

ВЕСТНИК ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА

УПРАВЛЕНИЕ,
ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА
И ИНФОРМАТИКА

TOMSK STATE UNIVERSITY
JOURNAL OF CONTROL AND COMPUTER SCIENCE

Научный журнал

2024

№ 67

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
Матрсова Анжела Юрьевна – проф., д-р техн. наук, НИ ТГУ
Назаров Анатолий Андреевич – проф., д-р техн. наук, НИ ТГУ
Пашинская Татьяна Юрьевна – д-р физ.-мат. наук, НИ ТГУ
Перепелкин Евгений Александрович – проф., д-р техн. наук, Санкт-Петербургский государственный университет аэрокосмического приборостроения (г. Санкт-Петербург)
Самуйлов Константин Евгеньевич – проф., д-р техн. наук, РУДН (г. Москва)
Семенкин Евгений Станиславович – проф., д-р техн. наук, СибГУ им. акад. М.Ф. Решетнева (г. Красноярск)
Сущенко Сергей Петрович – проф., д-р техн. наук, НИ ТГУ
Фархадов Маис Паша оглы – д-р техн. наук, Институт проблем управления РАН (г. Москва)
Харин Юрий Семенович – академик НАНБ, проф., д-р физ.-мат. наук, БГУ, Республика Беларусь (г. Минск)
Цициашвили Гурами Шалвович – проф., д-р физ.-мат. наук, ИПМ ДВО РАН (г. Владивосток)

РЕДАКЦИОННЫЙ СОВЕТ

Александр Николаевич Дудин – проф., д-р физ.-мат. наук, БГУ, Минск, Республика Беларусь
Ана Роза Ковали (Ana Rosa Cavalli) – проф., д-р. философии (PhD), University VII Paris, Франция
Енцо Орзингер (Enco Orzinger) – 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 Зорниан (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

Мансимов К.Б., Сулейманова Ш.Ш. Аналог принципа максимума Понтрягина в одной задаче оптимального управления с переменной структурой	4
Parshukov A.N. Technology for improving robust quality of control for one-dimensional linear discrete control systems with structural-parametric uncertainty	12

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

Имомов А.А., Тухтаев Э.Э. Локальная предельная теорема для ветвящихся случайных систем Гальтона–Ватсона с иммиграцией и бесконечной дисперсией	22
Кузнецов К.С., Амосова Е.В. Численное решение системы уравнений Навье–Стокса в случае сжимаемой среды с использованием нейронных сетей	31
Цициашвили Г.Ш., Харченко Ю.А. Стационарное распределение в системах массового обслуживания с переменной структурой	42

ОБРАБОТКА ИНФОРМАЦИИ

Абдушукуров А.А., Бозоров С.Б., Мансуров Д.Р. Оптимальное ядерное оценивание плотности методом статистического моделирования при случайном цензурировании наблюдений	49
Гаврилов К.В., Веретельникова Е.Л. Об одном способе выбора компромисса в семействе условно оптимальных оценок	60
Доронина Ю.В., Скатков А.В. Вероятностный подход при контроле состояния сложной технической аппаратуры	69
Осинцева Е.А., Чимитова Е.В. Оптимальное планирование эксперимента на основе винеровской деградационной модели с ковариатами с использованием условной информационной матрицы Фишера	83

ИНФОРМАТИКА И ПРОГРАММИРОВАНИЕ

Лещинский Д.В., Данилкин Е.А., Старченко А.В. Параллельная реализация микромасштабной модели турбулентного движения воздуха и переноса примеси с использованием технологии OpenMP	94
Tran V.T., Korikov A.M. Path planning for mobile objects based on modification of the probabilistic roadmap method	106

ПРОЕКТИРОВАНИЕ И ДИАГНОСТИКА ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

Матросова А.Ю., Пасько А.Д., До Там Дык. Построение эквивалентных схем в системах троирования	116
---	-----

CONTENTS

CONTROL OF DYNAMICAL SYSTEMS

Mansimov K.B., Suleymanov Sh.Sh. An analogue of Pontryagin's maximum principle in one problem optimal control with variable structure	4
Parshukov A.N. Technology for improving robust quality of control for one-dimensional linear discrete control systems with structural-parametric uncertainty	12

MATHEMATICAL MODELING

Imomov A.A., Tukhtaev E.E. Local limit theorem for branching random Galton-Watson systems with immigration and infinite variance	22
Kuznetsov K.S., Amosova E.V. Numerical solution of the system of Navier-Stokes equations in the case of a compressible medium using neural networks	31
Tsitsiashvili, G.Sh., Kharchenko Yu.N. Stationary distribution in queuing systems with a variable structure	42

DATA PROCESSING

Abdushukurov A.A., Bozorov S.B., Mansurov D.R. Optimal kernel density estimation by statistical modeling with random censoring of observations	49
Gavrilov K.V., Veretel'nikova E.L. On one way to choose a compromise in a family of conditionally optimal estimators	60
Doronina Yu.V., Skatkov A.V. A probabilistic approach to monitoring the condition of complex technical equipment	69
Osintseva E.A., Chimitova E.V. Optimal design of experiment based on the Wiener degradation model with covariate using conditional Fisher information matrix	83

INFORMATICS AND PROGRAMMING

Leshchinskiy D.V., Danilkin E.A., Starchenko A.V. Parallel implementation of microscale model of turbulent air movement and pollutant transport using OpenMP technology	94
Tran V.T., Korikov A.M. Path planning for mobile objects based on modification of the probabilistic roadmap method	106

DESIGNING AND DIAGNOSTICS OF COMPUTER SYSTEMS

Matrosova A.Yu., Pasko A.D., Do Tam Duc. Deriving equivalent circuits for TMR technique	116
---	-----

УПРАВЛЕНИЕ ДИНАМИЧЕСКИМИ СИСТЕМАМИ

CONTROL OF DYNAMICAL SYSTEMS

Научная статья

УДК 517.977.56

doi: 10.17223/19988605/67/1

Аналог принципа максимума Понтрягина в одной задаче
оптимального управления с переменной структуройКамиль Байрамали оглы Мансимов¹, Шабнам Шакир кызы Сулейманова²¹ Бакинский государственный университет, Баку, Азербайджан, kamilbmansimov@gmail.com^{1, 2} Институт систем управления Министерства науки и образования Азербайджана,
Баку, Азербайджан, kmansimov@mail.ru

Аннотация. Рассматривается одна задача оптимального управления системами с распределенными параметрами, описываемая в двух различных областях двумя системами Гурса–Дарбу при предположении произвольности областей управления. Критерий качества является терминального типа функционалом. На основе модифицированного варианта метода приращений доказано необходимое условие оптимальности в форме аналога принципа максимума Л.С. Понтрягина.

Ключевые слова: система Гурса–Дарбу; задача оптимального управления с переменной структурой; допустимое управление; необходимое условие оптимальности.

Для цитирования: Мансимов К.Б., Сулейманова Ш.Ш. Аналог принципа максимума Понтрягина в одной задаче оптимального управления с переменной структурой // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 67. С. 4–11. doi: 10.17223/19988605/67/1

Original article

doi: 10.17223/19988605/67/1

An analogue of Pontryagin's maximum principle in one problem
optimal control with variable structureKamil B. Mansimov¹, Shabnam Sh. Suleymanova²¹ Baku State University, Baku, Azerbaijan, kamilbmansimov@gmail.com^{1, 2} Institute of Control Systems of the Ministry of Science and Education of Azerbaijan,
Baku, Azerbaijan, kmansimov@mail.ru

Abstract. In this paper we consider one optimal control problem with distributed parameters described in two different domains by two Goursat-Darboux systems under the assumption that the control domains are arbitrary. The quality criterion is a terminal type functional. Based on a modified version of the increment method, a necessary condition for optimality is proved in the form of an analogue of Pontryagin's maximum principle.

Keywords: Goursat-Darboux systems; optimal control problem with variable structure; admissible control; necessary condition for optimality.

For citation: Mansimov, K.B., Suleymanov, Sh.Sh. (2024) An analogue of Pontryagin's maximum principle in one problem optimal control with variable structure. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 67. pp. 4–11. doi: 10.17223/19988605/67/1

Введение

Как отмечено, например, в работах [1–4], многие процессы, являясь сложными, носят многоэтапный характер.

Задачи оптимального управления подобными процессами называются ступенчатыми задачами оптимального управления [1–3], или задачами оптимального управления с переменной структурой [4].

В работах [1–4] и других изучен ряд задач оптимального управления с переменной структурой, описываемых обыкновенными дифференциальными уравнениями.

В работе [5] рассмотрена задача оптимального управления с переменной структурой, описываемая совокупностью обыкновенных дифференциальных уравнений и интегральных уравнений типа Вольтерра в предположении открытости областей управления. Получен аналог уравнения Эйлера и выведены необходимые условия оптимальности второго порядка.

Выводу необходимого условия оптимальности первого порядка и необходимых условий оптимальности второго порядка в задаче оптимального управления, описываемой системами гиперболических уравнений второго порядка с краевыми условиями Гурса (задача оптимального управления системами Гурса–Дарбу), посвящена работа [6]. Получены аналог уравнения Эйлера и аналог условия Лежандра–Клебша. Изучен случай вырождения аналога условия оптимальности Лежандра–Клебша (классически особый случай). Доказаны необходимые условия оптимальности особых в классическом смысле управлений.

В предлагаемой статье рассматривается одна задача оптимального управления с переменной структурой, описываемая в двух различных областях двумя гиперболическими уравнениями второго порядка с краевыми условиями Гурса (краевая задача Гурса–Дарбу) в предположении, что управляющие функции входят в класс измеримых и ограниченных вектор-функций, а области управления являются непустыми ограниченными и замкнутыми множествами

С использованием введенных сопряженных систем построена формула приращения первого порядка функционала качества.

С помощью аналога игольчатых вариаций Макшейна (см., напр.: [7–11]) доказано необходимое условие оптимальности в форме принципа максимума Л.С. Понтрягина [7–10].

В качестве следствия из него получен аналог дифференциального принципа максимума [9].

1. Постановка задачи

Пусть $D_i = [t_{i-1}, t_i] \times [x_0, x_1]$, $i = 1, 2$, – заданные прямоугольники, $U_1 \subset R^r, U_2 \subset R^q$ – заданные непустые, ограниченные и замкнутые множества, $u_1(t, x)(u_2(t, x))$ – $r(q)$ -мерная измеримая и ограниченная вектор-функция управляющих воздействий (допустимые управляющие функции), $\varphi_1(z)$ и $\varphi_2(y)$ – заданные имеющие непрерывные производные скалярные функции, $a(x)$ и $b_i(t)$, $i = 1, 2$, – заданные абсолютно непрерывные вектор-функции.

Допустим, что управляемый двухэтапный процесс в различных областях D_i , $i = 1, 2$, описывается краевыми задачами Гурса–Дарбу для нелинейных гиперболических уравнений второго порядка

$$z_{tx} = f_1(t, x, z, u_1), (t, x) \in D_1, \quad (1)$$

$$\begin{aligned} z(t_0, x) &= a(x), x \in X = [x_0, x_1], \\ z(t, x_0) &= b_1(t), t \in T_1 = [t_0, t_1], \end{aligned} \quad (2)$$

$$y_{tx} = f_2(t, x, y, u_2), (t, x) \in D_2, \quad (3)$$

$$\begin{aligned} y(t_1, x) &= G(z(t_1, x)), x \in X, \\ y(t, x_0) &= b_2(t), t \in T_2 = [t_1, t_2]. \end{aligned} \quad (4)$$

Здесь $f_1(t, x, z, u_1)(f_2(t, x, y, u_2))$ – заданная n -мерная вектор-функция, непрерывная в $D_1 \times R^n \times R^r (D_2 \times R^n \times R^q)$ вместе с частными производными по $z(y)$, $G(z)$ – заданная в R^n непрерывно-

дифференцируемая вектор-функция, управляющие функции $u_i(t, x), i = 1, 2$, удовлетворяют ограничениям

$$\begin{aligned} u_1(t, x) &\in U_1 \subset R^r, \quad t \in D_1, \\ u_2(t, x) &\in U_2 \subset R^q, \quad t \in D_2. \end{aligned} \quad (5)$$

Пару $(u_1(t, x), u_2(t, x))$ с вышеприведенными свойствами назовем допустимым управлением в рассматриваемой задаче (1)–(4).

Будем предполагать, что при заданном допустимом управлении $(u_1(t, x), u_2(t, x))$ краевые задачи (1), (2) и (3), (4) имеют единственные абсолютно непрерывные решения (см., напр.: [8–10]) $z(t, x)$ и $y(t, x)$ соответственно.

Рассмотрим задачу нахождения минимального значения терминального типа функционала

$$J(u_1, u_2) = \varphi_1(z(t_1, x_1)) + \varphi_2(y(t_2, x_1)), \quad (6)$$

определенного на абсолютно непрерывных решениях краевых задач (1), (2) и (3), (4), порожденных всевозможными допустимыми управлениями. Здесь $\varphi_1(x)$ и $\varphi_2(x)$ – заданные скалярные функции, имеющие непрерывные производные.

Допустимый процесс $(u_1(t, x), u_2(t, x), z(t, x), y(t, x))$, доставляющий минимальное значение функционалу (6), назовем оптимальным процессом, а соответствующее допустимое управление $(u_1(t, x), u_2(t, x))$ – оптимальным управлением.

Цель работы – вывод необходимого условия оптимальности типа принципа максимума Л.С. Понтрягина (см., напр.: [7–9]).

2. Построение формулы приращения функционала для двух допустимых управлений

Пусть $(u_1(t, x), u_2(t, x), z(t, x), y(t, x))$ и $(\bar{u}_1(t, x) = u_1(t, x) + \Delta u_1(t, x), \bar{u}_2(t, x) = u_2(t, x) + \Delta u_2(t, x), \bar{z}(t, x) = z(t, x) + \Delta z(t, x), \bar{y}(t, x) = y(t, x) + \Delta y(t, x))$ – два допустимых процесса.

Из этих обозначений ясно, что $\Delta z(t, x)$ и $\Delta y(t, x)$ будут решениями следующих краевых задач:

$$\Delta z_{tx} = f_1(t, x, \bar{z}, \bar{u}_1) - f_1(t, x, z, u_1), \quad (t, x) \in D_1, \quad (7)$$

$$\begin{aligned} \Delta z(t_0, x) &= 0, \quad x \in X, \\ \Delta z(t, x_0) &= 0, \quad t \in T_1, \end{aligned} \quad (8)$$

$$\Delta y_{tx} = f_2(t, x, \bar{y}, \bar{u}_2) - f_2(t, x, y, u_2), \quad (t, x) \in D_2, \quad (9)$$

$$\begin{aligned} \Delta y(t_1, x) &= G(\bar{z}(t_1, x)) - G(z(t_1, x)), \quad x \in X, \\ \Delta y(t, x_0) &= 0, \quad t \in T_2. \end{aligned} \quad (10)$$

Введем обозначения

$$\begin{aligned} H_1(t, x, z, u_1, \psi_1) &= \psi_1' f_1(t, x, z, u_1), \\ H_2(t, x, y, u_2, \psi_2) &= \psi_2' f_2(t, x, y, u_2), \\ \Delta_{\bar{u}_1} f_1[t, x] &\equiv f_1(t, x, z(t, x), \bar{u}_1(t, x)) - f_1(t, x, z(t, x), u_1(t, x)), \\ \Delta_{\bar{u}_2} f_2[t, x] &\equiv f_2(t, x, y(t, x), \bar{u}_2(t, x)) - f_2(t, x, y(t, x), u_2(t, x)), \\ \Delta_{\bar{u}_1(t, x)} H_1[t, x, \psi_1] &\equiv H_1(t, x, z(t, x), \bar{u}_1(t, x), \psi_1(t, x)) - H_1(t, x, z(t, x), u_1(t, x), \psi_1(t, x)), \\ \frac{\partial \Delta_{\bar{u}_1(t, x)} H_1[t, x, \psi_1]}{\partial z} &\equiv \frac{\partial H_1(t, x, z(t, x), \bar{u}_1(t, x), \psi_1(t, x))}{\partial z} - \frac{\partial H_1(t, x, z(t, x), u_1(t, x), \psi_1(t, x))}{\partial z}, \\ \frac{\partial H_1[t, x, \psi_1]}{\partial z} &\equiv \frac{\partial H_1(t, x, z(t, x), u_1(t, x), \psi_1(t, x))}{\partial z}, \end{aligned}$$

$$\frac{\partial H_2[t, x, \psi_2]}{\partial y} \equiv \frac{\partial H_2(t, x, y(t, x), u_2(t, x), \psi_2(t, x))}{\partial y},$$

$$M\left(\frac{\partial \psi_2}{\partial x}, z(t_1, x)\right) \equiv \frac{\partial \psi_2'(t_1, x)}{\partial x} G(z(t_1, x)),$$

$$N(\psi_2, z(t_1, x_1)) \equiv \psi_2'(t_1, x_1) G(z(t_1, x_1)).$$

Здесь для транспонирования используется знак ' (штрих), а $\psi_i(t, x), i=1,2$, – пока неизвестные n -мерные вектор-функции.

Учитывая тождества (7)–(10) и введенные обозначения, а также считая вектор-функции $\psi_i(t, x), i=1,2$, достаточно гладкими (например, абсолютно непрерывными) и применяя формулу Тейлора, приращение функционала качества (6) представляется в виде:

$$\begin{aligned} \Delta S(u_1, u_2) = S(\bar{u}_1, \bar{u}_2) - S(u_1, u_2) = & \frac{\partial \varphi_1'(z(t_1, x_1))}{\partial z} \Delta z(t_1, x_1) + o_1(\|\Delta z(t_1, x_1)\|) - \frac{\partial N'(\psi_2, z(t_1, x_1))}{\partial z} \Delta z(t_1, x_1) + \\ & - o_2(\|\Delta z(t_1, x_1)\|) + \int_{x_0}^{x_1} \frac{\partial M'(\frac{\partial \psi_2}{\partial x}, z(t_1, x))}{\partial z} \Delta z(t_1, x) dx + \int_{x_0}^{x_1} o_3(\|\Delta z(t_1, x)\|) dx + \\ & + \frac{\partial \varphi_2'(y(t_2, x_1))}{\partial y} \Delta y(t_2, x_1) + o_4(\|\Delta y(t_2, x_1)\|) + \int_{t_0}^{t_1} \int_{x_0}^{x_1} \frac{\partial^2 \psi_1'(t, x)}{\partial t \partial x} \Delta z(t, x) dx dt - \\ & - \int_{x_0}^{x_1} \frac{\partial \psi_1'(t_1, x)}{\partial x} \Delta z(t_1, x) dx - \int_{t_0}^{t_1} \frac{\partial \psi_1'(t, x_1)}{\partial t} \Delta z(t, x_1) dx + \psi_1'(t_1, x_1) \Delta z(t_1, x_1) + \\ & + \psi_2'(t_2, x_1) \Delta y(t_2, x_1) + \int_{t_1}^{t_2} \int_{x_0}^{x_1} \frac{\partial^2 \psi_2'(t, x)}{\partial t \partial x} \Delta y(t, x) dx dt + \\ & + \int_{x_0}^{x_1} \frac{\partial \psi_2'(t_1, x)}{\partial x} \Delta y(t_1, x) dx - \int_{t_1}^{t_2} \frac{\partial \psi_2'(t, x_1)}{\partial t} \Delta y(t, x_1) dx - \int_{x_0}^{x_1} \frac{\partial \psi_2'(t_2, x)}{\partial x} \Delta y(t_2, x) dx - \\ & - \int_{t_0}^{t_1} \int_{x_0}^{x_1} \Delta_{\bar{u}_1(t, x)} H_1[t, x, \psi_1] dx dt - \int_{t_0}^{t_1} \int_{x_0}^{x_1} \frac{\partial H_1[t, x, \psi_1]}{\partial z} \Delta z(t, x) dx dt - \int_{t_0}^{t_1} \int_{x_0}^{x_1} \frac{\partial \Delta_{\bar{u}_1(t, x)} H_1[t, x, \psi_1]}{\partial z} \Delta z(t, x) dx dt - \\ & - \int_{t_0}^{t_1} \int_{x_0}^{x_1} o_5(\|\Delta z(t, x)\|) dx dt - \int_{t_1}^{t_2} \int_{x_0}^{x_1} \Delta_{\bar{u}_2(t, x)} H_2[t, x, \psi_2] dx dt - \int_{t_1}^{t_2} \int_{x_0}^{x_1} \frac{\partial H_2[t, x, \psi_2]}{\partial y} \Delta y(t, x) dx dt - \\ & - \int_{t_1}^{t_2} \int_{x_0}^{x_1} \frac{\partial \Delta_{\bar{u}_2(t, x)} H_2[t, x, \psi_2]}{\partial z} \Delta y(t, x) dx dt - \int_{t_1}^{t_2} \int_{x_0}^{x_1} o_6(\|\Delta y(t, x)\|) dx dt. \end{aligned} \quad (11)$$

Здесь $\|\alpha\|$ – норма вектора $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)'$, определяемая формулой $\|\alpha\| = \sum_{i=1}^n |\alpha_i|$, а $o(\alpha)$ – величина

более высокого порядка малости, чем α , т.е. $\frac{o(\alpha)}{\alpha} \rightarrow 0$ при $\alpha \rightarrow 0$.

Пусть вектор-функции $\psi_i(t, x), i=1,2$, являются решениями соответственно следующих краевых задач:

$$\frac{\partial^2 \psi_1(t, x)}{\partial t \partial x} = \frac{\partial H_1[t, x]}{\partial z}, \quad (12)$$

$$\frac{\partial \psi_1(t, x_1)}{\partial t} = 0, \quad \frac{\partial \psi_1(t_1, x)}{\partial x} = \frac{\partial M(\frac{\partial \psi_2}{\partial x}, z(t_1, x))}{\partial z}, \quad (13)$$

$$\begin{aligned}\psi_1(t_1, x_1) &= -\frac{\partial \varphi(z(t_1, x_1))}{\partial z} + \frac{\partial N(\psi_2, z(t_1, x_1))}{\partial z}, \\ \frac{\partial^2 \psi_2(t, x)}{\partial t \partial x} &= \frac{\partial H_2[t, x, \psi_2]}{\partial y},\end{aligned}\quad (14)$$

$$\begin{aligned}\frac{\partial \psi_2(t_2, x)}{\partial x} &= 0, \quad \frac{\partial \psi_2(t, x_1)}{\partial t} = 0, \\ \frac{\partial \psi_2'(t_2, x_1)}{\partial x} &= -\frac{\partial \varphi_2(y(t_2, x_1))}{\partial y}.\end{aligned}\quad (15)$$

Тогда формула приращения (11) примет вид:

$$\begin{aligned}\Delta S(u_1, u_2) &= -\int_{t_0}^{t_1} \int_{x_0}^{x_1} \Delta_{\bar{u}_1(t, x)} H_1[t, x, \psi_1] dx dt - \int_{t_0}^{t_1} \int_{x_0}^{x_1} \frac{\partial \Delta_{\bar{u}_1(t, x)} H_1[t, x, \psi_1]}{\partial z} \Delta z(t, x) dx dt - \\ &- \int_{t_1}^{t_2} \int_{x_0}^{x_1} \Delta_{\bar{u}_2(t, x)} H_2[t, x, \psi_2] dx dt - \int_{t_1}^{t_2} \int_{x_0}^{x_1} \frac{\partial \Delta_{\bar{u}_2(t, x)} H_2[t, x, \psi_2]}{\partial z} \Delta y(t, x) dx dt + o_1(\|\Delta z(t_1, x_1)\|) - o_2(\|\Delta z(t_1, x_1)\|) + \\ &+ \int_{x_0}^{x_1} o_3(\|\Delta z(t_1, x)\|) dx + o_4(\|\Delta y(t_2, x_1)\|) - \int_{t_0}^{t_1} \int_{x_0}^{x_1} o_5(\|\Delta z(t, x)\|) dx dt - \int_{t_1}^{t_2} \int_{x_0}^{x_1} o_6(\|\Delta y(t, x)\|) dx dt.\end{aligned}\quad (16)$$

Для дальнейших изложений нам понадобятся оценки норм приращений состояний.

Из оценок, установленных и приведенных в работах [11, 12] и др., следует, что

$$\|\Delta z(t, x)\| \leq L_1 \int_{t_0}^{t_1} \int_{x_0}^{x_1} \|\Delta_{\bar{u}_1(\tau, s)} f_1[\tau, s]\| ds d\tau, \quad (17)$$

$$\|\Delta y(t, x)\| \leq L_2 \left[\int_{t_1}^t \int_{x_0}^x \|\Delta_{\bar{u}_2(\tau, s)} f_2[\tau, s]\| ds d\tau + \|\Delta z(t_1, x)\| \right], \quad (18)$$

где $L_i = \text{const} > 0$, $i = 1, 2$, – некоторые постоянные.

Учитывая неравенство (17), из неравенства (18) получим, что

$$\|\Delta y(t, x)\| \leq L_3 \left[\int_{t_1}^t \int_{x_0}^x \|\Delta_{\bar{u}_2(\tau, s)} f_2[\tau, s]\| ds d\tau + \int_{t_0}^{t_1} \int_{x_0}^{x_1} \|\Delta_{\bar{u}_1(\tau, s)} f_1[\tau, s]\| ds d\tau \right] \quad (19)$$

($L_3 = \text{const} > 0$ – некоторая постоянная).

3. Необходимое условие оптимальности

Данные оценки позволяют получить необходимое условие оптимальности в форме аналога принципа максимума Л.С. Понтрягина.

Имеет место

Теорема. Для оптимальности допустимого управления $(u_1(t, x), u_2(t, x))$ необходимо, чтобы соотношения

$$\max_{v_1 \in U_1} H_1(\theta, \xi, z(\theta, \xi), v_1, \psi_1(\theta, \xi)) = H_1(\theta, \xi, z(\theta, \xi), u_1(\theta, \xi), \psi_1(\theta, \xi)), \quad (20)$$

$$\max_{v_2 \in U_2} H_2(\theta, \xi, y(\theta, \xi), v_2, \psi_2(\theta, \xi)) = H_2(\theta, \xi, y(\theta, \xi), u_2(\theta, \xi), \psi_2(\theta, \xi)) \quad (21)$$

выполнялись для всех $(\theta, \xi) \in [t_0, t_1] \times [x_0, x_1]$ и $(\theta, \xi) \in [t_1, t_2] \times [x_0, x_1]$ соответственно.

Перейдем к доказательству теоремы.

Пусть $(\theta, \xi) \in [t_0, t_1] \times [x_0, x_1]$ – произвольная точка Лебега (см., напр.: [10]) оптимального управления $u_1(t, x)$, $v_1 \in U_1$ – произвольный вектор, $\varepsilon > 0$ – произвольное достаточно малое число такое, что $\theta + \varepsilon < t_1$, $\xi + \varepsilon < x_1$, а $u_2(t, x) = 0$, $(t, x) \in D_2$.

