10.1121/10.0003445
11. Li Q., Zhang X., Li H. Online direction of arrival estimation based on deep learning // *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2018. P. 2616–2620. doi: 10.1109/ICASSP.2018.8461386
12. Alpkocak A., Sis M. Computing impulse response of room acoustics using the ray-tracing method in time domain // *Archives of Acoustics*. 2010. V. 35 (4). P. 505–519. doi: 10.2478/v10168-010-0039-8
13. Ciaburro G., Iannace G. Acoustic characterization of rooms using reverberation time estimation based on supervised learning algorithm // *Applied Sciences*. 2021. V. 11 (4). Art.1661. https://doi: 10.3390/app11041661
14. Grumiaux P.A., Kitić S., Girin L., Guérin A. A survey of sound source localization with deep learning methods // *The Journal of the Acoustical Society of America*. 2022. V. 152 (1). P.107–151.
15. Siano D., Viscardi M., Panza M.A. Experimental acoustic measurements in far field and near field conditions: characterization of a beauty engine cover // *Recent Advances in Fluid Mechanics and Thermal Engineering*. 2014. P. 50–57. URL: https://core.ac.uk/download/pdf/81675242.pdf
16. Jacobsen F. Sound intensity // *Springer Handbook of Acoustics*. New York : Springer, 2014. P. 1093–1114.
17. Hacıhabiboğlu H. Theoretical analysis of open spherical microphone arrays for acoustic intensity measurements // *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. 2013. V. 22 (2). P. 465–476.

18. Ollivier B., Pepperell A., Halstead Z., Hioka Y. Noise robust bird call localisation using the generalised cross-correlation with phase transform in the wavelet domain // *The Journal of the Acoustical Society of America*. 2019. V. 146 (6). P. 4650–4663. doi: 10.1121/1.5138593
19. Щетинин Е.Ю., Севастьянов Л.А. О методах переноса глубокого обучения в задачах классификации биомедицинских изображений // *Информатика и ее применения*. 2021. Т. 15, № 4. С. 59–64.
20. Doan C.D., Liong S.Y. Generalization for multilayer neural network bayesian regularization or early stopping // *Proc. of Asia Pacific association of hydrology and water resources 2nd conference*. 2004. P. 5–8.
21. Шаход Д.М., Ибряева О.Л. Метод подавления акустического эха на основе рекуррентной нейронной сети и алгоритма кластеризации // *Вестник ЮУрГУ. Сер. Вычислительная математика и информатика*. 2022. Т. 11, № 2. С. 43–58. doi: 10.14529/cm
22. Habets E.A.P. Room impulse response (RIR) generator. 2020. URL: <https://github.com/ehabets/RIR-Generator> (accessed: 05.02.2024).
23. Allen J.B., Berkley D.A. Image method for efficiently simulating small-room acoustics // *The Journal of the Acoustical Society of America*. 1979. V. 65 (4). P. 943–950. doi: 10.1121/1.382599

References

1. Zhu, N. & Reza, T. (2019) A modified cross-correlation algorithm to achieve the time difference of arrival in sound source localization. *Measurement and Control*. 52(3-4). pp. 212–221. DOI: 10.1177/0020294019827977
2. Chiariotti, P., Martarelli, M. & Castellini, P. (2019) Acoustic beamforming for noise source localization – Reviews, methodology and applications. *Mechanical Systems and Signal Processing*. 120. pp. 422–448.
3. Zhong, Y., Xiang, J., Chen, X., Jiang, Y. & Pang, J. (2018) Multiple Signal Classification-Based Impact Localization in Composite Structures Using Optimized Ensemble Empirical Mode Decomposition. *Applied Sciences*. 8(9). Art. 1447. DOI: 10.3390/app8091447
4. Desai, D. & Mehendale, N. (2022) A Review on Sound Source Localization Systems. *Archives of Computational Methods in Engineering*. 29(7). pp. 4631–4642. DOI: 10.1007/s11831-022-09747-2
5. Nguyen, T.N.T., Gan, W.S., Ranjan, R. & Jones, D.L. (2020) Robust source counting and DOA estimation using spatial pseudo-spectrum and convolutional neural network. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. 28. pp. 2626–2637. DOI: 10.1109/TASLP.2020.3019646
6. Nguyen, T.N.T., Nguyen, N.K., Phan, H., Pham, L., Ooi, K., Jones, D.L. & Gan, W.S. (2021) A general network architecture for sound event localization and detection using transfer learning and recurrent neural network. *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. pp. 935–939. DOI: 10.1109/ICASSP39728.2021.9414602
7. Adavanne, S., Politis, A., Nikunen, J. & Virtanen, T. (2018) Sound event localization and detection of overlapping sources using convolutional recurrent neural networks. *IEEE Journal of Selected Topics in Signal Processing*. 13(1). pp. 34–48.
8. He, W., Motlicek, P. & Odobez, J.M. (2019) Adaptation of multiple sound source localization neural networks with weak supervision and domain-adversarial training. *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. pp. 770–774.
9. Perotin, L., Serizel, R., Vincent, E. & Guérin, A. (2018) CRNN-based joint azimuth and elevation localization with the Ambisonics intensity vector. *2018 16th International Workshop on Acoustic Signal Enhancement (IWAENC)*. pp. 241–245. DOI: 10.1109/IWAENC.2018.8521403
10. Liu, N., Chen, H., Songgong, K. & Li, Y. (2021) Deep learning assisted sound source localization using two orthogonal first-order differential microphone arrays. *The Journal of the Acoustical Society of America*. 149(2). pp. 1069–1084. DOI: 10.1121/10.0003445
11. Li, Q., Zhang, X. & Li, H. (2018) Online direction of arrival estimation based on deep learning. *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. pp. 2616–2620. DOI: 10.1109/ICASSP.2018.8461386
12. Alpkocak, A. & Sis, M. (2010) Computing impulse response of room acoustics using the ray-tracing method in time domain. *Archives of Acoustics*. 35(4). pp. 505–519. DOI: 10.2478/v10168-010-0039-8
13. Ciaburro, G. & Iannace, G. (2021) Acoustic characterization of rooms using reverberation time estimation based on supervised learning algorithm. *Applied Sciences*. 11(4). Art.1661. DOI: 10.3390/app11041661
14. Grumiaux, P.A., Kitić, S., Girin, L. & Guérin, A. (2022) A survey of sound source localization with deep learning methods. *The Journal of the Acoustical Society of America*. 152(1). pp. 107–151.
15. Siano, D., Viscardi, M. & Panza, M.A. (2014) Experimental acoustic measurements in far field and near field conditions: characterization of a beauty engine cover. *Recent Advances in Fluid Mechanics and Thermal Engineering*. 12. pp. 50–57.
16. Jacobsen, F. (2014) Sound Intensity. In: Rossing, T.D. (ed.) *Springer Handbook of Acoustics*. New York: Springer. pp. 1093–1114.
17. Hacıhabiboğlu, H. (2013) Theoretical analysis of open spherical microphone arrays for acoustic intensity measurements. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. 22(2). pp. 465–476.
18. Ollivier, B., Pepperell, A., Halstead, Z. & Hioka, Y. (2019) Noise robust bird call localisation using the generalised cross-correlation with phase transform in the wavelet domain. *The Journal of the Acoustical Society of America*. 146(6). pp. 4650–4663. DOI: 10.1121/1.5138593

19. Shchetinin, E.Yu. & Sevastyanov, L.A. (2021) On deep learning transfer methods in biomedical image classification problems. *Informatika i ee primeneniye*. 15(4). pp. 59–64.
20. Doan, C.D. & Liong, S.Y. (2004) Generalization for multilayer neural network bayesian regularization or early stopping. *Proceedings of Asia Pacific association of hydrology and water resources 2nd conference*. pp. 5–8.
21. Shakhod, Gh.M. & Ibryaeva, O.L. (2022) Metod podavleniya akusticheskogo ekha na osnove rekurrentnoy neyronnoy seti i algoritma klasterizatsii [Method of an Acoustic Echo Suppression Based on Recurrent Neural Network and Clustering]. *Vestnik YuUrGU. Ser. Vychislitel'naya matematika i informatika*. 11(2). pp. 43–58. DOI: 10.14529/cmse220204
22. Habets, E.A.P. (2020) *Room impulse response (RIR) generator*. [Online] Available from: <https://github.com/ehabets/RIR-Generator> (Accessed: 5th February 2024).
23. Allen, J.B. & Berkley, D.A. (1979) Image method for efficiently simulating small-room acoustics. *The Journal of the Acoustical Society of America*. 65(4). pp. 943–950. DOI: 10.1121/1.382599

Информация об авторах:

Шаход Джиах Михаил – аспирант кафедры систем автоматики, автоматизированного управления и проектирования Института космических и информационных технологий Сибирского федерального университета (Красноярск, Россия). E-mail: ghiathlovealaa@gmail.com

Агафонов Евгений Дмитриевич – доктор технических наук, профессор кафедры систем автоматики, автоматизированного управления и проектирования Института космических и информационных технологий Сибирского федерального университета (Красноярск, Россия). E-mail: evgeny.agafonov@mail.ru

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

Information about the authors:

Shahoud Ghiath M. (Post-graduate Student, Siberian Federal University, Krasnoyarsk, Russian Federation). E-mail: ghiathlovealaa@gmail.com

Agafonov Evgeniy D. (Doctor of Technical Sciences, Professor, Siberian Federal University, Krasnoyarsk, Russian Federation). E-mail: eagafonov@sfu-kras.ru

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

Поступила в редакцию 02.04.2024; принята к публикации 03.09.2024

Received 02.04.2024; accepted for publication 03.09.2024

ИНФОРМАТИКА И ПРОГРАММИРОВАНИЕ

INFORMATICS AND PROGRAMMING

Научная статья

УДК 004.9:336.71:332.1

doi: 10.17223/19988605/68/12

Об энергоэффективности и оперативности полиномиального и табличного методов приближения математических функций в системах с плавающей точкой**Андрей Александрович Чусов¹, Анна Андреевна Шкляр², Юлия Игоревна Ефимова³**^{1, 2, 3} *Дальневосточный федеральный университет, Владивосток, Россия*¹ *chusov.aa@dvfu.ru*² *shklyar.aan@dvfu.ru*³ *efimova.iui@dvfu.ru*

Аннотация. Представлено теоретическое и экспериментальное обоснование оперативности и энергоэффективности методов приближения некоторых элементарных функций, вычисляемых в условиях ограниченной точности вещественных чисел, аппроксимируемых с помощью длинной рациональной арифметики, в частности числами с плавающей запятой.

Рассмотренные методы аппроксимации вещественнозначных функций включают в себя полиномиальную аппроксимацию с помощью рядов Тэйлора, а также методы табличной аппроксимации CORDIC на основе векторного поворота (для тригонометрических функций) и итерационного вычисления разрядов значения логарифма с гарантиями корректности аппроксимации до ближайшего числа с плавающей точкой с приоритизацией значения с четным младшим разрядом мантиссы.

Методы аппроксимации исследованы на предмет их энергоэффективности и оперативности при заданных количестве разрядов нормализованной мантиссы гарантиях относительно верхних пределов абсолютных ошибок. Экспериментально показаны значительно более высокие энергоэффективность и оперативность табличных методов аппроксимации по сравнению с полиномиальными.

Ключевые слова: энергоэффективность; оперативность; аппроксимация; вещественная арифметика конечной точности; системы с плавающей точкой (запятой).

Благодарности: Данная работа выполнена в рамках государственного задания Минобрнауки России в сфере научной деятельности по проекту № FZNS-2023-0008.

Для цитирования: Чусов А.А., Шкляр А.А., Ефимова Ю.И. Об энергоэффективности и оперативности полиномиального и табличного методов приближения математических функций в системах с плавающей точкой // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 68. С. 112–122. doi: 10.17223/19988605/68/12

Original article

doi: 10.17223/19988605/68/12

On energy efficiency and performance of polynomial and LUT-based approximation of elementary functions in floating-point systems**Andrey A. Chusov¹, Anna A. Shklyar², Yulia I. Efimova³**^{1, 2, 3} *Far-Eastern Federal University, Vladivostok, Russian Federation*¹ *chusov.aa@dvfu.ru*

Abstract. The paper presents results of theoretical and experimental analysis of common methods used for polynomial and table-based numerical approximation of transcendental elementary functions calculated in a finite precision with guarantees on maximal absolute errors, and, particularly, with correct rounding in a given rational or floating-point system.

The methods of approximation of real-valued functions include Taylor-based polynomial approximation, also one based on the CORDIC “shift-and-add” LUT-based approximation as well as iterative evaluation of digits of a logarithm.

The methods are analyzed with the purpose to achieve their energy efficiency and speed given the constraint on maximal absolute errors which is that the result must be off by no more than a half of a unit in the last place of mantissa, with rounding to nearest even in the case of a floating-point system. Experimentally, LUT-based approximation outperforms polynomial approximation by a large margin, and asymptotically, the former consumes resources linearly with the required accuracy while the latter is polynomial and non-linear.

Keywords: energy efficiency; performance; approximation; finite precision real valued arithmetic; floating-point systems.

Acknowledgments: This work was carried out within the framework of the state assignment of the Ministry of Science and Higher Education of the Russian Federation in the field of scientific activity project No. FZNS-2023-0008.

For citation: Chusov, A.A., Shklyar, A.A., Efimova, Y.I. (2024) On energy efficiency and performance of polynomial and LUT-based approximation of elementary functions in floating-point systems. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 68. pp. 112–122. doi: 10.17223/19988605/68/12

Введение

Развитие вычислительной техники, математических и вычислительных методов обработки сигналов и данных связано с увеличением результативности численных вычислений, выполняемых за единицу ресурсопотребления. Наиболее дорогими из потребляемых ресурсов сегодня являются в первую очередь энергия, во вторую – оперативность. Действительно, развитие систем и методов высокопроизводительных вычислений обусловлено последние десятилетия, прежде всего, требованиями к допустимым энергозатратам и повышению результативности и оперативности [1–4]. При этом исследование энергоэффективности алгоритмов и реализаций представлено в литературе в гораздо меньшей степени по сравнению с оперативностью, измерение которой во многих случаях тривиально на всех уровнях программно-аппаратной абстракции. Измерение энергопотребления же во многих случаях доступно лишь на низких уровнях такой абстракции, его методы специфичны по отношению к конкретным вычислительным программным и аппаратным архитектурам, а доступ к этим методам ограничивается, поскольку они представляют собой сторонний канал атаки вычислительной системы [5].

Проблема энергоэффективности вещественных вычислений стоит настолько остро, что во многих случаях, описанных в литературе, гарантиями точности жертвуют в пользу меньшего энергопотребления при выполнении таких задач, как распознавание образов, классификация данных, машинное обучение [4, 6].

Поэтому авторами были исследованы и представлены в настоящей статье методы полиномиальной (по Тейлору), табличной и сдвиговой аппроксимации (CORDIC) тригонометрических и логарифмической функций с заданными требованиями к количеству корректных разрядов мантииссы, т.е. с корректным округлением результата до произвольной точности p , как определено в разд. 1, с формальным доказательством этой корректности.

Аппроксимация математических функций является ресурсозатратным процессом [4], поэтому в литературе представлены исследования методов понижения энергопотребления аппаратных и программных реализаций элементарных и составных операций [6] над вещественными данными. Например, в работе [7] предложен метод аппаратной реализации метода CORDIC для аппроксимации значений тригонометрических функций для выполнения дискретно-косинусного преобразования двумерных сигналов. Основная задача, решаемая в работе, состоит в обеспечении энергоэффективности преобра-

зования. В частности, авторами устанавливается масштабируемость энергопотребления с ростом количества итераций метода CORDIC.

В статье [4] описывается аппаратная организация устройств вещественной арифметики суперкомпьютера «Космодром», пиковое энергопотребление которого составляет 28 Гфлоп/Дж: показан метод реализации арифметических устройств на кристалле с точки зрения того, насколько плотно можно упаковать аппаратные устройства вещественной арифметики и памяти и как это влияет на энергопотребление; также представлено экспериментальное измерение энергопотребления во время выполнения одной операции объединенного умножения-сложения FMA (fused multiply-add).

В работе [8] предлагается и исследуется энергоэффективная аппаратная реализация умножения чисел с плавающей запятой, при котором двоичная мантисса разбивается на группы по три бита, т.е. умножение производится над разрядами из $\mathbb{Z}_8$ мантиссы. Приведенные авторами результаты применения алгоритма для компрессии JPEG и обратного дискретно-косинусного преобразования показывают некоторую масштабируемость ошибки восстановленных данных относительно исходных вследствие информационных потерь, обусловленных ограниченной точностью мантиссы, а также энергопотребление, выраженное средней потребляемой мощностью, в связи с требуемой точностью.