Тогда специальное приращение допустимой управляющей функции $u_1(t, x)$ можно определить по формуле

$$\Delta u_1(t, x; \varepsilon) = \begin{cases} v_1 - u_1(t, x), & (t, x) \in [\theta, \theta + \varepsilon) \times [\xi, \xi + \varepsilon), \\ 0, & (t, x) \in D_1 \setminus [\theta, \theta + \varepsilon) \times [\xi, \xi + \varepsilon). \end{cases} \quad (22)$$

Здесь $\setminus$ означает разность множеств D_1 и $[\theta, \theta + \varepsilon) \times [\xi, \xi + \varepsilon)$.

Через $(\Delta z(t, x; \varepsilon), \Delta y(t, x; \varepsilon))$ обозначим специальное приращение состояния $(z(t, x), y(t, x))$, отвечающее специальному приращению (22) управляющей функции $u_1(t, x)$.

Из оценок (17) и (19) следует, что

$$\|\Delta z(t, x; \varepsilon)\| \leq L_4 \varepsilon^2, (t, x) \in D_1,$$

$$\|\Delta y(t, x; \varepsilon)\| \leq L_5 \varepsilon^2, (t, x) \in D_2,$$

где $L_{i+3} = \text{const} > 0, i = 1, 2$, – некоторые постоянные.

Поэтому из формулы приращения (16) следует, что вдоль оптимального управления $(u_1(t, x), u_2(t, x))$

$$S(u_1(t, x) + \Delta u_1(t, x; \varepsilon), u_2(t, x)) - S(u_1(t, x), u_2(t, x)) = -\varepsilon^2 \Delta_{v_1} H_1[\theta, \xi, \psi_1] + o(\varepsilon^2) \geq 0.$$

Отсюда получаем неравенство

$$\Delta_{v_1} H_1[\theta, \xi, \psi_1] \leq 0.$$

Этим условие максимума (20) доказано.

Для доказательства соотношения (21) предположим, что $(\theta, \xi) \in [t_1, t_2) \times [x_0, x_1)$ – произвольная точка Лебега управления $u_2(t, x)$, $v_2 \in U_2$ – произвольный вектор, $\mu > 0$ – произвольное малое число такое, что $\theta + \mu < t_2, \xi + \varepsilon < x_1$, а $u_1(t, x) = 0, (t, x) \in D_1$.

Тогда специальное приращение допустимой управляющей функции $u_2(t, x)$ можно определить по формуле

$$\Delta u_2(t, x; \mu) = \begin{cases} v_2 - u_2(t, x), & (t, x) \in [\theta, \theta + \mu) \times [\xi, \xi + \mu), \\ 0, & (t, x) \in D_2 \setminus [\theta, \theta + \mu) \times [\xi, \xi + \mu). \end{cases} \quad (23)$$

Учитывая формулу (23), из оценки (19) получаем, что

$$\|\Delta y(t, x; \mu)\| \leq L_6 \mu^2, (t, x) \in D_2. \quad (24)$$

Учитывая оценку (24) и формулу (23), из формулы приращения (16) получаем, что вдоль оптимального управления $(u_1(t, x), u_2(t, x))$

$$\begin{aligned} & S(u_1(t, x), u_2(t, x) + \Delta u_2(t, x; \mu)) - S(u_1(t, x), u_2(t, x)) = \\ & = -\mu^2 [H_2(\theta, \xi, y(\theta, \xi), v_2, \psi_2(\theta, \xi)) - H_2(\theta, \xi, y(\theta, \xi), u_2(\theta, \xi), \psi_2(\theta, \xi))] + o(\mu^2) \geq 0. \end{aligned}$$

Из последнего разложения следует, что

$$\Delta_{v_2} H_2[\theta, \xi, \psi_2] \leq 0.$$

Этим соотношение (21) и, следовательно, теорема доказаны.

Замечание. Из доказанной теоремы можно получить аналог дифференциального принципа максимума.

Пусть функции $f_1(t, x, z, u_1), f_2(t, x, y, u_2)$ непрерывно-дифференцируемы также по u_1 и u_2 соответственно, а множества $U_i, i = 1, 2$, выпуклы.

При сделанных предположениях из аналога принципа максимума Понтрягина получается

Следствие (дифференциальный принцип максимума). Вдоль оптимального управления $(u_1(t, x), u_2(t, x))$ выполняются условия

$$\frac{\partial H_1'(\theta, \xi, z(\theta, \xi), u_1(\theta, \xi), \psi_1(\theta, \xi))}{\partial u_1} u_1(\theta, \xi) = \max_{v_1 \in U_1} \frac{\partial H_1'(\theta, \xi, z(\theta, \xi), v_1, \psi_1(\theta, \xi))}{\partial u_1} v_1,$$

$$(\theta, \xi) \in [t_0, t_1] \times [x_0, x_1],$$

$$\frac{\partial H_2'(\theta, \xi, y(\theta, \xi), u_2(\theta, \xi), \psi_2(\theta, \xi))}{\partial u_2} u_2(\theta, \xi) = \max_{v_2 \in U_2} \frac{\partial H_2'(\theta, \xi, y(\theta, \xi), v_2, \psi_2(\theta, \xi))}{\partial u_2} v_2,$$

$$(\theta, \xi) \in [t_1, t_2] \times [x_0, x_1].$$

Доказательство следствия проводится по схеме, аналогичной схеме из работ [9, 11].

Заключение

В работе рассматривается задача оптимального управления с переменной структурой с распределенными параметрами, описываемая в двух областях различными краевыми условиями Гурса для двух гиперболических уравнений второго порядка (задача оптимального управления системами Гурса–Дарбу). Введена сопряженная система в классической форме, и построена формула приращения первого порядка функционала качества, носящая конструктивный характер.

Применяя двумерный аналог игольчатой вариации Макшейна, доказано необходимое условие оптимальности в форме аналога условия максимума Л.С. Понтрягина.

Из доказанного условия оптимальности получен аналог дифференциального условия максимума.

Список источников

1. Величенко В.В. Оптимальное управление составными системами // Доклады АН СССР. 1967. Т. 176, № 4. С. 754–756.
2. 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.
3. Розова В.Н. Оптимальное управление ступенчатыми системами с неинтегральным функционалом // Вестник РУДН. Сер. Прикладная и компьютерная математика. 2002. № 1 (1). С. 131–136.
4. Никольский М.С. Об одной вариационной задаче с переменной структурой // Вестник Московского университета. Сер. Вычислительная математика и кибернетика. 1987. № 2. С. 36–41.
5. Мансимов К.Б., Алибберов А.А. Необходимые условия оптимальности второго порядка в одной задаче управления с переменной структурой // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2018. № 43. С. 4–15.
6. Мансимов К.Б., Сулейманова Ш.Ш. К оптимальности особых в классическом смысле управлений в одной задаче оптимального управления системами с переменной структурой // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2018. № 44. С. 10–24. doi: 10.17223/19988605/44/2
7. Понтрягин Л.С., Болтянский В.Г., Гамкрелидзе Р.В., Мищенко Е.Ф. Математическая теория оптимальных процессов. М. : Наука, 1984. 384 с.
8. Габасов Р., Кириллова Ф.М., Альсевич В.В., Калинин А.И. и др. Методы оптимизации. Минск : Четыре четверти, 2011. 472 с.
9. Габасов Р., Кириллова Ф.М. Принцип максимума в теории оптимального управления. М. : URSS, 2013. 272 с.
10. Плотников В.И., Сумин В.И. Методы оптимального управления системами математической физики. Горький : Изд-во Горьковского гос. ун-та, 1986. 87 с.
11. Мансимов К.Б. Марданов М.Дж. Качественная теория оптимального управления системами Гурса–Дарбу. Баку : ЭЛМ, 2010. 360 с.
12. Васильев О.В., Срочко В.А., Терлецкий В.А. Методы оптимизации и их приложения. Новосибирск : Наука, Сиб. отд-ние, 1990. 151 с.

References

1. Velichenko, V.V. (1967) Optimal'noe upravlenie sostavnymi sistemami [Optimal control of composite systems]. DAN. USSR. 176(4), pp. 754–756.
2. Ismaylov, 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.
3. 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.

4. Nikolsky, M.S. (1987) Ob odnoy variatsionnoy zadache s peremennoy strukturoy [On a variational problem with a variable structure]. *Vestnik Moskovskogo universiteta. Ser. Vychislitel'naya matematika i kibernetika*. 2. pp. 36–41.
5. Mansimov, K.B. & Alikberov, A.A. (2018) The second order necessary optimality conditions in a control problem with variable structure. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 43. pp. 4–15. DOI: 10.17223/19988605/43/1
6. Mansimov, K.B. & Suleymanova, Sh.Sh. (2018) On the optimality of controls that are special in the classical sense in one problem of optimal control of systems with a variable structure. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 44. pp. 10–24. DOI: 10.17223/19988605/44/2
7. Pontryagin, L.S., Boltyansky, V.G., Gamkrelidze, R.V. & Mishchenko, E.F. (1984) *Matematicheskaya teoriya optimal'nykh protsessov* [Mathematical Theory of Optimal Processes]. Moscow: Nauka.
8. Gabasov, R., Kirillova, F.M., Alseich, V.V. & Kalinin, A.I. (2011) *Metody optimizatsii* [Optimization Methods]. Minsk: Chetyre chetverti.
9. Gabasov, R. & Kirillova, F.M. (2013) *Printsip maksimuma v teorii optimal'nogo upravleniya* [The maximum principle in the theory of optimal control]. Moscow: URSS.
10. Plotnikov, V.I. & Sumin, V.I. (1986) *Metody optimal'nogo upravleniya sistemami matematicheskoy fiziki* [Methods for optimal control of systems of mathematical physics]. Gorky: Gorky State University.
11. Mansimov, K.B. & Mardanov, M. Dzh. (2010) *Kachestvennaya teoriya optimal'nogo upravleniya sistemami Gursa–Darbu* [Qualitative Theory of Optimal Control of Goursat–Darboux Systems]. Baku: ELM.
12. Vasiliev, O.V., Srochko, V.A. & Terletsky, V.A. (1990) *Metody optimizatsii i ikh prilozheniya* [Optimization Methods and their Applications]. Novosibirsk: Nauka.

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

Сулейманова Шабнам Шакир кызы – диссертант лаборатории «Методы управления в сложных динамических системах» Института систем управления Министерства науки и образования Азербайджана (Баку, Азербайджан). E-mail: kmansimov@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, Head of the Laboratory “Control Methods in Complex Dynamic Systems” of the Institute of Control Systems of the Ministry of Science and Education of Azerbaijan, Baku, Azerbaijan; Head of the Department of Mathematical Cybernetics, Baku State University, Baku, Azerbaijan). E-mail: kamilbmansimov@gmail.com

Suleymanova Shabnam Sh. (Dissertation Candidate at the Laboratory “Control methods in complex dynamic systems” of the Institute of Control Systems of the Ministry of Science and Education of Azerbaijan, Baku, Azerbaijan). E-mail: kmansimov@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 03.06.2024

Received 28.12.2023; accepted for publication 03.06.2024

Original article

UDK 517.9

doi: 10.17223/19988605/67/2

Technology for improving robust quality of control for one-dimensional linear discrete control systems with structural-parametric uncertainty

Andrej N. Parshukov

Industrial University of Tyumen, Tyumen, Russian Federation, anparshukov@mail.ru

Abstract. The article considers one-dimensional linear automatic control systems in discrete time. It is assumed that the operators of the control object contain parametric and/or structural uncertainty. The purpose of the study is to develop a method for the synthesis of modal controllers that provide the maximum robust quality of control of a closed system. The main result is presented in the form of an algorithm for improving the robust quality of control. The efficiency of the algorithm is illustrated by an example.

Keywords: structural perturbations; modal controller; robust control quality.

For citation: Parshukov, A.N. (2024) Technology for improving robust quality of control for one-dimensional linear discrete control systems with structural-parametric uncertainty. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 67. pp. 12–21. doi: 10.17223/19988605/67/2

Научная статья

doi: 10.17223/19988605/67/2

Технология повышения свойств робастного качества управления для одномерных линейных дискретных систем управления со структурно-параметрической неопределенностью

Андрей Николаевич Паршуков

Тюменский индустриальный университет, Тюмень, Россия, anparshukov@mail.ru

Аннотация. Рассмотрены одномерные линейные системы автоматического управления в дискретном времени. Предполагается, что операторы объекта управления содержат параметрическую и / или структурную неопределенность. Цель исследования состоит в разработке метода синтеза модальных регуляторов, обеспечивающих максимальное робастное качество управления замкнутой системой. Основной результат оформлен в виде алгоритма повышения робастного качества управления. Эффективность алгоритма проиллюстрирована примером.

Ключевые слова: структурные возмущения; модальный регулятор; робастное качество управления.

Для цитирования: Паршуков А.Н. Технология повышения свойств робастного качества управления для одномерных линейных дискретных систем управления со структурно-параметрической неопределенностью // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 67. С. 12–21. doi: 10.17223/19988605/67/2

Introduction

Currently, in industry, technological processes with different tempo dynamics are increasingly considered as control objects [1, 2]. A feature of these processes is that their mathematical models are restored with

an accuracy of belonging to some families of linear difference equations with different orders (structural-parametric uncertainty). When synthesizing a controller, structural perturbations, as a rule, are not taken into account [3]. Since the properties of stability and quality of control of the system are determined by the location of the zeros of the characteristic polynomial, the question arises: for which operators of structural perturbations in the control object, the closed system will still retain the properties of stability (robust stability) and quality of control (robust quality of control)?

In the modal control scheme [4. P. 8–21] the quality of control is given as an area S on the complex plane; the area S determines the desired location of the zeros of the characteristic polynomial. Consequently, the issues of studying (testing) robust stability and robust quality of control can be considered from a unified standpoint: is it required to check whether the zeros of a given family of polynomials belong to the area S ?

The problem of studying robust stability and robust quality of control is widely presented in the literature. There are three main directions in which this problem is solved: 1) the principle of zero exclusion [5–8]; 2) H^∞ theory [9, 10]; 3) LMI method [11–13]. At the same time, the main attention is paid to the study of robust stability, and the methods for studying the robust quality of control have not yet been sufficiently developed.

In this paper, we develop the results obtained by us in [14]. The article [14] developed a method for checking the robust quality of control. In the same place, a function was introduced that allows one to compare the robust properties of various dynamic control systems. In this paper, based on the introduced function (measure of robust properties), an algorithm for variation of the modal controller coefficients is developed that consistently increases the property of the robust quality of control.

Further, the following notation is adopted: $\doteq$ is equal by definition; R^n is the n -dimensional space of real numbers; C^n is the n -dimensional space of complex numbers; C^1 is the complex plane; T is the transposition operation; E is the identity matrix of the corresponding dimension; $\mathbf{0}$ is a column vector consisting entirely of zeros; j is the imaginary unit; s is a variable (generally complex); s^* is the number complex conjugate of s ; t is discrete time ($t = 0, \pm 1, \pm 2, \dots$); z is advance operator by one cycle; S is the area on C^1 ; ∂S is the boundary of the area S ; $\text{int } S$ is the interior of the area S .

Let $f(\mathbf{x})$ be a function of the vector argument $\mathbf{x}$ defined on the domain $X \subset R^n$; denote

$$|f|_+ \doteq \max_{\mathbf{x} \in X} |f(\mathbf{x})|, \quad |f|_- \doteq \min_{\mathbf{x} \in X} |f(\mathbf{x})|.$$

Operator of the form

$$f(n, z) = \sum_{i=0}^n f_i \cdot z^i, \quad (1)$$

is called a polynomial operator of degree n . Replacing z with s in (1), where s is a variable, we obtain an algebraic polynomial

$$f(n, s) = \sum_{i=0}^n f_i \cdot s^i.$$

The set of zeros of the polynomial $f(n, s)$ is denoted by $\Lambda(f)$:

$$\Lambda(f) \doteq \{\lambda_i \in C^1 : f(n, \lambda_i) = 0, \quad i \in \overline{1, n}\}.$$

Family of operators of the form

$$f(n, F, z) = \left\{ f(n, \mathbf{f}, z) = \sum_{i=0}^n f_i \cdot z^i : \mathbf{f} \doteq (f_0, \dots, f_n) \in F \right\}, \quad (2)$$

$$F = \{\mathbf{f} \in R^{n+1} : f_i \in [f_i^0 - \Delta f_i; f_i^0 + \Delta f_i],$$

$$f_n^0 \neq 0, \quad \Delta f_i \geq 0, \quad i = \overline{0, n}\},$$

is called an *interval polynomial operator* of degree n .

Replacing z with s in (2), we obtain an *interval polynomial*:

$$f(n, F, s) = \left\{ f(n, \mathbf{f}, s) = \sum_{i=0}^n f_i \cdot s^i : \mathbf{f} \in F \right\}. \quad (3)$$

Interval polynomial (3) can be represented as

$$f(n, F, s) = f^0(n, \mathbf{f}^0, s) + f(n, \Delta F, s), \quad (4)$$

where

$$f^0(n, \mathbf{f}^0, s) = \sum_{i=0}^n f_i^0 \cdot s^i, \quad \mathbf{f}^0 \doteq (f_0^0, \dots, f_n^0),$$

ordinary polynomial, and

$$f(n, \Delta F, s) = \left\{ f(n, \delta \mathbf{f}, s) = \sum_{i=0}^n \delta f_i \cdot s^i : \delta \mathbf{f} \doteq (\delta f_0, \dots, \delta f_n) \in \Delta F \right\},$$

$$\Delta F = \left\{ \delta \mathbf{f} \in R^{n+1} : \delta f_i \in [-\Delta f_i, \Delta f_i], \quad i = \overline{0, n} \right\},$$

interval polynomial with symmetric coefficient uncertainty bounds.

Let $\mathbf{A}$ be a matrix whose coefficients are complex numbers, then denote by $\mathbf{A}_{\text{Re}}$ and $\mathbf{A}_{\text{Im}}$ the matrices whose coefficients are, respectively, the real and imaginary parts of the coefficients of the matrix $\mathbf{A}$; in this sense, we denote

$$\mathbf{A}_{\text{Re}} \doteq \text{Re}(\mathbf{A}), \quad \mathbf{A}_{\text{Im}} \doteq \text{Im}(\mathbf{A}).$$

Let's call the *generalized gradient* of the function $f(\mathbf{x})$ a vector that coincides with the ordinary gradient $\nabla_{\mathbf{x}} f(\mathbf{x})$ at those points where the ordinary gradient exists, and is redefined as a zero vector at those points where the ordinary gradient does not exist, that is:

$$\nabla_{\mathbf{x}}^g f(\mathbf{x}) \doteq \begin{cases} \nabla_{\mathbf{x}} f(\mathbf{x}), & \exists \nabla_{\mathbf{x}} f(\mathbf{x}), \\ \mathbf{0}, & \nexists \nabla_{\mathbf{x}} f(\mathbf{x}), \end{cases}$$

here $\nabla_{\mathbf{x}}^g$ is the generalized gradient.

1. Method of synthesis of a modal controller for a control object under conditions of structural-parametric disturbances

Let a linear one-dimensional discrete control object is given by a model of the form:

$$v(l, V, z) \cdot a(n, A, z) y(t) = w(h, W, z) \cdot b(m, B, z) u(t), \quad n > m, \quad l \geq h, \quad (5)$$

here u is the input (control) signal, y is the output (control) signal, a, b, v, w are interval operators of the form (4) such that

$$a_n^0 = 1, \quad \Delta a_n = 0, \quad v_0^0 = 1, \quad \Delta v_0 = 0, \quad w_0^0 = 1, \quad \Delta w_0 = 0.$$

Model

$$a^0(n, \mathbf{a}^0, z) y(t) = b^0(m, \mathbf{b}^0, z) u(t), \quad (6)$$

belonging to the family of models (5), we will call "nominal"; interval operators a, b are "basic dynamics"; operators v, w are "structural perturbations".

The quality of control will be assigned in the form of an area S , which determines the admissible location of zeros of the characteristic polynomial on C^1 . We assume that the area S satisfies the following conditions: it is located inside a circle of unit radius; simply connected; also holds $s^* \in S$ for any point $s \in S$. In addition, based on the meaning of the problem, we require that the operators of structural perturbations v, w satisfy the expressions:

$$\Lambda(v) \subset \text{int } S, \quad \Lambda(w) \subset \text{int } S. \quad (7)$$

Note that the fulfillment of expressions (7) is easy to check using the methods described in [8. P. 7–10].

Following the method of modal control [14. P. 6–7], the controller for the nominal model (6) is sought in the form of a difference equation of the $n - 1$ degree:

$$\beta(n-1, z) u(t) = \alpha(n-1, z) y(t) + \chi(n-1, z) g(t), \quad \beta_{n-1} = 1, \quad (8)$$

here g is a given program signal. In the modal control scheme, the coefficients of the operators β and α of the controller (8) are calculated from the condition of conversion to the identity of the equation

$$a^{st}(2n-1, s) = a^0(n, \mathbf{a}^0, s) \beta(n-1, s) - b^0(m, \mathbf{b}^0, s) \alpha(n-1, s),$$

$$a_{2n-1}^{st} = 1, \quad (9)$$

where a^{st} is the given characteristic polynomial of the standard control system (hereinafter referred to as the "standard"); the choice of standard is limited by the condition

$$\Lambda(a^{st}) \subset \text{int } S. \quad (10)$$

It is obvious that the choice of the polynomial χ does not affect the properties of robust stability and the robust quality of control, so the issue of calculating the polynomial χ in the modal control method is not considered. In this sense, the vector of coefficients of the polynomials β and α is called the vector of coefficients of the controller and is denoted

$$\mathbf{r} \doteq \text{col}(\mathbf{r}_1, \mathbf{r}_2),$$

$$\mathbf{r}_1 \doteq \text{col}(\beta_0, \beta_1, \dots, \beta_{n-2}), \quad \mathbf{r}_2 \doteq \text{col}(\alpha_0, \alpha_1, \dots, \alpha_{n-1});$$

After closing the original object (5) by the controller synthesized according to the scheme (8)–(10), it is easy to obtain the following expression for the characteristic polynomial of the closed system

$$a^c(2n+l-1, A, B, V, W, s) = \{a^c(2n+l-1, \mathbf{a}, \mathbf{b}, \mathbf{v}, \mathbf{w}, s) : \\ \forall \mathbf{a} \in A, \mathbf{b} \in B, \mathbf{v} \in V, \mathbf{w} \in W\}, \quad (11)$$

$$a^c(2n+l-1, \mathbf{a}, \mathbf{b}, \mathbf{v}, \mathbf{w}, s) \doteq v(l, \mathbf{v}, s) \cdot a(n, \mathbf{a}, s) \cdot \beta(n-1, s) - \\ - w(h, \mathbf{w}, s) \cdot b(m, \mathbf{b}, s) \cdot \alpha(n-1, s).$$

A closed system with a characteristic polynomial of the form (11) has a robust control quality if the set

$$\Lambda(a^c) = \{\lambda_i : \exists \mathbf{a} \in A, \exists \mathbf{b} \in B, \exists \mathbf{v} \in V,$$

$$\exists \mathbf{w} \in W, \quad a^c(2n+l-1, \mathbf{a}, \mathbf{b}, \mathbf{v}, \mathbf{w}, \lambda_i) = 0, \quad i = \overline{1, 2n+l-1}\}$$

lies inside the area S , i.e.

$$\Lambda(a^c) \subset \text{int } S. \quad (12)$$

In the presence of structural disturbances in the control object, it is impossible to guarantee in advance that the modal controller calculated by formulas (8)–(10) will ensure the fulfillment of condition (12). Thus, the problem of synthesizing a modal controller in the presence of structural disturbances in the control object consists of the following steps:

- 1) synthesis of a modal controller for a nominal object (6) (according to formulas (8)–(10);
- 2) subsequent verification of the fulfillment of condition (12) for a given family of characteristic polynomials (11) (the **problem of testing the robust quality of control**).

Theorem 1 [14. P. 8]. *Let the family of polynomials (11) satisfy conditions (7), (10) and*

$$\rho(s) \doteq \rho_1(s) - \rho_2(s) > 0, \quad \forall s \in \partial S, \quad (13)$$

where

$$\rho_1(s) \doteq \left| v^0(l, \mathbf{v}^0, s) + v(l, \delta \mathbf{v}, s) \right|_- \cdot \left| a^{st}(2n-1, s) + \right. \\ \left. + a(n-1, \delta \mathbf{a}, s) \cdot \beta(n-1, s) - b(m, \delta \mathbf{b}, s) \cdot \alpha(n-1, s) \right|_-, \\ \rho_2(s) \doteq \left| v^0(l, \mathbf{v}^0, s) + v(l, \delta \mathbf{v}, s) - w^0(h, \mathbf{w}^0, s) - w(h, \delta \mathbf{w}, s) \right|_+ \cdot \\ \cdot \left| b^0(m, \mathbf{b}^0, s) + b(m, \delta \mathbf{b}, s) \right|_+ \cdot \left| \alpha(n-1, s) \right|. \quad (14)$$

Then the family of polynomials (11) satisfies (12).

An algorithm for checking the robust quality of control based on Theorem 1 is described in [14. P. 9]. The function ρ , defined by formulas (13)–(14), makes it possible to compare various closed systems with each other, thus, the function ρ is a measure of robust properties. The introduction of the measure ρ makes it possible to pose the problem of increasing the robust properties of a closed system.

2. Method for improving the properties of robust control quality

2.1. Statement of the problem of improving the properties of the robust quality of control

Let us assume that the boundary ∂S is replaced by a finite number of points [8. P. 10]:

$$\Omega = \{s_i \in \partial S, \quad i \in \overline{1, N}\}, \quad (N < \infty),$$

here the points s_i included in the set Ω are chosen on the boundary ∂S based on the condition:

$$\max_{s \in \partial S} \Delta \arg(\omega - s) \Big|_{\omega=s_i}^{\omega=s_{i+1}} \leq \frac{1}{2n+l-1} \cdot \frac{\pi}{2} \quad \text{radian}.$$

Suppose that for a given control object (5)–(7) a controller (8) is synthesized that satisfies conditions (9)–(10). Note that in the synthesis scheme of the modal controller (5)–(10), freedom is allowed in choosing the standard a^{st} within the limits of condition (10). The new chosen standard can be considered as a result of the variation of the vector of controller coefficients (8):

$$\mathbf{r}(k+1) = \mathbf{r}(k) + \delta \mathbf{r}(k), \quad k = 0, 1, 2, \dots, K < \infty. \quad (15)$$

The present work is devoted to the search for the vector $\mathbf{r}^*$ providing a local maximum of the function ρ in the space of the vector of the controller coefficients. The local maximum is sought in the vicinity of the vector $\mathbf{r}(0)$.

Consider the procedure for searching for the vector $\mathbf{r}^*$ as an iterative process, let us denote the number of iteration as k (see expression (15)). Further, for convenience of presentation, we will indicate the iteration number as the first of the arguments of functions, vectors and matrices: for example, the vector $\mathbf{r}$, the polynomial $a^{\text{st}}(2n-1, s)$, the functions $\rho(s)$, ρ , and so on will be written, respectively, $\mathbf{r}(k)$, $a^{\text{st}}(k, 2n-1, s)$, $\rho(k, s)$, $\rho(k)$.

According to the known vector $\mathbf{r}(k)$, the polynomial $a^{\text{st}}(k, 2n-1, s)$ is determined by the formula

$$\begin{aligned} a^{\text{st}}(k, 2n-1, s) = & a^0(n, \mathbf{a}^0, s) \beta(k, n-1, s) - \\ & - b^0(m, \mathbf{b}^0, s) \alpha(k, n-1, s), \end{aligned} \quad (16)$$

the polynomials $\beta(k, n-1, s)$ and $\alpha(k, n-1, s)$ can be expressed in terms of the components of the vector $\mathbf{r}(k)$:

$$\beta(k, n-1, s) = s^{n-1} + \boldsymbol{\sigma}_1 \cdot \mathbf{r}_1(k), \quad \alpha(k, n-1, s) = \boldsymbol{\sigma}_2 \cdot \mathbf{r}_2(k),$$

here we use the notation

$$\boldsymbol{\sigma}_1 \doteq (1, s, \dots, s^{n-2}), \quad \boldsymbol{\sigma}_2 \doteq (1, s, \dots, s^{n-1}).$$

Suppose that at the current k iteration $\mathbf{r}(k)$, $a^{\text{st}}(k, 2n-1, s)$, $\rho(k)$ are defined, and for the polynomial $a^{\text{st}}(k, 2n-1, s)$ done

$$\Lambda(a^{\text{st}}(k)) \subset \text{int } S. \quad (17)$$

It should be noted that for $k = 0$ condition (17) is satisfied.

At the next $k+1$ iteration, we will look for a vector $\delta \mathbf{r}(k)$ such that:

1) after substituting the vector $\delta \mathbf{r}(k)$ into (15) for the polynomial $a^{\text{st}}(k+1, 2n-1, s)$ done

$$\Lambda(a^{\text{st}}(k+1)) \subset \text{int } S. \quad (18)$$

2) done

$$\rho(k+1) > \rho(k). \quad (19)$$

2.2. Algorithm for improving the property of robust control quality

Denote by Θ the set of points s at which the function $\rho(k, s)$ takes the minimum value, the total number of minimum points at the k iteration will be denoted by $N(k)$, meaning $N(k) \leq N$. The set Θ can be written:

$$\Theta = \left\{ s_i = \arg \min_{s \in \Omega} \rho(k, s), \quad i = \overline{1, N(k)} \right\}.$$

For each point $s_i \in \Theta$, we find the nominated generalized gradient of the function $\rho(k, s_i)$ with respect to $\mathbf{r}(k)$, we denote such a gradient by $\mathbf{p}(k, i)$:

$$\mathbf{p}(k, i) \doteq \nabla_{\mathbf{r}}^g \rho(k, s_i) / \left| \nabla_{\mathbf{r}}^g \rho(k, s_i) \right|. \quad (20)$$

It's not hard to get

$$\begin{aligned} \nabla_{\mathbf{r}}^g \rho(k, s) &= \nabla_{\mathbf{r}}^g \rho_1(k, s) - \nabla_{\mathbf{r}}^g \rho_2(k, s), \\ \nabla_{\mathbf{r}}^g \rho_1(k, s) &= \left| v^0(l, \mathbf{v}^0, s) + v(l, \delta \mathbf{v}, s) \right|_- \cdot \nabla_{\mathbf{r}}^g \eta_1(k, s), \\ \nabla_{\mathbf{r}}^g \rho_2(k, s) &= \left(\mathbf{0} \in R^{n-1}, \quad \nabla_{\mathbf{r}_2}^g \rho_2(k, s) \right), \\ \nabla_{\mathbf{r}_2}^g \rho_2(k, s) &= \left| v^0(l, \mathbf{v}^0, s) + v(l, \delta \mathbf{v}, s) - w^0(h, \mathbf{w}^0, s) - \right. \\ &\quad \left. - w(h, \delta \mathbf{w}, s) \right|_+ \cdot \left| b^0(m, \mathbf{b}^0, s) + b(m, \delta \mathbf{b}, s) \right|_+ \cdot \nabla_{\mathbf{r}_2}^g \eta_2(k, s), \end{aligned} \quad (21)$$

here the designations are introduced

$$\begin{aligned} \theta &\doteq a^{\text{st}}(k, 2n-1, s) + a(n-1, \delta \mathbf{a}, s) \cdot s^{n-1}, \\ \boldsymbol{\mu} &\doteq (a(n-1, \delta \mathbf{a}, s) \cdot \boldsymbol{\sigma}_1, \quad -b(m, \delta \mathbf{b}, s) \cdot \boldsymbol{\sigma}_2), \\ \eta_1(k, s) &\doteq |\theta + \boldsymbol{\mu} \cdot \mathbf{r}(k)|, \quad \eta_2(k, s) \doteq |\boldsymbol{\sigma}_2 \cdot \mathbf{r}_2(k)|. \end{aligned}$$

The gradients of the functions η_1 and η_2 are

$$\nabla_{\mathbf{r}} \eta_1(k, s) = \frac{(\theta_{\text{Re}} + \boldsymbol{\mu}_{\text{Re}} \cdot \mathbf{r}) \boldsymbol{\mu}_{\text{Re}} + (\theta_{\text{Im}} + \boldsymbol{\mu}_{\text{Im}} \cdot \mathbf{r}) \boldsymbol{\mu}_{\text{Im}}}{|\eta_1(k, s)|},$$

at those points Θ where $|\eta_1(k, s)| \neq 0$, and

$$\nabla_{\mathbf{r}_2} \eta_1(k, s) = \frac{\boldsymbol{\sigma}_{2, \text{Re}} \cdot \mathbf{r}_2 \cdot \boldsymbol{\sigma}_{2, \text{Re}} + \boldsymbol{\sigma}_{2, \text{Im}} \cdot \mathbf{r}_2 \cdot \boldsymbol{\sigma}_{2, \text{Im}}}{|\eta_2(k, s)|},$$

at those points Θ where $|\eta_2(k, s)| \neq 0$. Thus,

$$\begin{aligned} \nabla_{\mathbf{r}}^g \rho_1(k, s) &= \left| v^0(l, \mathbf{v}^0, s) + v(l, \delta \mathbf{v}, s) \right|_- \cdot \\ &\quad \cdot \begin{cases} \nabla_{\mathbf{r}}^g \eta_1(k, s), & |\eta_1(k, s)| \neq 0, \\ \mathbf{0} \in R^{2n-1}, & |\eta_1(k, s)| = 0, \end{cases} \\ \nabla_{\mathbf{r}_2}^g \rho_2(k, s) &= \left| v^0(l, \mathbf{v}^0, s) + v(l, \delta \mathbf{v}, s) - w^0(h, \mathbf{w}^0, s) - \right. \\ &\quad \left. - w(h, \delta \mathbf{w}, s) \right|_+ \cdot \left| b^0(m, \mathbf{b}^0, s) + b(m, \delta \mathbf{b}, s) \right|_+ \cdot \\ &\quad \cdot \begin{cases} \left(\mathbf{0} \in R^{n-1} \right), & |\eta_2(k, s)| \neq 0, \\ \nabla_{\mathbf{r}_2}^g \eta_2(k, s), & |\eta_2(k, s)| = 0. \end{cases} \end{aligned} \quad (22)$$

The normalized generalized gradient $\mathbf{p}(k, i)$ is calculated by formulas (20)–(23). It follows from these formulas that the vector $\mathbf{p}(k, i)$ is not defined only when

$$|\eta_1(k, s_i)| = 0, \quad \text{и} \quad |\eta_2(k, s_i)| = 0.$$

Denote by I the index set of points $s_i \in \Theta$ at which the vector $\mathbf{p}(k, i)$ exists:

$$I = \{i \in \overline{1, N(k)} : \exists \mathbf{p}(k, i)\}.$$

The set of vectors $\mathbf{p}(k, i)$ calculated for all points of the set Θ will be denoted by P :

$$P = \{\mathbf{p}(k, i) : i \in I\}.$$

Consider the case when the set I is empty

$$I = \emptyset. \quad (24)$$

This case means that at all points of the set Θ the gradient of the function $\rho(k, s_i)$ is either equal to zero or does not exist; therefore, at the k iteration, the local maximum of the function is reached and the problem of improving the robust control quality is solved.

Let us consider the case when the set I is not empty, then the set P is the direction of growth of the function $\rho(k)$. A common direction for all elements of the set P (of course, if it exists at all) can be found, for example, using the “linear classifier” algorithm [15. P. 113–115]. We will not give the entire algorithm of the linear classifier here, but we will introduce the concepts necessary for the functioning of the algorithm. The algorithm assumes that the set I is not empty; admissible error value ε_0 ($\varepsilon_0 > 0$).

If there is a common vector $\tau \in R^{2n-1}$ such that

$$\{\tau \cdot p(k, i) > \varepsilon_0, \quad \forall i \in I\}, \quad (25)$$

then such a vector can be found in no more than ε_0^{-2} iterations of the algorithm (Novikov's theorem [15. P. 28]). If for the specified number of iterations of the linear classification algorithm it is not possible to obtain a vector τ that satisfies (25), then, according to the Novikov theorem, such a vector does not exist. As applied to the problem we are considering, this means that the function $\rho(k)$ has reached a local maximum, and our algorithm must be stopped. Therefore, it is further assumed that the vector τ has been found.

The vector $\delta r(k)$ in expression (15) is sought in the form:

$$\delta r(k) = \gamma \cdot \tau, \quad \gamma > 0, \quad (26)$$

here τ is the general direction of increase of the function $\rho(k, s_i)$ at the points of the set Θ if the number of such points is more than one, and

$$\tau = p(k, i),$$

if the set Θ consists of one point s_i .

The coefficient of proportionality γ in (26) is chosen based on conditions (18)–(19). At the first stage, for each point $s_i \in \Omega$, one should determine the value of the parameter γ in (26), at which the zeros of the polynomial $a^{st}(k+1, 2n-1, s)$ reach the point s_i for the first time. We denote such a value by $\gamma(s_i)$.

For further presentation, it is necessary to obtain an explicit expression for the polynomial $a^{st}(k+1, 2n-1, s)$. By analogy with formula (16), we can write

$$\begin{aligned} a^{st}(k+1, 2n-1, s) &= a^0(n, a^0, s) \beta(k+1, n-1, s) - \\ &- b^0(m, b^0, s) \alpha(k+1, n-1, s), \end{aligned}$$

here

$$\begin{aligned} \beta(k+1, n-1, s) &= s^{n-1} + \sigma_1 \cdot r_1(k+1), \\ \alpha(k+1, n-1, s) &= \sigma_2 \cdot r_2(k+1), \\ \text{col}(r_1(k+1), r_2(k+1)) &= r(k+1), \end{aligned}$$

the vector $r(k+1)$ is determined by formulas (15) and (26)

$$r(k+1) = r(k) + \gamma \cdot \tau.$$

After reducing similar ones, we get the following expression for the polynomial $a^{st}(k+1, 2n-1, s)$:

$$a^{st}(k+1, 2n-1, s) = a^{st}(k, 2n-1, s) + \gamma \cdot \Delta a^{st}(k, 2n-2, s), \quad (27)$$

$$\Delta a^{st}(k, 2n-2, s) = (a^0(n, a^0, s) \cdot \sigma_1, -b^0(m, b^0, s) \cdot \sigma_2) \cdot \tau.$$

The desired value $\gamma(s_i)$ can be found from the condition that the polynomial $a^{st}(k+1, 2n-1, s)$ vanishes at the point s_i , which corresponds to the system of equations:

$$\begin{cases} a_{\text{Re}}^{st}(k, 2n-1, s_i) + \gamma \cdot \Delta a_{\text{Re}}^{st}(k, 2n-2, s_i) = 0, \\ a_{\text{Im}}^{st}(k, 2n-1, s_i) + \gamma \cdot \Delta a_{\text{Im}}^{st}(k, 2n-2, s_i) = 0. \end{cases} \quad (28)$$

The compatibility condition for system (28) is established by the Kronecker-Capelli theorem. If the system of equations (28) is consistent, then its solution has the form:

$$\gamma(s_i) = \begin{cases} \frac{-a_{\text{Re}}^{\text{st}}(k, 2n-1, s_i)}{\Delta a_{\text{Re}}^{\text{st}}(k, 2n-2, s_i)}, & \Delta a_{\text{Re}}^{\text{st}}(k, 2n-2, s_i) \neq 0, \\ \frac{-a_{\text{Im}}^{\text{st}}(k, 2n-1, s_i)}{\Delta a_{\text{Im}}^{\text{st}}(k, 2n-2, s_i)}, & \Delta a_{\text{Im}}^{\text{st}}(k, 2n-2, s_i) \neq 0. \end{cases} \quad (29)$$

It is easy to verify that the simultaneous execution

$$\Delta a_{\text{Re}}^{\text{st}}(k, 2n-2, s_i) = 0, \quad \text{and} \quad \Delta a_{\text{Im}}^{\text{st}}(k, 2n-2, s_i) = 0,$$

is equivalent to a violation of condition (17) and is thus excluded.

Note that when searching for $\gamma(s_i)$, the following situations are possible:

1) the number $\gamma(s_i)$ found by formula (29) can be negative, such a solution in the framework of the problem we are considering does not make sense;

2) the zeros of the polynomial $a^{\text{st}}(k+1, 2n-1, s)$ do not intersect the boundaries of the area S at the point s_i for any γ .

In these cases, the desired value $\gamma(s_i)$ will be taken equal to $+\infty$.

Define the function

$$\gamma_1 = \min_{s_i \in \Theta} \gamma^+(s_i),$$

$$\gamma^+(s_i) = \begin{cases} \gamma(s_i), & \exists \gamma(s_i) \wedge \gamma(s_i) > 0, \\ +\infty & \exists \gamma(s_i) \vee \gamma(s_i) < 0. \end{cases}$$

Based on the above, we can conclude that condition (18) is satisfied over the entire interval $\Gamma \doteq [0, \gamma_1]$.

The proof of this statement for $\gamma = 0$ is obvious, and for $\gamma = \gamma_1$ the zeros of the polynomial (27) for the first time come to the boundary of the area S .

On the interval Γ , one should find such a value of the parameter γ (we denote it by γ^*) for which condition (19) is satisfied for $\rho(k+1)$. The existence of γ^* follows from the existence of the vector τ , the general direction of increase of the function $\rho(k, s)$. The search for γ^* can be performed using the segment dichotomy method.

The above reasoning is the rationale for the algorithm given below for improving the property of robust control quality. The stopping condition of the algorithm is the achievement of a local maximum of the function ρ in the space of vectors $\mathbf{r}$. The achievement of the local maximum of the function ρ at the $k+1$ iteration corresponds to the fulfillment of at least one of the conditions:

- 1) the set I is empty, that (24) is satisfied;
- 2) there is no vector τ such that it satisfies inequalities (25).

As a result, we can formulate the following algorithm.

Algorithm for improving the property of robust control quality.

For the algorithm to function, it is necessary to define $\mathbf{r}(0)$, $a^{\text{st}}(0, 2n-1, s)$, Ω , $\rho(0, s_i)$, $\rho(0)$, ε_0 .

Accept $k = 0$.

Step 1. Form sets Θ и I .

Step 2. Check condition (24). If this condition is met, then go to stop, otherwise go to the next step.

Step 3. Using the linear classifier algorithm, find the vector τ that satisfies the system of inequalities (26).

If direction τ is found, go to the next step, if direction τ does not exist, go to stop.

Step 4. Calculate the value of γ^* for which conditions (18)–(19) are satisfied.

Step 5. Calculate vector

$$\mathbf{r}(k+1) = \mathbf{r}(k) + \gamma^* \cdot \tau.$$

$$\text{Accept } \mathbf{r}^* = \mathbf{r}(k+1).$$

Step 6. Accept $k = k+1$. Go to step 1.

Stop.

The convergence of this algorithm to the local maximum point of the function ρ was justified above. The number of iterations of the algorithm is finite, since the value of the function ρ is bounded.

3. An example of improving the properties of robust control quality

Let the control object be given by

$$\begin{aligned} & (v^0(2, z) + \Delta v(2, z)) \cdot (a^0(2, z) + \Delta a(1, z)) y(t) = \\ & = (w^0(2, z) + \Delta w(2, z)) \cdot (b^0(1, z) + \Delta b(1, z)) u(t), \end{aligned} \quad (30)$$

here

$$a^0(2, z) = z^2 - 1,918z + 0,923, \quad b^0(1, z) = 0,232z - 0,179,$$

$$\Delta a(1, z) = 10^{-4}[-1; 1]z + 10^{-4}[-2; 2],$$

$$\Delta b(1, z) = 10^{-2}[-1; 1]z + 10^{-2}[-1; 1].$$

$$v^0(2, z) = 1,830z^2 - 2,707z + 1,$$

$$w^0(2, z) = 1,848z^2 - 2,717z + 1,$$

$$\Delta v(2, z) = 10^{-4}[-5; 5]z^2 + 10^{-3}[-1; 1]z,$$

$$\Delta w(2, z) = 10^{-4}[-5; 5]z^2 + 10^{-3}[-1, 5; 1, 5]z.$$

Zeros of polynomials $v^0(2, z)$ and $w^0(2, z)$:

$$\Lambda(v^0) = \{0,715, \quad 0,764\} \subset C^1,$$

$$\Lambda(w^0) = \{0,735 + 0,027j, \quad 0,735 - 0,027j\} \subset C^1.$$

Requirements for the quality of control are set by the area

$$S = \{s : \eta_2 \leq |s| \leq \eta_1, \quad |\arg(s)| \leq \varphi\},$$

where $\eta_1 = 0,607$, $\eta_2 = 0,961$, $\varphi = \pi/4$ radian. Sets $\Lambda(v^0)$ and $\Lambda(w^0)$ lie inside the area S .

In accordance with the above synthesis scheme, the modal controller has the structure of a first-order difference equation:

$$(z + \beta_0)u(t) = (\alpha_1 z + \alpha_0)y(t) + \chi_0 g(t), \quad \mathbf{r} = \text{col}(\beta_0, \alpha_0, \alpha_1). \quad (31)$$

In [14], for the control object (30) and the standard

$$a^{\text{st}}(3, z) = z^3 - 2,715z^2 + 2,456z - 0,741.$$

we have calculated a modal controller of the form (31) with a vector of coefficients [14. P. 10]:

$$\mathbf{r}(0) = \text{col}(-0,833, \quad 0,157, \quad -0,158).$$

Let us set the problem: it is required to find a vector $\mathbf{r}^*$ such that it provides a local maximum of the function $\rho(\mathbf{r})$ in the vicinity of the vector $\mathbf{r}(0)$.

Solution. As a result of applying the above algorithm, after 12 iterations, we obtain the vector

$$\mathbf{r}^* = \text{col}(-0,838, \quad 0,135, \quad -0,134).$$

For the vector $\mathbf{r}(0)$ the value of ρ is equal to $6,136 \cdot 10^{-4}$, for the vector $\mathbf{r}^*$ the value of ρ is equal to $7,894 \cdot 10^{-4}$. Note that $\rho(\mathbf{r}^*) > \rho(\mathbf{r}(0))$ holds. The required vector $\mathbf{r}^*$ is found.

Conclusions

In this article, we have developed an algorithm for improving the robust quality of control for a closed control system consisting of a linear one-dimensional control object with structural-parametric uncertainty and a modal controller. The efficiency of the algorithm is illustrated by an example. The application of this algorithm will make it possible to obtain closed automatic control systems with higher indicators of robust stability margins and robust control quality. The design and implementation of such systems will lead to significant savings in the operation of control systems under conditions of uncertainty.

References

1. Krushel, E.G. & Stepanchenko, O.V. (2006) *Synthesis and modeling of digital control systems with a double time scale*. Moscow: Mashinostroenie-1.
2. Yurkevich, V.D. (2000) *Synthesis of non-linear non-stationary control systems with different-tempo processes*. St. Petersburg: Nauka.
3. Parshukov, A.N. (2021) Method of synthesis of a reduced-order modal regulator. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 56. pp. 12–19. DOI: 10.17223/19988605/56/2
4. Parshukov, A.N. (2009) *Methods of synthesis of modal regulators*. Tyumen: TyumGNGU.
5. Barmish, B.R. (1994) *New tools for robustness of linear systems*. New York: MacMillan.
6. Polyak, B.T. & Tsytkin, Ya.Z. (1990) Frequency criteria of robust stability and aperiodicity of linear systems. *Automation and Remote Control*. 51(9). pp. 1192–1201.
7. Polyak, B.T. & Shcherbakov, P.S. (2002) *Robust stability and control*. Moscow: Nauka.
8. Parshukov, A.N. (2022) Method of synthesis of a modal regulator for a linear discrete-time system with interval uncertainty of coefficients. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 60. pp. 4–12. DOI: 10.17223/19988605/60/1
9. Doyle, J.C., Glover, K., Khargonekar, P.P. & Francis, B.A. (1989) State-space solutions to standard H_2 and H_∞ control problems. *IEEE Transactions on Automatic Control*. 34(8). pp. 831–847.
10. Kogan, M.M. (2016) Optimal Discrete-time H_∞/γ_0 Filtering and Control under Unknown Covariances. *International Journal of Control*. 89(4). pp. 691–700.
11. Chestnov, V.N. (2015) H_∞ -approach to controller synthesis under parametric uncertainty and polyharmonic external disturbances. *Automation and Remote Control*. 76(6). pp. 1036–1048. DOI: 10.1134/S0005117915060077
12. Selvesyuk, N.I. (2016) Analytical synthesis of robust controllers of the given accuracy under external perturbations. *Journal of Computer and Systems Sciences International*. 55(4). pp. 558–569. DOI: 10.1134/S1064230716040109
13. Balandin, D.V., Kogan, M.M., Krivdina, L.N. & Fedyukov, A.A. (2014) Design of Generalized Discretetime H_∞ -optimal Control over Finite and Infinite Intervals. *Automation and Remote Control*. 75(1). pp. 1–17. DOI: 10.1134/S0005117914010019
14. Parshukov, A.N. (2024) Method for testing robust control quality for linear one-dimensional discrete control systems. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 66. pp. 4–12. DOI: 10.17223/19988605/66/1
15. Vapnik, V.N. & Chervonenkis, A.Ya. (1974) *Pattern recognition theory (statistical learning problems)*. Moscow: Nauka.

Information about the author:

Parshukov Andrej N. (Candidate of Technical Sciences, Associate Professor, Industrial University of Tyumen, Tyumen, Russian Federation). E-mail: anparshukov@mail.ru

The author declares no conflicts of interests.

Информация об авторе:

Паршуков Андрей Николаевич – доцент, кандидат технических наук, доцент кафедры «Электроэнергетика» Тюменского индустриального университета (Тюмень, Россия). E-mail: anparshukov@mail.ru

Автор заявляет об отсутствии конфликта интересов.

Received 29.11.2023; accepted for publication 03.06.2024

Поступила в редакцию 29.11.2023; принята к публикации 03.06.2024

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

MATHEMATICAL MODELING

Научная статья

УДК 519.218.2

doi: 10.17223/19988605/67/3

**Локальная предельная теорема для ветвящихся случайных систем
Гальтона–Ватсона с иммиграцией и бесконечной дисперсией****Аъзам Абдурахимович Имомов¹, Эркин Эгамбердиевич Тухтаев²**^{1,2} *Каршинский государственный университет, Карши, Узбекистан*¹ *imomov_azam@mail.ru*² *tukhtaev_erkina@mail.ru*

Аннотация. Рассматривается стохастическая модель развития популяции, которая представляет собой критическую случайную ветвящуюся систему Гальтона–Ватсона с иммиграцией. Рассмотрен случай, когда математическое ожидание закона притока частиц-иммигрантов и дисперсия закона размножения частиц-аборигенов имеют бесконечные значения. Исследованы асимптотические свойства вероятностей перехода в случае, когда состояния системы невозвратны. Для этого случая установлена локальная предельная теорема.

Ключевые слова: стохастическая модель популяции; ветвящиеся системы; иммиграция; переходные вероятности; медленно меняющаяся функция; предельные теоремы.

Для цитирования: Имомов А.А., Тухтаев Э.Э. Локальная предельная теорема для ветвящихся случайных систем Гальтона–Ватсона с иммиграцией и бесконечной дисперсией // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 67. С. 22–30. doi: 10.17223/19988605/67/3

Original article

doi: 10.17223/19988605/67/3

**Local limit theorem for branching random Galton-Watson systems
with immigration and infinite variance****A'zam A. Imomov¹, Erkin E. Tukhtaev²**^{1,2} *Karshi State University, Karshi, Uzbekistan*¹ *imomov_azam@mail.ru*² *tukhtaev_erkina@mail.ru*

Abstract. The paper considers a stochastic population growth model, which is the critical Galton-Watson stochastic branching system allowing immigration. We consider the case when the mathematical expectation of the law of influx of immigrant particles and the variance of the law of reproduction of aboriginal particles have infinite values. The asymptotic properties of transition probabilities are studied in the case when the states of the system are irreversible. For this case, a local limit theorem has been established.

Keywords: stochastic population model; branching system; immigration; transition probabilities; slow variation; limit theorems.

For citation: Imomov, A.A., Tukhtaev, E.E. (2024) Local limit theorem for branching random Galton-Watson systems with immigration and infinite variance. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 67. pp. 22–30. doi: 10.17223/19988605/67/3

Введение

В работе мы рассматриваем известную стохастическую модель развития популяции, называемую ветвящейся случайной системой Гальтона–Ватсона (Г–В) с иммиграцией. Эволюция этой системы происходит по следующей схеме. Вначале система пуста. В некоторый момент времени начинается приток однотипных частиц в систему из фиктивного источника, управляемый дискретным случайным законом $\{h_j, j \in \mathbb{N}_0\}$, здесь $\mathbb{N}_0 := \{0\} \cup \mathbb{N}$ и $\mathbb{N}$ – множество натуральных чисел. Каждая частица в любой момент времени $n \in \mathbb{N}$ независимо от своей предыстории и от числа существующих частиц производит частиц $j \in \mathbb{N}_0$ с вероятностью p_j . Одновременно продолжается приток иммигрантов по закону $\{h_j, j \in \mathbb{N}_0\}$. Частицы, попадающие в систему, претерпевают дальнейшие изменения в соответствии с распределением вероятностей $\{p_j\}$. Таким образом, эволюция системы происходит в соответствии с случайными законами $\{p_j\}$ и $\{h_j\}$. Обозначив X_n количество частиц в системе Г–В с иммиграцией в момент времени n , мы можем записать следующее рекуррентное соотношение:

$$X_0 = 0, \quad X_n = \xi_{n1} + \xi_{n2} + \dots + \xi_{nX_{n-1}} + \eta_n, \quad (1)$$

где ξ_{nk} и η_k – независимые случайные величины с распределением вероятностей $P\{\xi_{nk} = j\} = p_j$ и $P\{\eta_k = j\} = h_j$ соответственно для произвольного $k \in \mathbb{N}$. С этого момента мы предполагаем, что

$$p_0 > 0, \quad m := \sum_{j \in \mathbb{N}_0} jp_j < \infty \quad \text{и} \quad \sum_{j \in \mathbb{N}_0} h_j = 1.$$

Последовательность числа поколений $\{X_n\}$ образует однородно-дискретную во времени неразложимую цепь Маркова с пространством $\mathcal{S} \subset \mathbb{N}_0$ состоящим из одного класса сообщающихся состояний.