В работе [9] также описан алгоритм CORDIC, однако для увеличения скорости сходимости при векторном повороте CORDIC для вычисления синуса и косинуса углов поворота на первых итерациях применяются полиномы Тейлора степеней, не превышающих 4, что позволило добиться масштабируемости.

Помимо этого, ряды Тейлора часто используются для аппроксимации сигналов, математических функций и других процессов. В настоящей статье аппроксимация Тейлора применена для приближения тригонометрических функций и сравнения эффективности с CORDIC.

Для реализации функции двоичного логарифмирования использован итерационный алгоритм, предложенный в [10]. Авторы этой работы представили результаты измерений точности метода, а настоящая статья дополняет их измерениями оперативности и энергоэффективности для произвольной разрядности аргумента.

В настоящей статье исследуется проблема энергоэффективности и оперативности методов приближения функций рационального аргумента, включая числовые элементы множеств чисел с плавающей точкой произвольной заданной точности, при обеспечении гарантий результативности, которой оказывается абсолютная ошибка приближения, выраженная в единицах на последнем месте мантиссы [6, 11].

1. Представление вещественных чисел и гарантия точности

В настоящей статье исследуются энергоэффективность и оперативность нескольких методов аппроксимации функций вида $f: \mathbb{R} \rightarrow \mathbb{R}$ функциями и значениями произвольной, но фиксированной точности с гарантиями относительно максимальных абсолютных ошибок так, что если x – безошибочное рациональное значение, представленное в ограниченной точности, и $f_r(x)$ аппроксимирует $f(x)$, то

$$|f_r(x) - f(x)| < D_{\max}, \quad (1)$$

где $D_{\max}$ опосредованно задается требованием к точности результата.

А именно, если требуется точность в p разрядах и $b^E \leq f(x) < b^{E+1}$, где b – основание системы счисления и E – целочисленная экспонента, задающая позицию плавающей запятой, то $D_{\max} = b^{1-p} \times f(x)$.

Для того чтобы иметь возможность *точно* представить любое рациональное число, анализ, представленный в настоящей работе, а также реализация длинной арифметики опираются на представление вещественных чисел в виде пятерок целочисленных значений, S , I , N , D и E , описывающих рациональные элементы множества

$$R_{b^p} = \left\{ (-1)^S \cdot \left(I + \frac{N}{D} \right) \cdot b^E \mid b \geq 2 \right\}, \quad (2)$$

где $S \in \mathbb{Z}_2$ – бит знака, I , N и D являются соответственно целой частью, числителем и знаменателем вещественной части мантиссы, причем все значения принадлежат $\mathbb{Z}_{b^p}$, такому что $p \geq 1$ ниже в статье

называется точностью представления вещественного числа. Значение $E \in Z_{b^p E} = [-b^{pE-1}, b^{pE-1} - 1] \cong \mathbb{Z}_b^p$ ниже в статье называется экспонентой. Очевидно, что если N зафиксировать нулевым, а действительные значения аппроксимировать до p разрядов I (с некоторой экспонентой E), то полученная в результате числовая система будет являться системой с плавающей запятой.

Для того чтобы нормализовать и сделать единственным представление числовых значений из (2), устанавливаются следующие ограничения: либо $x = 0$, $S = I = N = E = 0$ и $D = 1$, либо $x \neq 0$, $D \bmod b \neq 0$, $\gcd(N, D) = 1$ и $I \bmod b \neq N \bmod b$.

Кроме указанной системы R_b^p длинных вещественных чисел дополнительно формулируется множество чисел с плавающей запятой с точностью мантииссы в p разрядов

$$F_{b^p} = \{(-1)^S \cdot M \cdot b^{E+1-p} \mid b \geq 2 \wedge M \in Z_{b^p}\}. \quad (3)$$

В настоящей статье опущено рассмотрение денормализованных чисел в системах с плавающей запятой, а также бесконечностей и нечисел, которые представимы и как элементы множества (2) (при $D = 0$), и как элементы некоторой системы с плавающей запятой.

Таким образом, в настоящей статье полагается, что для любого ненулевого элемента F_b^p выполняется неравенство

$$b^{p-1} \leq M \leq b^p - 1. \quad (4)$$

Единица в правой части неравенства является минимальным значением, представимым ненулевой мантииссой M , и называется единицей на последнем месте (ulp – Unit in the Last Place); в настоящей статье единица на последнем месте соответствует определению Харрисона, но не Голдберга. Численно $1\text{ulp} = b^{E+1-p}$.

Таким образом, можно однозначно определить известное понятие корректности округления [12]: действительное $\hat{x} \in \mathbb{R}$ округлено *корректно* числом с плавающей запятой $x \in F_b^p$, если $|\hat{x} - x| \leq 0,5\text{ulp} = 1/2 \cdot b^{E+1-p}$.

Из этого определения в настоящей статье выводится производное определение корректно округления значения из $\mathbb{R}$ элементом R_b^p – через отображение $R_b^p \rightarrow F_b^p$, которое реализуется длинным эвклидовым делением значения $ID + N$ на D в (2) с тем, чтобы получить p -разрядное частное M , удовлетворяющее (4), и остаток r . Путем сравнения $2r$ и D вычисляется направление корректного округления: если $2r > D$, значение M увеличивается на единицу (то есть на 1ulp), после чего вероятна ренормализация результата; иначе, если $2r < D$, то M уже является корректно округленной мантииссой; в противном случае, т.е. если $2r = D$, корректно округленной мантииссой оказывается M или $M + 1$ – та из них, у которой младший разряд чётен.

Тогда, если $x = (-1)^S \times M \times b^E \in F_{b^{p_F}}^p$ является образом x_R отображения $R_{b^p} \rightarrow F_{b^{p_F}}^p : x_R \mapsto x$, а x_R , в свою очередь, является образом отображения $\mathbb{R} \rightarrow R_{b^p} : \hat{x} \mapsto x_R$, и при этом $|\hat{x} - x| \leq 0,5\text{ulp}$, то x_R считается корректным округлением $\hat{x}$.

Две указанные числовые системы, R_b^p и F_b^p , образуют собой инструмент для реализации арифметических вычислений произвольной точности и для исследования эффективности указанных ниже методов аппроксимации.

1.1. Аппроксимация полиномом Тейлора

Приближение полиномами Тейлора опирается на равенство аналитической функции $f(x)$ значению $f(x) = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{df}{dx} \Big|_{x=a} \frac{(x-a)^k}{k!} = \lim_{n \rightarrow \infty} P_n(x)$. При конечном n дифференцируемая n раз функция $f(x)$ тогда приближается к значению полинома $P_n(x)$, т.е.

$$f(x) \approx P_n. \quad (5)$$

В настоящей статье, в частности, рассматривается аппроксимация синуса и косинуса, она осуществляется соответственно следующим образом:

$$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^n \cdot x^{2n}}{(2n+1)!} = \sum_{j=0}^n y_j; \quad \cos(x) = \sum_{n=0}^{\infty} \frac{(-1)^n \cdot x^{2n}}{(2n)!} = \sum_{j=0}^n t_j; \quad (6)$$

где $y_{j+1} = y_j \cdot \frac{-x^2}{(2j) \cdot (2j+1)}$, $t_{j+1} = -t_j \cdot \frac{x^2}{2n+1}$, $y_0 = x$, $x \in R_{(b^p)}$, $t_0 = 1$.

Для того чтобы гарантировать точность аппроксимации, сформулируем следующую теорему:

Теорема 1. Если в (6) $y_n < \delta$, то $|\sin(x) - P_n(x)| < \delta$; если в (6) $t_n < \delta$, то $|\cos(x) - P_n(x)| < \delta$.

Доказательство. Если $x \leq j$, то если $j > 0$, $\frac{x^{j+1}}{(j+1)!} < \frac{x^j}{j!}$, поэтому магнитуа элементов ряда

Тейлора, $\frac{(-1)^n x^{2n}}{(2n+1)!}$ в случае синуса и $\frac{(-1)^n x^{2n}}{(2n)!}$ в случае косинуса, строго убывает с ростом n , а знак элемента меняется, поэтому по теореме о двух милиционерах если

$$|y_n| < \delta \text{ или } |t_n| < \delta, \quad (7)$$

то абсолютная ошибка приближения не превысит δ , и поэтому в настоящей работе алгоритм останавливается, когда достигается условие (7), где δ – абсолютная ошибка. Q.E.D.

1.2. Аппроксимация сдвигово-аддитивным методом CORDIC

Алгоритм реализации метода CORDIC Джека Вандера основан на следующей итерации:

$$\begin{cases} x_{n+1} = x_n - d_n \cdot y_n \cdot 2^{-n}; \\ y_{n+1} = y_n - d_n \cdot x_n \cdot 2^{-n}; \\ z_{n+1} = z_n - d_n \cdot \arctg(2^{-n}), \end{cases} \quad (8)$$

где члены $\arctg(2^{-n})$ предварительно вычисляются и сохраняются, d_i равны $+1$ или -1 .

В режиме вращения значение d_n выбирается эквивалентным знаком z_n , то есть $+1$, если $z_n \geq 0$, иначе -1 . Если модуль z_0 меньше, чем

$$\sum_{k=0}^{\infty} \arctg(2^{-k}) = 1,7432866204723400035..., \quad (9)$$

то получается выражение

$$\lim_{n \rightarrow \infty} \begin{pmatrix} x_n \\ y_n \end{pmatrix} = K \times \begin{pmatrix} x_0 \cdot \cos(z_0) - y_0 \cdot \sin(z_0) \\ x_0 \cdot \sin(z_0) - y_0 \cdot \cos(z_0) \end{pmatrix}, \quad (10)$$

где $K = \lim_{n \rightarrow \infty} \left(\prod_{j=0}^n \sqrt{1 + 2^{-2j}} \right) \approx 1,646760258121$ – инвариантный масштабирующий коэффициент.

В настоящей статье приведены результаты исследования алгоритма приближения функций синус и косинус числа θ с помощью метода CORDIC, который в этом случае основан на разложении $z_0 = \theta$ на сумму углов $w_n = \arctg(2^{-n})$ с использованием невосстанавливающего алгоритма, который

дает разложение $\theta = \sum_{k=0}^{\infty} d_k w_k$, где $d_k = \pm 1$, $w_k = \arctg(2^{-k})$. Последовательным вращением вектор (x_0, y_0)

поворачивается до тех пор, пока не будет получен вектор (x_n, y_n) при помощи поворота на угол $d_{n-1}w_{n-1}$. Это дает невосстанавливающее разложение, показанное в формуле

$$t_0 = 0; \quad t_{n+1} = t_n + d_n w_n. \quad (11)$$

Тогда n -е вращение имеет вид:

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} \cos(d_n w_n) & -\sin(d_n w_n) \\ \sin(d_n w_n) & \cos(d_n w_n) \end{pmatrix} \cdot \begin{pmatrix} x_n \\ y_n \end{pmatrix} \quad (12)$$

Учитывая, что $d_n = \pm 1$, $\cos(d_n w_n) = \cos(w_n)$ и $\sin(d_n w_n) = d_n \cdot \sin(w_n)$, а также используя соотношение $\operatorname{tg}(w_n) = 2^{-n}$, получаем выражение

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \cos(w_n) \cdot \begin{pmatrix} 1 & -d_n \cdot 2^{-n} \\ d_n \cdot 2^{-n} & 1 \end{pmatrix} \cdot \begin{pmatrix} x_n \\ y_n \end{pmatrix}. \quad (13)$$

В данной формуле умножение на $\cos(w_n) = 1/\sqrt{1+2^{-2n}}$ является единственным безошибочным умножением, поскольку в системе счисления 2 умножение на 2^{-n} сводится к поразрядному сдвигу. Следовательно, чтобы избежать этого умножения, вместо (13) выполняется

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} 1 & -d_n \cdot 2^{-n} \\ d_n \cdot 2^{-n} & 1 \end{pmatrix} \cdot \begin{pmatrix} x_n \\ y_n \end{pmatrix}. \quad (14)$$

Теорема 2. Если $d_{n+1} \neq d_n$, $|x_{n+1} - x_n| < \delta$ и $|y_{n+1} - y_n| < \delta$, то $|x_n - \cos\theta| < \delta$ и $|y_n - \sin\theta| < \delta$.

Доказательство. Условие $d_{n+1} \neq d_n$ означает, что $x_n < \cos\theta \leq x_{n+1}$ или, наоборот, $x_{n+1} \leq \cos\theta < x_n$, а также что $y_n < \sin\theta \leq y_{n+1}$ или $y_{n+1} \leq \sin\theta < y_n$. Но по условию теоремы $|x_{n+1} - x_n| < \delta$ и $|y_{n+1} - y_n| < \delta$, следовательно $|x_n - \cos\theta| < \delta$ и $|y_n - \sin\theta| < \delta$. Q.E.D.

Эта теорема дает возможность гарантировать максимальную абсолютную ошибку приближения значений синуса и косинуса путем подстановки $D_{\max}$ вместо δ .

В реализации, которая использована в настоящей статье, используется заранее генерируемая таблица поиска (LUT; Lookup Table) значений углов для параметризации синуса и косинуса, которые при каждой последующей итерации уменьшаются в 2 раза. Генерация LUT легко распараллеливается и может быть осуществлена как при компиляции программы, так и путем предварительной инициализации при запуске или первом обращении к реализации CORDIC.

1.3. Итерационный метод вычисления функции логарифм

Вычисление логарифма по основанию b с гарантиями точности опирается на следующую теорему:

Теорема 3. Пусть $x \in F_b^p$. Тогда если x не является целой степенью b , то $\log_b x$ никогда не будет находится на расстоянии $0,5\operatorname{ulp}$ от ближайших чисел с плавающей запятой.

Доказательство. Поскольку $x \in F_b^p$, то оно точно представимо выражением

$$d_0, d_1 \dots d_{p-1} \cdot b^E = (d_0 b^{p-1} + d_1 b^{p-2} + \dots + d_{p-1} b^0) \cdot b^{E+1-p}.$$

Предположим, что $\log_b x - r(\log_b x) = 0,5\operatorname{ulp}$, где $r: \mathbb{R} \rightarrow F_b^p$ – операция округления действительного аргумента до ближайшего элемента из F_b^p .

Тогда $\log_b x = kb^t$, где $k \in \mathbb{Z}$ и $t \in \mathbb{Z}$. То есть

$\log_b(d_0 b^{p-1} + d_1 b^{p-2} + \dots + d_{p-1} b^0) + E + 1 - p = kb^t \Leftrightarrow \log_b(d_0 b^{p-1} + d_1 b^{p-2} + \dots + d_{p-1} b^0) = kb^t - E - 1 + p = L \in \mathbb{Z}$, но это возможно, только если x является целой степенью b (а именно b^L), что противоречит условию теоремы. Q.E.D.

Следствие. Для вычисления логарифма по основанию b с точностью p и корректным округлением достаточно использования не более одного защитного разряда, т.е. логарифм достаточно вычислить с точностью $p + 1$, поскольку один защитный разряд однозначно скажет о направлении корректного округления.

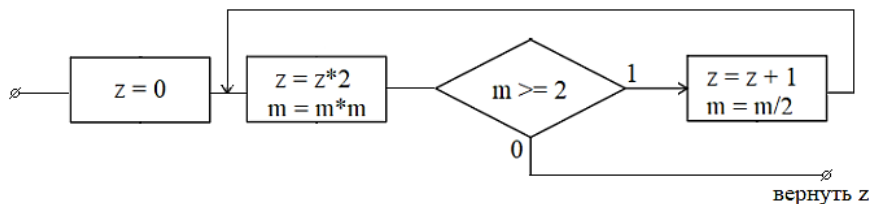


Рис. 1 Итерационное вычисление разрядов двоичного логарифма
Fig. 1. Iterative calculation of binary logarithm digits

В этой связи для точного вычисления логарифма достаточно использовать итерационные методы, например описанные в [10], для получения $p + 1$ разряда, которые потом достаточно округлить до p .

В рамках настоящей работы был реализован и исследован итерационный алгоритм, представленный на рис. 1.