Введем в рассмотрение производящие функции

$$h(s) := \sum_{j \in \mathbb{N}_0} h_j s^j \quad \text{и} \quad f(s) := \sum_{j \in \mathbb{N}_0} p_j s^j \quad \text{для} \quad s \in [0, 1].$$

Обозначим p_{ij} вероятности перехода системы $\{X_n\}$ из состояния $i \in \mathcal{S}$ в состояние $j \in \mathcal{S}$ за один шаг. Тогда из (1) следует, что эти вероятности равны коэффициентам при s^i в разложении в степенной ряд функции $h(s)(f(s))^i$. Введем в рассмотрение переходные вероятности состояний системы за n шагов

$$p_{ij}^{(n)} := P\{X_{n+k} = j \mid X_k = i\} = P\{X_n = j \mid X_0 = i\}$$

для любого $k \in \mathbb{N}$. Стандартные аналитические вычисления показывают, что соответствующая производящая функция имеет следующий вид:

$$\mathcal{P}_n^{(i)}(s) := \sum_{j \in \mathcal{S}} p_{ij}^{(n)} s^j = (f_n(s))^i \prod_{k=0}^{n-1} h(f_k(s)), \quad (2)$$

где функция $f_n(s)$ представляет собой n -кратную итерацию производящей функции $f(s)$, т.е. $f_{n+1}(s) = f(f_n(s)) = f_n(f(s))$ (см.: [1]). Таким образом, вероятности $\{p_{ij}^{(n)}\}$ полностью определяются вероятностями $\{p_j\}$ и $\{h_j\}$. Учитывая представление (2), нетрудно вычислить математическое ожидание

$$E_i X_n = \sum_{j \in \mathcal{S}} jp_{ij}^{(n)} = \begin{cases} \left(\frac{\alpha}{m-1} + i \right) m^n - \frac{\alpha}{m-1} & \text{если } m \neq 1, \\ \alpha n + i & \text{если } m = 1, \end{cases}$$

где $\alpha = h'(1-)$. Отсюда видно, что классификация состояний цепи $\{X_n\}$ зависит от значения параметра $m = f'(1-)$ – среднего числа потомков одной частицы. Система $\{X_n\}$ называется докритической, критической и надкритической, если $m < 1$, $m = 1$ и $m > 1$ соответственно.

Вышеописанный популяционный процесс называем системой Г–В с иммиграцией. Она впервые была рассмотрена С. Хиткотом [2] в 1965 г. До этого, в 1957 г., Б. Севастьяновым [3] введено понятие и начато исследование ветвящихся случайных систем с иммиграцией в непрерывном случае. После публикации С. Хиткота [2] модель Г–В с иммиграцией стали исследовать многие известные ученые. В частности, задачу существования инвариантных мер относительно переходных вероятностей $p_{ij}^{(n)}$ и свойства состояний пространства $\mathcal{S}$ при бесконечном росте времени изучали Е. Сенета [4, 5], А. Пэйкс [6, 7]. В этих и в большинстве работ того периода в основном требовались моментные условия более высокого порядка для законов превращения $\{p_j\}$ и иммиграции частиц $\{h_j\}$.

Однако в работе [8], опубликованной В. Золотарёвым в 1957 г., были установлены принципиально новые результаты для марковских ветвящихся систем с непрерывным временем в обход условия конечности моментов целого порядка. Эта работа является одной из первых, в которых эффективно используется концепция теории медленно меняющихся (ММ) функций в смысле Карамата. Перспективность подхода В. Золотарёва подтвердилась в дальнейших исследованиях Р. Слэйка [9] и Е. Сенеты [10] в случае процессов Г–В без иммиграции и в работах А. Пэйкса [11, 12] для системы Г–В с иммиграцией (см. также: [13–18]).

1. Вспомогательные результаты

Рассмотрим критический случай.

А. Пэйкс [12] доказал следующую теорему.

Теорема P1 [12, Р. 80]. Если $m = 1$, тогда

$$p_{00}^{(n)} \sim K \exp \left\{ \int_1^{e^n} \frac{\ln h\{1 - \varphi(y)\}}{y} dy \right\} \quad \text{при } n \rightarrow \infty,$$

где $\varphi(y)$ – убывающая ММ-функция. Если

$$\sum_{m=0}^{\infty} \left[(1 - h(f_m(0))) (1 - f'(f_m(0))) \right] < \infty,$$

тогда

$$p_{00}^{(n)} \sim K_1 \exp \left\{ \int_0^{f_n(0)} \frac{\ln h(y)}{f(y) - y} dy \right\} \quad \text{при } n \rightarrow \infty.$$

Здесь K и K_1 – не имеющие явного вида абсолютные константы.

Допускаем, что производящая функция $f(s)$ имеет следующий вид:

$$f(s) - s = (1 - s)^{1+\nu} \mathcal{L}\left(\frac{1}{1-s}\right) \quad [f_\nu] \quad (1)$$

для $s \in [0, 1)$, где $0 < \nu < 1$ и $\mathcal{L}(x)$ – ММ-функция на бесконечности. Наличие условия $[f_\nu]$ влечет $2b := f''(1-) = \infty$, а в случае $0 < b < \infty$ это условие имеет место с $\nu = 1$ и $\mathcal{L}(t) \rightarrow b$ при $t \rightarrow \infty$. Также предполагается, что производящая функция $h(s)$ закона притока частиц-иммигрантов допускает следующее представление:

$$1 - h(s) = (1 - s)^\delta l\left(\frac{1}{1-s}\right) \quad [h_\delta] \quad (2)$$

для $s \in [0, 1)$, где $0 < \delta < 1$ и $l(x)$ – ММ-функция на бесконечности.

В упомянутой выше работе А. Пэйкса [12] было выделено следующее асимптотическое разбиение пространства состояний в зависимости от соотношений параметров v и δ : если $\delta > v$, то $\mathcal{S}$ эргодично, если $\delta < v$ и $\delta(1+v) > v$, то $\mathcal{S}$ нуль возвратно, а если $\delta(1+v) < v$, то $\mathcal{S}$ невозвратно.

В работе авторов [19] рассматривался случай $\delta > v$, и в этом случае утверждения Теоремы P1 были улучшены путем наложения дополнительных условий на функции $\mathcal{L}(x)$ и $l(x)$. Там же найдено асимптотическое представление производящей функции $\mathcal{P}_n^{(i)}(s)$ для всех $s \in [0,1)$. Ниже приводятся условия, принятые в статье [19]. Поскольку $\mathcal{L}(x)$ – ММ-функция на бесконечности, то $\mathcal{L}(\lambda x)/\mathcal{L}(x) \rightarrow 1$ при $x \rightarrow \infty$ для любого $\lambda \in (0, \infty)$. Тогда можно записать

$$\frac{\mathcal{L}(\lambda x)}{\mathcal{L}(x)} - 1 = \alpha_\lambda(x), \quad \mathcal{L}_\alpha$$

где $\alpha_\lambda(x) \rightarrow 0$ при $x \rightarrow \infty$. В дальнейшем будем считать, что задана положительная функция $g(x)$ такая, что $g(x) \rightarrow 0$ и $\alpha_\lambda(x) = o(g(x))$ при $x \rightarrow \infty$. В этом случае $\mathcal{L}(x)$ называется ММ-функцией на бесконечности с остатком $\alpha_\lambda(x)$ [14. Р. 185, SR3]. Выбираем $g(x) = \mathcal{O}(\mathcal{L}(x)/x^v)$. Тогда получаем

$$\alpha_\lambda(x) = o\left(\frac{\mathcal{L}(x)}{x^v}\right) \quad \text{при } x \rightarrow \infty. \quad (3)$$

Точно так же на функцию $l(x)$ накладывается следующее условие:

$$\frac{l(\lambda x)}{l(x)} - 1 = \beta_\lambda(x), \quad [l_\beta]$$

где

$$\beta_\lambda(x) = o\left(\frac{l(x)}{x^\delta}\right) \quad \text{при } x \rightarrow \infty \quad (4)$$

для любого $\lambda \in (0, \infty)$.

Далее для удобства изложения введем вспомогательную функцию

$$\Lambda(y) := \frac{f(1-y) - (1-y)}{y} = y^v \mathcal{L}\left(\frac{1}{y}\right) \quad (5)$$

для $y \in (0,1]$. Тогда условие $[f_v]$ принимает следующий вид:

$$f(s) - s = (1-s)\Lambda(1-s). \quad f_v$$

Ниже приведем несколько полезных для наших дальнейших рассуждений результатов, установленных в предыдущих работах авторов.

Лемма 1 [16. Р. 52]. Если выполнено условие $[f_v]$, то

$$1 - f_n(s) = \frac{\mathcal{N}(n)}{(vn)^{1/v}} \cdot \left[1 - \frac{U_n(s)}{vn}\right],$$

где $\mathcal{N}(\cdot)$ – ММ-функция на бесконечности такая, что

$$\mathcal{N}(n) \mathcal{L}^{1/v}\left(\frac{(vn)^{1/v}}{\mathcal{N}(n)}\right) \rightarrow 1$$

при $n \rightarrow \infty$, и функция $U_n(s)$ имеет следующие свойства:

1) $U_n(s) = U(s)(1 + o(1))$ при $n \rightarrow \infty$, и функция $U(s)$ удовлетворяет уравнению Абеля

$$U(f(s)) = U(s) + 1;$$

2) $\lim_{s \uparrow 1} U_n(s) = vn$ при любом $n \in \mathbb{N}$;

3) $U_n(0) = 0$ при любом $n \in \mathbb{N}$.

Теорема 1 [19]. Если условия $[f_v]$ и $[\mathcal{L}_\alpha]$ вместе с (3) выполняются, то

1) производящая функция $U(s)$ имеет следующий вид:

$$U(s) = \int_0^s \frac{\psi(y)}{(1-y)\Lambda(1-y)} dy, \quad (6)$$

где функция $\psi(s)$ непрерывна для $s \in [0,1]$ и $f'(s) \leq \psi(s) \leq 1$;

2) для всех $s \in [0,1]$ производная $U'(s)$ имеет следующий вид:

$$U'(s) = \frac{\psi(s)}{(1-s)\Lambda(1-s)}, \quad (7)$$

где $\psi(s) = 1 + \mathcal{O}(\Lambda(1-s))$ при $s \uparrow 1$.

Известно, что в рассматриваемом случае $f_n(s) \rightarrow 1$ при $n \rightarrow \infty$. Поэтому, ввиду формулы (2), достаточно рассмотреть случай $i=0$ для изучения асимптотики системы. Обозначим $\mathcal{P}_n(s) := \mathcal{P}_n^{(0)}(s)$. Следующая теорема обобщает утверждения Теоремы P1.

Теорема 2 [19]. Пусть имеется система Г–В с иммиграцией с условием $[f_v]$.

1) если условие $[\mathcal{L}_\alpha]$ выполняется вместе с (3), то

$$\mathcal{P}_n(s) = \Delta_n(s) \exp \left\{ \int_s^{f_n(s)} \frac{\psi(y) \ln h(y)}{(1-y)\Lambda(1-y)} dy \right\} \quad (8)$$

где функции $\psi(s)$ определена в Теореме 1 и

$$\frac{h(s)}{h(f_n(s))} \leq \Delta_n(s) \leq 1, \quad (9)$$

2) если дополнительно выполняются условия $[h_\delta]$ и $\delta > v$, то

$$\mathcal{P}_n(s) = \Delta_n(s) \omega_n(s) \exp \left\{ - \int_s^{f_n(s)} \frac{1-h(y)}{(1-y)\Lambda(1-y)} \psi(y) dy \right\}, \quad (10)$$

здесь

$$\exp \left\{ - \frac{1}{h(s)} \int_s^{f_n(s)} \frac{[1-h(y)]^2}{(1-y)\Lambda(1-y)} dy \right\} \leq \omega_n(s) \leq 1. \quad (11)$$

Лемма 2 ([19]). Справедливы следующие утверждения:

- условие $[\mathcal{L}_\alpha]$ с (3) равносильно $\mathcal{L}(x) = C_\mathcal{L} + o(x^{-v})$ при $x \rightarrow \infty$;
- условие $[l_\beta]$ с (4) равносильно $l(x) = C_l + o(x^{-\delta})$ при $x \rightarrow \infty$,

где $C_\mathcal{L}$ и C_l – положительные константы.

Следующая лемма указывает на «асимптотическую нечувствительность» ММ-функций при интегрировании регулярно меняющихся функций вдоль положительной бесконечной оси.

Лемма 3 [20, Р. 143]. Пусть $L(t)$ – ММ-функция на бесконечности. Тогда для любого $\sigma < 0$ и любой постоянной $c \geq 1$ справедливо следующее асимптотическое представление:

$$\int_c^t y^{-(1+\sigma)} L(y) dy = \frac{1}{|\sigma|} t^{|\sigma|} L(t) (1 + o(1)) \quad \text{при } t \rightarrow \infty.$$

Доказательства приведенных выше утверждений полностью и подробно приведены в указанной литературе.

2. Локальная предельная теорема

Перейдем к изложению и доказательству основного результата статьи. Рассмотрим случай $\delta < v$. Введем следующие обозначения:

$$L(x) := \frac{l(x)}{\mathcal{L}(x)} \quad \text{и} \quad \tau(n) := \frac{(vn)^{1/\nu}}{\mathcal{N}(n)},$$

где функция $\mathcal{N}(n)$ определена в Лемме 1.

Теорема 3. Пусть выполнены условия $[f_\nu]$, $[h_\delta]$ и $\gamma := \delta - \nu < 0$. Тогда

$$-(\tau(n))^{-|\gamma|} \cdot \ln p_{00}^{(n)} = \frac{1}{|\gamma|} L(\tau(n)) + O\left(\frac{1}{n^{1-\delta/\nu}}\right) \quad \text{при } n \rightarrow \infty.$$

Доказательство. Первичные формальные рассуждения приводят нас к соотношению (10). Из этого соотношения при $s = 0$, учитывая условия $[f_\nu]$ и $[h_\delta]$, получаем следующее равенство:

$$p_{00}^{(n)} = \omega_n \Delta_n \exp \left\{ - \int_0^{f_n(0)} (1-y)^{\gamma-1} L\left(\frac{1}{1-y}\right) dy \right\}, \quad (12)$$

где $\omega_n := \omega_n(0)$, $\Delta_n := \Delta_n(0)$. Введя обозначения $Q_n := 1 - f_n(0)$ и $x := 1/(1-y)$, преобразуем равенство (12) к следующему виду:

$$p_{00}^{(n)} = \omega_n \Delta_n \exp \left\{ - \int_1^{1/Q_n} x^{-(1+\gamma)} L(x) dx \right\} =: \omega_n \Delta_n \exp \{ -\mathcal{J}_n \}, \quad (13)$$

где $\mathcal{J}_n := \int_1^{1/Q_n} x^{-(1+\gamma)} L(x) dx$.

Вычислим интеграл $\mathcal{J}_n$. Подынтегральное выражение удовлетворяет условиям Леммы 3. Поэтому справедливо следующее соотношение:

$$\mathcal{J}_n = \int_1^{1/Q_n} x^{-(1+\gamma)} L(x) dx = \frac{1}{|\gamma|} \cdot \frac{1}{Q_n^{|\gamma|}} L\left(\frac{1}{Q_n}\right) (1 + o(1))$$

при $n \rightarrow \infty$. С другой стороны, согласно Лемме 1, $\tau(n)Q_n \rightarrow 1$, или же $\tau(n) \sim 1/Q_n$, при $n \rightarrow \infty$. Следовательно, имеем

$$\mathcal{J}_n = \frac{1}{|\gamma|} \cdot (\tau(n))^{|\gamma|} \cdot L(\tau(n)) (1 + o(1)) \quad \text{при } n \rightarrow \infty. \quad (14)$$

По определению функции $\tau(n)$ справедливо соотношение

$$\tau(n) = O(n^{1/\nu}) \quad \text{при } n \rightarrow \infty.$$

Отсюда следует, что интеграл $\mathcal{J}_n$ расходится.

Соотношения (13) и (14) дают следующее асимптотическое представление:

$$\ln p_{00}^{(n)} = \ln \omega_n \Delta_n - \frac{1}{|\gamma|} \cdot (\tau(n))^{|\gamma|} \cdot L(\tau(n)) (1 + o(1)) \quad (15)$$

при $n \rightarrow \infty$. Умножая на $-(\tau(n))^{-|\gamma|}$ в обеих частях (15), получаем

$$-(\tau(n))^{-|\gamma|} \cdot \ln p_{00}^{(n)} \sim \frac{1}{|\gamma|} \cdot L(\tau(n)) - (\tau(n))^{-|\gamma|} \cdot \ln \omega_n \Delta_n \quad (16)$$

при $n \rightarrow \infty$. Оценим $\omega_n \Delta_n$. Положим $s = 0$ в неравенствах (9) и (11). В этом случае из неравенства (11) получаем оценку

$$\exp \left\{ - \frac{1}{h_0} \frac{1}{\gamma + \delta} L_l(1) \right\} \leq \omega := \lim_{n \rightarrow \infty} \omega_n \leq 1, \quad (17)$$

где $L_l(x) \sim L(x)l(x)$ при $x \rightarrow \infty$. Кроме того, как показано в [14, Р. 179], существует $\Delta := \lim_{n \rightarrow \infty} \Delta_n$, для которого верна оценка

$$h_0 \leq \Delta \leq 1. \quad (18)$$

Из (17) и (18) следует, что выражение $\omega_n \Delta_n$ равномерно ограничено при $n \rightarrow \infty$. Тогда в силу определения функции $\tau(n)$ следует соотношение

$$(\tau(n))^{-|\gamma|} \ln \omega_n \Delta_n = \mathcal{O}\left(\frac{1}{n^{|\gamma|/n}}\right) \quad (19)$$

при $n \rightarrow \infty$. Из (16) и (19) получаем, что

$$-(\tau(n))^{-|\gamma|} \ln p_{00}^{(n)} = \frac{1}{|\gamma|} L(\tau(n)) + \mathcal{O}\left(\frac{1}{n^{|\gamma|/v}}\right)$$

при $n \rightarrow \infty$. Найденное соотношение равносильно утверждению теоремы.

Заключение

Рассмотренная в статье случайная ветвящаяся система Гальтона–Ватсона с иммиграцией является одной из наиболее подходящих моделей для моделирования задач, связанных с развитием популяции. Аналогичные задачи возникают, например, в демографии, когда приток иммигрантов поддерживается в целях балансирования человеческих ресурсов.

Наш метод, основанный на свойствах медленно меняющихся функций Караматы, позволяет обойти условия конечности высоких моментов закона размножения частиц-аборигенов и закон притока частиц-иммигрантов. В нашем случае важна информация об асимптотическом поведении вероятностей перехода. Несмотря на то, что мы нашли асимптотическое выражение только для локальной вероятности $p_{00}^{(n)}$, на самом деле этого достаточно, поскольку теорема о монотонной сходимости отношений $p_{ij}^{(n)} / p_{00}^{(n)}$ позволяет оценить вероятность перехода для любых $i, j \in S$ [17].

Иногда появляется необходимость в аналоге результата, полученного в Теореме 2, который можно установить и для других популяционных моделей. Например, более сложные проблемы возникают, когда частицы в системе имеют разную продолжительность жизни. Для этого случая подходит модель развития популяции, называемая «ветвящиеся системы (процессы) с превращениями, зависящими от возраста частиц». В этой связи наши следующие работы будут посвящены исследованию таких типов моделей развития популяции.

Список источников

1. Pakes A.G. Limit theorems for the simple branching process allowing immigration, I. The case of finite offspring mean // Adv. Appl. Prob. 1979. V. 11. P. 31–62.
2. Heathcote C.R. A branching process allowing immigration // J. Royal Stat. Soc. 1965. V. B-27. P. 138–143.
3. Севастьянов Б.А. Предельные теоремы для ветвящихся случайных процессов специального вида // Теория вероятностей и ее применения. 1957. Т. 2, вып. 3. С. 339–348.
4. Seneta E. Functional equations and the Galton-Watson process // Adv. Appl. Prob. 1969. V. 1. P. 1–42.
5. Seneta E. The stationary distribution of a branching process allowing immigration: a remark on the critical case // J R. Stat. Soc. 1968. V. B-30 (1). P. 176–179.
6. Pakes A.G. On the critical Galton-Watson process with immigration // Jour. Austral. Math. Soc. 1971. V. 12. P. 476–482.
7. Pakes A.G. Branching processes with immigration // Jour. Appl. Prob. 1971. V. 8 (1). P. 32–42.
8. Zolotarev V.M. More exact statements of several theorems in the theory of branching processes // Theory Prob. and Appl. 1957. V. 2. P. 256–266.
9. Slack R.S. A branching process with mean one and possible infinite variance // Wahrscheinlichkeitstheor. und Verw. Geb. 1968. V. 9. P. 139–145.
10. Seneta E. Regularly varying functions in the theory of simple branching process // Adv. Appl. Prob. 1974. V. 6. P. 408–420.
11. Pakes A.G. Revisiting conditional limit theorems for the mortal simple branching process // Bernoulli. 1999. V. 5 (6). P. 969–998.
12. Pakes A.G. Some results for non-supercritical Galton-Watson process with immigration // Math. Biosci. 1975. V. 24. P. 71–92.
13. Bingham N.H., Goldie C.M., Teugles J.L. Regular Variation. Cambridge : Cambridge University Press, 1987.
14. Pakes A.G. Some new limit theorems for the critical branching process allowing immigration // Stoch. Proc. Appl. 1975. V. 3. P. 175–185.
15. Imomov A.A. On a limit structure of the Galton-Watson branching processes with regularly varying generating functions // Prob. and Math. Stat. 2019. V. 39 (1). P. 61–73. doi:10.19195/0208-4147.39.1.4

16. Imomov A.A., Tukhtaev E.E. On application of slowly varying functions with remainder in the theory of Galton-Watson branching process // Jour. Siber. Fed. Univ. Math. Phys. 2019. V. 12 (1). P. 51–57. doi:10.17516/1997-1397-2019-12-1-51-57
17. Imomov A.A. On long-time behaviors of states of Galton-Watson branching processes allowing immigration // Jour. Siber. Fed. Univ. Math. Phys. 2015. V. 8 (4). P. 394–405.
18. Imomov A.A. Limit theorem for the joint distribution in the Q-processes // Jour. Siber. Fed. Univ. Math. Phys. 2014. V. 7 (3). P. 289–296.
19. Imomov A.A., Tukhtaev E.E. On asymptotic structure of critical Galton-Watson branching processes allowing immigration with infinite variance // Stochastic Models. 2023. V. 39 (1). P. 118–140.
20. Imomov A.A., Meyliyev A.Kh. On asymptotic structure of continuous-time Markov branching processes allowing immigration without higher-order moments // Ufimsk. Mat. Zh. 2021. V. 13 (1). P. 137–147.

References

1. Pakes, A.G. (1979) Limit theorems for the simple branching process allowing immigration. I. The case of finite offspring mean. *The Advances in Applied Probability*. 11. pp. 31–62.
2. Heathcote, C.R. (1965) A branching process allowing immigration. *The Journal of the Royal Statistical Society*. B-27. pp. 138–143.
3. Sevastyanov, B.A. (1957) Predel'nye teoremy dlya vetvyashchikhsya sluchaynykh protsessov spetsial'nogo vida [Limit theorems for branching random processes of a special type]. *Teoriya veroyatnostey i ee primeneniya* [Probability Theory and its Applications]. 2(3). pp. 360–374.
4. Seneta, E. (1969) Functional equations and the Galton-Watson process. *The Advances in Applied Probability*. 1. pp. 1–42.
5. Seneta, E. (1968) The stationary distribution of a branching process allowing immigration: a remark on the critical case. *The Journal of the Royal Statistical Society*. B-30(1). pp. 176–179.
6. Pakes, A.G. (1971) On the critical Galton-Watson process with immigration. *The Journal of the Australian Mathematical Society*. 12. pp. 476–482.
7. Pakes, A.G. (1971) Branching processes with immigration. *Journal of Applied Probability*. 8(1). pp. 32–42.
8. Zolotarev, V.M. (1957) More exact statements of several theorems in the theory of branching processes. *Theory of Probability and Its Applications*. 2. pp. 256–266.
9. Slack, R.S. (1968) A branching process with mean one and possible infinite variance. *Wahrscheinlichkeitstheorie und verwandte Gebiete*. 9. pp. 139–145.
10. Seneta, E. (1974) Regularly varying functions in the theory of simple branching process. *The Advances in Applied Probability*. 6. pp. 408–420.
11. Pakes, A.G. (1999) Revisiting conditional limit theorems for the mortal simple branching process. *Bernoulli*. 5(6). pp. 969–998.
12. Pakes, A.G. (1975) Some results for non-supercritical Galton-Watson process with immigration. *Mathematical Biosciences*. 24. pp. 71–92.
13. Bingham, N.H., Goldie C.M. & Teugles, J.L. (1987) *Regular Variation*. Cambridge University Press.
14. Pakes, A.G. (1975) Some new limit theorems for the critical branching process allowing immigration. *Stochastic Processes and their Applications*. 3. pp. 175–185.
15. Imomov, A.A. (2019) On a limit structure of the Galton-Watson branching processes with regularly varying generating functions. *Probability and Mathematical Statistics*. 39(1). pp. 61–73. DOI: 10.19195/0208-4147.39.1.4
16. Imomov, A.A. & Tukhtaev, E.E. (2019) On application of slowly varying functions with remainder in the theory of Galton-Watson branching process. *Journal of Siberian Federal University. Mathematics and Physics*. 12(1). pp. 51–57. DOI: 10.17516/1997-1397-2019-12-1-51-57
17. Imomov, A.A. (2015) On long-time behaviors of states of Galton-Watson branching processes allowing immigration. *Journal of Siberian Federal University. Mathematics and Physics*. 8(4). pp. 394–405.
18. Imomov, A.A. (2014) Limit theorem for the joint distribution in the Q-processes. *Journal of Siberian Federal University. Mathematics and Physics*. 7(3). pp. 289–296.
19. Imomov, A.A. & Tukhtaev, E.E. (2023) On asymptotic structure of critical Galton-Watson branching processes allowing immigration with infinite variance. *Stochastic Models*. 39(1). pp. 118–140.
20. Imomov, A.A. & Meyliyev, A.Kh. (2021) On asymptotic structure of continuous-time Markov branching processes allowing immigration without higher-order moments. *Ufimskiy matematicheskiy zhurnal*. 13(1). pp. 137–147.

Информация об авторах:

Имомов Аъзам Абдурахимович – доктор физико-математических наук, заведующий кафедрой алгебры и геометрии Каршинского государственного университета (Карши, Узбекистан). E-mail: imomov_azam@mail.ru

Тухтаев Эркин Эгамбердиевич – аспирант Каршинского государственного университета (Карши, Узбекистан). E-mail: tukhtaev_erkin@mail.ru

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

Information about the authors:

Imomov A'zam A. (Doctor of Physical and Mathematical sciences, Karshi State University, Karshi, Uzbekistan). E-mail: imomov_azam@mail.ru

Tukhtaev Erkin E. (Post-graduate Student, Karshi State University, Karshi, Uzbekistan). E-mail: tukhtaev_erkin@mail.ru

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

Поступила в редакцию 12.02.2024; принята к публикации 03.06.2024

Received 12.02.2024; accepted for publication 03.06.2024

Научная статья

УДК 517.95

doi: 10.17223/19988605/67/4

Численное решение системы уравнений Навье–Стокса в случае сжимаемой среды с использованием нейронных сетей

Кирилл Сергеевич Кузнецов¹, Елена Владимировна Амосова²

^{1,2} Дальневосточный федеральный университет, Владивосток, Россия

^{1,2} Институт прикладной математики ДВО РАН, Владивосток, Россия

¹ kuznetsovks17@gmail.com

² el_amosova@mail.ru

Аннотация. Рассматриваются вопросы использования метода Physics Informed Neural Networks (PINN) для численного решения нестационарной нелинейной системы дифференциальных уравнений в частных производных, описывающей процесс движения одномерного теплопроводного газа. Используемый подход основан на том, что нейронная сеть приближает решение системы дифференциальных уравнений, при этом учитывая физику моделируемого процесса. Обучение нейронной сети происходит на основе минимизации квадратичного функционала, построенного на невязке дифференциальных уравнений, граничных и начальных условий. Обсуждаются различные виды приближения исходных уравнений в случае, когда оператор по времени непрерывен или дискретен. Выполнен анализ результатов моделирования.

Ключевые слова: уравнения Навье–Стокса; нейронные сети; моделирование процессов газовой динамики.

Благодарности: Работа выполнена в рамках государственного задания ИПМ ДВО РАН (№ 075-00459-24-00) и при поддержке Министерства науки и высшего образования РФ (проект № 075-02-2024-1440).

Для цитирования: Кузнецов К.С., Амосова Е.В. Численное решение системы уравнений Навье–Стокса в случае сжимаемой среды с использованием нейронных сетей // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 67. С. 31–41. doi: 10.17223/19988605/67/4

Original article

doi: 10.17223/19988605/67/4

Numerical solution of the system of Navier-Stokes equations in the case of a compressible medium using neural networks

Kirill S. Kuznetsov¹, Elena V. Amosova²

^{1,2} Far Eastern Federal University, Vladivostok, Russia

^{1,2} Institute for Applied Mathematics FEB RAS, Vladivostok, Russia

¹ kuznetsovks17@gmail.com

² el_amosova@mail.ru

Abstract. The use of the Physics Informed Neural Networks (PINN) method for the numerical solution of a nonstationary nonlinear system of partial differential equations describing the process of motion of a one-dimensional heat-conducting gas is considered. The approach is based on the fact that a neural network approximates the solution of a system of differential equations, while taking into account the physics of the simulated process. The neural network is trained by minimizing a quadratic functional built on the difference between the predicted values and residuals of differential equations, boundary and initial conditions. Different types of approximation of the original equations are discussed in the case when the operator is continuous or discrete in time. An analysis of the presented methods is carried out. Their advantages and disadvantages are indicated.

Keywords: Navier-Stokes equations; simulation of gas dynamics processes; neural networks.

Acknowledgments: The work was carried out with the state assignment of the Institute for Applied Mathematics, Far Eastern Branch of the Russian Academy of Sciences (No. 075-00459-24-00) and with the support of the Ministry of Science and Higher Education of the Russian Federation (project No. 075-02-2024-1440).

For citation: Kuznetsov, K.S., Amosova, E.V. (2024) Numerical solution of the system of Navier-Stokes equations in the case of a compressible medium using neural networks. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 67. pp. 31–41. doi: 10.17223/19988605/67/4

Введение

Численное решение системы уравнений Навье–Стокса актуально в силу ее широкого практического применения. Особый интерес вызывает численное решение подобных задач в приложении к сжимаемым средам вследствие их математической сложности.

Поведение решений уравнений Навье–Стокса «в целом» по времени для сжимаемой среды исчерпывающе изучено только в одномерном случае [1, 2].

В работе [3] впервые описан метод решения дифференциальных уравнений в частных производных при помощи нейронных сетей. Авторы предложили подход, в основе которого лежит принцип приближения неизвестной функции нейронной сетью путем минимизации квадратичного функционала, включающего невязку уравнений, начальные и граничные условия. Впоследствии метод получил широкое распространение под названием Physics Informed Neural Networks (PINN) и активно используется в последние годы для решения различных задач гидродинамики [4–7]. В работе [3] рассмотрен способ временной дискретизации с использованием методов Рунге–Кутты высокого порядка, что позволяет снизить требования к вычислительным мощностям при расчете без потери точности решения.

В работах [3–6] рассматривается решение уравнений Навье–Стокса. В [3] приводится решение задачи восстановления неизвестных параметров и поля давления по известным значениям скорости среды. В [5] исследуется решение уравнений Рейнольдса в двумерном стационарном случае при числах Рейнольдса в диапазоне 10^2 – $2 \cdot 10^5$. В работах проведено сравнение полученных при помощи приближения решений дифференциальных уравнений с экспериментальными данными.

В данной работе подход, развитый в [3–6], используется для решения системы уравнений Навье–Стокса в случае сжимаемой среды. Исследуется применение PINN как в случае модели с непрерывным временем процесса, так и для модели с дискретизацией по времени. Отметим, что уравнения Навье–Стокса в случае сжимаемой и в случае несжимаемой среды отличаются с точки зрения классификации дифференциальных уравнений, в силу чего к ним не могут применяться одни и те же способы и методы решения. Другой особенностью данной работы является исследование возможности решения системы уравнений для неоднородной среды, характеризуемой большими числами Рейнольдса и Пекле, $Re \approx 10^7$, $Pe \approx 10^7$, которые соответствуют реальным физическим характеристикам движения газа. Показано, что применение временной дискретизации позволяет снизить требования к вычислительным ресурсам.

1. Математическая модель

Математическую модель, описывающую нестационарное движение вязкого политропного газа в одномерном пространстве через интервал $\Omega_0 = \{x: 0 < x < L\}$, ограниченный с двух сторон при $t > 0$, вместе с граничными условиями можно выразить следующей системой уравнений [1, 2]:

$$\rho \left[\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} \right] = \mu \frac{\partial^2 u}{\partial x^2} - \frac{\partial}{\partial x} (R\rho\theta), \quad (1)$$

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x} (u\rho) = 0, \quad (2)$$

$$c_v \rho \left[\frac{\partial \theta}{\partial t} + u \frac{\partial \theta}{\partial x} \right] = \kappa \frac{\partial^2 \theta}{\partial x^2} + \mu \left(\frac{\partial u}{\partial x} \right)^2 - R\rho\theta \frac{\partial u}{\partial x}, \quad (3)$$

$$u|_{t=0}=u_0(x), \quad \rho|_{t=0}=\rho_0(x), \quad \theta|_{t=0}=\theta_0(x), \quad (4)$$

$$u|_{x=0}=u_1(t), \quad \rho|_{x=0}=\rho_1(t), \quad \theta|_{x=0}=\theta_1(t), \quad (5)$$

$$u|_{x=L}=u_2(t), \quad \theta|_{x=L}=\theta_2(t), \quad x \in [0, L], \quad t \in [0, T], \quad (6)$$

где u, ρ, θ – неизвестные функции скорости, плотности и температуры газа соответственно, μ – динамическая вязкость газа, R – универсальная газовая постоянная, c_v – удельная теплоемкость газа, κ – коэффициент теплопроводности газа, $u_0, \rho_0, \theta_0, u_1, \rho_1, \theta_1, u_2, \theta_2$ – заданные функции.

Решать задачу необходимо в безразмерных переменных. Введем новые переменные

$$\tilde{x} = \frac{x}{L}, \quad \tilde{t} = \frac{t}{T}, \quad \tilde{u} = \frac{u}{u_s}, \quad \tilde{\rho} = \frac{\rho}{\rho_s}, \quad \tilde{\theta} = \frac{\theta}{\theta_s},$$

где u_s, ρ_s, θ_s – заданные значения.

После подстановки введем новую замену:

$$\tilde{x} = x, \quad \tilde{t} = t, \quad \tilde{u} = u, \quad \tilde{\rho} = \rho, \quad \tilde{\theta} = \theta,$$

Тогда модель (1)–(6) в безразмерных переменных будет иметь вид:

$$\rho \left[\frac{\partial u}{\partial t} + Sh u \frac{\partial u}{\partial x} \right] = \frac{Sh}{Re} \frac{\partial^2 u}{\partial x^2} - Sh k \frac{\partial}{\partial x} (\rho \theta), \quad (7)$$

$$\frac{\partial \rho}{\partial t} + Sh \frac{\partial}{\partial x} (u \rho) = 0, \quad (8)$$

$$\rho \left[\frac{\partial \theta}{\partial t} + Sh u \frac{\partial \theta}{\partial x} \right] = \frac{Sh}{Pe} \frac{\partial^2 \theta}{\partial x^2} + \frac{Sh}{Re} \frac{\pi}{k} \left(\frac{\partial u}{\partial x} \right)^2 - Sh \pi \rho \theta \frac{\partial u}{\partial x}, \quad (9)$$

$$u|_{t=0}=u_0^s(x), \quad \rho|_{t=0}=\rho_0^s(x), \quad \theta|_{t=0}=\theta_0^s(x), \quad (10)$$

$$u|_{x=0}=u_1^s(t), \quad \rho|_{x=0}=\rho_1^s(t), \quad \theta|_{x=0}=\theta_1^s(t), \quad (11)$$

$$u|_{x=1}=u_2^s(t), \quad \theta|_{x=1}=\theta_2^s(t), \quad x \in [0, 1], \quad t \in [0, 1], \quad (12)$$

где Sh – число Струхала, Pe – число Пекле, Re – число Рейнольдса, k, π – безразмерные коэффициенты, $u_0^s, u_1^s, u_2^s, \rho_0^s, \rho_1^s, \theta_0^s, \theta_1^s, \theta_2^s$ – безразмерные функции граничных и начальных условий. Данные константы и функции определяются по формулам

$$Sh = \frac{Tu_s}{L}, \quad Pe = \frac{u_s \rho_s c_v L}{\kappa}, \quad Re = \frac{u_s \rho_s L}{\mu}, \quad k = \frac{R \theta_s}{u_s^2}, \quad \pi = \frac{R}{c_v},$$

$$u_0^s(x) = \frac{u_0(x)}{u_s}, \quad u_1^s(x) = \frac{u_1(x)}{u_s}, \quad u_2^s(x) = \frac{u_2(x)}{u_s}, \quad \rho_0^s(x) = \frac{\rho_0(x)}{\rho_s},$$

$$\rho_1^s(x) = \frac{\rho_1(x)}{\rho_s}, \quad \theta_0^s(x) = \frac{\theta_0(x)}{\theta_s}, \quad \theta_1^s(x) = \frac{\theta_1(x)}{\theta_s}, \quad \theta_2^s(x) = \frac{\theta_2(x)}{\theta_s}.$$

2. Метод решения

Как указано во введении, для решения используется подход, основанный на приближении неизвестных функций нейронными сетями. Для его использования необходимо определить квадратичный функционал, построенный на невязках дифференциальных уравнений, а также граничных и начальных условиях.

В работе рассматривается метод нахождения решения как для непрерывной по времени модели, так и для дискретной. В первом случае рассматривается решение системы (7)–(12) в непрерывном времени. Во втором случае уравнения (7)–(12) заменяются дискретными аналогами и поиск решения происходит с использованием неявного метода Рунге–Кутты высокого порядка.

Отметим, что в работах [3–6] для приближения решения системы уравнений используется одна нейронная сеть. В данной работе для приближения каждой из неизвестных функций применяется

собственная нейронная сеть. Основная идея заключается в том, что у каждой из функций u , ρ , θ своя зависимость от пространственно-временных переменных. В отличие от работ [3, 6], где используется архитектура с большим количеством слоев и нейронов в них, в текущей работе используется несколько нейронных сетей малой архитектуры, которые будут описаны далее.

Для каждой нейронной сети определяется свой собственный функционал, включающий невязку дифференциальных уравнений всей системы, а также граничные и начальные условия только для приближаемой функции.

Отметим, что решение задачи выполнено в безразмерных переменных.

2.1. Описание метода для непрерывной модели

Для использования метода в случае непрерывной модели определим функционалы качества J_u , J_ρ , J_θ для неизвестных u , ρ , θ , соответствующие системе (7)–(12):

$$J_u = \frac{W_r}{N_r} \sum_{i=1}^{N_r} \left[\left[r_1 \left(\bar{x}_i^r \right) \right]^2 + \left[r_2 \left(\bar{x}_i^r \right) \right]^2 + \left[r_3 \left(\bar{x}_i^r \right) \right]^2 \right] + \frac{W_0}{N_0} \sum_{i=1}^{N_0} \left[u_p \left(\bar{x}_i^0 \right) - u_0 \left(\bar{x}_i^0 \right) \right]^2 + \frac{W_b}{N_b} \sum_{i=1}^{N_b} \left[u_p \left(\bar{x}_i^1 \right) - u_1 \left(\bar{x}_i^1 \right) \right]^2 + \frac{W_b}{N_b} \sum_{i=1}^{N_b} \left[u_p \left(\bar{x}_i^2 \right) - u_2 \left(\bar{x}_i^2 \right) \right]^2, \quad (13)$$

$$J_\rho = \frac{W_r}{N_r} \sum_{i=1}^{N_r} \left[\left[r_1 \left(\bar{x}_i^r \right) \right]^2 + \left[r_2 \left(\bar{x}_i^r \right) \right]^2 + \left[r_3 \left(\bar{x}_i^r \right) \right]^2 \right] + \frac{W_0}{N_0} \sum_{i=1}^{N_0} \left[\rho_p \left(\bar{x}_i^0 \right) - \rho_0 \left(\bar{x}_i^0 \right) \right]^2 + \frac{W_b}{N_b} \sum_{i=1}^{N_b} \left[\rho_p \left(\bar{x}_i^1 \right) - \rho_1 \left(\bar{x}_i^1 \right) \right]^2, \quad (14)$$

$$J_\theta = \frac{W_r}{N_r} \sum_{i=1}^{N_r} \left[\left[r_1 \left(\bar{x}_i^r \right) \right]^2 + \left[r_2 \left(\bar{x}_i^r \right) \right]^2 + \left[r_3 \left(\bar{x}_i^r \right) \right]^2 \right] + \frac{W_0}{N_0} \sum_{i=1}^{N_0} \left[\theta_p \left(\bar{x}_i^0 \right) - \theta_0 \left(\bar{x}_i^0 \right) \right]^2 + \frac{W_b}{N_b} \sum_{i=1}^{N_b} \left[\theta_p \left(\bar{x}_i^1 \right) - \theta_1 \left(\bar{x}_i^1 \right) \right]^2 + \frac{W_b}{N_b} \sum_{i=1}^{N_b} \left[\theta_p \left(\bar{x}_i^2 \right) - \theta_2 \left(\bar{x}_i^2 \right) \right]^2. \quad (15)$$

Здесь u_p , ρ_p , θ_p – приближения неизвестных функций нейронными сетями, r_1 , r_2 , r_3 – невязки уравнений (1)–(3), N_r , N_0 , N_b – общее количество точек для невязки уравнений (7)–(9), начального, а также граничных условий соответственно, $\bar{x}_i^r = (x_i^r, t_i^r)$, $\bar{x}_i^0 = (x_i^0, t_i^0)$, $\bar{x}_i^1 = (x_i^1, t_i^1)$, $\bar{x}_i^2 = (x_i^2, t_i^2)$ – наборы данных, соответствующие области невязки уравнений (7)–(9), области в момент времени $t = 0$, а также левой и правой границам соответственно, W_r , W_0 , W_b – весовые коэффициенты для слагаемых невязки, начальных и граничных условий соответственно.

В данной работе при обучении нейронных сетей используются следующие коэффициенты:

$$W_r = 1, \quad W_0 = 10, \quad W_b = 10.$$

Их использование необходимо для того, чтобы слагаемые для граничных и начальных условий в (13)–(15) были на порядок больше слагаемого невязки. В противном случае при обучении нейронная сеть будет принимать неизвестные функции константами и решение задачи станет невозможным. Тем не менее для оценки точности решения в работе используется функционал J , который считается как сумма трех функционалов (13)–(15) с единичными весовыми коэффициентами:

$$J = J_u |_{W_r=W_0=W_b=1} + J_\rho |_{W_r=W_0=W_b=1} + J_\theta |_{W_r=W_0=W_b=1}. \quad (16)$$

Таким образом, решение прямой задачи заключается в решении задач минимизации функционалов:

$$J_\theta \rightarrow \min, \quad J_\rho \rightarrow \min, \quad J_u \rightarrow \min. \quad (17)$$

2.2. Описание метода для дискретной модели

Для модели с дискретизацией по времени будем использовать неявный метод Рунге–Кутты высокого порядка для решения. Впервые методы Рунге–Кутты в методе PINN были использованы в работе [3], где было показано, что их использование для решения дискретизированных по времени задач возможно и позволяет получить решение с малой погрешностью метода, снизив при этом размерность пространства задачи.

Неявный метод Рунге–Кутты порядка s для функции двух переменных $y(x, t)$ можно описать следующим образом:

$$y_{n+1} = y_n + \tau \sum_{j=1}^s b_j k_j, \quad (18)$$

$$k_i = f(x, t_n + c_i \tau, y_n + \tau \sum_{j=1}^s a_{ij} k_j), \quad i = 1, \dots, s, \quad (19)$$

где $y_n = y(x, t_n)$, $y_{n+1} = y(x, t_n + h)$, τ – заданный шаг.

Воспользуемся дискретизацией неизвестных функций при помощи конечной разности вперед:

$$\frac{\partial u}{\partial t} = \frac{u_{n+1} - u_n}{\tau}, \quad \frac{\partial \rho}{\partial t} = \frac{\rho_{n+1} - \rho_n}{\tau}, \quad \frac{\partial \theta}{\partial t} = \frac{\theta_{n+1} - \theta_n}{\tau}.$$

Для неизвестных данных задачи (7)–(12) метод Рунге–Кутты примет вид:

$$u_{n+1} = u_n + \tau \sum_{j=1}^s b_j k_j^u, \quad k_i^u = f^u(x, t_n + c_i \tau, u_n + \tau \sum_{j=1}^s a_{ij} k_j^u, \rho_n + \tau \sum_{j=1}^s a_{ij} k_j^\rho, \theta_n + \tau \sum_{j=1}^s a_{ij} k_j^\theta), \quad (20)$$

$$\rho_{n+1} = \rho_n + \tau \sum_{j=1}^s b_j k_j^\rho, \quad k_i^\rho = f^\rho(x, t_n + c_i \tau, u_n + \tau \sum_{j=1}^s a_{ij} k_j^u, \rho_n + \tau \sum_{j=1}^s a_{ij} k_j^\rho), \quad (21)$$

$$\theta_{n+1} = \theta_n + \tau \sum_{j=1}^s b_j k_j^\theta, \quad k_i^\theta = f^\theta(x, t_n + c_i \tau, u_n + \tau \sum_{j=1}^s a_{ij} k_j^u, \rho_n + \tau \sum_{j=1}^s a_{ij} k_j^\rho, \theta_n + \tau \sum_{j=1}^s a_{ij} k_j^\theta), \quad (22)$$

где f^u, f^ρ, f^θ , как следует из (7)–(12):

$$f^u(x, t, u, \rho, \theta) = -Sh u \frac{\partial u}{\partial x} + \frac{Sh}{\rho Re} \frac{\partial^2 u}{\partial x^2} - \frac{Sh k}{\rho} \frac{\partial}{\partial x}(\rho \theta),$$

$$f^\rho(x, t, u, \rho) = -Sh \frac{\partial}{\partial x}(u \rho),$$

$$f^\theta(x, t, u, \rho, \theta) = -Sh u \frac{\partial \theta}{\partial x} + \frac{1}{\rho} \frac{Sh}{Pe} \frac{\partial^2 \theta}{\partial x^2} + \frac{1}{\rho} \frac{Sh}{Re k} \left(\frac{\partial u}{\partial x} \right)^2 - Sh \pi \theta \frac{\partial u}{\partial x}.$$

Поскольку начальные значения u_n, ρ_n, θ_n считаются заданными, удобно выразить формулы (20)–(22) через них. Тогда функционалы качества можно выразить как

$$\begin{aligned} J_u^{rk} &= W_r \sum_{i=1}^{N_0} \left[\left[u_p \left(\begin{smallmatrix} -0 \\ x_i \end{smallmatrix} \right) - u_n \left(\begin{smallmatrix} -0 \\ x_i \end{smallmatrix} \right) \right]^2 + \left[\rho_p \left(\begin{smallmatrix} -0 \\ x_i \end{smallmatrix} \right) - \rho_n \left(\begin{smallmatrix} -0 \\ x_i \end{smallmatrix} \right) \right]^2 + \left[\theta_p \left(\begin{smallmatrix} -0 \\ x_i \end{smallmatrix} \right) - \theta_n \left(\begin{smallmatrix} -0 \\ x_i \end{smallmatrix} \right) \right]^2 \right] + \\ &\quad + W_b \sum_{j=1}^{s+1} \sum_{i=1}^{N_b} \left[u_p \left(\begin{smallmatrix} -1 \\ x_{ij} \end{smallmatrix} \right) - u_1 \left(\begin{smallmatrix} -1 \\ x_{ij} \end{smallmatrix} \right) \right]^2 + W_b \sum_{j=1}^{s+1} \sum_{i=1}^{N_b} \left[u_p \left(\begin{smallmatrix} -2 \\ x_{ij} \end{smallmatrix} \right) - u_2 \left(\begin{smallmatrix} -2 \\ x_{ij} \end{smallmatrix} \right) \right]^2, \\ J_\rho^{rk} &= W_r \sum_{i=1}^{N_0} \left[\left[u_p \left(\begin{smallmatrix} -0 \\ x_i \end{smallmatrix} \right) - u_n \left(\begin{smallmatrix} -0 \\ x_i \end{smallmatrix} \right) \right]^2 + \left[\rho_p \left(\begin{smallmatrix} -0 \\ x_i \end{smallmatrix} \right) - \rho_n \left(\begin{smallmatrix} -0 \\ x_i \end{smallmatrix} \right) \right]^2 + \left[\theta_p \left(\begin{smallmatrix} -0 \\ x_i \end{smallmatrix} \right) - \theta_n \left(\begin{smallmatrix} -0 \\ x_i \end{smallmatrix} \right) \right]^2 \right] + \\ &\quad + W_b \sum_{j=1}^{s+1} \sum_{i=1}^{N_1} \left[\rho_p \left(\begin{smallmatrix} -1 \\ x_{ij} \end{smallmatrix} \right) - \rho_1 \left(\begin{smallmatrix} -1 \\ x_{ij} \end{smallmatrix} \right) \right]^2, \\ J_\theta^{rk} &= W_r \sum_{i=1}^{N_0} \left[\left[u_p \left(\begin{smallmatrix} -0 \\ x_i \end{smallmatrix} \right) - u_n \left(\begin{smallmatrix} -0 \\ x_i \end{smallmatrix} \right) \right]^2 + \left[\rho_p \left(\begin{smallmatrix} -0 \\ x_i \end{smallmatrix} \right) - \rho_n \left(\begin{smallmatrix} -0 \\ x_i \end{smallmatrix} \right) \right]^2 + \left[\theta_p \left(\begin{smallmatrix} -0 \\ x_i \end{smallmatrix} \right) - \theta_n \left(\begin{smallmatrix} -0 \\ x_i \end{smallmatrix} \right) \right]^2 \right] + \end{aligned}$$

$$+ W_b \sum_{j=1}^{s+1} \sum_{i=1}^{N_1} 2 \left[\theta_p \left(\frac{-1}{x_{ij}} \right) - \theta_1 \left(\frac{-1}{x_{ij}} \right) \right] + W_b \sum_{j=1}^{s+1} \sum_{i=1}^{N_2} 2 \left[\theta_p \left(\frac{-2}{x_{ij}} \right) - \theta_2 \left(\frac{-2}{x_{ij}} \right) \right]^2,$$

где $\bar{x}_{ij}^0 = (x_i^r, 0)$, $\bar{x}_{ij}^{b_1} = (x_i^{b_1}, t_n + c_j \tau)$, $\bar{x}_{ij}^{b_2} = (x_i^{b_2}, t_n + c_j \tau)$.

В данном случае прямая задача сводится к решению задачи о минимизации функционалов:

$$J_\theta^{rk} \rightarrow \min, \quad J_\rho^{rk} \rightarrow \min, \quad J_u^{rk} \rightarrow \min. \quad (23)$$

Для оценки качества решения аналогично функционалу (16) будем использовать следующий функционал качества:

$$J^{rk} = J_u^{rk} |_{W_r=W_b=1} + J_\rho^{rk} |_{W_r=W_b=1} + J_\theta^{rk} |_{W_r=W_b=1}. \quad (24)$$

В качестве метода решения задачи с дискретизацией будем использовать метод Гаусса–Лежандра. Метод является неявным и А-устойчивым [7]. Для расчета коэффициентов c_i в (19) используются корни смещенного полинома Лежандра $P(2x - 1)$, а коэффициенты a_{ij} и b_j в (18) и (19) рассчитываются по следующим формулам:

$$a_{ij} = \int_0^{c_j} l_j(x) dx, \quad b_j = \int_0^1 l_j(x) dx,$$

где $l_j(x) = \prod_{j \neq i} \frac{x - x_j}{x_j - x_i}, i=1, \dots, s$.

Таким образом, можно вывести метод сколь угодно высокого порядка точности. В текущей работе используется метод Гаусса–Лежандра порядка $s = 100$. Решение задачи на нормированном по времени интервале $t \in [0, 1]$ будем осуществлять в два шага:

1. Решение задачи на интервале $t \in [0, 0,5]$. Начальные условия $u_0^s, \rho_0^s, \theta_0^s$ используются в качестве u_n, ρ_n, θ_n в (20)–(22). В результате решения задачи (23) с шагом $\tau = 0,5$ получим решение $u_{n+1}, \rho_{n+1}, \theta_{n+1}$ в момент времени $t = 0,5$.

2. Найденные на предыдущем шаге $u_{n+1}, \rho_{n+1}, \theta_{n+1}$ будем использовать как u_n, ρ_n, θ_n в (20) – (22). Таким образом, найдя решение с шагом $\tau=0.5$ еще раз, получим решение $u_{n+1}, \rho_{n+1}, \theta_{n+1}$ в момент времени $t = 1,0$.

Помимо этого, стоит отметить, что согласно (18)–(19) мы также получаем решение в ста промежуточных временных интервалах $t_n + c_i \tau$ как на первом шаге, так и на втором. Погрешность двух шагов в данном случае будет равна $2\tau^{2s} = 2\tau^{200} = 2 \cdot 0,5^{200} = 1,24 \cdot 10^{-60}$.