2. Методы измерения эффективности

Современные центральные процессоры предоставляют собственное программно-аппаратное обеспечение измерений времени и энергопотребления. Доступ к времени осуществляется тривиально, но измерение энергопотребления осуществляется через использование специализированных регистров MSR (Model Specific Registers) процессоров Intel и AMD. Поддержка этих регистров действительно специфична и осуществляется в общем случае по-разному на разных моделях процессоров даже одного производителя. Поэтому в настоящем разделе приведено описание обобщенного метода измерения энергопотребления, использующего относительно распространенные MSR процессоров Intel и AMD, который был использован для получения энергетических характеристик исследованных алгоритмов. Измерение энергопотребления программ осуществлялось путем обращения к регистрам MSR RDMSR $\times 86$ и $\times 86-64$ по следующим сдвигам.

В случае микропроцессоров Intel для чтения энергии, затраченной центральным процессором, начиная с некоторого момента в прошлом и до обращения к регистру, используется сдвиг 0x611 (MSR_PKG_ENERGY_STATUS). Регистром задается целочисленное значение, выраженное в $\mathbb{Z}_2^{32}$ в единицах измерения, которые соотносятся с джоулями соответственно битам 7–25 регистра MSR_RAPL_POWER_UNIT. В регистре MSR_RAPL_POWER_UNIT со сдвигом 0x606 хранится информация о единице измерения потребленной энергии, выраженная пятибитовым целочисленным логарифмом по основанию 1/2 из соответствующего количества джоулей.

Аналогично процессорами AMD определяются регистры AMD_MSR_PACKAGE_ENERGY (сдвиг 0xC001029B) с целочисленным 32-битовым значением потребленной энергии, выраженным в единицах измерения, равных значению 2^{-t} джоулей, где t – значение, хранимое в битах 8–12 регистра AMD_MSR_PWR_UNIT (сдвиг 0xC0010299).

Доступ к регистрам MSR как к средствам практически неконтролируемого доступа и анализа вычислительных систем в целях информационной безопасности в значительной степени ограничивается операционными системами Windows и Linux. Последняя предоставляет привилегированный доступ к регистрам, отображая их в файловую систему по адресу /dev/cpu/<номер ядра>/msr. В случае симметричной мультипотокости MSR каждого логического ядра предоставляется отдельно. В Windows аналогичного механизма нет, и для получения доступа к MSR необходимо выполнять инструкцию RDMSR из пространства ядра операционной системы путем реализации программного драйвера.

Реализация такого драйвера для процессоров Intel, разработанного авторами и использованного для снятия измерений, представленных в настоящей статье, доступна читателю в GIT-репозитории по Интернет-адресу <https://gitlab.com/lpsztemp/chsvreadenergy.git>.

3. Экспериментальное исследование энергоэффективности и оперативности

В настоящем разделе представлены результаты экспериментальных измерений эффективности, а именно оперативности и энергопотребления, перечисленных выше методов аппроксимации действительностнозначных функций значениями из F_b^p , т.е. с гарантией безошибочного значения мантиссы в p старших разрядах и ошибкой аппроксимации не более $0,5ulp$.

3.1. Метод Тейлора

На рис. 2 показан график зависимости времени от требуемой и гарантируемой точности вычисления синуса $\pi/6$ вблизи нуля. Наблюдается ожидаемое подобие обоих графиков. В частности, график демонстрирует показательный рост потребления времени и энергии.

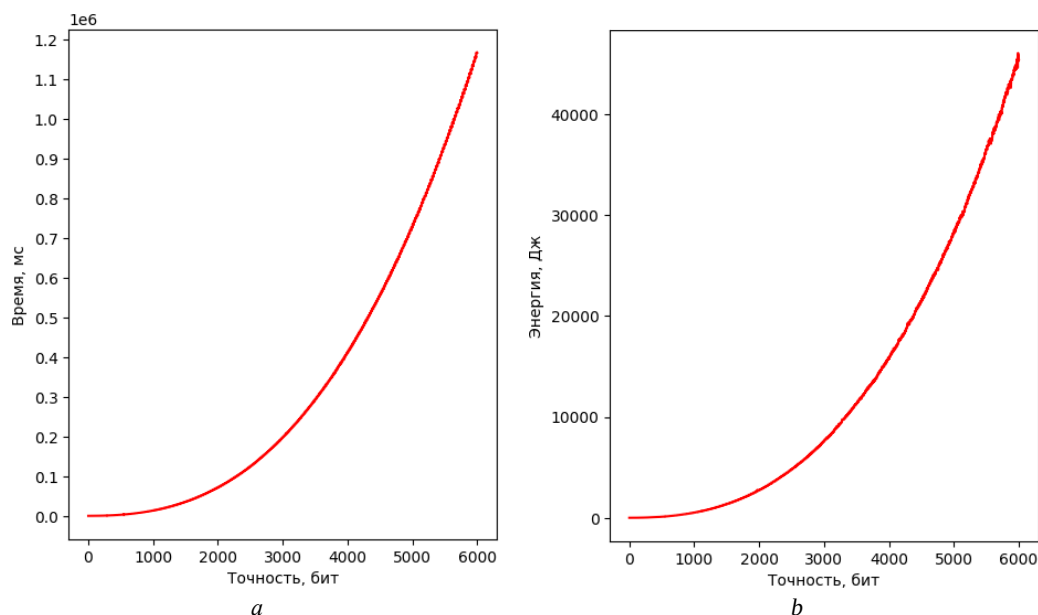


Рис. 2. График зависимости времени (a) и потребляемой энергии (b) от точности для функции синус
Fig. 2. Graph of the dependence of time (a) and energy consumption (b) on accuracy for sine functions

3.2. Метод CORDIC

Далее рассмотрим метод CORDIC. Так как в нашей реализации данный метод выполняет одновременно расчет и для функции синус, и для функции косинус, приведен только график синуса. На рис 3, a и 3, b соответственно показаны зависимости времени вычислений и потребляемой энергии от требуемой точности для функции синус методом CORDIC. Из графика видно, что зависимость имеет линейный вид, что говорит о меньшем ресурсопотреблении данного метода в сравнении с методом Тейлора, особенно при высоких требованиях к точности.

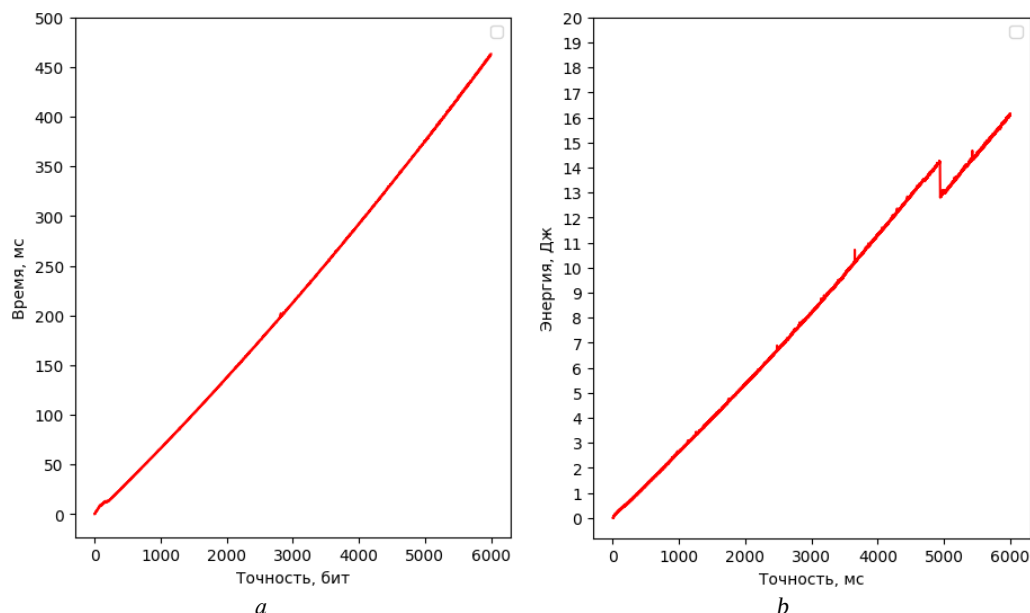


Рис. 3. График зависимости времени (a) и потребляемой энергии (b) от точности для функций синус / косинус методом CORDIC

Fig. 3. Graph of the dependence of time (a) and energy consumption (b) on accuracy for sine/cosine functions using the CORDIC method

Наглядно показано превосходство метода CORDIC по сравнению с методом Тейлора.

3.3. Итерационный метод вычисления функции логарифм

Далее рассмотрим итерационный метод для вычисления функции логарифм.

На рис. 4 показаны графики зависимости времени и энергопотребления реализации метода от требуемой точности значения двоичного логарифма.

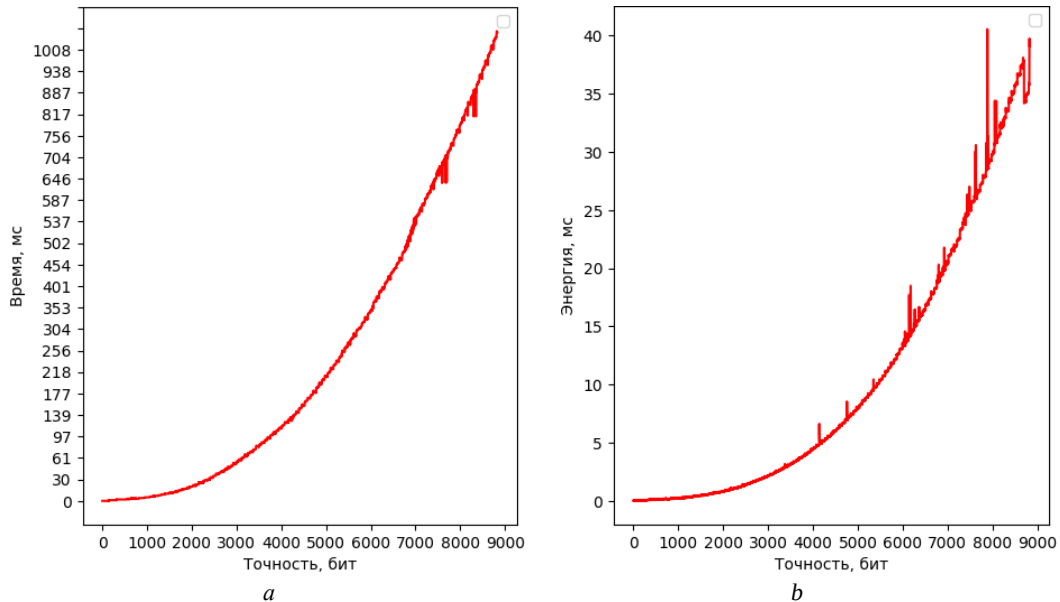


Рис. 4. График зависимости времени (а) и потребляемой энергии (б) от точности для функции логарифм
Fig. 4. Graph of the dependence of time (a) and energy consumption (b) on accuracy for the logarithm function

Данный метод рассчитан с гарантиями точности до 9 000 бит. Наглядно показано превосходство метода CORDIC начиная с точности в 6 000 бит, хотя при меньшей точности итерационный метод оказывается более эффективным по сравнению с методом CORDIC.

Заключение

В настоящей статье предложено формальное обоснование гарантий точности некоторых пространственных методов аппроксимации действительных чисел и действительных функций с помощью методов конечной машинной арифметики произвольной точности, а также экспериментальное обоснование энергоэффективности этих методов на основе доступа к аппаратным ресурсам микропроцессоров Intel $\times 86-64$.

В частности, выполнены формальное обоснование и экспериментальная реализация вещественных чисел и операций с произвольной конечной гарантированной точностью на основе длинной целочисленной арифметики; выполнена реализация ряда известных методов, таких как полиномиальный, аддитивно-сдвиговой и итерационный, для приближенного вычисления математических функций действительного аргумента с помощью функций рационального аргумента, приближающего действительный с гарантиями точности; представлен и использован метод измерения оперативности и энергоэффективности в дополнение к результативности.

Результаты измерений эффективности согласуются с прогнозируемым поведением методов аппроксимации. В результате экспериментов с гарантируемой точностью мантиссы в 6 000 бит метод CORDIC оказался эффективнее метода приближения рядами Тейлора в 2 518 раз по временным показателям и в 2 817 раз по энергетическим показателям. Временные затраты и энергопотребление метода CORDIC растут в асимптоте линейно, хотя метод и требует предрасчета корректирующих коэффициентов или значений углов, что соответствует полученным результатам экспериментальных измерений, представленных в статье для больших значений требуемой точности аппроксимации.

Список источников

1. Inacio E.C., Dantas M.R. A survey into performance and energy efficiency in HPC, cloud and big data environments *International Journal of Networking and Virtual Organisations*. 2015. V. 14 (4). P. 299–318. doi: 10.1504/IJNVO.2014.067878
2. Jarus M., Varrette S., Oleksiak A., Bouvry P. Performance Evaluation and Energy Efficiency of High-Density HPC Platforms Based on Intel, AMD and ARM Processors // *Energy Efficiency in Large Scale Distributed Systems. EE-LSDS 2013 / J.M. Pierson, G. Da Costa, L. Dittmann (eds). Berlin ; Heidelberg : Springer, 2013. doi: 10.1007/978-3-642-40517-4_16 (Lecture Notes in Computer Science; v. 8046).*
3. Schöne R., Velten M., Hackenberg D., Ilsche T. Energy Efficiency Features of the Intel Alder Lake Architecture // *Proc. of the 15th ACM/SPEC International Conference on Performance Engineering (ICPE '24). New York : Association for Computing Machinery, 2024. P. 95–106. doi: 10.1145/3629526.3645040*
4. Zaruba F., Schuiki F., Mach S., Benini L. The Floating Point Trinity: A Multi-modal Approach to Extreme Energy-Efficiency and Performance // *Proc. 26th IEEE International Conference on Electronics, Circuits and Systems (ICECS). Genoa, 2019. P. 767–770. doi: 10.1109/ICECS46596.2019.8964820*
5. Kogler A., Weber D., Haubenwallner M., Lipp M., Gruss D., Schwarz M. Finding and Exploiting CPU Features using MSR Templating // *Proc. IEEE Symposium on Security and Privacy (SP). San Francisco, CA, 2022. P. 1474–1490. doi: 10.1109/SP46214.2022.9833599*
6. Imani M., Peroni D., Rosing T. CFPU: Configurable Floating Point Multiplier for Energy-Efficient Computing // *Proc. of the 54th Annual Design Automation Conference 2017 (DAC '17). New York : Association for Computing Machinery, 2017. Art. 76. P. 1–6. doi: 10.1145/3061639.3062210*
7. Satpute V.R., Nawandar N.K. CORDIC based novel energy-efficient approximate DCT architecture for error-resilient applications // *Proc. 2016 11th International Conference on Industrial and Information Systems (ICIIS). Roorkee, 2016. P. 655–660. doi: 10.1109/ICIINFS.2016.8263020*
8. Ding R., Guo Y., Sun H., Kimura S. Energy-Efficient Approximate Floating-Point Multiplier Based on Radix-8 Booth Encoding // *Proc. IEEE 14th International Conference on ASIC (ASICON). Kunming, 2021. P. 1–4. doi: 10.1109/ASICON52560.2021.9620455.*
9. Wadkar S.S., Das B.P., Meher P.K. Low Latency Scaling-Free Pipeline CORDIC Architecture Using Augmented Taylor Series // *Proc. IEEE International Symposium on Smart Electronic Systems (iSES) (Formerly iNiS). Rourkela, 2019. P. 312–315. doi: 10.1109/iSES47678.2019.00077*
10. Mansour A., Sayed A. A new floating point implementation of base 2 logarithm // *Int. J. Inf. Commun. Eng.* 2014. V. 8 (10). P. 1352–1355.
11. Muller J.M. et al. *Handbook of floating-point arithmetic*. Basel : Birkhäuser, 2018. P. 44–91.
12. Lafage V. Revisiting “What every computer scientist should know about floating-point arithmetic” // *arXiv*. 2020. 4 Dec. P. 1–13. doi: 10.48550/arXiv.2012.02492. URL: https://www.researchgate.net/publication/346668652_Revisiting_What_Every_Computer_Scientist_Should_Know_About_Floating-point_Arithmetic