3. Генерация данных и обучение нейронных сетей

Для обучения нейронных сетей необходимо сгенерировать базу данных. В данном случае под базой данных подразумевается набор точек, в которых будут вычислены невязки уравнений (7)–(9), а также граничные и начальные условия (10)–(12).

В случае непрерывной модели будем воспринимать время как пространственное измерение. Тогда получим двумерную плоскость ХОТ, в которой необходимо сгенерировать обучающую выборку. Для расчета невязки уравнений было сгенерировано 100×100 точек в области $\Omega = \{(x, t): 0 < t \leq 1; 0 < x < 1\}$ при помощи случайной выборки из равномерного распределения. Для расчета граничных условий было аналогичным образом сгенерировано по 200 точек в начале процесса в области $\Omega_0 = \{(x, t): t = 0; 0 \leq x \leq 1\}$, на левой и правой границах, в $\Omega_1 = \{(x, t): 0 \leq t \leq 1; x = 0\}$ и $\Omega_2 = \{(x, t): 0 \leq t \leq 1; x = 1\}$ соответственно.

Для формирования переходных слоев в решении нейронной сетью выборка точек коллокации в начале обучения генерируется на некотором расстоянии от границ. По мере обучения сети отступ уменьшается путем деления пополам до тех пор, пока не станет равным шагу пространственной сети. На рис. 1 показана сгенерированная обучающая выборка на расстоянии шага пространственной сетки от границ.

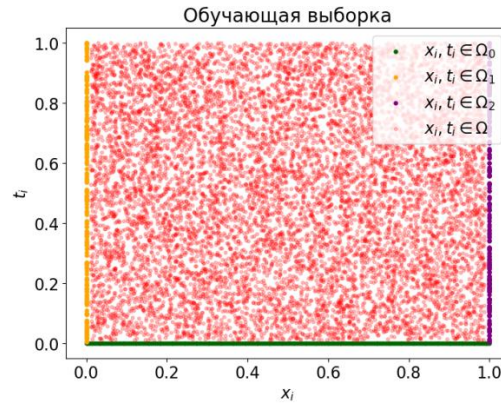


Рис. 1. Обучающая выборка для непрерывной модели

Fig. 1. Training dataset for continuous model

В случае дискретной модели необходимо сгенерировать 250 точек в пространстве $\Omega_0 = \{x: 0 < x < 1\}$, а также по одной точке на левой и правой границах, $x = 0$ и $x = 1$ соответственно. В данном случае также точки коллокации генерируются на некотором расстоянии от границ, после чего отступ уменьшается, пока не станет равен шагу пространственной сетки.

Для обучения нейронных сетей в двух случаях были выбраны архитектуры нейронных сетей, как показано в таблице, где $R_{tanh}(x) = (e^x - e^{-x})/(e^x + e^{-x})$, $R_{swish}(x) = x/(1 + e^{-x})$, $R_{linear}(x) = x$.

Архитектуры нейронных сетей

№ слоя	Описание слоя	Функция активации на выходе слоя	Число нейронов в слое	
			для непрерывного случая	для дискретного случая
1	Входной слой	—	2	1
2	Скрытый слой	R_{tanh}	64	256
3	Скрытый слой	R_{tanh}	32	128
4	Скрытый слой	R_{swish}	16	16
5	Выходной слой	R_{linear}	1	101

Использование функций активации, указанных в таблице, показало наилучший результат при решении задачи, что также соответствует данным работы [8].

Обучение нейронных сетей происходило на базе библиотеки Tensorflow [9] языка программирования Python. Для минимизации функций потерь использовался алгоритм Adam [10] с шагом обучения $lr = 2 \cdot 10^{-4}$. Все остальные параметры алгоритма являются стандартными для библиотеки Tensorflow. Нейронные сети заданной архитектуры последовательно обучаются на сгенерированных в данном разделе данных путем минимизации соответствующих функционалов: (17) для случая без дискретизации и (23) для случая с дискретизацией.

В зависимости от задачи возможно изменение как архитектуры нейронных сетей, показанной в таблице, так и размера обучающей выборки. При расчете более простых систем число скрытых слоев нейронной сети и нейронов в них для функций активации гиперболического тангенса может быть уменьшено, как и размер базы данных. При расчете более сложных систем, напротив, число скрытых слоев возможно увеличить. Тем не менее авторы полагают, что для многих задач архитектура, описанная в таблице, будет оптимальной. Чрезмерное увеличение сложности моделей может привести к переобучению нейронных сетей, а чрезмерное увеличение базы данных — к нехватке оперативной памяти при вычислениях. Оптимальной точностью при решении авторы считают значение функционалов $J = 10^{-4}$, $J^k = 10^{-4}$.

4. Численный эксперимент

В рамках текущей работы было проведено несколько численных экспериментов с различными параметрами, отвечающими физическим параметрам газов. В разделе будет представлено по одному эксперименту, соответствующему разделам 2.1, 2.2.

4.1. Эксперимент для непрерывной модели

Первый вычислительный эксперимент, для случая непрерывной модели, был решен с использованием метода, описанного в разделе 2.1. Для него были выбраны характеристики, отвечающие параметрам углекислого газа, а также следующие граничные и начальные условия:

$$L = 100[\text{м}], \quad T = 3600[\text{с}], \quad u_s = 20\left[\frac{\text{м}}{\text{с}}\right], \quad \rho_s = 2\left[\frac{\text{кг}}{\text{м}^3}\right], \quad \theta_s = 323[\text{К}], \quad (25)$$

$$\kappa = 0,017\left[\frac{\text{Вт}}{\text{м} \cdot \text{К}}\right], \quad \mu = 5 \cdot 10^{-5}[\text{Па} \cdot \text{с}], \quad c_v = 700\left[\frac{\text{Дж}}{\text{кг} \cdot \text{К}}\right], \quad R = 8,314\left[\frac{\text{Дж}}{\text{моль} \cdot \text{К}}\right], \quad (26)$$

$$u_0^s(x) = 0,5 + 0,1 \cdot x^2, \quad u_1^s(x) = 0,5 + 0,1 \cdot t, \quad u_2^s(x) = 0,6 + 0,1 \cdot t, \quad (27)$$

$$\rho_0^s(x) = 0,9 - 0,1 \cdot x^2, \quad \rho_1^s(x) = 0,9 - 0,1 \cdot t, \quad (28)$$

$$\theta_0^s(x) = 1,0 - 0,1 \cdot x^2, \quad \theta_1^s(x) = 1,0 - 0,1 \cdot t, \quad \theta_2^s(x) = 0,9 - 0,1 \cdot t. \quad (29)$$

После обучения нейронных сетей было получено решение, представленное на рис. 2, *a*, *b*, *c*, *d* в моменты времени $t = 0$ с, $t = 60$ с, $t = 1\,200$ с и $t = 3\,600$ с соответственно.

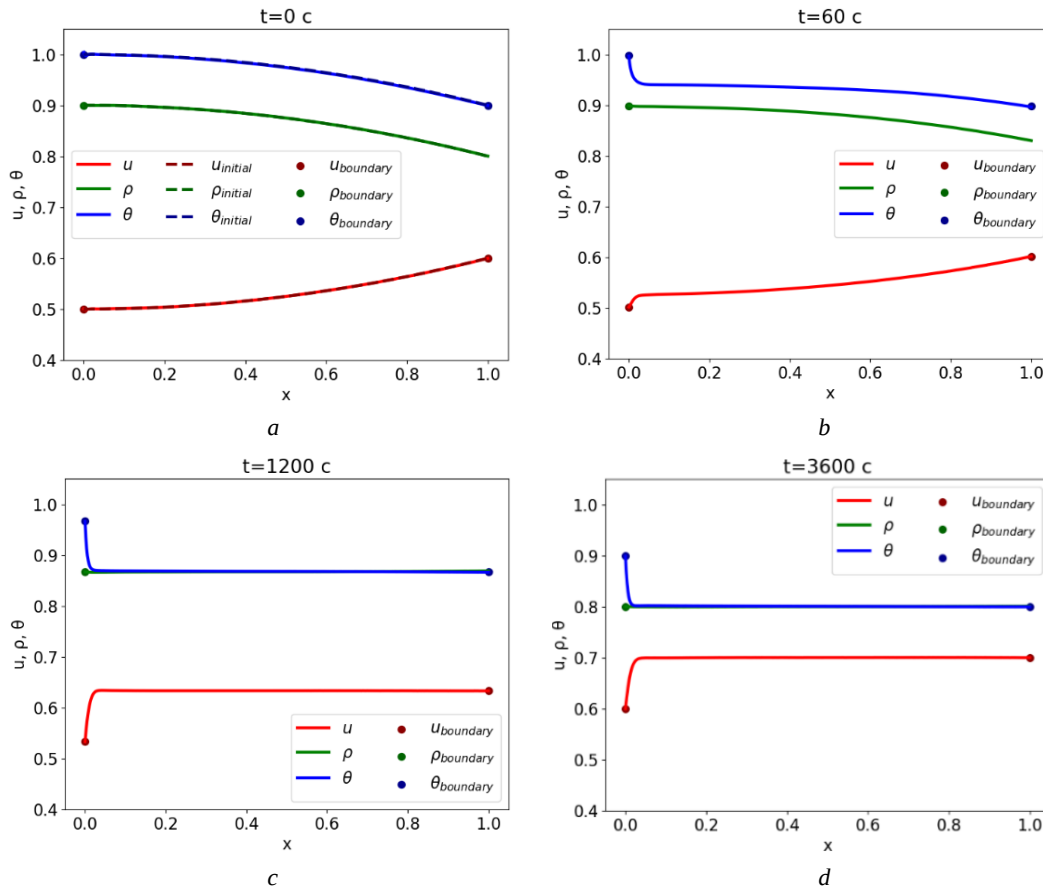


Рис. 2. Решение задачи (17) с параметрами (25)–(29). Непрерывной линией указано приближение функции нейронной сетью. Точки на границах обозначают граничные условия

Fig. 2. Solution of problem (17) with parameters (25)–(29). The continuous line indicates the approximation of the function by a neural network. Dots on the boundaries indicate boundary conditions

Из рис. 2 видно, что полученная функция полностью соответствует начальным и граничным условиям. На левой границе также были сформированы переходные слои, что обусловливается малыми коэффициентами при вторых производных уравнений (7), (9). Низкое значение функционала (16) $J = 2 \cdot 10^{-5}$, полученное при решении, говорит о соответствии полученного приближения функций задаче (7)–(12). На расчет в численном эксперименте было затрачено 1,37 часа.

4.2. Эксперимент для дискретной модели

Второй вычислительный эксперимент, для случая с дискретизацией по времени, был решен с использованием метода, описанного в разделе 2.2. Для него были выбраны физические параметры среды, соответствующие газу азоту, а также следующие начальные и граничные условия:

$$L = 50[\text{м}], \quad T = 1[\text{с}], \quad u_s = 10\left[\frac{\text{м}}{\text{с}}\right], \quad \rho_s = 1,35\left[\frac{\text{кг}}{\text{м}^3}\right], \quad \theta_s = 333[\text{К}], \quad (30)$$

$$\kappa = 0,022\left[\frac{\text{Вт}}{\text{м} \cdot \text{К}}\right], \quad \mu = 1,665 \cdot 10^{-5}[\text{Па} \cdot \text{с}], \quad c_v = 1050\left[\frac{\text{Дж}}{\text{кг} \cdot \text{К}}\right], \quad R = 8,314\left[\frac{\text{Дж}}{\text{моль} \cdot \text{К}}\right], \quad (31)$$

$$u_0^s(x) = 0,65 + 0,1 \cdot x^2, \quad u_1^s(x) = 0,65 + 0,1 \cdot t, \quad u_2^s(x) = 0,75 + 0,1 \cdot t, \quad (32)$$

$$\rho_0^s(x) = 0,9 - 0,1 \cdot x^2, \quad \rho_1^s(x) = 0,9 - 0,1 \cdot t, \quad (33)$$

$$\theta_0^s(x) = 1,0 - 0,1 \cdot x^2, \quad \theta_1^s(x) = 1,0 - 0,1 \cdot t, \quad \theta_2^s(x) = 0,9 - 0,1 \cdot t. \quad (34)$$

На первом шаге, согласно методу, решение было получено в моменты времени $t = 0$ с, $t = 0,5$ с, а также во всех ста моментах времени, соответствующих коэффициентам c_j в таблице Бутчера для метода Гаусса–Лежандра порядка $s = 100$. На втором шаге было получено решение в моментах времени $t = 0,5$, $t = 1,0$ и ста промежуточных моментах времени, соответствующих коэффициентам c_j . Полученные решения в моменты времени $t = 0$ с, $t = 0,5$ с, $t = 0,677$ с и $t = 1,0$ с представлены на рисунках 3, *a*, *b*, *c*, *d* соответственно.

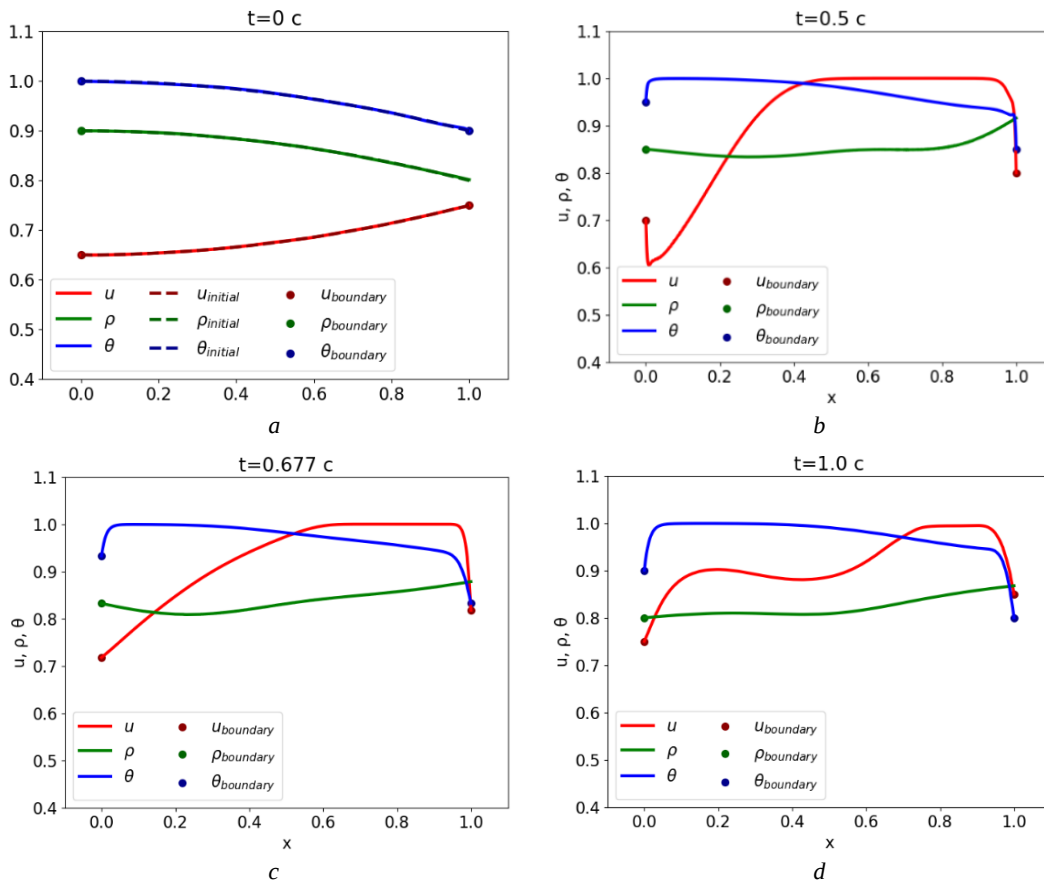


Рис. 3. Решение задачи (23) с параметрами (30)–(34). Непрерывной линией указано приближение функции нейронной сетью. Точки на границах обозначают граничные условия

Fig. 3. Solution of problem (23) with parameters (30)–(34). The continuous line indicates the approximation of the function by a neural network. Dots on the boundaries indicate boundary conditions

В ходе решения функционал (28) достиг значения $J^k = 1 \cdot 10^{-4}$. На решение было потрачено 2 часа. Как следует из рис. 3, приближения, полученные нейронными сетями, также соответствуют гранич-

ным и начальным условиям. Помимо этого, в решении формируются зоны резкого изменения функций, расположенные возле границ, что обуславливается малыми коэффициентами при вторых производных в уравнении движения (7) и уравнении теплопроводности (9).

Отметим, что на рис. 2, а и 3, а линии начальных условий совпадают с графиками процессов для скорости, плотности и температуры, найденными с помощью нейронной сети в начальный момент времени, что характеризует высокую точность решения задачи.

4.3. Сравнение методов

При сравнении двух методов, описанных в разделах 2.1 и 2.2, было замечено, что при решении одномерных нестационарных задач метод для непрерывной модели показал себя более эффективно с точки зрения времени вычислений. Также было установлено, что для модели с дискретизацией по времени приближение неизвестной функции к начальному условию второго шага может быть затруднительно, если в решении, полученном в конечный момент времени на первом шаге, присутствуют зоны быстрого роста функции. Отметим, что применение метода, описанного в разделе 2.2, является затруднительным при $Sh \gg 1$. Авторы полагают, что причиной этого может являться алгоритм прямого накопления для обучения нейронных сетей. В то же время для непрерывной модели нейронные сети обучаются при помощи метода обратного распространения ошибки.

Тем не менее применение метода, описанного в разделе 2.2, позволяет снизить размерность задачи, что является важным критерием при расчете задач в двумерном и трехмерном случаях, когда необходимо генерировать огромную обучающую выборку, следствием чего могут стать нехватка оперативной памяти и невозможность применения алгоритма.

Стоит отметить, что описанный в разделе 2 подход с приближением каждой неизвестной отдельной нейронной сетью не является громоздким с вычислительной точки зрения при увеличении размерности пространства в задаче, поскольку компоненты неизвестных функций обладают похожей зависимостью и могут быть приближены одной нейронной сетью.

В дальнейшем планируется качественный анализ системы (7)–(12).

Заключение

В данной работе был проведен анализ численного решения системы уравнений Навье–Стокса в случае сжимаемой среды при помощи нейронных сетей. Исследована возможность применения метода PINN с использованием без дискретизации по времени и с ней. Во втором случае для решения системы с дискретизацией был использован метод Рунге–Кутты высокого порядка. Показано, что использование данных методов является численно эффективным и требует сравнительно небольших вычислительных ресурсов. Помимо этого, использование PINN не требует линеаризации и позволяет решать задачу с наличием переходных слоев в решении на равномерной сетке.

В работе представлен эффективный вычислительный алгоритм для решения задач с наличием зон быстрого изменения функции. Кроме того, авторы надеются, что используемый подход обладает потенциалом при решении задач большей размерности и сложности. Применение временной дискретизации позволяет снизить требования к вычислительным ресурсам, что актуально для двумерных и трехмерных задач.

Список источников

1. Кажихов А.В. Избранные труды. Математическая гидродинамика. Новосибирск : Изд-во Ин-та гидродинамики, 2008. 419 с.
2. Рождественский Б.Л., Яненко Н.Н. Система квазилинейных уравнений и их приложения к газовой динамике. М. : Наука, 1978. 688 с.
3. Raissi M., Perdikaris P., Karniadakis G.E. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations // Journal of Computational Physics. 2019. V. 50 (3). P. 686–707. doi: 10.1016/j.jcp.2018.10.045
4. Almajid M., Abu-Elseid M. Prediction of porous media fluid flow using physics informed neural networks // Journal of Petroleum Science and Engineering. 2021. V. 208. Art. 109205. doi: 10.1016/j.petrol.2021.109205

5. Eivazi H., Tahani M., Schlatter P., Vinuesa R. Physics-informed neural networks for solving Reynolds-averaged Navier–Stokes equations // *Physics of Fluids*. 2022. V. 34 (7). Art. 075117. doi: 10.1063/5.0095270
6. Jin X., Cai S., Li H., Karniadakis G.E. NSFnets (Navier-Stokes flow nets): Physics-informed neural networks for the incompressible Navier-Stokes equations // *Journal of Computational Physics*. 2021. V. 426. Art. 109951. doi: 10.1016/j.jcp.2020.109951
7. Butcher J.C. Practical Runge-Kutta methods for scientific computation // *The ANZIAM Journal*. 2009. V. 50 (3). P. 333–342.
8. Ramachandran P., Zoph B., Le Q.V. Searching for activation functions // arXiv. 2021.
9. Abadi M. et al. TensorFlow: a system for large-scale machine learning // 12th USENIX Symposium on Operating Systems Design and Implementation Proc. 2016. P. 1–21.
10. Kingma D.P., Ba J. Adam: a method for stochastic optimization // arXiv. 2014.

References

1. Kazhikhov, A.V. (2008) *Izbrannye trudy. Matematicheskaya gidrodinamika [Selected Works. Mathematical Hydrodynamics]*. Novosibirsk: Institute of Hydrodynamics.
2. Rozhdestvensky, B.L. & Yanenko, N.N. (1978) *Sistema kvazilineynykh uravneniy i ikh prilozheniya k gazovoy dinamike [The system of quasilinear equations and their applications to gas dynamics]*. Moscow: Nauka.
3. Raissi, M., Perdikaris, P. & Karniadakis, G.E. (2019) Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*. 50(3). pp. 686–707. DOI: 10.1016/j.jcp.2018.10.045
4. Almajid, M. & Abu-Elseid, M. (2022) Prediction of porous media fluid flow using physics informed neural networks. *Journal of Petroleum Science and Engineering*. 208. Art. 109205. DOI: 10.1016/j.petrol.2021.109205
5. Eivazi, H., Tahani, M., Schlatter, P. & Vinuesa, R. (2022) Physics-informed neural networks for solving Reynolds-averaged Navier–Stokes equations. *Physics of Fluids*. 34(7). DOI: 10.1063/5.0095270
6. Jin, X., Cai, S., Li, H. & Karniadakis, G.E. (2021) NSFnets (Navier-Stokes flow nets): Physics-informed neural networks for the incompressible Navier-Stokes equations. *Journal of Computational Physics*. 426. Art. 109951. DOI: 10.1016/j.jcp.2020.109951
7. Butcher, J.C. (2009) Practical Runge-Kutta methods for scientific computation. *The ANZIAM Journal*. 50(3). pp. 333–342.
8. Ramachandran, P., Zoph, B. & Le, Q.V. (2021) *Searching for activation functions*. arXiv.
9. Abadi, M. et al. (2016) TensorFlow: A system for large-scale machine learning. *12th USENIX Symposium on Operating Systems Design and Implementation Proceedings*. pp. 1–21.
10. Kingma, D.P. & Ba, J. (2014) *Adam: A method for stochastic optimization*. arXiv.

Информация об авторах:

Кузнецов Кирилл Сергеевич – аспирант Института математических и компьютерных технологий Дальневосточного федерального университета; ассистент департамента математического и компьютерного моделирования Дальневосточного федерального университета (Владивосток, Россия); младший научный сотрудник Института прикладной математики ДВО РАН (Владивосток, Россия). E-mail: kuznetsovks17@gmail.com

Амосова Елена Владимировна – доцент, кандидат физико-математических наук, профессор кафедры математического и компьютерного моделирования Дальневосточного федерального университета (Владивосток, Россия); научный сотрудник Института прикладной математики ДВО РАН (Владивосток, Россия). E-mail: el_amosova@mail.ru

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

Information about the authors:

Kuznetsov Kirill S. (Post-Graduate Student of the Institute of Mathematical and Computer Technologies of the Far Eastern Federal University; Assistant at the Department of Mathematical and Computer Modeling of the Far Eastern Federal University, Vladivostok, Russian Federation; Junior Research Assistant at the Institute of Applied Mathematics, Far Eastern Branch of the Russian Academy of Sciences, Vladivostok, Russian Federation). E-mail: kuznetsovks17@gmail.com

Amosova Elena V. (Associate Professor, Candidate of Physical and Mathematical Sciences, Professor of the Department of Mathematical and Computer Modeling of the Far Eastern Federal University, Vladivostok, Russian Federation; Researcher at the Institute of Applied Mathematics, Far Eastern Branch of the Russian Academy of Sciences, Vladivostok, Russian Federation). E-mail: el_amosova@mail.ru

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

Поступила в редакцию 26.09.2023; принята к публикации 03.06.2024

Received 26.09.2023; accepted for publication 03.06.2024

Научная статья

УДК 519.217.2

doi: 10.17223/19988605/67/5

Стационарное распределение в системах массового обслуживания с переменной структурой

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. Для построения математической модели системы массового обслуживания с переменной структурой в статье используется конструкция дискретных эргодических марковских процессов с непрерывным временем. Такая модель представляет собой объединение нескольких моделей систем обслуживания. Множество состояний объединенной системы является объединением множеств состояний объединяемых систем в отличие от классических систем обслуживания, в которых множество состояний системы является прямым произведением множеств состояний объединяемых систем. Переходные интенсивности между состояниями различных систем выбираются так, чтобы предельное распределение объединенной системы было смесью предельных распределений объединяемых систем с различными весами, определяемыми вводимыми переходными интенсивностями. В результате процесс функционирования объединенной системы получается переключением процессов, соответствующих различным объединяемым системам, в определенные моменты времени.

Ключевые слова: системы массового обслуживания; переменная структура; переходные интенсивности; смесь предельных распределений.

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

Для цитирования: Цициашвили Г.Ш., Харченко Ю.А. Стационарное распределение в системах массового обслуживания с переменной структурой // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 67. С. 42–48. doi: 10.17223/19988605/67/5

Original article

doi: 10.17223/19988605/67/5

Stationary distribution in queuing systems with a variable structure

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 construction of discrete ergodic Markov processes with continuous time is using to construct a mathematical model of a queuing system with a variable structure. This model is a combination of several models of queuing systems. The set of states of the combined system is a combination of sets of states of the combined systems, unlike classical queuing systems, in which the set of states of the system is a direct product of the sets of states of the combined systems. The transition intensities between the states of different systems are choosing so that the limit distribution of the combined system is a mixture of the limit distributions of the combined systems with different weights determined by the introduced transition intensities. As a result, the process of functioning of the combined system is obtaining by switching intensities corresponding to different combined systems at certain points in time.

Keywords: queuing systems; variable structure; switching intensities; mixtures of limit distributions.

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) Stationary distribution in queuing systems with a variable structure. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 67. pp. 42–48. doi: 10.17223/19988605/67/5

Введение

В последнее время в связи с развитием теоретико-информационного подхода к моделированию процессов самоорганизации в открытых физических системах со стохастичностью появилась возможность объединения динамической и информационной (управляющей) частей их описания в единую систему, когда в условиях конкуренции информационных и физических процессов эволюционное развитие термодинамической системы в значительной мере начинает определяться и ее информационными свойствами, включая информационное отношение к внешнему окружению (см., напр.: [1–4]).

В частности, при достижении сверхкритических значений управляющих параметров («информаторов», по терминологии Г. Хакена [5]) в самоорганизующихся открытых системах происходят вынужденные переходы между стационарными состояниями, при которых возникают упорядоченные пространственно-временные макроскопические образования (диссипативные структуры), сохраняющиеся только при наличии подкачки энергии, вещества и тому подобного из окружающей среды.

В мировой литературе хорошо представлены традиционные термодинамические и статистические аспекты спонтанного возникновения когерентных структур, т.е. процессов самоорганизации, при необратимых процессах в нелинейных термодинамических системах. Диссипация играет при образовании макроскопических структур конструктивную роль при микро- и макроуровневых методах описания (см., напр.: [6, 7]).

Различающая информация не только дает возможность адекватно описывать процессы хаотизации и самоорганизации в классических термодинамических системах (см., напр.: [8]), но и позволяет исследовать вопрос о взаимодополняемости термодинамических и информационных потоков, вводимых в рассмотрение при моделировании спонтанных и вынужденных переходов между стационарными состояниями открытых сплошных сред, находящихся вдали от термодинамического равновесия. Аналогичные исследования проводились в рамках исследования экономического равновесия в пространственных системах [9].

Следует отметить, что наряду с термодинамическими системами близкими по постановке являются системы управления с меняющейся структурой и скользящими режимами [10–13], модели газовой динамики с распадом разрыва [14]. В теории массового обслуживания, по-видимому, наиболее близкой к рассматриваемой в статье модели является модель альтернирующего входного потока с последовательно меняющимися ON- и OFF-периодами [15]. Общим для этих систем (моделей) является переключение от одной системы (управления, движения газа, периода во входном потоке) к другой.

Данная идея положена в основу построения модели массового обслуживания с переменной структурой, рассмотренной в данной работе. Поэтому возникает необходимость попытаться распространить приведенный синергетический подход на системы массового обслуживания. Для построения математической модели СМО с внешним управлением используется конструкция дискретных эргодических марковских процессов с непрерывным временем, а сама модель представляет собой объединение нескольких СМО, причем множество состояний объединенной системы является объединением множеств состояний объединяемых систем. В этом состоит отличие от классических СМО, в которых множество состояний системы является прямым произведением множеств состояний объединяемых систем [16, 17].

Однако именно конструкция с объединением множеств состояний дает возможность ввести управляющие параметры («информаторы»). Таковыми являются переходные интенсивности между некоторыми состояниями, принадлежащими множествам состояний разных СМО, входящих в объединенную систему. При этом на эти интенсивности накладываются условия локального равновесия.

При исследовании свойств построенной СМО используются теоремы эргодичности для дискретных марковских процессов и стационарные уравнения Колмогорова–Чепмена. Это предполагает, что функционирование объединенной СМО определяется не протоколом перемещения заявок в системе, а интенсивностями переходов между состояниями объединенной системы и весами объединяемых систем. В результате получается достаточно общая конструкция СМО с внешним управлением, которая позволяет создавать новые модели СМО.

Основные результаты

В монографии [18] доказаны следующие утверждения. Назовем однородный марковский дискретный процесс $X(t)$ с множеством состояний X и переходными интенсивностями $\lambda(x, y)$, $x, y \in X$, неприводимым, если существует такое состояние $x^* \in X$, которое достижимо из любого другого состояния. То есть для любого $x \neq x^*$ либо $\lambda(x, x^*) > 0$, либо найдется такая цепочка состояний $x_1, \dots, x_n$, что $\lambda(x, x_1) > 0, \lambda(x_1, x_2) > 0, \dots, \lambda(x_{n-1}, x_n) > 0, \lambda(x_n, x^*) > 0$.

Теорема 1. *Неприводимый однородный марковский процесс тогда и только тогда является эргодическим, когда система уравнений*

$$\sum_{y \neq x} u(y) \lambda(x, y) = \sum_{y \neq x} u(y) \lambda(y, x), \quad x \in X, \quad (1)$$

имеет хотя бы одно решение $u(x)$, $x \in X$, для которого сумма $\sum_{x \in X} |u(x)| < \infty$.

Рассмотрим сначала два дискретных, однородных и неприводимых (а значит, эргодических) марковских процесса $X_k(t)$, $t \geq 0$, $k = 1, 2$, с множествами состояний X_k , переходными интенсивностями $\lambda_k(x_k, y_k) \geq 0$ и стационарными вероятностями $P_k(x_k)$, $x_k, y_k \in X_k$. Из стационарных уравнений Колмогорова–Чепмена следуют равенства

$$\sum_{y_k \in X_k} P_k(x_k) \lambda_k(x_k, y_k) = \sum_{y_k \in X_k} P_k(y_k) \lambda_k(y_k, x_k), \quad x_k \in X_k. \quad (2)$$

Теорема 2. *Предположим, что между некоторыми состояниями $x_1^* \in X_1$, $x_2^* \in X_2$ этих двух процессов введены переходные интенсивности $\Lambda(x_1^*, x_2^*) > 0$, $\Lambda(x_2^*, x_1^*) > 0$, удовлетворяющие при некоторых c_1, c_2 , $0 < c_1, c_2 < 1$, $c_1 + c_2 = 1$, условию локального равновесия*

$$c_1 P_1(x_1^*) \Lambda(x_1^*, x_2^*) = c_2 P_2(x_2^*) \Lambda(x_2^*, x_1^*). \quad (3)$$

Тогда процесс $X(t)$ с множеством состояний $X_1 \cup X_2$ и переходными интенсивностями

$$\lambda_1(x_1, y_1), \quad x_1, y_1 \in X_1; \quad \lambda_2(x_2, y_2), \quad x_2, y_2 \in X_2; \quad \Lambda(x_1^*, x_2^*), \quad \Lambda(x_2^*, x_1^*),$$

также является эргодическим, и его стационарные вероятности имеют вид:

$$\pi(x_1) = c_1 P_1(x_1), \quad x_1 \in X_1; \quad \pi(x_2) = c_2 P_2(x_2), \quad x_2 \in X_2, \quad (4)$$

и удовлетворяют стационарным уравнениям Колмогорова–Чепмена

$$\begin{aligned} \pi(x_1) \sum_{y_1 \in X_1} \lambda_1(x_1, y_1) &= \sum_{y_1 \in X_1} \pi(y_1) \lambda_1(y_1, x_1), \quad x_1 \in X_1, \quad x_1 \neq x_1^*, \\ \pi(x_1^*) \left(\sum_{y_1 \in X_1} \lambda_1(x_1^*, y_1) + \Lambda(x_1^*, x_2^*) \right) &= \sum_{y_1 \in X_1} \pi(y_1) \lambda_1(y_1, x_1^*) + \pi(x_2^*) \Lambda(x_2^*, x_1^*) \\ \pi(x_2) \sum_{y_2 \in X_2} \lambda_2(x_2, y_2) &= \sum_{y_2 \in X_2} \pi(y_2) \lambda_2(y_2, x_2), \quad x_2 \in X_2, \quad x_2 \neq x_2^*, \\ \pi(x_2^*) \left(\sum_{y_2 \in X_2} \lambda_2(x_2^*, y_2) + \Lambda(x_2^*, x_1^*) \right) &= \sum_{y_2 \in X_2} \pi(y_2) \lambda_2(y_2, x_2^*) + \pi(x_1^*) \Lambda(x_1^*, x_2^*). \end{aligned} \quad (5)$$

Доказательство. Действительно, из формул (2), (4) следуют равенства

$$\pi(x_k) \sum_{y_k \in X_k} \lambda_k(x_k, y_k) = \sum_{y_k \in X_k} \pi(y_k) \lambda_k(y_k, x_k), \quad k = 1, 2.$$

Поэтому для получения равенств (5) достаточно выполнения равенства

$$\pi(x_1^*)\Lambda(x_1^*, x_2^*) = \pi(x_2^*)\Lambda(x_2^*, x_1^*),$$

вытекающего из равенств (3), (4).

Следствие 1. Пусть имеются $m > 2$ дискретных, однородных и неприводимых (а значит, эргодических) марковских процесса $X_k(t)$, $t \geq 0$, с множествами состояний X_k , переходными интенсивностями $\lambda_k(x_k, y_k) \geq 0$ и стационарными вероятностями $P_k(x_k)$, $x_k, y_k \in X_k$, $k = 1, \dots, m$. Предположим, что при некоторых $c_1, \dots, c_m$, $0 < c_1, \dots, c_m < 1$, $\sum_{k=1}^m c_k = 1$ выполняются равенства

$$\pi(x_k) = c_k P_k(x_k), x_k \in X_k, k = 1, \dots, m, \quad (6)$$

$$\pi(x_1^*)\Lambda(x_1^*, x_2^*) = \pi(x_2^*)\Lambda(x_2^*, x_3^*), \pi(x_2^*)\Lambda(x_2^*, x_3^*) = \pi(x_3^*)\Lambda(x_3^*, x_4^*), \dots, \quad (7)$$

$$\pi(x_{m-1}^*)\Lambda(x_{m-1}^*, x_m^*) = \pi(x_m^*)\Lambda(x_m^*, x_1^*), \pi(x_m^*)\Lambda(x_m^*, x_1^*) = \pi(x_1^*)\Lambda(x_1^*, x_2^*).$$

Тогда дискретный марковский процесс $X(t)$ с множеством состояний $X = \bigcup_{k=1}^m X_k$ и переходными интенсивностями

$$\lambda_k(x_k, y_k), x_k, y_k \in X_k, k = 1, \dots, m, \Lambda(x_1^*, x_2^*), \Lambda(x_2^*, x_3^*), \dots, \Lambda(x_{m-1}^*, x_m^*), \Lambda(x_m^*, x_1^*),$$

является однородным, неприводимым (и значит, эргодическим), и его предельное распределение совпадает с $\pi(x)$, $x \in X$.

Замечание 1. Константы $c_1, \dots, c_m$, $0 < c_1, \dots, c_m < 1$, $\sum_{k=1}^m c_k = 1$ определяются только $m - 1$ переменными. Поэтому константы $\Lambda(x_1^*, x_2^*), \Lambda(x_2^*, x_3^*), \dots, \Lambda(x_{m-1}^*, x_m^*), \Lambda(x_m^*, x_1^*)$ удовлетворяют равенствам

$$\Lambda(x_2^*, x_3^*) = \frac{\pi(x_1^*)\Lambda(x_1^*, x_2^*)}{\pi(x_2^*)}, \Lambda(x_3^*, x_4^*) = \frac{\pi(x_2^*)\Lambda(x_2^*, x_3^*)}{\pi(x_3^*)}, \dots, \\ \Lambda(x_m^*, x_1^*) = \frac{\pi(x_{m-1}^*)\Lambda(x_{m-1}^*, x_m^*)}{\pi(x_m^*)}, \Lambda(x_1^*, x_2^*) = \frac{\pi(x_m^*)\Lambda(x_m^*, x_1^*)}{\pi(x_1^*)}.$$

Причем по заданным константам $c_1, \dots, c_m$ (всего их $m - 1$ независимых), по $\pi(x_1^*), \dots, \pi(x_m^*)$ и по $\Lambda(x_1^*, x_2^*)$ можно определить все остальные переходные интенсивности $\Lambda(x_2^*, x_3^*), \dots, \Lambda(x_{m-1}^*, x_m^*), \Lambda(x_m^*, x_1^*)$. Таким образом, эти переходные интенсивности определяются ровно m константами $c_1, \dots, c_{m-1}, \Lambda(x_1^*, x_2^*)$.

Замечание 2. Наиболее интересный случай в Следствии 1, когда выполняются равенства $c_k = 1/m$, $k = 1, \dots, m$. В этом случае равенства (7) преобразуются в следующие:

$$P_1(x_1^*)\Lambda(x_1^*, x_2^*) = P_2(x_2^*)\Lambda(x_2^*, x_3^*), P_2(x_2^*)\Lambda(x_2^*, x_3^*) = P_3(x_3^*)\Lambda(x_3^*, x_4^*), \dots, \\ P_{m-1}(x_{m-1}^*)\Lambda(x_{m-1}^*, x_m^*) = P_m(x_m^*)\Lambda(x_m^*, x_1^*), P_m(x_m^*)\Lambda(x_m^*, x_1^*) = P_1(x_1^*)\Lambda(x_1^*, x_2^*).$$

Пример. Предположим, что дискретные марковские процессы $X_1(t)$, $X_2(t)$, рассмотренные в Теореме 2, описывают число заявок в одноканальных системах массового обслуживания $M | M | 1 | \infty$ с пуассоновскими входными потоками с интенсивностями λ_1, λ_2 и с интенсивностями обслуживания μ_1, μ_2 , $\rho_1 = \lambda_1/\mu_1 < 1$, $\rho_2 = \lambda_2/\mu_2 < 1$. Тогда дискретные марковские процессы $X_1(t)$, $X_2(t)$ имеют предельные распределения

$$P_1(x_1 = k) = (1 - \rho_1)\rho_1^k, P_2(x_2 = l) = (1 - \rho_2)\rho_2^l, k, l = 0, 1, \dots$$

Выделим состояния $x_1^* = 0$, $x_2^* = 0$, выберем c_1, c_2 , $0 < c_1, c_2$, $c_1 + c_2 = 1$, и определим интенсивности переходов $\Lambda_1 = \Lambda(x_1^* = 0, x_2^* = 0)$, $\Lambda_2 = \Lambda(x_2^* = 0, x_1^* = 0)$ из условия (6): $c_1(1 - \rho_1)\Lambda_1 = c_2(1 - \rho_2)\Lambda_2$. Таким образом, построен процесс $X(t)$ с множеством состояний $X = X_1 \cup X_2$, положительные интенсивности переходов между состояниями определяются рис. 1.

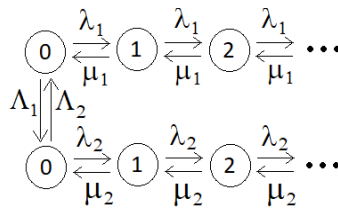


Рис. 1. Положительные интенсивности перехода между состояниями множества $X = X_1 \cup X_2$ дискретного марковского процесса $X(t)$

Fig. 1. Positive transition intensities between the states of the set $X = X_1 \cup X_2$ of a discrete Markov process $X(t)$

Вычисление предельного распределения дискретного марковского процесса с переходными интенсивностями, определяемыми на рис. 1, является достаточно простой задачей. Но гораздо важнее добиться того, чтобы сохранялись пропорции между предельными вероятностями в состояниях множества X_1 (в состояниях множества X_2).

Замечание 3. Нетрудно проверить, что в примере, если в начальный момент времени 0 первая система находится в состоянии $X(0) = x_1^* = 0$, то в любой момент $t^* > t_0$ перехода из состояния $x_1^* = 0$ в состояние $x_2^* = 0$ вторая система в последний момент своего функционирования перед моментом t^* будет находиться в состоянии $x_2^* = 0$. Иными словами, если $T^* = \sup(t : t < t^*, X(t) \in X_2)$, то $X(T^*) = x_2^* = 0$. Аналогично в любой момент t_* перехода из состояния $x_2^* = 0$ в состояние $x_1^* = 0$ первая система в последний момент своего функционирования перед моментом t_* будет находиться в состоянии $x_1^* = 0$.

Замечание 4. Количество таких примеров можно существенно расширить, используя известные формулы для предельных распределений в различных системах массового обслуживания, количество заявок в которых описывается процессами гибели и рождения [18]. Аналогично для этой цели можно также использовать случайные процессы, описывающие число заявок в узлах открытой [16] или замкнутой [17] сети массового обслуживания, предельные распределения в которых подчиняются мультипликативным теоремам.

Заключение

В настоящей работе строится объединение нескольких систем массового обслуживания путем введения дополнительных переходных интенсивностей между выделенными состояниями объединяемых систем. В результате такого построения предельное распределение объединенной системы становится смесью предельных распределений объединяемых систем с некоторыми весами. Эти веса определяются переходными интенсивностями, введенными между специально выделенными состояниями объединяемых систем. Функционирование так построенной объединенной системы содержит отрезки времени, когда эта объединенная система работает как одна из объединяемых систем обслуживания до очередного момента переключения на другую систему обслуживания. Такая конструкция позволяет строить объединение не только отдельных систем, но и отдельных сетей массового обслуживания (открытых или замкнутых). Использование предложенной модели предполагается распространить на построение вероятностной модели скользящего режима в теории управления.

Список источников

1. Зубарев Д.П., Морозов В.Г., Рёпке Г. Статистическая механика неравновесных процессов. М. : Физматлит, 2002. Т. 1. 431 с.
2. Климонтович Ю.Л. Статистическая теория открытых систем. М. : Янус, 1995. Т. 1. 624 с.
3. Зарипов Р.Г. Принципы неэкстенсивной статистической механики и геометрия мер беспорядка и порядка. New Казань : Изд-во Казан. гос. техн. ун-та, 2010. 404 с.
4. Колесниченко А.В. Континуальные модели природных и космических сред: проблемы термодинамического конструирования. М. : ЛЕНАНД, 2017. 400 с.

5. Хакен Г. Информация и самоорганизация: макроскопический подход к сложным системам. М. : УРСС : ЛЕНАНД, 2014. 320 с.
6. Пригожин И., Кондепуди Д. Современная термодинамика: от тепловых двигателей до диссипативных структур. М. : Мир, 2002. 461 с.
7. Marov M.Ya., Kolesnichenko A.V. *Turbulence and Self-Organization. Modeling Astrophysical Objects*. New York et al. : Springer, 2013. 657 p.
8. Kullback S., Leibler R.A. On information and sufficiency // *Ann. Math. Statist.* 1951. V. 22. P. 79–86.
9. Ибрагимов Н.М. Концепция экономического равновесия в пространственных системах // *Мир экономики и управления*. 2021. Т. 21, № 4. С. 5–23. doi: 10.25205/2542-0429-2021-21-4-5-23
10. Zadeh M.K., Gavagsaz-Ghoachani R., Martin J.P. et al. Discrete-time tool for stability analysis of DC power electronics-based cascaded systems // *IEEE Trans. Power Electron.* 2017. V. 32 (1). P. 652–667.
11. Уткин В.И. Системы с переменной структурой: состояние проблемы, перспективы // *Автоматика и телемеханика*. 1983. № 9. С. 5–25.
12. Utkin V., Poznyak A., Orlov Yu., Polyakov A. Conventional and high order sliding mode control // *Journal of the Franklin Institute*. 2020. V. 357 (15). P. 10244–10261. doi: 10.1016/j.jfranklin.2020.06.018
13. Sira-Ramirez H., Gomez-Leon B.C., Aguilar-Orduna M.A. A Geometric Algebra Approach to Invariance Control in Sliding Regimes for Switched Systems // *Advances in Applied Clifford Algebras*. 2023. Vol. 33 (3). Art. 35.
14. Годунов С.К., Рябенский В.С. Уравнения математической физики. М. : Наука, 1971. 416 с.
15. Mikosch T., Resnick S., Rootzen H., Stegeman A. Is network traffic approximated by stable Levy motion or fractional Brownian motion? // *Annals of Applied Probability*. 2002. V. 12 (1). P. 23–68.
16. Jackson J.R. Networks of Waiting Lines // *Oper. Res.* 1957. V. 5 (4). P. 518–521.
17. Gordon K.D., Newell G.F. Closed Queuing Systems with Exponential Servers // *Oper. Research*. 1967. V. 15 (2). P. 254–265.
18. Ивченко Г.И., Каштанов В.А., Коваленко И.Н. Теория массового обслуживания. М. : Высшая школа, 1982. 256 с.

References

1. Zubarev, D.P., Morozov, V.G. & Repke, G. (2002) *Statisticheskaya mekhanika neravnovesnykh protsessov* [Statistical mechanics of nonequilibrium processes]. Vol. 1. Moscow: Fizmatlit.
2. Klimontovich, Yu.L. (1995) *Statisticheskaya teoriya otkrytykh sistem* [Statistical theory of open systems]. Vol. 1. Moscow: Yanus.
3. Zaripov, R.G. (2010) *Printsipy neekstensivnoy statisticheskoy mekhaniki i geometriya mer besporyadka i poryadka* [Principles of nonextensive statistical mechanics and geometry of measures of disorder and order]. Kazan: Kazan State Technical University.
4. Kolesnichenko, A.V. (2017) *Kontinual'nye modeli prirodnykh i kosmicheskikh sred: Problemy termodinamicheskogo konstruirovaniya* [Continuum models of natural and space environments: Problems of thermodynamic design]. Moscow: LENAND.
5. Haken, G. (2014) *Informatsiya i samoorganizatsiya: makroskopicheskiy podkhod k slozhnym sistemam* [Information and self-organization: a macroscopic approach to complex systems]. Translated from English. Moscow: URSS: LENAND.
6. Prigozhin, I. & Kondepudi, D. (2002) *Sovremennaya termodinamika. Ot teplovykh dvigateley do dissipativnykh struktur* [Modern thermodynamics. From heat engines to dissipative structures]. Moscow: Mir.
7. Marov, M.Ya. & Kolesnichenko, A.V. (2013) *Turbulence and Self-Organization. Modeling Astrophysical Objects*. New York; Heidelberg; Dordrecht; London: Springer.
8. Kullback, S. & Leibler, R.A. (1951) On information and sufficiency. *Annual Mathematical Statistics*. 22. pp. 79–86.
9. Ibragimov, N.M. (2021) The concept of economic equilibrium in spatial systems. *Mir ekonomiki i upravleniya – World of Economics and Management*. 21(4). pp. 5–23. DOI: 10.25205/2542-0429-2021-21-4-5-23
10. Zadeh, M.K., Gavagsaz-Ghoachani, R., Martin, J.P. et al. (2017) Discrete-time tool for stability analysis of DC power electronics-based cascaded systems. *IEEE Trans. Power Electron.* 32(1). pp. 652–667.
11. Utkin, V.I. (1983) Sistemy s peremennoy strukturoy: sostoyanie problemy, perspektivy [Systems with variable structure: the state of the problem, prospects]. *Avtomatika i telemekhanika*. 9. pp. 5–25.
12. Utkin, V., Poznyak, A., Orlov, Yu. & Polyakov, A. (2020) Conventional and high order sliding mode control. *Journal of the Franklin Institute*. 357(15). pp. 10244–10261. DOI: 10.1016/j.jfranklin.2020.06.018
13. Sira-Ramirez, H., Gomez-Leon, B.C. & Aguilar-Orduna, M.A. (2023) A Geometric Algebra Approach to Invariance Control in Sliding Regimes for Switched Systems. *Advances in Applied Clifford Algebras*. 33(3). Art. 35.
14. Godunov, S.K. & Ryabenky, V.S. (1971) *Uravneniya matematicheskoy fiziki* [Equations of mathematical physics]. Moscow: Nauka.
15. Mikosch, T., Resnick, S., Rootzen, H. & Stegeman, A. (2002) Is network traffic approximated by stable Levy motion or fractional Brownian motion? *Annals of Applied Probability*. 12(1). pp. 23–68.
16. Jackson, J.R. (1957) Networks of Waiting Lines. *Operations Research*. 5(4). pp. 518–521.
17. Gordon, K.D. & Newell, G.F. (1967) Closed Queuing Systems with Exponential Servers. *Operations Research*. 15(2). pp. 254–265.
18. Ivchenko, G.I., Kashtanov, V.A. & Kovalenko, I.N. (1982) *Teoriya massovogo obsluzhivaniya* [Queuing Theory]. Moscow: Vysshaya shkola.

Информация об авторах:

Цициашвили 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 Yuri 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 03.06.2024

Received 12.04.2024; accepted for publication 03.06.2024

ОБРАБОТКА ИНФОРМАЦИИ

DATA PROCESSING

Научная статья

УДК 519.234

doi: 10.17223/19988605/67/6

Оптимальное ядерное оценивание плотности методом статистического моделирования при случайном цензурировании наблюдений**Абдурахим Ахмедович Абдушукуров¹, Сухроб Баходирович Бозоров²,
Дилшод Равилович Мансуров³**¹ *Московский государственный университет, филиал в г. Ташкенте, Ташкент, Узбекистан*² *Гулистанский государственный университет, Гулистан, Узбекистан*³ *Навоийский государственный педагогический институт, Навои, Узбекистан*¹ *a_abdushukurov@rambler.ru*² *suxrobbek_8912@mail.ru*³ *mathematicianmd@gmail.com*

Аннотация. Рассматриваются задачи непараметрического оценивания вероятности безотказной работы и функции плотности распределения продолжительности работы некоторого физического устройства. Используются случайно цензурированные справа наблюдения. Строятся оценки сложной структуры: множительная, экспоненциальная, степенная, а также оценка с модификацией. Предлагаются новые оценки для плотности распределения с использованием ядерного метода. Приведены таблицы оптимального выбора параметров сглаживания, полученные с использованием статистического моделирования.

Ключевые слова: множительная оценка; экспоненциальная оценка; степенные оценки; ядерное оценивание; случайное цензурирование.

Для цитирования: Абдушукуров А.А., Бозоров С.Б., Мансуров Д.Р. Оптимальное ядерное оценивание плотности методом статистического моделирования при случайном цензурировании наблюдений // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 67. С. 49–59. doi: 10.17223/19988605/67/6

Original article

doi: 10.17223/19988605/67/6

**Optimal kernel density estimation by statistical modeling
with random censoring of observations****Abdurakhim A. Abdushukurov¹, Sukhrob B. Bozorov², Dilshod R. Mansurov³**¹ *Moscow State University branch in Tashkent, Tashkent, Uzbekistan*² *Gulistan State University, Gulistan, Uzbekistan*³ *Navoi State Pedagogical Institute, Navoi, Uzbekistan*¹ *a_abdushukurov@rambler.ru*² *suxrobbek_8912@mail.ru*³ *mathematicianmd@gmail.com*

Abstract. The problems of nonparametric estimation of the probability of failure-free operation and the density distribution function of the operating time of some physical device are considered. Randomly right-censored observa-

tions are used. Estimates of a complex structure are constructed: multiplying, exponential, as well as power estimates and an estimate with modification. New estimates for the distribution density using the kernel method are proposed. Tables of the optimal choice of smoothing parameters obtained using statistical modeling are presented.

Keywords: multiplier estimation; exponential estimation; power estimators; kernel estimation; random censoring.