References

1. Inacio, E.C. & Dantas, M.R. (2015) A survey into performance and energy efficiency in HPC, cloud and big data environments. *International Journal of Networking and Virtual Organisations*. 14(4). pp. 299–318. DOI: 10.1504/IJNVO.2014.067878
2. Jarus, M., Varrette, S., Oleksiak, A. & Bouvry, P. (2013) Performance Evaluation and Energy Efficiency of High-Density HPC Platforms Based on Intel, AMD and ARM Processors. In: Pierson, JM., Da Costa, G. & Dittmann, L. (eds) *Energy Efficiency in Large Scale Distributed Systems. EE-LSDS 2013. Lecture Notes in Computer Science*. Vol. 8046. Berlin; Heidelberg: Springer. DOI: 10.1007/978-3-642-40517-4_16
3. Schöne, R., Velten, M., Hackenberg, D. & Ilsche, T. (2024) Energy Efficiency Features of the Intel Alder Lake Architecture. *Proceedings of the 15th ACM/SPEC International Conference on Performance Engineering (ICPE '24)*. Association for Computing Machinery. New York, NY. USA. pp. 95–106. DOI: 10.1145/3629526.3645040
4. Zaruba, F., Schuiki, F., Mach, S. & Benini, L. (2019) The Floating Point Trinity: A Multi-modal Approach to Extreme Energy-Efficiency and Performance. *Proceedings 2019 26th IEEE International Conference on Electronics, Circuits and Systems (ICECS). Genoa. Italy*. pp. 767–770. DOI: 10.1109/ICECS46596.2019.8964820
5. Kogler, A., Weber, D., Haubenwallner, M., Lipp, M., Gruss, D. & Schwarz, M. (2022) Finding and Exploiting CPU Features using MSR Templating. *Proceedings 2022 IEEE Symposium on Security and Privacy (SP)*. San Francisco. CA. USA. pp. 1474–1490. DOI: 10.1109/SP46214.2022.9833599
6. Imani, M., Peroni, D. & Rosing, T. (2017) CFPU: Configurable Floating Point Multiplier for Energy-Efficient Computing. *Proceedings of the 54th Annual Design Automation Conference 2017 (DAC '17)*. Association for Computing Machinery. New York. NY. USA. Art. 76. pp. 1–6. DOI: 10.1145/3061639.3062210
7. Satpute, V.R. & Nawandar, N.K. (2016) CORDIC based novel energy-efficient approximate DCT architecture for error-resilient applications. *Proceedings 11th International Conference on Industrial and Information Systems (ICIIS)*. Roorkee. India. 2016. pp. 655–660. DOI: 10.1109/ICIINFS.2016.8263020
8. Ding, R., Guo, Y., Sun, H. & Kimura, S. (2021) Energy-Efficient Approximate Floating-Point Multiplier Based on Radix-8 Booth Encoding. *Proceedings IEEE 14th International Conference on ASIC (ASICON). Kunming. China*. pp. 1–4. DOI: 10.1109/ASICON52560.2021.9620455

9. Wadkar, S.S., Das, B.P. & Meher, P.K. (2019) Low Latency Scaling-Free Pipeline CORDIC Architecture Using Augmented Taylor Series. *Proceedings IEEE International Symposium on Smart Electronic Systems (iSES) (Formerly iNiS)*. Rourkela, India. pp. 312–315. DOI: 10.1109/iSES47678.2019.00077
10. Mansour, A. & Sayed, A. (2014) A new floating point implementation of base 2 logarithm. *International Journal of Information and Communication Engineering*. 8(10). pp. 1352–1355.
11. Muller, J.M. et al. (2018) *Handbook of floating-point arithmetic*. Basel, Switzerland: Birkhäuser. pp. 44–91.
12. Lafage, V. (2020) *Revisiting “What every computer scientist should know about floating-point arithmetic.”* arXiv. 2020. 4th December. pp. 1–13. DOI: 10.48550/arXiv.2012.02492

Информация об авторах:

Чусов Андрей Александрович – кандидат технических наук, доцент департамента электроники, телекоммуникации и приборостроения Политехнического института Дальневосточного федерального университета (Владивосток, Россия). E-mail: chusov.aa@dvfu.ru.

Шкляр Анна Андреевна – магистрант департамента электроники, телекоммуникации и приборостроения Политехнического института Дальневосточного федерального университета (Владивосток, Россия). E-mail: shklyar.aan@dvfu.ru

Ефимова Юлия Игоревна – аспирант департамента электроники, телекоммуникации и приборостроения Политехнического института Дальневосточного федерального университета (Владивосток, Россия). E-mail: iui.efimova@gmail.com

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

Information about the authors:

Chusov Andrey A. (Candidate of Technical Sciences, Associate Professor, Far-Eastern Federal University, Vladivostok, Russian Federation). E-mail: chusov.aa@dvfu.ru

Shklyar Anna A. (Master student, Far-Eastern Federal University, Vladivostok, Russian Federation). E-mail: shklyar.aan@dvfu.ru

Efimova Julia I. (Post-graduate Student, Far-Eastern Federal University, Vladivostok, Russian Federation). E-mail: efimova.iui@dvfu.ru

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

Received 22.06.2024; accepted for publication 03.09.2024

Поступила в редакцию 22.06.2024; принята к публикации 03.09.2024

ПРОЕКТИРОВАНИЕ И ДИАГНОСТИКА
ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

DESIGNING AND DIAGNOSTICS OF COMPUTER SYSTEMS

Научная статья

УДК 681.518.5+004.052.32

doi: 10.17223/19988605/68/13

Синтез самопроверяемых цифровых устройств на основе взвешенных кодов
с суммированием, проверочные символы которых описываются
самодвойственными булевыми функциями

Дмитрий Викторович Ефанов

*Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, Россия;**Российский университет транспорта, Москва, Россия;**Ташкентский государственный транспортный университет, Ташкент, Узбекистан,**TrES-4b@yandex.ru*

Аннотация. Рассматриваются коды с суммированием в кольце вычетов по заданному модулю M взвешенных разрядов информационного вектора, проверочные символы которых описываются самодвойственными булевыми функциями. Доказано, что такие коды могут быть построены только при условии $M \in \{2, 4, 8, \dots\}$ для некоторых массивов весовых коэффициентов. Приведено разделение частного случая взвешенных кодов с суммированием в кольце вычетов по модулю $M = 4$ при числе информационных символов $m = 4$ на особые классы. Определено, что только два варианта взвешивания весовыми коэффициентами из массивов $[1, 1, 2, 3]$ и $[1, 2, 2, 2]$ без учета перестановок весов дают коды со свойством самодвойственности функций, описывающих проверочные символы. Приведены структуры организации схем встроенного контроля на основе взвешенных кодов с суммированием в кольце вычетов по модулю $M = 4$ при числе информационных символов $m = 4$.

Ключевые слова: самопроверяемые устройства; контроль вычислений на выходах логических схем; контроль вычислений по признаку самодвойственности; взвешенные коды с суммированием; особые классы взвешенных кодов с суммированием.

Для цитирования: Ефанов Д.В. Синтез самопроверяемых цифровых устройств на основе взвешенных кодов с суммированием, проверочные символы которых описываются самодвойственными булевыми функциями // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 68. С. 123–139. doi: 10.17223/19988605/68/13

Original article

doi: 10.17223/19988605/68/13

The self-checking digital devices synthesis based on weight-based sum code,
which check symbols are described by self-dual Boolean functions

Dmitry V. Efanov

*Peter the Great St. Petersburg Polytechnic University, Saint Petersburg, Russian Federation;**Russian University of Transport, Moscow, Russian Federation;**Tashkent State Transport University, Tashkent, Uzbekistan,**TrES-4b@yandex.ru*

Abstract. The proposed article considers sum code in a residue ring over a given modulo M of the data vector weighted bits. Test symbols are described by self-dual Boolean functions. It is proven that such codes can be constructed

only under the condition $M \in \{2, 4, 8, \dots\}$ for certain arrays of weighting coefficients. The division into special classes of the weight-based sum code special case in the residue ring modulo $M=4$ is given. Cases with the number of data bits $m=4$ are considered. It has been determined that only two options for weighing with weight coefficients from arrays $[1, 1, 2, 3]$ and $[1, 2, 2, 2]$, without considering permutations of weights, give codes with the function self-duality property that describe check symbols. The concurrent error-detection circuit structure based on weight-based sum code in a residue ring by modulo $M=4$ with the $m=4$ data bits is given.

Keywords: self-testing devices; control of calculations at the logic circuit outputs; control of calculations based on self-duality; weight-based sum codes; weight-based sum codes special classes.

For citation: Efanov, D.V. (2024) The self-checking digital devices synthesis based on weight-based sum code, which check symbols are described by self-dual Boolean functions. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika* – Tomsk State University Journal of Control and Computer Science. 68. pp. 123–139. doi: 10.17223/19988605/68/13

Введение

При синтезе отказоустойчивых и самопроверяемых устройств и вычислительных систем зачастую используются методы теории информации и кодирования [1–4]. Их применение позволяет на практике строить менее избыточные структуры, чем при реализации структур с использованием кратного резервирования [5, 6].

Одним из часто используемых классов избыточных кодов является класс с суммированием единичных информационных разрядов по произвольному модулю M , называемых еще модульными, или остаточными кодами [7, 8]. Некоторые из данных кодов имеют особые свойства и тем самым особое значение в решении задач синтеза устройств. Например, для модульных кодов со значениями модуля $M \in \{2^1, 2^2, \dots, 2^{\lceil \log_2(m+1) \rceil - 1}\}$, где m – число разрядов в информационном векторе, при $m \geq M$

удается формировать все 2^k , $k = \log_2 M$, контрольных вектора, что позволяет довольно просто обеспечивать полную проверку тестера такого кода. Поэтому такие модульные коды получили распространение при решении задач построения отказоустойчивых и самопроверяемых устройств [1]. Среди модульных кодов с суммированием с модулями $M \in \{2^1, 2^2, \dots, 2^{\lceil \log_2(m+1) \rceil - 1}\}$ выделяются коды со значениями $M = 4$ и $M = 8$, которые наряду с кодами со значениями $M = 12$, принято называть кодами Боуза–Лина [9]. Коды Боуза–Лина со значениями $M = 4$ и $M = 8$ имеют постоянное малое число проверочных символов, $k = 2$ и $k = 3$ соответственно, вне зависимости от числа m информационных символов. Поэтому при их использовании вносится небольшая избыточность в структуру схем встроенного контроля (СВК) [5].

Для заданных m и M строится только один код Боуза–Лина. Однако множество кодов можно расширить за счет простой модификации, что важно и с практической точки зрения с позиции вносимой структурной избыточности в синтезируемые устройства. Модификация может быть основана на предварительном установлении неравноправия между разрядами информационного вектора путем приписывания им весовых коэффициентов в виде натуральных чисел. После такого взвешивания суммируются уже весовые коэффициенты единичных разрядов в кольце вычетов по заданному модулю M . Сам способ был предложен в той же классической статье Джея Бергера, где описан код, названный впоследствии его именем [10]. Только в предложенном Бергером варианте суммирование осуществляется по модулю, превышающему сумму всех весовых коэффициентов информационного вектора.

Среди более современных работ на тему использования взвешенных кодов с суммированием при синтезе самопроверяемых и отказоустойчивых структур можно выделить следующие публикации. В [11] рассматривается задача применения взвешенных кодов с суммированием для организации контроля вычислений на выходах многоуровневых логических схем. Показано, что общая вероятность обнаружения ошибок, вызванных внутренними неисправностями, с применением данных кодов выше, чем при использовании классических кодов. В [12] уже рассматриваются взвешенные коды с суммированием по модулю M . Показано, что такие коды могут строиться для рассматриваемой

схемы с учетом ее структуры с оценкой вероятности существования путей распространения сигналов от внутренних элементов к выходам, что дает возможность высокого покрытия (около 99%) ошибок на выходах. В статьях [13, 14] исследуются характеристики обнаружения ошибок взвешенными кодами с суммированием на случай их возникновения только в разрядах информационных векторов при безошибочности разрядов контрольного вектора, что актуально в тех их приложениях, в которых информационные и проверочные символы вычисляются физически различными устройствами. В [15, 16] обсуждаются вопросы построения различных взвешенных кодов с суммированием, в том числе в кольце вычетов по заданному модулю M , и анализируются характеристики обнаружения ими ошибок при разнообразных комбинациях весовых коэффициентов.

Исследования автора настоящей статьи показывают, что некоторые взвешенные коды с суммированием в кольце вычетов по заданному модулю M обладают особыми свойствами: выходы их кодеров описываются самодвойственными и «близкими» к ним булевыми функциями [17]. Такое особое свойство ряда взвешенных кодов с суммированием позволяет использовать их при синтезе самопроверяемых устройств. В настоящей статье описываются некоторые результаты исследований применения взвешенных кодов с суммированием в кольце вычетов по модулю $M = 4$ – взвешенных кодов Боуза–Лина с модулем $M = 4$ – при синтезе структур с контролем вычислений по двум диагностическим признакам.

1. Взвешенные коды с суммированием

Взвешенные коды с суммированием строятся по следующему алгоритму.

Алгоритм 1. Правила получения разрядов контрольных векторов взвешенных кодов с суммированием:

1. Задается значение m .
2. Выбирается значение модуля $M \in \mathbb{N}$, $M \geq 2$.
3. Формируется массив весовых коэффициентов $[w_m, w_{m-1}, \dots, w_2, w_1]$, $w_i \in \{1, 2, \dots, M-1\}$, $i = \overline{1, m}$.
4. Определяется сумма весовых коэффициентов единичных разрядов информационного вектора – число $W_M = \left(\sum_{i=1}^m f_i w_i \right) (\bmod M)$, где f_i – значение i -го разряда информационного вектора $\langle f_m f_{m-1} \dots f_2 f_1 \rangle$.
5. Число W_M представляется в двоичном виде и записывается в $k = \lceil \log_2 M \rceil$ разрядов контрольного вектора.

Обозначим взвешенные коды с суммированием как $WS(m, k, M)$ коды. Отдельно при необходимости будем указывать массив весовых коэффициентов.

Приведем пример получения значений разрядов контрольных векторов для информационного вектора $\langle 1011 \rangle$ $WS(4, 2, 4)$ кода с массивом весовых коэффициентов $[w_4, w_3, w_2, w_1] = [1, 1, 2, 3]$:

$$W_4 = \left(\sum_{i=1}^{m=4} f_i w_i \right) (\bmod 4) = (1 \cdot 3 + 1 \cdot 2 + 0 \cdot 1 + 1 \cdot 1) (\bmod 4) = 6 (\bmod 4) = 2, [2]_{10} = [10]_2.$$

Чем больше значение M , тем больше число вариантов взвешивания разрядов информационного вектора при построении кода. Различные массивы весовых коэффициентов дают разные по характеристикам коды.

Рассмотрим $WS(m, k, M)$ коды и, в частности, более подробно взвешенные коды Боуза–Лина $WS(4, 2, 4)$.

2. Особые классы взвешенных кодов Боуза–Лина со значением модуля $M = 4$

Исследования показывают, что среди множества $WS(m, k, M)$ кодов можно выделить такие, которые обладают особыми свойствами – самодвойственности функций, описывающих проверочные символы кодовых слов, а также «близости» к данному классу функций.

Утверждение 1. Не существует $WS(m, k, M)$ кодов с $M \notin \{2^1, 2^2, \dots, 2^{\left\lceil \log_2 \left(\sum_{i=1}^m w_i + 1 \right) \right\rceil}\}$, все проверочные символы которых описываются самодвойственными булевыми функциями.

Доказательство. Предположим, что утверждение 1 неверно, и можно построить $WS(m, k, M)$ код с $M \notin \{2^1, 2^2, \dots, 2^{\left\lceil \log_2 \left(\sum_{i=1}^m w_i + 1 \right) \right\rceil}\}$, все проверочные символы которого описываются самодвойственными булевыми функциями. Эта особенность означает, что как минимум для двух информационных векторов $\langle 00 \dots 00 \rangle$ и $\langle 11 \dots 11 \rangle$ будут получены контрольные векторы, разряды в которых полностью отличаются. Так как код строится путем суммирования весовых коэффициентов разрядов, то для информационного вектора $\langle 00 \dots 00 \rangle$ всегда получается суммарный вес, равный 0, или двоичный вектор $\langle 00 \dots 00 \rangle$. Это означает, что если все функции, описывающие проверочные символы $WS(m, k, M)$ кода самодвойственны, то для информационного вектора $\langle 11 \dots 11 \rangle$ суммарный вес в кольце вычетов по модулю $M \notin \{2^1, 2^2, \dots, 2^{\left\lceil \log_2 \left(\sum_{i=1}^m w_i + 1 \right) \right\rceil}\}$ должен давать такое число, которому соответствует контрольный вектор $\langle 11 \dots 11 \rangle$. В качестве суммарного веса в кольце вычетов по заданному модулю M может получиться только число из множества $\{0, 1, \dots, M-1\}$. Двоичному числу $\langle 11 \dots 11 \rangle$ соответствует только такое значение вычета, которое описывается формулой $2^j - 1, j \in \mathbb{N}$. Но все числа такого вида получаются в случае $M-1$. Или, иначе, $M-1 = 2^j - 1 \Leftrightarrow M = 2^j$. Отсюда становится ясным, что невозможно при условии $M \notin \{2^1, 2^2, \dots, 2^{\left\lceil \log_2 \left(\sum_{i=1}^m w_i + 1 \right) \right\rceil}\}$ получить контрольные векторы $\langle 00 \dots 00 \rangle$ и $\langle 11 \dots 11 \rangle$ для информационных векторов $\langle 00 \dots 00 \rangle$ и $\langle 11 \dots 11 \rangle$, а значит, и функции, описывающие проверочные символы такого кода, одновременно не могут оказаться самодвойственными. ■

Таким образом, при построении СВК с контролем вычислений по признаку самодвойственности целесообразно сосредоточиться на рассмотрении $WS(m, k, M)$ кодов, для которых $M \in \{2^1, 2^2, \dots, 2^{\left\lceil \log_2 \left(\sum_{i=1}^m w_i + 1 \right) \right\rceil}\}$.

$WS(m, 1, 2)$ код является кодом паритета [6]. Для него условия, при которых все выходы описываются самодвойственными булевыми функциями, установлены в [18].

Интересными с практической точки зрения могут оказаться приложения $WS(m, 2, 4)$ кодов, поскольку данные коды имеют малое число проверочных символов, а число способов построения их – значительно и определяется количеством вариантов взвешивания разрядов информационного вектора.

Остановимся на частном случае $WS(m, 2, 4)$ кодов при $m = 4$. Их особенностью является то, что они могут рассматриваться как «базовые» коды для построения СВК. СВК на основе $WS(4, 2, 4)$ кода строится для групп выходов по 4 – при способе дополнения информационных векторов для кодовых слов разделимых кодов в СВК, и для групп выходов по 6 – при способе, подразумевающем логическую коррекцию сигналов.

Путем перебора всех возможных значений весовых коэффициентов без учета их расположения в массиве (это справедливо при рассмотрении полного множества кодовых слов) можно построить 15 $WS(4, 2, 4)$ кодов. Среди них будет только два кода, обладающих свойством самодвойственности функций, описывающих оба проверочных символа. Кроме того, могут быть выделены $WS(4, 2, 4)$ коды, обладающие проверочными символами со «схожими» с самодвойственными функциями.

Исследования показывают, что среди 15 помехозащищенных $WS(4, 2, 4)$ кодов существует только два, кодеры которых являются самодвойственными.

Утверждение 2. Кодеры $WS(4, 2, 4)$ кодов с весовыми коэффициентами $[w_4, w_3, w_2, w_1] = [1, 1, 2, 3]$ и $[w_4, w_3, w_2, w_1] = [1, 2, 2, 2]$ являются самодвойственными.

Доказательство. Для доказательства данного факта синтезируем кодеры на полных, полу- и четвертьсумматорах (сумматоры по модулю $M = 2$) [19]. При синтезе представим весовые коэффициенты в виде двоичных чисел, в которых значения единичных разрядов будут заменены значениями

функций, описывающих соответствующие разряды информационного вектора. Для кода с массивом весовых коэффициентов $[1, 1, 2, 3]$ имеем: $\langle 0f_4 \rangle$, $\langle 0f_3 \rangle$, $\langle f_2 0 \rangle$, $\langle f_1 f_1 \rangle$. Для кода с массивом весовых коэффициентов $[1, 2, 2, 2]$ имеем: $\langle 0f_4 \rangle$, $\langle f_3 0 \rangle$, $\langle f_2 0 \rangle$, $\langle f_1 0 \rangle$. Далее синтезируем схему сложения двоичных чисел $\langle 0f_4 \rangle$, $\langle 0f_3 \rangle$, $\langle f_2 0 \rangle$ и $\langle f_1 f_1 \rangle$ в кольце вычетов по модулю $M = 4$ (рис. 1, а). Аналогично те же процедуры выполним для второго случая (рис. 1, б).

Сфокусируемся на структуре кодера для первого случая. На первом выходе сумматора формируется значение младшего разряда суммы всех весов по формуле $g_1 = f_4 \oplus f_3 \oplus f_1$. На втором выходе сумматора вычисляется мажоритарная функция $a = f_4 f_3 \vee f_4 f_1 \vee f_3 f_1$. Она, в свою очередь, входит как аргумент в функцию $g_2 = f_2 \oplus f_1 \oplus a$. Функция сложения по модулю $M = 2$ от нечетного числа аргументов, от которых она зависит существенно, является самодвойственной [18, 20]. Мажоритарная функция также является самодвойственной [18]. Отсюда следует, что g_1 – самодвойственная. Поскольку класс самодвойственных функций замкнут [21], то функция $g_2 = f_2 \oplus f_1 \oplus a$ в силу самодвойственности функции $a = f_4 f_3 \vee f_4 f_1 \vee f_3 f_1$ также является самодвойственной. Отсюда следует, что кодер данного $WS(4, 2, 4)$ кода представляет собой самодвойственное устройство.

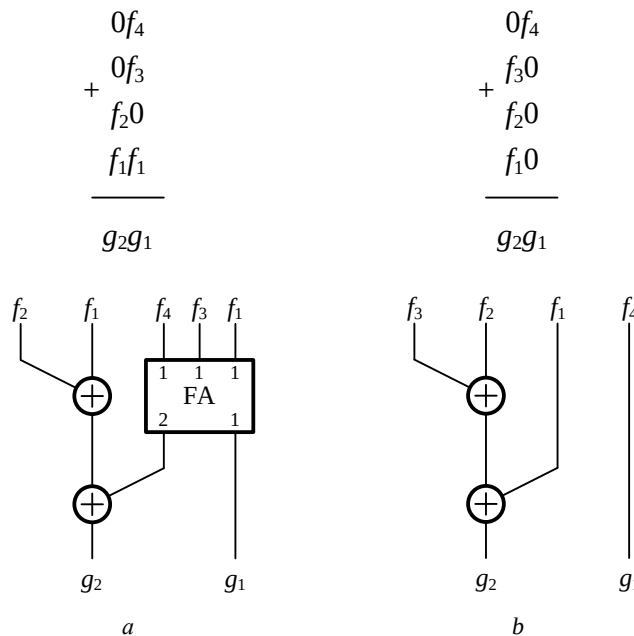


Рис. 1. Кодеры $WS(4, 2, 4)$ кодов: а – для $[w_4, w_3, w_2, w_1] = [1, 1, 2, 3]$; б – для $[w_4, w_3, w_2, w_1] = [1, 2, 2, 2]$

Fig. 1. $WS(4, 2, 4)$ codes encoders: a) for $[w_4, w_3, w_2, w_1] = [1, 1, 2, 3]$; b) for $[w_4, w_3, w_2, w_1] = [1, 2, 2, 2]$

Рассмотрим структуру кодера для второго случая. Функция $g_1 = f_4$, а $g_2 = f_3 \oplus f_2 \oplus f_1$. Обе функции являются самодвойственными. Кодер данного $WS(4, 2, 4)$ кода также представляет собой самодвойственное устройство. ■

Аналогично нетрудно показать, что кодеры всех остальных $WS(4, 2, 4)$ кодов самодвойственными не являются: для каких-то кодеров одна из функций является самодвойственной, а для каких-то – нет.

Классифицируем $WS(4, 2, 4)$ коды по тому, к каким типам устройств относятся их кодеры, выделяя самодвойственные (SD-функции, от англ. self-dual) и «близкие» к ним по свойствам – самоквазидвойственные¹ (SQD-функции, от англ. self-quasidual) выходы. Таким образом, классифицируем $WS(4, 2, 4)$ коды по особым классам в соответствии с описанными признаками (рис. 2). Особый класс обозначим как p_2/p_1 группу $WS(4, 2, 4)$ кодов, для которых функция g_2 имеет признак p_2 и функция g_1

¹ SQD-функция – это функция, которая квазидвойственна самой себе. Квазидвойственная функция схожа с двойственной, однако на инверсных наборах она принимает одинаковые значения, в отличие от двойственной. В некоторых источниках SQD-функции определены как самоантидвойственные, например в [17].

имеет признак p_1 , $p_1, p_2 \in \{SD, SQD\}$; если функция не принадлежит к классу SD или SQD, то отсутствие какого-либо из указанных признаков укажем знаком «-».

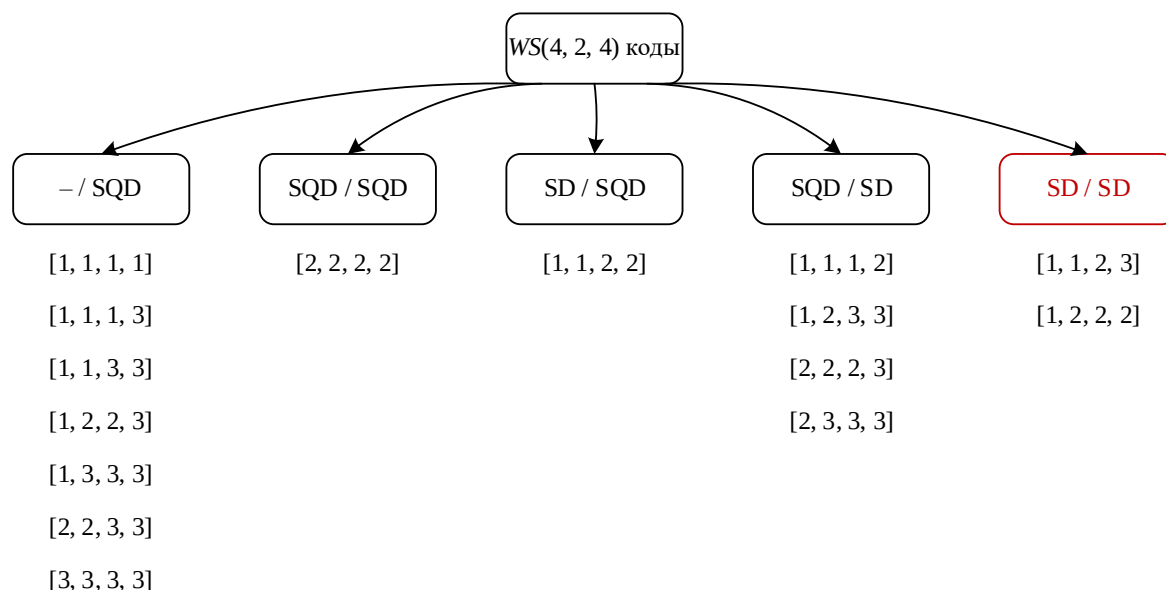


Рис. 2. Разделение WS(4, 2, 4) кодов на особые классы

Fig. 2. Division of WS(4, 2, 4) codes into special classes

Таблица 1

Характеристики обнаружения ошибок SD/SD-WS(4, 2, 4) кодами
и классификации информационных векторов между всеми контрольными векторами

[w ₄ , w ₃ , w ₂ , w ₁]											
[1, 1, 2, 3]						[1, 2, 2, 2]					
Распределения информационных векторов между всеми контрольными векторами											
	W ₄						W ₄				
	0	1	2	3			0	1	2	3	
	Контрольные векторы						Контрольные векторы				
	00	01	10	11			00	01	10	11	
	Информационные векторы						Информационные векторы				
	0000	0011	0010	0001			0000	1000	0001	1001	
	0101	0100	0111	0110			0011	1011	0010	1010	
	1001	1000	1011	1010			0101	1101	0100	1100	
1110	1101	1100	1111	0110	1110	0111	1111				
Характеристика необнаруживаемых ошибок в кодовых словах											
240v _[1, 1, 2, 3] = 44v ₂ + 20σ ₂ + 16v ₃ + 80α ₃ + + 20v ₄ + 28σ ₄ + 16α ₄ + 2v ₆ + 4σ ₆ + 10α ₆						240v _[1, 2, 2, 2] = 64v ₂ + 48σ ₂ + 28v ₄ + 36σ ₄ + 48α ₄ + 2v ₆ + 14α ₆					
Характеристика необнаруживаемых ошибок в информационных векторах											
48η _[1, 1, 2, 3] = 16v ₂ + 8σ ₂ + 4v ₃ + 20α ₃						48η _[1, 2, 2, 2] = 24v ₂ + 24σ ₂					

Примечание. Для кодов дано распределение необнаруживаемых ошибок в кодовых словах в виде: $Nv_{[w_4, w_3, w_2, w_1]} = Nv_d + N\sigma_d + N\alpha_d$, где N – число необнаруживаемых ошибок: v – всего в кодовых словах, η – всего в информационных векторах, v_d – число монотонных необнаруживаемых ошибок кратностью d , σ_d – число симметричных необнаруживаемых ошибок кратностью d , α_d – число асимметричных необнаруживаемых ошибок кратностью d .

Выделение особых классов WS(4, 2, 4) кодов важно с практической точки зрения при организации схем встроенного контроля (СВК) по признакам самодвойственности и / или самоквазидвойственности вычисляемых кодерами функций. Прежде всего, важной является группа SD/SD. Обозначим коды, входящие в данную особую группу, SD/SD-WS(4, 2, 4) кодами. Однако с учетом того, что для получения самодвойственной функции из несамодвойственной достаточно всего одной дополнительной переменной [22, 23], важны и остальные классы WS(4, 2, 4) кодов, каждый из которых может быть применен при реализации контроля вычислений с некоторой модификацией кодера [20].

При синтезе СВК с применением разделимых кодов важным является распределение информационных векторов между всеми контрольными векторами, поскольку такое распределение позволяет установить используемые кодовые слова данных кодов с учетом условий их формирования на каждой входной комбинации. Также важным будет являться распределение необнаруживаемых ошибок по видам и кратностям, поскольку потребуется покрыть все виды ошибок на выходах контролируемых устройств.

В табл. 1 приведены необходимые для синтеза СВК сведения для двух рассматриваемых SD/SD-WS(4, 2, 4) кодов.

Рассмотрим два способа организации СВК при использовании SD/SD-WS(4, 2, 4) кодов.

3. Структуры организации СВК на основе SD/SD-WS(4, 2, 4) кодов

С использованием SD/SD-WS(4, 2, 4) кодов могут быть организованы СВК для комбинационных составляющих цифровых устройств по двум представленным далее структурам.

Структура рис. 3 строится исключительно для четырех выходов самодвойственного устройства $F(x)$. Выходы устройства подключаются ко входам СВК, содержащей два функциональных блока. Блок $G(f)$ является кодером SD/SD-WS(4, 2, 4) кода и формирует контрольные векторы по значениям информационных векторов, вычисляемых блоком $F(x)$. Оба разряда контрольного вектора описываются самодвойственными булевыми функциями и контролируются с помощью устройства 2SDC1, осуществляющего контроль самодвойственности каждого из сигналов от кодера $G(f)$. 2SDC1 реализуется на основе двух элементарных тестеров самодвойственного сигнала SDC (от англ. self-dual checker) [20] и одного модуля сжатия парафазных сигналов TRC (от англ. two-rail checker) [24]. Выходы последнего являются и контрольными выходами СВК z^0 и z^1 , на которых при исправности всех блоков и отсутствии ошибок в вычислениях формируется парафазный сигнал; нарушение же парафазности свидетельствует об ошибке в вычислениях, вызванной сбоями или неисправностями устройства $F(x)$ или одного из блоков СВК.

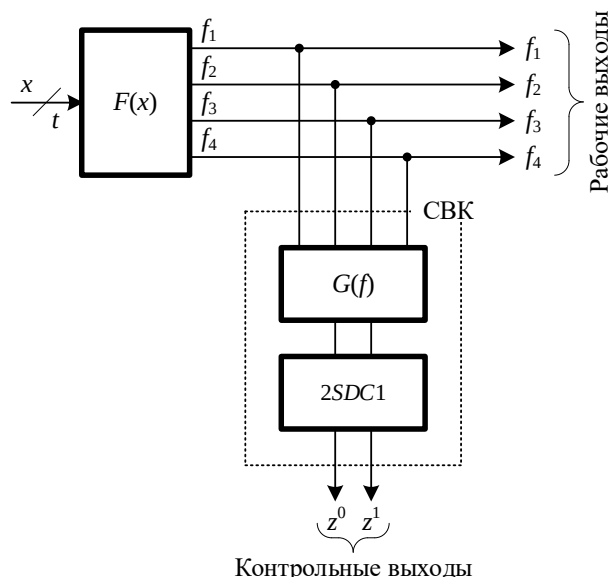


Рис. 3. Структура организации СВК на основе SD/SD-WS(4, 2, 4) кода со сжатием сигналов и контроля вычислений по признаку самодвойственности функций

Fig. 3. Structure of the concurrent checking organization based on SD/SD-WS(4, 2, 4) code with compression of signals and control of calculations based on self-duality of functions

При синтезе СВК для полного покрытия ошибок на выходах устройства $F(x)$ требуется учитывать особенности его структуры и характеристики обнаружения ошибок в информационных векторах SD/SD-WS(4, 2, 4) кодами (см. табл. 1), а также свойства самодвойственных функций для обнаружения ошибок [25]. При этом для многовыходных устройств осуществляется покрытие полного множе-

ства выходов группами по четыре выхода в каждой (для повышения обнаруживающих способностей одни и те же выходы могут повторяться в разных группах).

Структура на рис. 4 реализуется несколько сложнее. При этом она может быть синтезирована как для самодвойственного, так и для несамодвойственного устройства. Данная структура строится для устройства с шестью выходами. В СВК значения с выходов объекта диагностирования благодаря наличию блока коррекции сигналов (БКС) корректируются в кодовые слова $\langle h_6 h_5 h_4 h_3 h_2 h_1 \rangle$ SD/SD-WS(4, 2, 4) кода. БКС образован двухвходовыми сумматорами по модулю $M = 2$ (элементами XOR), на первые выходы которых поступают сигналы от объекта диагностирования, а на вторые – от блока контрольной логики $G(x)$. Благодаря тому что кодовое слово $\langle h_6 h_5 h_4 h_3 h_2 h_1 \rangle$ принадлежит именно SD/SD-WS(4, 2, 4) коду, может быть организован контроль вычислений по двум диагностическим признакам: по принадлежности формируемого вектора $\langle h_6 h_5 h_4 h_3 h_2 h_1 \rangle$ множеству кодовых слов SD/SD-WS(4, 2, 4) кода и по принадлежности к классу самодвойственных каждой функции, описывающей информационные и проверочные символы. Это осуществляется с помощью модуля контроля вычислений (МКВ). В нем выделяется тестер кода TSC (от англ. totally self-checking checker), схема контроля самодвойственности $6SDC1$ и модуль TRC . TSC состоит из кодера SD/SD-WS(4, 2, 4) кода, модуля TRC и двух инверторов для обеспечения функционирования модуля TRC в парафазной логике. $6SDC1$ представляет собой устройство, включающее в себя либо 6 элементарных тестеров самодвойственности SDC и компаратор для сравнения 6 сигналов, синтезируемый на основе пяти модулей TRC , либо схему сжатия самодвойственных сигналов в один самодвойственный сигнал и один модуль SDC . Особенности SDC и схем сжатия самодвойственных сигналов описаны в [18, 20]. Последний каскад МКВ образует единственный модуль TRC , осуществляющий сжатие контрольных выходов от двух подсхем контроля в один контрольный сигнал z^0 и z^1 .

Такое обустройство СВК позволяет повышать число тестовых комбинаций среди рабочих и, соответственно, увеличивать показатели наблюдаемости ошибок в структуре [20].

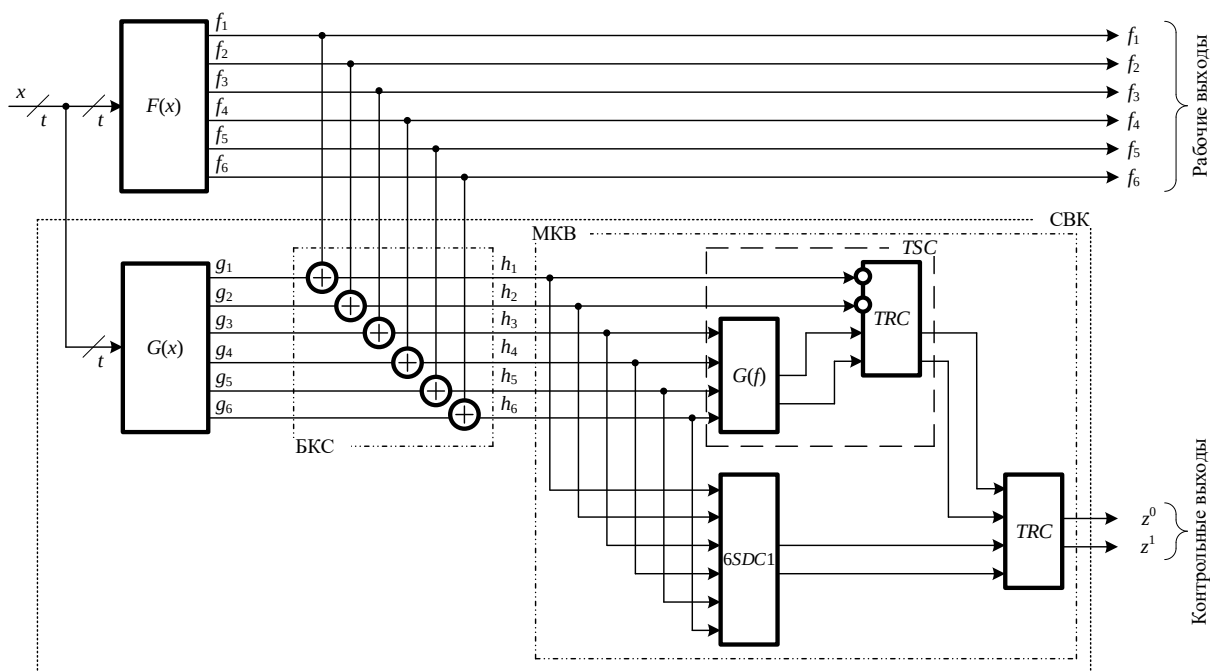


Рис. 4. Структура организации СВК на основе SD/SD-WS(4, 2, 4) кода с контролем вычислений по двум диагностическим признакам

Fig. 4. Structure of the concurrent checking organization based on SD/SD-WS(4, 2, 4) code with control of calculations based on two diagnostic criteria

Для работы структур рис. 3 и рис. 4 требуется импульсный режим работы и подача на входы прямой и ортогональной по всем переменным (инверсной) комбинаций [23, 26]. Это позволяет подавать комбинации парами – рабочая и контрольная. Тем самым в ходе работы последовательно фор-

мируются прямая и инверсная комбинации, что требует определенных принципов доопределения векторов $\langle h_6 \ h_5 \ h_4 \ h_3 \ h_2 \ h_1 \rangle$ на входных комбинациях: на прямой и инверсной комбинациях должны формироваться кодовые слова с расстоянием Хэмминга $d = 6$. При выборе алгоритма доопределения функций, определяющих информационные и проверочные символы SD/SD-WS(4, 2, 4) кода, это необходимо учитывать.

4. Алгоритмы синтеза СВК

Структура СВК, основанная на сжатии сигналов с использованием кодера SD/SD-WS(4, 2, 4) кода, реализуется по следующим правилам.

Алгоритм 2. Правила синтеза СВК для самодвойственного устройства на основе схемы сжатия с применением кодера SD/SD-WS(4, 2, 4) кода:

1. Выходы устройства разбиваются на группы по четыре выхода в каждой.
2. Для каждой группы строится структура, приведенная на рис. 4, на основе выбранного SD/SD-WS(4, 2, 4) кода (для разных групп можно применить разные SD/SD-WS(4, 2, 4) коды).
3. Выходы отдельных схем контроля объединяются на входах самопроверяемого компаратора.

При использовании алгоритма 2 не осуществляется целенаправленный выбор групп выходов объекта диагностирования с учетом характеристик обнаружения ошибок SD/SD-WS(4, 2, 4) кодами, поэтому требуется проверка покрытия ошибок и их обнаружения с помощью самодвойственного контроля. Несмотря на то, что ошибки могут оказаться в классе необнаруживаемых SD/SD-WS(4, 2, 4) кодом, они могут быть обнаружены на прямой и инверсной входных комбинациях при наличии искажения на одной из них.

Структура СВК с контролем вычислений по двум диагностическим признакам строится сложнее.

Алгоритм 3. Правила синтеза СВК для самодвойственного устройства на основе ЛКС с применением SD/SD-WS(4, 2, 4) кода и контролем вычислений по двум диагностическим признакам:

1. Выходы устройства разбиваются на группы по шесть выходов в каждой.
2. Строится таблица, описывающая сигналы на входах СВК, внутренних линиях и ее выходах.
3. Осуществляется доопределение значений функций h_i , $i = \overline{1, 6}$, с учетом таких особенностей:

– на парах входных комбинаций, соответствующих десятичным номерам $(j, (2^t - 1) - j)$, $j = \overline{0, 2^{t-1} - 1}$,

должны присутствовать кодовые слова SD/SD-WS(4, 2, 4) кода, имеющие расстояния Хэмминга $d = 6$;

– для каждого элемента XOR должен формироваться проверяющий тест $\{00, 01, 10, 11\}$ [27];

– должны хотя бы по разу сформироваться все контрольные векторы SD/SD-WS(4, 2, 4) кода, что требуется для проверки тестера.

4. Вычисляются функции логической коррекции сигналов g_i : $g_i = f_i \oplus h_i$, $i = \overline{1, 6}$.

5. Проверяется формирование множества тестовых комбинаций для элементов XOR и тестера (если это не было выполнено автоматически при доопределении значений функций логической коррекции сигналов).

6. Реализуется блок МКВ в виде типового устройства.

7. Выходы отдельных схем контроля объединяются на входах самопроверяемого компаратора.

При использовании алгоритма 3 ключевым шагом является выполнение п. 3, поскольку именно от процедуры доопределения значений функций h_i будет зависеть сложность реализации блока контрольной логики $G(x)$. Здесь существует внушительное количество способов доопределения: можно использовать как все возможные кодовые слова SD/SD-WS(4, 2, 4) кода, так и некоторое их подмножество с учетом выполнения требований п. 3 алгоритма 3. При доопределении только первая или вторая половина таблицы относительно ее середины дополняется вариативно: если заполнена какая-либо строка первой половины, то строка, соответствующая инверсной входной комбинации, заполняется автоматически с учетом условий п. 3 алгоритма 3. Отсюда и из таблиц распределений информационных векторов между контрольными векторами (см. табл. 1) следует, что при доопределении каждой из 2^{t-1} строк существует максимум по 2^4 вариантов доопределения, а общее их количество

определяется величиной $2^{t-1}2^4 = 2^{t+3}$. Например, если устройство имеет $t = 4$ входа, то существует $2^7 = 128$ способов доопределения и, как следствие, 128 различных блоков $G(x)$, позволяющих синтезировать структуру рис. 4.

Если алгоритм 2 представляется простым, то алгоритм 3 требует процедур построчного анализа булевых векторов, формируемых на выходах объекта диагностирования, и булевых векторов, формируемых на выходах БКС с последующим вычислением функций логической коррекции сигналов. Приведем здесь алгоритм доопределения функций логической коррекции сигналов с использованием SD/SD-WS(4, 2, 4) кодов, подразумевающий на этапах доопределения значений функций логической коррекции обеспечение тестируемости элементов СВК. Данный алгоритм является развитием алгоритма из [28], описанного для применения кодов Боуза–Лина при несамодвойственном контроле вычислений.

Алгоритм 4. Правила доопределения функций логической коррекции сигналов до функций, описывающих символы SD/SD-WS(4, 2, 4) кодов с массивом весовых коэффициентов [1, 1, 2, 3]:

1. Из анализа таблицы распределения информационных векторов между контрольными векторами следует (см. табл. 1), что в каждом столбце формируются по разу информационные векторы $\langle 00 \sim \rangle$, $\langle 01 \sim \rangle$, $\langle 10 \sim \rangle$, $\langle 11 \sim \rangle$, что позволяет положить значения функций, описывающих два старших разряда¹ на первой половине таблицы истинности, равными $h_6 = 0$, $h_5 = 0$, а на второй – равными $h_6 = 1$, $h_5 = 1$.

2. Проверяется наличие тестовых комбинаций для элементов преобразования рабочих функций, соответствующих старшим разрядам, для чего достаточно, чтобы на первой половине и на второй половине таблицы истинности формировались хотя бы по разу значения 0 и 1 для f_6 и f_5 .

3. Таким образом, на каждой строке можно доопределить оставшиеся значения разрядов информационного и контрольного векторов четырьмя способами. При этом от заполнения оставшихся информационных символов или оставшихся проверочных символов будут напрямую зависеть два недоопределенных символа кодового слова.

4. Как только доопределяется значение какой-либо функции на какой-либо строке, определяется значение функции логической коррекции сигналов на этой строке: $g_i = f_i \oplus h_i$.

5. Для обеспечения тестируемости элементов XOR на первой половине таблицы истинности (для строк с номерами $\{0, 1, \dots, 2^{t-1} - 1\}$) выполняется следующее:

5.1. Для каждой из функций f_4 и f_3 находится по два значения 0 и по два значения 1.

5.2. Для двух нулей и для двух единиц на разных строках формируются значения $h_4 = 0$, $h_4 = 1$ и $h_3 = 0$, $h_3 = 1$, что обеспечивает формирование теста для двух элементов преобразования.

5.3. Если после этого оказываются заполненными оба значения h_3 и h_4 на какой-либо строке, то вычисляются и значения h_2 и h_1 .

5.4. Далее для каждой из функций f_2 и f_1 находится по два значения 0 и по два значения 1.

5.5. Проверяется наличие для двух нулей и для двух единиц на разных строках формирования значений $h_2 = 0$, $h_2 = 1$ и $h_1 = 0$, $h_1 = 1$.

5.6. Если какое-либо из значений на предыдущем шаге не сформировалось, то аналогично п. 5.2 значения h_2 и h_1 доопределяются до формирования по разу значений $h_2 = 0$, $h_2 = 1$ и $h_1 = 0$, $h_1 = 1$, что дает проверяющий тест для оставшихся элементов преобразования.

6. Для обеспечения тестируемости TSC на первой половине таблицы истинности (для строк с номерами $\{0, 1, \dots, 2^{t-1} - 1\}$) выполняется следующее:

6.1. Проверяется формирование как минимум по разу контрольных векторов $\langle h_2 h_1 \rangle = \{00, 01, 10, 11\}$.

6.2. Если хотя бы по разу контрольные векторы не сформировались, то на оставшихся незаполненных строках получают требуемые контрольные векторы и, как следствие, доопределяют значения еще на ряде строк.

¹ Можно обратить внимание на то, что для кода с массивом весовых коэффициентов [1, 2, 2, 2] положить равными нулю следует младшие разряды информационного вектора: $h_4 = 0$, $h_3 = 0$ (см. табл. 1).

7. На противоположных строках таблицы истинности относительно ее середины значения функций заполняются антисимметрично (противоположными сигналами).

8. Оставшиеся строки на первой половине таблицы истинности заполняют произвольно с учетом формирования наиболее простых функций логической коррекции сигналов, а затем на противоположных строках относительно середины таблицы истинности – антисимметрично.

9. Если какое-то из условий не удовлетворяется, то выполняют замену столбцов путем их перестановки, если это возможно. Если это оказывается невозможным после всех перестановок, то требуется выделить «шестерок» групп выходов объекта диагностирования иначе. Если и это не дает результата, то построить самопроверяемую СВК не удастся.

10. Далее синтезируются блок $G(x)$ и сама СВК в выбранном элементном базисе.

Продemonстрируем работу алгоритма 4 на примере организации СВК для произвольного шестивыходного устройства с четырьмя входами. Устройство задано в табл. 2. Там же даны и значения функций логической коррекции сигналов и значения разрядов кодовых слов SD/SD-WS(4, 2, 4) кода с массивом весовых коэффициентов [1, 1, 2, 3].

Таблица 2

Описание исходного устройства и части сигналов для тестирования элементов СВК

№	x_4	x_3	x_2	x_1	f_6	f_5	f_4	f_3	f_2	f_1	h_6	h_5	h_4	h_3	h_2	h_1	g_6	g_5	g_4	g_3	g_2	g_1
0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0
1	0	0	0	1	0	1	0	0	0	1	0	0	1	1	0	1	0	1	1	1	0	0
2	0	0	1	0	0	1	0	1	0	1	0	0	1	0	1	0	0	1	1	1	1	1
3	0	0	1	1	0	0	1	1	1	0	0	0	0	1	1	1	0	0	1	0	0	1
4	0	1	0	0	0	1	1	0	0	0	0	0	1	*	*	*	0	1	0	*	*	*
5	0	1	0	1	1	1	0	0	1	0	0	0	*	*	*	*	1	1	*	*	*	*
6	0	1	1	0	0	1	0	0	1	1	0	0	*	*	*	*	0	1	*	*	*	*
7	0	1	1	1	0	1	0	0	1	0	0	0	*	*	*	*	0	1	*	*	*	*
8	1	0	0	0	1	0	1	1	1	1	1	1	*	*	*	*	0	1	*	*	*	*
9	1	0	0	1	1	0	0	1	0	0	1	1	*	*	*	*	0	1	*	*	*	*
10	1	0	1	0	0	1	0	1	0	0	1	1	*	*	*	*	1	0	*	*	*	*
11	1	0	1	1	0	0	1	0	0	0	1	1	0	*	*	*	1	1	1	*	*	*
12	1	1	0	0	1	0	1	0	0	1	1	1	1	0	0	0	0	1	0	0	0	1
13	1	1	0	1	0	1	1	0	0	1	1	1	0	1	0	1	1	0	1	1	0	0
14	1	1	1	0	0	1	0	1	1	0	1	1	0	0	1	0	1	0	0	1	0	0
15	1	1	1	1	1	0	0	1	1	0	1	1	1	1	1	1	0	1	1	0	0	1

Примечание. Знаком «*» отмечены те клетки таблицы, значения которых на первоначальном этапе не определены.

Заполняем все столбцы h_6 и h_5 значениями 0 на первой половине таблицы и значениями 1 – на второй. Проверяем условия формирования проверяющих тестов для элементов преобразования. Вычисляем значения функций g_6 и g_5 . Рассматриваем столбец h_4 : находим строки, соответствующие двум нулям и двум единицам. Это строки с номерами 0, 1, 3, 4. Вычисляем значения функции g_4 на этих строках. Рассматриваем столбец h_3 : находим строки, где записаны два нуля и две единицы для f_3 и заполняем их. Это строки с номерами 0, 1, 2, 3. Вычисляем значения функции g_3 на этих строках. Формируем на заполненных значениями строках h_4 и h_3 значения h_2 и h_1 и g_2 и g_1 . В данном случае это строки с номерами 0, 1, 3. Проверяем формирование контрольных векторов SD/SD-WS(4, 2, 4) кода: формируются контрольные векторы <00>, <01>, <11>. Доопределяем значение h_4 на строке с № 2 таким образом, чтобы была сформирована недостающая комбинация <10>. Доопределяем значения функций логической коррекции. Таким образом, указанные шаги обеспечивают тестируемость элементов коррекции и тестера.

Далее заполняем оставшиеся строки таким образом, чтобы функции логической коррекции сигналов были простыми, при этом кодовые слова на выходах БКС должны принадлежать SD/SD-WS(4, 2, 4) коду с массивом весовых коэффициентов [1, 1, 2, 3]. Результат заполнения представлен в табл. 3.

Для построения блока $G(x)$ функции g_i оптимизируются как система.

Описание работы устройства и СВК на требуемых линиях

№	x ₄	x ₃	x ₂	x ₁	f ₆	f ₅	f ₄	f ₃	f ₂	f ₁	h ₆	h ₅	h ₄	h ₃	h ₂	h ₁	g ₆	g ₅	g ₄	g ₃	g ₂	g ₁
0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	1	0
1	0	0	0	1	0	1	0	0	0	1	0	0	1	1	0	1	0	1	1	1	0	0
2	0	0	1	0	0	1	0	1	0	1	0	0	1	0	1	0	0	1	1	1	1	1
3	0	0	1	1	0	0	1	1	1	0	0	0	0	1	1	1	0	0	1	0	0	1
4	0	1	0	0	0	1	1	0	0	0	0	0	1	0	1	0	0	1	0	0	1	0
5	0	1	0	1	1	1	0	0	1	0	0	0	0	0	0	0	1	1	0	0	1	0
6	0	1	1	0	0	1	0	0	1	1	0	0	0	0	0	0	0	1	0	0	1	1
7	0	1	1	1	0	1	0	0	1	0	0	0	0	0	0	0	0	1	0	0	1	0
8	1	0	0	0	1	0	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	0
9	1	0	0	1	1	0	0	1	0	0	1	1	1	1	1	1	0	1	1	0	1	1
10	1	0	1	0	0	1	0	1	0	0	1	1	1	1	1	1	1	0	1	0	1	1
11	1	0	1	1	0	0	1	0	0	0	1	1	0	1	0	1	1	1	1	1	0	1
12	1	1	0	0	1	0	1	0	0	1	1	1	1	0	0	0	0	1	0	0	0	1
13	1	1	0	1	0	1	1	0	0	1	1	1	0	1	0	1	1	0	1	1	0	0
14	1	1	1	0	0	1	0	1	1	0	1	1	0	0	1	0	1	0	0	1	0	0
15	1	1	1	1	1	0	0	1	1	0	1	1	1	1	1	1	0	1	1	0	0	1

Оценим сложность реализации устройства с помощью программного средства SIS и библиотеки функциональных элементов stdcell2_2.genlib [29, 30]. Из анализа структуры рис. 4 следует, что показатель сложности ее реализации определяется величиной

$$L_{CED} = L_{F(x)} + L_{G(x)} + 6L_{XOR} + L_{G(f)} + L_{6SDC1} + 2L_{NOT} + 2L_{TRC}, \quad (1)$$

где $L_{F(x)}$ – показатель сложности реализации устройства $F(x)$, $L_{G(x)}$ – показатель сложности реализации устройства $G(x)$, $L_{G(f)}$ – показатель сложности кодера, L_{6SDC1} – показатель сложности реализации схемы контроля самодвойственности, L_{INV} – показатель сложности реализации инвертора, L_{TRC} – показатель сложности реализации TRC .

Блок $F(x)$ синтезируем по таблице его описания (см. табл. 2) без оптимизации структуры (исходное устройство). Его сложность в условных единицах stdcell2_2.genlib равна $L_{F(x)} = 1\,144$ усл. ед. Блок $G(x)$ синтезируем по полученной таблице (см. табл. 3) и оптимизируем его структуру с помощью команды full_simplify. Показатель сложности его реализации равен $L_{G(x)} = 880$ усл. ед. Сложность одного элемента XOR равна $L_{XOR} = 40$. Структура кодера приведена на рис. 1, а и оценивается величиной в 256 усл. ед. Схему контроля самодвойственности реализуем путем каскадного соединения схемы сжатия самодвойственных сигналов (см. структуру схемы сжатия: [18. § 4.2; 31. § 9.10.2]) и одного тестера самодвойственности. Схема сжатия сигналов содержит три трехвходовых элемента XOR и оценивается показателем в 240 усл. ед. Тестер самодвойственности оценивается показателем в 96 усл. ед. Вся схема контроля самодвойственности имеет показатель сложности реализации $L_{6SDC1} = 336$ усл. ед. Инвертор имеет показатель сложности $L_{INV} = 16$ усл. ед., а модуль TRC – $L_{TRC} = 192$ усл. ед. Итого по формуле (1) имеем

$$L_{CED} = 1144 + 880 + 6 \cdot 40 + 256 + 336 + 2 \cdot 16 + 2 \cdot 192 = 3\,272.$$

Сравним с показателем сложности реализации устройства с СВК на основе дублирования [32]:

$$L_D = 2L_{F(x)} + nL_{NOT} + (n-1)L_{TRC}, \quad (2)$$

где n – число выходов устройства $F(x)$.

В нашем случае формула (2) дает следующий результат:

$$L_D = 2 \cdot 1144 + 6 \cdot 16 + (6-1) \cdot 192 = 3\,344.$$

Устройство, синтезированное по предложенному методу, имеет показатель сложности реализации, составляющий примерно 97,847% от показателя сложности реализации устройства с СВК на основе дублирования. Выигрыш в показателе сложности является небольшим, однако нужно учитывать, что за счет импульсного режима работы повышается контролепригодность структуры!

Конечно, здесь можно учесть и то, что для создания режима автоматической импульсной работы потребуется установка на входах структуры рис. 4 схемы преобразования входных векторов в последовательности прямоугольных импульсов. Она состоит из генератора прямоугольных импульсов и элементов XOR для каждого входа [23, 26]. Это тоже внесет свой вклад в усложнение структуры и в рассматриваемом случае приведет к превышению показателя сложности реализации в сравнении с дублированием. Однако превышение окажется несущественным на фоне значительного улучшения в показателях контролепригодности¹. Более того, для устройств со сложными структурами и большим числом выходов разница между дублированием будет значительнее за счет увеличения сложности компаратора в структуре СВК на основе дублирования.

Таким образом, предложенный метод в ряде случаев будет давать более простые по показателям структурной избыточности СВК, чем при дублировании, и при этом будет улучшен показатель контролепригодности за счет увеличения числа тестовых комбинаций среди рабочих.

Заключение

Использование SD/SD-WS(4, 2, 4) кодов позволяет организовывать СВК с контролем вычислений по двум диагностическим признакам – принадлежности формируемых в СВК кодовых слов SD/SD-WS(4, 2, 4) кодам и принадлежности каждой функции, описывающей символы кодовых слов, классу самодвойственных. Это, как показано в ряде работ, например в [20, 33], дает возможность синтеза полностью самопроверяемых цифровых устройств с улучшенными показателями контролепригодности в части наблюдаемости.

В статье описан один из алгоритмов доопределения функций логической коррекции сигналов, который подразумевает получение значений построчно с учетом формирования проверяющих тестов для элементов преобразования и тестера в СВК. В этом можно усмотреть и преимущество, и недостаток. Преимущество заключается непосредственно в формировании проверяющих тестов для блоков СВК. Недостаток же связан с необходимостью построчного анализа таблиц и, возможно, перестановкой выходов и выборов выходов для формирования их группы. Следует также отметить возможность организации СВК по представленному методу для объектов диагностирования с произвольными структурами. При этом чем больше число входов объекта диагностирования, тем проще обеспечить полную самопроверяемость СВК.

Интересно, что для организации СВК по двум диагностическим признакам можно применить и не только SD/SD-WS(4, 2, 4) коды – можно применить любые из представленных в классификации, данной на рис. 2, кодов с учетом данных статьи [20]. Для этого несамодвойственный выход следует преобразовать в самодвойственный с использованием всего одного элемента XOR и альтернативного сигнала от генератора, подаваемого на второй его вход. Тем самым производится модификация кодера и самого WS(4, 2, 4) кода в δ -код, проверочные символы которого описываются самодвойственными булевыми функциями. Далее можно для нового δ -кода использовать аналог представленного в настоящей статье для SD/SD-WS(4, 2, 4) кода подхода.

При синтезе СВК с учетом диагностического признака самодвойственности могут быть использованы как WS(m , 2, 4) коды с большими m , так и другие WS(m , k , M) коды, проверочные символы которых описываются самодвойственными булевыми функциями, со значениями

$$M \in \{2^1, 2^2, \dots, 2^{\left\lceil \log_2 \left(\sum_{i=1}^m w_i + 1 \right) \right\rceil}\}.$$

Для построения таких кодов могут быть выделены соответствующие массивы весовых коэффициентов среди всего их многообразия. К примеру, для случая $m = 5$ и $M = 4$ (их можно применять при контроле вычислений на «семерках» выходов устройств) массивы [1, 1, 1, 1, 3], [1, 1, 1, 2, 2], [1, 1, 3, 3, 3], [1, 2, 2, 3, 3], [2, 2, 2, 2, 3] и [3, 3, 3, 3, 3] дают самодвойственные кодеры WS(5, 2, 4) кодов.

¹ Структуру дублирования можно синтезировать и для условий импульсного режима работы, что так же, как и в рассматриваемом случае применения модульных взвешенных кодов с суммированием, потребует наличия линии элементов XOR на входах и генератора прямоугольных импульсов.

Для дальнейших исследований интерес представляют вопросы моделирования самопроверяемых устройств с контролем вычислений по двум диагностическим признакам, а также постановка экспериментов с тестовыми схемами [34].

Модульные взвешенные коды с суммированием представляют интересный класс равномерных делимых кодов, использование которых позволяет синтезировать полностью самопроверяемые цифровые устройства, в том числе для блоков и узлов критического применения, где важными являются требования к контролепригодности в условиях редкой смены входных воздействий [35, 36].

Список источников

1. Согомонян Е.С., Слабаков Е.В. Самопроверяемые устройства и отказоустойчивые системы. М. : Радио и связь, 1989, 208 с.
2. Микони С.В. Общие диагностические базы знаний вычислительных систем. СПб. : СПИИРАН, 1992. 234 с.
3. Lala P.K. Self-Checking and Fault-Tolerant Digital Design. San Francisco : Morgan Kaufmann Publishers, 2001. 216 p.
4. Дрозд А.В., Харченко В.С., Антошук С.Г., Дрозд Ю.В., Дрозд М.А., Сулима Ю.Ю. Рабочее диагностирование безопасных информационно-управляющих систем / под ред. А.В. Дрозда, В.С. Харченко. Харьков : Нац. аэрокосмический ун-т им. Н.Е. Жуковского «ХАИ», 2012. 614 с.
5. Сапожников В.В., Сапожников Вл.В., Ефанов Д.В. Коды с суммированием для систем технического диагностирования. М. : Наука, 2020. Т. 1: Классические коды Бергера и их модификации. 383 с.
6. Сапожников В.В., Сапожников Вл.В., Ефанов Д.В. Коды с суммированием для систем технического диагностирования. М. : Наука, 2021. Т. 2: Взвешенные коды с суммированием. 455 с.
7. Слабаков Е.В. Построение полностью самопроверяемых комбинационных устройств с использованием остаточных кодов // Автоматика и телемеханика. 1979. № 10. С. 133–141.
8. Sapozhnikov V., Sapozhnikov V.I., Efanov D. Modular Sum Code in Building Testable Discrete Systems // Proc. of 13th IEEE East-West Design & Test Symposium (EWDTS'2015), Batumi, Georgia, September 26–29. 2015. P. 181–187. doi: 10.1109/EWDTS.2015.7493133
9. Piestrak S.J. Design of Self-Testing Checkers for Unidirectional Error Detecting Codes. Wrocław : Oficyna Wydawnicza Politechniki Wrocławskiej, 1995. 111 p.
10. Berger J.M. A Note on Error Detection Codes for Asymmetric Channels // Information and Control. 1961. V. 4, is.1. P. 68–73. doi: 10.1016/S0019-9958(61)80037-5
11. Das D., Toubia N.A. Weight-Based Codes and Their Application to Concurrent Error Detection of Multilevel Circuits // Proc. of 17th IEEE Test Symposium, California, USA, 1999. P. 370–376, doi: 10.1109/VTEST.1999.766691
12. Das D., Toubia N.A., Seuring M., Gossel M. Low Cost Concurrent Error Detection Based on Modulo Weight-Based Codes // Proc. of the 6th IEEE International On-Line Testing Workshop (IOLTW 2000), 3–5 July 2000, Palma de Mallorca, Spain. IEEE Computer Society, 2000. P. 171–176. doi: 10.1109/OLT.2000.856633
13. Сапожников В.В., Сапожников Вл.В., Ефанов Д.В. Взвешенные коды с суммированием для организации контроля логических устройств // Электронное моделирование. 2014. Т. 36, № 1. С. 59–80.
14. Efanov D.V., Sapozhnikov V.V., Sapozhnikov V.I. Using Codes with Summation of Weighted Bits to Organize Checking of Combinational Logical Devices // Automatic Control and Computer Sciences. 2019. V. 53, is. 1. P. 1–11. doi: 10.3103/S0146411619010061
15. Сапожников В.В., Сапожников Вл.В., Ефанов Д.В. Модульно-модифицированные взвешенные коды с суммированием, обнаруживающие любые ошибки нечетных кратностей // Электронное моделирование. 2018. Т. 40, № 3. С. 41–61. doi: 10.15407/emodel.40.03.041
16. Ефанов Д.В., Сапожников В.В., Сапожников Вл.В. Коды с суммированием с фиксированными значениями кратностей обнаруживаемых монотонных и асимметричных ошибок для систем технического диагностирования // Автоматика и телемеханика. 2019. № 6. С. 121–141.
17. Шалыто А.А. Логическое управление. Методы аппаратной и программной реализации. СПб. : Наука, 2000. 780 с.
18. Сапожников В.В., Сапожников Вл.В., Гессель М. Самодвойственные дискретные устройства. СПб. : Энергоатомиздат, С.-Петербург, отд-ние, 2001. 331 с.
19. Harris D.M., Harris S.L. Digital Design and Computer Architecture. Morgan Kaufmann, 2012. 569 p.
20. Ефанов Д.В., Погодина Т.С. Исследование свойств самодвойственных комбинационных устройств с контролем вычислений на основе кодов Хэмминга // Информатика и автоматизация. 2023. Т. 22, № 2. С. 349–392. doi: 10.15622/ia.22.2.5
21. Яблонский С.В., Гаврилов Г.П., Кудрявцев В.Б. Функции алгебры логики и классы Поста. М. : Наука, 1966. 120 с.
22. Аксёнова Г.П. Восстановление в дублированных устройствах методом инвертирования данных // Автоматика и телемеханика. 1987. № 10. С. 144–153.
23. Gössel M., Ocheretny V., Sogomonyan E., Marienfeld D. New Methods of Concurrent Checking. Ed. 1. Dordrecht : Springer Science + Business Media B.V., 2008, 184 p.
24. Carter W.C., Duke K.A., Schneider P.R. Self-Checking Error Checker for Two-Rail Coded Data. United States Patent Office, filed July 25, 1968, ser. No. 747533, patented Jan. 26, 1971, New York. 10 p.

25. Гессель М., Морозов А.А., Сапожников В.В., Сапожников Вл.В. Построение самопроверяемых комбинационных схем на основе свойств самодвойственных функций // Автоматика и телемеханика. 2000. № 2. С. 151–163.
26. Гессель М., Мошанин В.И., Сапожников В.В., Сапожников Вл.В. Обнаружение неисправностей в самопроверяемых комбинационных схемах с использованием свойств самодвойственных функций // Автоматика и телемеханика. 1997. № 12. С. 193–200.
27. Аксенова Г.П. Необходимые и достаточные условия построения полностью проверяемых схем свертки по модулю 2 // Автоматика и телемеханика. 1979. № 9. С. 126–135.
28. Ефанов Д.В. Синтез самопроверяемых комбинационных устройств на основе метода логической коррекции сигналов с применением кодов Боуза–Лина // Информационные технологии. 2023. Т. 29, № 10. С. 503–511. doi: 10.17587/it.29.503-511
29. Sentovich E.M., Singh K.J., Lavagno L., Moon C., Murgai R., Saldanha A., Savoj H., Stephan P.R., Brayton R.K., Sangiovanni-Vincentelli A. SIS: A System for Sequential Circuit Synthesis. Electronics Research Laboratory, Department of Electrical Engineering and Computer Science, University of California, Berkeley, 1992. 45 p.
30. Sentovich E.M., Singh K.J., Moon C., Savoj H., Brayton R.K., Sangiovanni-Vincentelli A. Sequential Circuit Design Using Synthesis and Optimization // Proc. IEEE International Conference on Computer Design: VLSI in Computers & Processors, 11–14 October 1992. Cambridge, MA, 1992. P. 328–333, doi: 10.1109/ICCD.1992.276282
31. Сапожников В.В., Сапожников Вл.В., Ефанов Д.В. Основы теории надежности и технической диагностики. СПб. : Лань, 2019. 588 с.
32. Efanov D.V., Sapozhnikov V.V., Sapozhnikov V.I. The Self-Checking Concurrent Error-Detection Systems Synthesis Based on the Boolean Complement to the Bose-Lin Codes with the Modulo Value $M = 4$ // Electronic Modeling. 2021. V. 43, is. 1. P. 28–45. doi: 10.15407/emodel.43.01.028
33. Ефанов Д.В., Погодина Т.С. Самодвойственные цифровые устройства с контролем вычислений по кодам Сяо // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2023. № 63. С. 118–136. doi: 10.17223/19988605/63/14
34. Collection of Digital Design Benchmarks. URL: <https://ddd.fit.cvut.cz/www/prj/Benchmarks/>
35. Drozd A., Kharchenko V., Antoshchuk S., Sulima J., Drozd M. Checkability of the Digital Components in Safety-Critical Systems: Problems and Solutions // Proc. of 9th IEEE East-West Design & Test Symposium (EWDTS'2011). Sevastopol, 2011. P. 411–416. doi: 10.1109/EWDTS.2011.6116606
36. Drozd O., Perebeinos I., Martynyuk O., Zashcholkina K., Ivanova O., Drozd M. Hidden Fault Analysis of FPGA Projects for Critical Applications // Proc. of the IEEE International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET), 25–29 February 2020. Lviv-Slavsko, 2020. Art. 142. doi: 10.1109/TCSET49122.2020.235591

References

1. Sogomonyan, E.S. & Slabakov, E.V. (1989) *Samoproveryaemye ustroystva i otkazoustoychivye sistemy* [The Self-Checking Devices and Fault-Tolerant Systems]. Moscow: Radio i svyaz'.
2. Mikoni, S.V. (1992) *Obshchie diagnosticheskie bazy znaniy vychislitel'nykh sistem* [General diagnostic knowledge bases of computer systems]. St. Petersburg: SPIIRAS.
3. Lala, P.K. (2001) *Self-Checking and Fault-Tolerant Digital Design*. San Francisco: Morgan Kaufmann Publishers.
4. Drozd, A.V., Kharchenko, V.S., Antoshchuk, S.G., Drozd, Yu.V., Drozd M.A. & Sulima Yu.Yu. (2012) *Rabochee diagnostirovanie bezopasnykh informatsionno-upravlyayushchikh sistem* [Working Diagnostics of Safe Information and Control Systems]. Kharkov: N.E. Zhukovsky National Aerospace University.
5. Sapozhnikov, V.V., Sapozhnikov, V.I. & Efanov, D.V. (2020) *Kody s summirovaniem dlya sistem tekhnicheskogo diagnostirovaniya* [Sum Codes for Technical Diagnostics Systems]. Vol. 1. Moscow: Nauka.
6. Sapozhnikov, V.V., Sapozhnikov, V.I. & Efanov, D.V. (2021) *Kody s summirovaniem dlya sistem tekhnicheskogo diagnostirovaniya* [Sum Codes for Technical Diagnostics Systems]. Vol. 2. Moscow: Nauka.
7. Slabakov, E.V. (1979) Design of completely Self-Checking Combinational Devices with the Use of Residual Codes. *Automation and Remote Control*. 10. pp. 133–141.
8. Sapozhnikov, V., Sapozhnikov, V.I., & Efanov, D. (2015) Modular Sum Code in Building Testable Discrete Systems. *Proceedings of 13th IEEE East-West Design & Test Symposium (EWDTS'2015)*. Batumi, Georgia, September 26–29, 2015. pp. 181–187. DOI: 10.1109/EWDTS.2015.7493133
9. Piestrak, S.J. (1995) *Design of Self-Testing Checkers for Unidirectional Error Detecting Codes*. Wrocław: Oficyna Wydawnicza Politechniki Wrocławskiej.
10. Berger, J.M. (1961) A Note on Error Detection Codes for Asymmetric Channels. *Information and Control*. 4(1). pp. 68–73. DOI: 10.1016/S0019-9958(61)80037-5
11. Das, D. & Touba, N.A. (1999) Weight-Based Codes and Their Application to Concurrent Error Detection of Multilevel Circuits. *Proceedings of 17th IEEE Test Symposium*. California, USA. pp. 370–376. DOI: 10.1109/VTEST.1999.766691
12. Das, D., Touba, N.A., Seuring, M. & Gossel, M. (2000) Low Cost Concurrent Error Detection Based on Modulo Weight-Based Codes. *Proceedings of the IEEE 6th International*. July 3–5, 2000. Palma de Mallorca, Spain. pp. 171–176. DOI: 10.1109/OLT.2000.856633.

13. Sapozhnikov, V.V., Sapozhnikov, V.I. & Efanov, D.V. (2014) Vzveshennyye kody s summirovaniem dlya organizatsii kontrolya logicheskikh ustroystv [Weight-Based Sum Codes for Organizing Control of Logical Devices]. *Elektronnnoye modelirovaniye*. 36(1). pp. 59–80.
14. Efanov, D.V., Sapozhnikov, V.V. & Sapozhnikov, V.I. (2019) Using Codes with Summation of Weighted Bits to Organize Checking of Combinational Logical Devices. *Automatic Control and Computer Sciences*. 53(1). pp. 1–11. DOI: 10.3103/S0146411619010061
15. Sapozhnikov, V.V., Sapozhnikov, V.I. & Efanov, D.V. (2018) Modul'no-modifitsirovannyye vzveshennyye kody s summirovaniem, obnaruzhivayushchie lyubyye oshibki nechetnykh kratnostey [Modular-Modified Weighted Summation Codes Detecting Any Errors of Odd Multiplicities]. *Elektronnnoye modelirovaniye*. 40(3). pp. 41–61. DOI: 10.15407/emodel.40.03.041
16. Efanov, D.V., Sapozhnikov, V.V. & Sapozhnikov, V.I. (2019) Sum Codes with Fixed Values of Multiplicities for Detectable Unidirectional and Asymmetrical Errors for Technical Diagnostics of Discrete Systems. *Automation and Remote Control*. 6. pp. 121–141.
17. Shalyto, A.A. (2000) *Logicheskoye upravleniye. Metody apparatnoy i programmnoy realizatsii* [Logic Control. Methods of Hardware and Software Implementation]. St. Petersburg: Nauka.
18. Sapozhnikov, V.V., Sapozhnikov, V.I. & Göessel, M. (2001) *Samodvoystvennyye diskretnyye ustroystva* [Self-dual Digital Devices]. St. Petersburg: Energoatomizdat.
19. Harris, D.M., & Harris, S.L. (2012) *Digital Design and Computer Architecture*. Morgan Kaufmann.
20. Efanov, D.V. & Pogodina, T.S. (2023) Issledovaniye svoystv samodvoystvennykh kombinatsionnykh ustroystv s kontrolem vychisleniy na osnove kodov Khemminga [Properties Investigation of Self-Dual Combinational Devices with Calculation Control Based on Hamming Codes]. *Informatika i avtomatizatsiya*. 22(2). pp. 349–392. DOI: 10.15622/ia.22.2.5
21. Yablonsky, S.V., Gavrilov, G.P. & Kudryavtsev, V.B. (1966) *Funktsii algebry logiki i klassy Posta* [Logic algebra functions and Post classes]. Moscow: Nauka.
22. Aksenova, G.P. (1987) Vosstanovleniye v dublirovannykh ustroystvakh metodom invertirovaniya dannykh [Restoration in Duplicated Units by the Method of Data Inversion]. *Automation and Remote Control*. 10. pp. 144–153.
23. Göessel, M., Ocheretny, V., Sogomonyan, E. & Marienfeld, D. (2008) *New Methods of Concurrent Checking*. 1st ed. Dordrecht: Springer Science+Business Media B.V.
24. Carter, W.C., Duke, K.A. & Schneider, P.R. (1971) *Self-Checking Error Checker for Two-Rail Coded Data*. United States Patent Office, filed July 25, 1968, ser. No. 747533, patented Jan. 26, 1971. N. Y.
25. Göessel, M., Morozov, A.A., Sapozhnikov, V.V. & Sapozhnikov, V.I. (2000) Self-Testing Combinational Circuits: Their Design Through the Use of the Properties of Self-Dual Functions. *Automation and Remote Control*. 2. pp. 151–163.
26. Göessel, M., Moshanin, V.I., Sapozhnikov, V.V. & Sapozhnikov, V.I. (1997) Fault Detection in Self-Test Combination Circuits Using the Properties of Self-Dual Functions. *Automation and Remote Control*. 12. pp. 193–200.
27. Aksenova, G.P. (1979) Necessary and Sufficient Conditions for Design of Completely Checkable Module 2 Convolution Circuits. *Automation and Remote Control*. 9. pp. 126–135.
28. Efanov, D.V. (2023) Sintez samoproveryaemykh kombinatsionnykh ustroystv na osnove metoda logicheskoy korrektsii signalov s primeneniem kodov Bouza–Lina [Self-Checking Combinational Devices Synthesis Based on the Boolean Signal Correction Method Using Bose-Lin Codes]. *Informatsionnyye tekhnologii*. 29(10). pp. 503–511. DOI: 10.17587/it.29.503-511
29. Sentovich, E.M., Singh, K.J., Lavagno, L., Moon, C., Murgai, R., Saldanha, A., Savoj, H., Stephan, P.R., Brayton, R.K. & Sangiovanni-Vincentelli, A. (1992) *SIS: A System for Sequential Circuit Synthesis*. Electronics Research Laboratory, Department of Electrical Engineering and Computer Science, University of California, Berkeley. 4th May.
30. Sentovich, E.M., Singh, K.J., Moon, C., Savoj, H., Brayton, R.K. & Sangiovanni-Vincentelli, A. (1992) Sequential Circuit Design Using Synthesis and Optimization. *Proceedings IEEE International Conference on Computer Design: VLSI in Computers & Processors*. October 11–14, 1992. Cambridge, MA, USA. pp. 328–333. DOI: 10.1109/ICCD.1992.276282
31. Sapozhnikov, V.V., Sapozhnikov, V.I. & Efanov, D.V. (2019) *Osnovy teorii nadezhnosti i tekhnicheskoy diagnostiki* [Fundamentals of Reliability Theory and Technical Diagnostics]. St. Petersburg: Lan'.
32. Efanov, D.V., Sapozhnikov, V.V. & Sapozhnikov, V.I. (2021) The Self-Checking Concurrent Error-Detection Systems Synthesis Based on the Boolean Complement to the Bose-Lin Codes with the Modulo Value $M=4$. *Electronic Modeling*. 43(1). pp. 28–45. DOI: 10.15407/emodel.43.01.028
33. Efanov, D.V. & Pogodina, T.S. (2023) Self-Dual Digital Devices with Calculations Testing by Hsiao Codes. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravleniye, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 63. pp. 118–136. DOI: 10.17223/19988605/63/14
34. *Collection of Digital Design Benchmarks*. [Online] Available from: <https://ddd.fit.cvut.cz/www/prj/Benchmarks/>
35. Drozd, A., Kharchenko, V., Antoshchuk, S., Sulima, J. & Drozd, M. (2011) Checkability of the Digital Components in Safety-Critical Systems: Problems and Solutions. *Proceedings of 9th IEEE East-West Design & Test Symposium (EWDTS'2011)*. Sevastopol, Ukraine, 2011. pp. 411–416. DOI: 10.1109/EWDTS.2011.6116606
36. Drozd, O., Perebeinos, I., Martynyuk, O., Zashcholkina, K., Ivanova, O. & Drozd, M. (2020) Hidden Fault Analysis of FPGA Projects for Critical Applications. *Proceedings of the IEEE International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)*. February 25–29, 2020. Lviv-Slavsko, Ukraine. Art. 142. DOI: 10.1109/TCSET49122.2020.235591

Информация об авторе:

Ефанов Дмитрий Викторович – доктор технических наук, профессор, действительный член Международной академии транспорта, член Института инженеров электротехники и электроники (IEEE); профессор Высшей школы транспорта Института машиностроения, материалов и транспорта Санкт-Петербургского политехнического университета Петра Великого (СПбПУ Петра Великого) (Санкт-Петербург, Россия); профессор кафедры «Автоматика, телемеханика и связь на железнодорожном транспорте» Российского университета транспорта (МИИТ) (Москва, Россия); кафедра «Автоматика и телемеханика» Ташкентского государственного транспортного университета (Ташкент, Узбекистан). E-mail: TrES-4b@yandex.ru

Авторы заявляют об отсутствии конфликта интересов.

Information about the author:

Efanov Dmitry V. (Doctor of Technical Sciences, Professor, Full-member of the International Transport Academy, IEEE member, Professor of Transport Higher School of Mechanical Engineering, Material and Transport Institute at Peter the Great Saint Petersburg Polytechnic University, St. Petersburg, Russian Federation; Professor of Automation, Remote Control and Communication on Railway Transport Department, Russian University of Transport Moscow, Russian Federation; Professor of Automation and Remote Control Department of Tashkent State Transport University, Tashkent, Uzbekistan). E-mail: TrES-4b@yandex.ru

The author declares no conflicts of interests.

Received 16.04.2024; accepted for publication 03.09.2024

Поступила в редакцию 16.04.2024; принята к публикации 03.09.2024

Научный журнал

**ВЕСТНИК
ТОМСКОГО
ГОСУДАРСТВЕННОГО
УНИВЕРСИТЕТА**

**УПРАВЛЕНИЕ,
ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА
И ИНФОРМАТИКА**

**TOMSK STATE UNIVERSITY
JOURNAL OF CONTROL AND COMPUTER SCIENCE**

2024. № 68

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 23.09.2024 г. Формат 60x84^{1/8}.
Гарнитура Times. Усл. печ. л. 16,3.
Тираж 250 экз. Заказ № 6024. Цена свободная.

Дата выхода в свет 26.09.2024 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