For citation: Abdushukurov, A.A., Bozorov, S.B., Mansurov, D.R. (2024) Optimal kernel density estimation by statistical modeling with random censoring of observations. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 67. pp. 49–59. doi: 10.17223/19988605/67/6

Введение

При анализе числовых данных типа времени жизни (в медико-биологических испытаниях индивидуума на выживаемость или в промышленных экспериментах на продолжительность безотказной работы технических устройств) интерес представляют неотрицательные случайные величины и их числовые характеристики. В этой связи рассмотрим случайную величину X на вероятностном пространстве $(\Omega, \mathcal{A}, P)$ со значениями в измеримом пространстве $(\mathbb{R}^+; \mathcal{B}(\mathbb{R}^+))$ с абсолютно непрерывным распределением относительно меры Лебега на $\mathbb{R}^+$. Пусть $F(t) = P(X \leq t)$, $t \in \mathbb{R}^+$, – функция распределения случайной величины X . Тогда мера dF имеет производную Радона–Никодима f . Функция выживания, означающая вероятность безотказной работы испытываемого объекта, определяется выражением

$$1 - F(t) = P(X > t) = \int_t^{+\infty} f(u) du, \quad t \in \mathbb{R}^+.$$

Плотность интенсивности отказов и функция кумулятивной (интегральной) интенсивности определяются формулами

$$\lambda(t) = \frac{f(t)}{1 - F(t)}, \quad \Lambda(t) = \int_0^t \lambda(u) du, \quad t \in \mathbb{R}^+.$$

Поскольку

$$\lambda(t) = \lim_{h \rightarrow 0} \frac{1}{h} P(X \leq t + h | X \geq t), \quad t \in \mathbb{R}^+,$$

функцию $\lambda(t)$ также можно интерпретировать как скорость смертности в момент времени t . Справедливы следующие равенства [1–3]:

$$1 - F(t) = \exp(-\Lambda(t)), \quad \Lambda(t) = -\log(1 - F(t)), \quad t \in \mathbb{R}^+.$$

Во многих практических ситуациях невозможно полностью наблюдать случайную величину X . Тогда анализ данных опирается на неполные, а точнее – на цензурированные данные. В данной работе рассматриваются цензурированные справа наблюдения. Различные модели цензурирования приведены в исследованиях [3–13].

1. Математическая модель

Пусть $X_1, X_2, \dots$ и $Y_1, Y_2, \dots$ – две независимые последовательности независимых неотрицательных случайных величин, определенных на одном вероятностном пространстве $(\Omega, \mathcal{A}, P)$ с абсолютно непрерывными функциями распределений $F(t)$ и $G(t)$, $t \in \mathbb{R}^+$, с соответствующими плотностями $f(t)$ и $g(t)$. При случайном цензурировании справа наблюдаются следующие пары:

$$\mathbb{C}^{(n)} = \{(Z_i, \delta_i), 1 \leq i \leq n\},$$

где $Z_i = \min(X_i, Y_i)$ и $\delta_i = I(X_i \leq Y_i)$ – индикаторная случайная величина, показывающая, что Z_i цензурирована ($\delta_i = 0$) или не цензурирована ($\delta_i = 1$). Через H обозначим функцию распределения $Z = \min(X, Y)$:

$$H(t) = P(Z \leq t) = 1 - (1 - F(t))(1 - G(t)) = H^{(0)}(t) + H^{(1)}(t), t \in \mathbb{R}^+,$$

где $H^{(m)}(t) = P(Z \leq t, \delta = m)$, $m = 0, 1$, – субраспределения, для которых справедливы выражения

$$H^{(0)}(t) = \int_0^t (1 - F(u)) dG(u), H^{(1)}(t) = \int_0^t (1 - G(u)) dF(u), t \in \mathbb{R}^+. \quad (1)$$

Таким образом, случайные величины X_i наблюдаемы лишь в случае $\delta_i = 1$. Тогда общее их число в выборке $\mathbb{C}^{(n)}$ равно случайной величине $v(n) = \delta_1 + \delta_2 + \dots + \delta_n$, имеющей биномиальное распределение $Bi(n; p)$, где

$$p = P(X \leq Y) = H^{(1)}(\infty) = \int_0^{+\infty} (1 - G(u)) dF(u) \in (0, 1).$$

В этой модели задача состоит в оценивании распределения F , используя выборку $\mathbb{C}^{(n)}$. Естественно, в качестве искомой оценки для F сперва рассмотрим эмпирическую статистику

$$\hat{F}_n(t) = \frac{1}{v(n)} \sum_{i=1}^n \delta_i \cdot I(Z_i \leq t), t \in \mathbb{R}^+, \quad (2)$$

использующую только наблюдаемые X_i в выборке $\mathbb{C}^{(n)}$. В [3] установлено, что

$$\hat{F}_n(t) \xrightarrow[n \rightarrow \infty]{p} \hat{F}(t) = \frac{H^{(1)}(t)}{p}, t \in \mathbb{R}^+.$$

Видно, что $\hat{F}(t) \neq F(t)$, и равенство $\hat{F}(t) = F(t)$ возможно лишь в случае отсутствия цензурирования (т.е. при $G \equiv 0$). Следовательно, $\hat{F}_n$ является несостоятельной статистикой F .

2. Оценивание функции распределения

Анализ работ показал, что существующие к настоящему времени оценки для F имеют более сложную структуру, чем эмпирическая функция распределения. Впервые при случайном цензурировании справа PL (Product-Limit) оценка Каплана–Мейера [6] была предложена в 1958 г. Она имеет множительную структуру:

$$F^{KM}(t) = \begin{cases} 1 - \prod_{\{i: Z_{(i)} \leq t\}} \left(1 - \frac{\delta_{(i)}}{n - i + 1}\right), & t \leq Z_{(n)}, \\ 1, & t > Z_{(n)}, \delta_{(n)} = 1, \\ \text{неопределена,} & t > Z_{(n)}, \delta_{(n)} = 0, \end{cases} \quad (3)$$

где $Z_{(1)} \leq Z_{(2)} \leq \dots \leq Z_{(n)}$ – порядковые статистики, построенные по Z -выборке $\{Z_i, 1 \leq i \leq n\}$, а $\{\delta_{(i)}, 1 \leq i \leq n\}$, – сопровождающие их индикаторы. Оценка (3) имеет большое влияние на развитие статистической теории цензурированных наблюдений. Как видно из формулы определения оценки (3), она в последней точке $t > Z_{(n)}$ зависит от цензурирования. Тем не менее в литературе встречаются разнообразные модификации этой оценки.

Однако (3) – не единственная оценка для F , построенная по выборке $\mathbb{C}^{(n)}$. Для введения других оценок нам необходимы некоторые дополнительные обозначения и определения. Оценим субраспределения $H^{(0)}(t)$ и $H^{(1)}(t)$ эмпирическими функциями

$$H_n^{(m)}(t) = \frac{1}{n} \sum_{i=1}^n I(Z_i \leq t, \delta_i = m), \quad m = 0, 1. \quad (4)$$

Введем кумулятивные функции интенсивности для G, F и H

$$\begin{aligned} \Lambda^{(0)}(t) &= \int_0^t \frac{dG(u)}{1-G(u)} = \int_0^t \frac{dH^{(0)}(u)}{1-H(u)}, \\ \Lambda^{(1)}(t) &= \int_0^t \frac{dF(u)}{1-F(u)} = \int_0^t \frac{dH^{(1)}(u)}{1-H(u)}, \\ \Lambda(t) &= \Lambda^{(0)}(t) + \Lambda^{(1)}(t) = \int_0^t \frac{dH(u)}{1-H(u)}. \end{aligned} \quad (5)$$

Введем их оценки

$$\begin{aligned} \Lambda_n^{(0)}(t) &= \int_0^t \frac{dH_n^{(0)}(u)}{1-H_n(u) + \frac{1}{n}}, \quad \Lambda_n^{(1)}(t) = \int_0^t \frac{dH_n^{(1)}(u)}{1-H_n(u) + \frac{1}{n}}, \\ \Lambda_n(t) &= \int_0^t \frac{dH_n(u)}{1-H_n(u) + \frac{1}{n}} = \Lambda_n^{(0)}(t) + \Lambda_n^{(1)}(t), \quad H_n(t) = H_n^{(0)}(t) + H_n^{(1)}(t), \end{aligned} \quad (6)$$

где слагаемое $\frac{1}{n}$ в знаменателе для оценок (6) добавлено для предотвращения деления на ноль при их вычислении.

Согласно формулам (5) и (6), естественной оценкой экспоненциального вида является следующая оценка Альтшулера–Бреслоу [10, 11]:

$$\begin{aligned} F_n^{AB}(t) &= 1 - \exp(-\Lambda_n^{(1)}(t)) = 1 - \exp\left\{-\sum_{i=1}^n \frac{I(Z_i \leq t, \delta_i = 1)}{n \left[1 - H_n(Z_i) + \frac{1}{n}\right]}\right\} = \\ &= 1 - \exp\left(-\sum_{\{i: Z_{(i)} \leq t\}} \frac{\delta_{(i)}}{n-i+1}\right) = 1 - \left[\prod_{\{i: Z_{(i)} \leq t\}} \exp\left(-\frac{1}{n-i+1}\right)\right]^{\delta_{(i)}}, \quad t \in \mathbb{R}^+. \end{aligned} \quad (7)$$

Для сравнения с (7) используем следующую модификацию оценки (3):

$$\hat{F}^{KM}(t) = \begin{cases} 1 - \prod_{\{i: Z_{(i)} \leq t\}} \left(1 - \frac{\delta_{(i)}}{n-i+1}\right), & t \leq Z_{(n)}, \\ 1, & t > Z_{(n)}, \end{cases} = \begin{cases} \prod_{\{i: Z_{(i)} \leq t\}} \left(\frac{n-i}{n-i+1}\right)^{\delta_{(i)}}, & t \leq Z_{(n)}, \\ 1, & t > Z_{(n)}. \end{cases} \quad (8)$$

Поскольку при $x \approx 0$, $e^{-x} \approx 1 - x$, из формул (7) и (8) имеем аппроксимацию

$$-\log(1 - \hat{F}_n^{KM}(t)) \approx \Lambda_n^{(1)}(t), \quad t \in \mathbb{R}^+, \quad (9)$$

что показывает близость этих оценок при больших n . Автором [8, 9] была предложена оценка степенного типа

$$F_n^{(1)}(t) = 1 - [1 - H_n(t)]^{R_n(t)} = \begin{cases} 0, & t < Z_{(1)}, \\ 1 - \left(\frac{n-i}{n}\right)^{R_n(t)}, & Z_{(i)} \leq t < Z_{(i+1)}, \quad 1 \leq i \leq n-1, \\ 1, & t \geq Z_{(n)}, \end{cases} \quad (10)$$

где $R_n(t) = \Lambda_n^{(1)}(t) \cdot [\Lambda_n(t)]^{-1}$ – оценка дроби $R(t) = \Lambda^{(1)}(t) \cdot [\Lambda(t)]^{-1}$. В работах [3, 8, 9] показано, что оценка (10) идентифицируема с рассматриваемой моделью случайного цензурирования справа, т.е. для оценок структуры (10) имеет место аналог равенства

$$(1 - F(t))(1 - G(t)) = 1 - H(t), \quad t \in \mathbb{R}^+. \quad (11)$$

Действительно, если $G_n^{(1)}(t) = 1 - [1 - H_n(t)]^{1-R_n(t)}$ – оценка для $G(t)$ вида (10), то очевидно

$$(1 - G_n^{(1)}(t))(1 - F_n^{(1)}(t)) = 1 - H_n(t), \quad t \in \mathbb{R}^+, \quad n \geq 1. \quad (12)$$

Другими словами, оценки вида (10) определены во всех точках $\mathbb{R}^+$.

Теперь рассмотрим условную вероятность

$$p(t) = P(\delta = 1 | Z = t) = E[\delta | Z = t], \quad t \in \mathbb{R}^+. \quad (13)$$

Легко видеть, что справедливы равенства

$$H^{(1)}(t) = \int_0^t p(u) dH(u), \quad \Lambda^{(1)}(t) = \int_0^t p(u) d\Lambda(u), \quad t \in \mathbb{R}^+.$$

Следовательно,

$$R(t) = [\Lambda(t)]^{-1} \cdot \int_0^t p(u) d\Lambda(u), \quad t \in \mathbb{R}^+. \quad (14)$$

Так как вероятность (13) фактически является функцией регрессии, то ее естественно оценивать следующей статистикой Надарая–Ватсона [14, 15]:

$$p_n(t) = \left[\frac{1}{nh(n)} \sum_{i=1}^n \delta_i \cdot k\left(\frac{t - Z_i}{h(n)}\right) \right] \cdot \left[\frac{1}{nh(n)} \sum_{i=1}^n k\left(\frac{t - Z_i}{h(n)}\right) \right]^{-1}, \quad t \in \mathbb{R}^+. \quad (15)$$

Пусть $\hat{R}_n(t) = [\Lambda_n(t)]^{-1} \cdot \int_0^t p_n(u) d\Lambda_n(u)$, $t \in \mathbb{R}^+$, – оценка для дроби (14), получаемая по формулам (6)

(третья) и (15). Тогда в формуле (10), заменив $R_n(t)$ на $\hat{R}_n(t)$, получаем еще одну новую степенную оценку структуры (10):

$$F_n^{(2)}(t) = 1 - [1 - H_n(t)]^{\hat{R}_n(t)} = \begin{cases} 0, & t < Z_{(1)}, \\ 1 - \left(\frac{n-i}{n}\right)^{\hat{R}_n(t)}, & Z_{(i)} \leq t < Z_{(i+1)}, \quad 1 \leq i \leq n-1, \\ 1, & t \geq Z_{(n)}. \end{cases} \quad (16)$$

Приведенный на рис. 1 график оценок построен с использованием объема выборки $n = 100; 1000$ с уровнем цензурирования 30%.

Как видно из рис. 1, а, если выбраны правильная функция ядра и соответствующий параметр ширины окна, то оценка $F_n^{(2)}(t)$ лучше, чем $F_n^{(1)}(t)$. При объеме выборки $n = 1000$ (см. рис. 1, b) оценки асимптотически неразличимы. Однако само изображение не дает достаточной информации для выводов. Поэтому, чтобы узнать, какая из этих оценок «лучше», мы создадим табл. 1–3 на основе $N = 2000$ экспериментов значения выражения равномерной метрики

$$\sup_{t \geq 0} |F_n(t) - F(t)|, \quad \text{при } F_n(t) = F_n^{KM}(t), F_n^{(1)}(t), F_n^{(2)}(t)$$

на разных уровнях цензурирования с объемом $n = 100, 200, 500, 1000, 2000$.

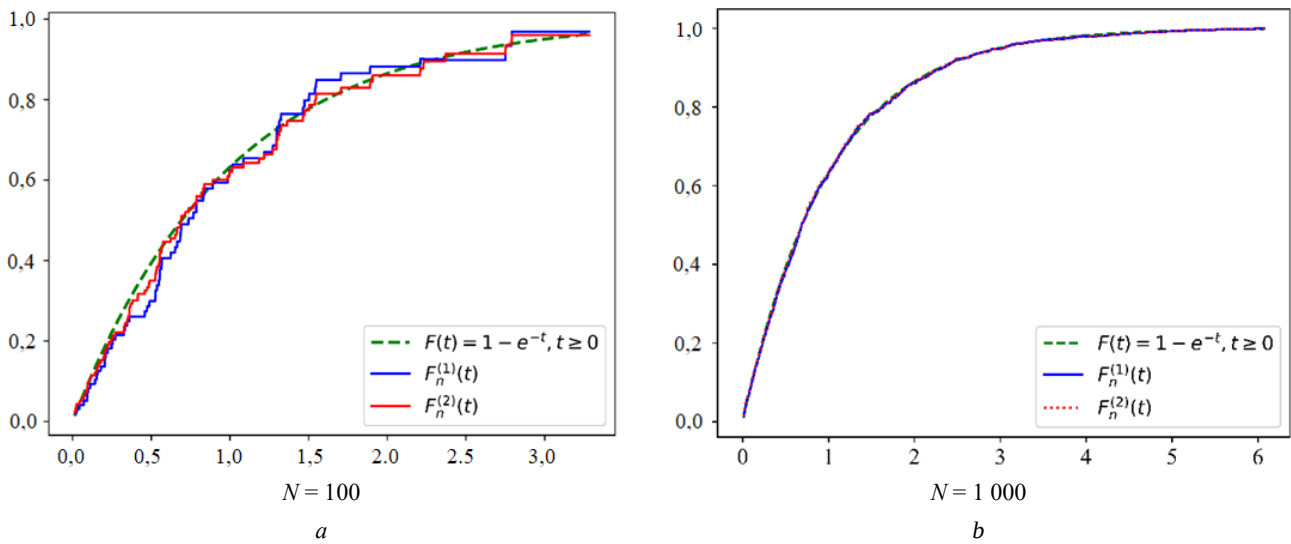
Рис. 1. Оценки $F_n^{(1)}(t)$ и $F_n^{(2)}(t)$ Fig. 1. Estimations $F_n^{(1)}(t)$ and $F_n^{(2)}(t)$

Таблица 1

Отклонение оценки $F_n^{KM}(t)$ от теоретического распределения

Теоретическое распределение $F(t) = 1 - e^{-t}, t \geq 0$					
n	$n = 100$	$n = 200$	$n = 500$	$n = 1\,000$	$n = 2\,000$
% цензур.					
0%	0,08525	0,06047	0,03847	0,02735	0,01942
10%	0,08703	0,06173	0,03948	0,02793	0,01977
20%	0,09014	0,06405	0,04079	0,02889	0,02041
30%	0,09482	0,06781	0,04317	0,03052	0,02166
40%	0,10512	0,07564	0,04846	0,03456	0,02453
50%	0,12387	0,09182	0,06106	0,04431	0,03233
60%	0,15258	0,11939	0,08502	0,06528	0,05020
70%	0,23185	0,18872	0,14288	0,11633	0,09412
80%	0,36522	0,31959	0,26567	0,23120	0,20212

Таблица 2

Отклонение оценки $F_n^{(1)}(t)$ от теоретической функции распределения

Теоретическое распределение $F(t) = 1 - e^{-t}, t \geq 0$					
n	$n = 100$	$n = 200$	$n = 500$	$n = 1\,000$	$n = 2\,000$
% цензур.					
0%	0,08514	0,06041	0,03798	0,02696	0,01891
10%	0,08576	0,06065	0,03891	0,02778	0,01946
20%	0,08620	0,06132	0,03941	0,02832	0,01977
30%	0,08721	0,06303	0,04036	0,03006	0,02124
40%	0,08845	0,06424	0,04257	0,03195	0,02207
50%	0,10239	0,07618	0,05093	0,03661	0,02596
60%	0,14121	0,09419	0,07887	0,05906	0,03708
70%	0,18716	0,13896	0,12348	0,10134	0,07319
80%	0,21807	0,19675	0,16330	0,14907	0,12187

Отклонение оценки $F_n^{(2)}(t)$ от теоретической функции распределения

Теоретическое распределение $F(t) = 1 - e^{-t}, t \geq 0$					
$\% \text{ цензур.}$ \ n	$n = 100$	$n = 200$	$n = 500$	$n = 1\,000$	$n = 2\,000$
0%	0,08512	0,06023	0,03742	0,02716	0,01911
10%	0,08561	0,06044	0,03812	0,02767	0,01932
20%	0,08617	0,06126	0,03871	0,02830	0,01956
30%	0,08712	0,06212	0,03945	0,02902	0,02017
40%	0,08785	0,06345	0,04105	0,03112	0,02197
50%	0,09034	0,07332	0,04789	0,03538	0,02487
60%	0,12056	0,08919	0,07012	0,05423	0,02718
70%	0,17120	0,11998	0,10256	0,09558	0,07018
80%	0,19823	0,18678	0,14993	0,13554	0,11956

Сравнивая каждые столбец и строку табл. 1–3, видим, что оценка (10) ближе к функции распределения, чем оценка (3). Это особенно заметно, когда объем выборки небольшой и уровень цензурирования высок. Однако сравнивая табл. 2 и 3, легко видеть, что оценка (16) лучше, чем (10).

3. Оценивание плотности вероятности

Согласно определению субплотности $h^{(1)}(t) = \frac{dH^{(1)}(t)}{dt}$ в (1) следует, что $(1 - G(t))f(t) = h^{(1)}(t)$,

т.е. $f(t) = \frac{h^{(1)}(t)}{1 - G(t)}$.

Отсюда имеем две оценки для основной плотности $f(t)$:

$$f_n^{(k)}(t) = \frac{h_n^{(1)}(t)}{1 - G_n^{(k)}(t) + \frac{1}{n}}, \quad k = 1, 2, \quad (17)$$

где

$$h_n^{(1)}(t) = \frac{1}{nh(n)} \sum_{i=1}^n \delta_i \cdot k \left(\frac{t - Z_i}{h(n)} \right) \quad (18)$$

есть ядерная оценка субплотности $h^{(1)}(t)$ (числитель оценки (15)),

$$G_n^{(1)}(t) = 1 - [1 - H_n(t)]^{1 - R_n(t)}, \quad G_n^{(2)}(t) = 1 - [1 - H_n(t)]^{1 - \hat{R}_n(t)}$$

есть степенные оценки для $G(t)$ структур (10) и (16) соответственно, а слагаемое $\frac{1}{n}$ в знаменателе в (17) добавлено для предотвращения операции деления на ноль при их вычислении. На рис. 2 приведена оценка $G_n^{(k)}(t)$.

Кроме того, для плотности $f(t)$ вычислим следующие гистограммные оценки, использующие степенные оценки для $F(t)$, вычисляемые по формулам (10) и (16):

$$\hat{f}_n^{(k)}(t) = \frac{F_n^{(k)}(t + h(n)) - F_n^{(k)}(t - h(n))}{2h(n)}, \quad k = 1, 2,$$

где $h(n) = n^{-1/5}$.

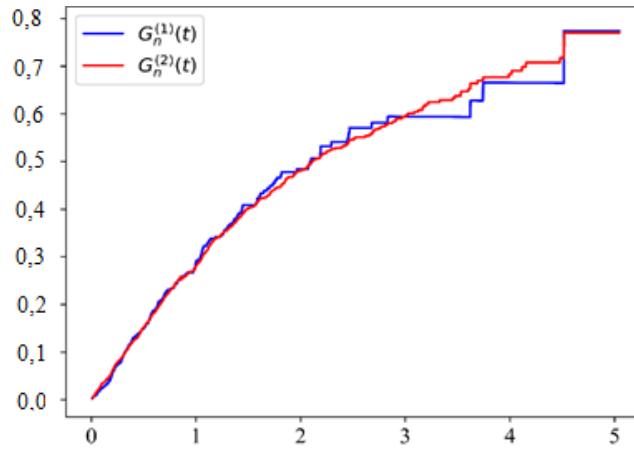
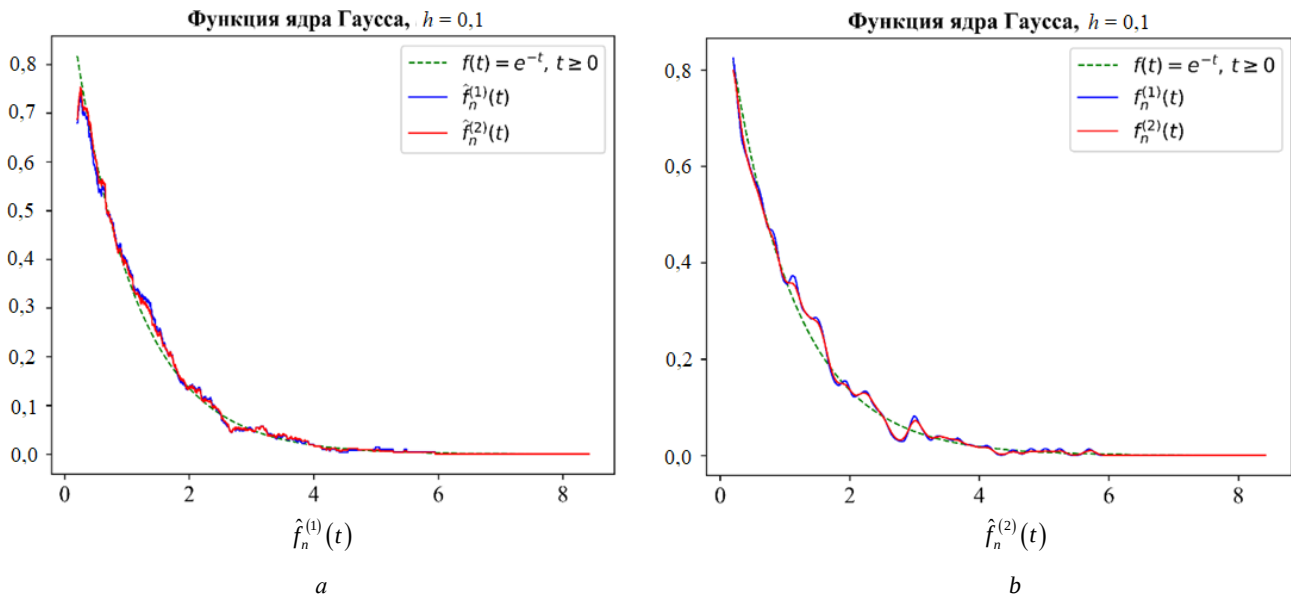
Рис. 2. Оценки $G_n^{(k)}(t)$ Fig. 2. Estimations $G_n^{(k)}(t)$

Рисунок 3 построен с использованием выборки объема $n = 100$ с уровнем цензурирования 30%. Как видно из рис. 3, если выбраны правильная функция ядра и соответствующие параметры ширины окна, оценка $\hat{f}_n^{(2)}(t)$ по сравнению с оценкой $\hat{f}_n^{(1)}(t)$ ближе к $f(t)$.

Рис. 3. Оценки $\hat{f}_n^{(1)}(t)$ и $\hat{f}_n^{(2)}(t)$ Fig. 3. Estimations $\hat{f}_n^{(1)}(t)$ and $\hat{f}_n^{(2)}(t)$

То же самое можно сказать и об оценках $f_n^{(k)}(t)$. Это можно объяснить преимуществом оценки $F_n^{(2)}(t)$.

Теперь, чтобы проверить, как оценка $f_n^{(2)}(t)$ зависит от параметра ширины окна и существует ли оптимальный параметр ширины окна, мы можем составить табл. 4 из средних значений

$$\sup_{t \geq 0} |f_n^{(2)}(t) - f(t)|$$

в 20 000 экспериментах.

Таблица 4

Зависимости оценки $f_n^{(2)}(t)$ от параметров

h	% цензурирования						
	0%	10%	20%	30%	40%	50%	60%
0,01	0,03413	0,03783	0,04328	0,05058	0,06089	0,07820	0,10967
0,02	0,01655	0,01839	0,02107	0,02429	0,02870	0,03591	0,04857
0,03	0,01233	0,01369	0,01563	0,01784	0,02108	0,02622	0,03484
0,04	0,01102	0,01194	0,01388	0,01577	0,01875	0,02327	0,03071
0,05	0,01081	0,01219	0,01357	0,01541	0,01843	0,02316	0,03007
0,06	0,01113	0,01230	0,01398	0,01589	0,01908	0,02281	0,03112
0,07	0,01174	0,01301	0,01480	0,01683	0,02029	0,02504	0,03313
0,08	0,01254	0,01393	0,01587	0,01807	0,02186	0,02697	0,03574
0,09	0,01346	0,01500	0,01712	0,01952	0,02366	0,02922	0,03876
0,1	0,01447	0,01617	0,01849	0,02110	0,02565	0,03169	0,04208
0,11	0,01554	0,01741	0,01994	0,02280	0,02776	0,03435	0,04562
0,12	0,01667	0,01872	0,02147	0,02459	0,02997	0,03713	0,04933
0,13	0,01783	0,02006	0,02305	0,02644	0,03225	0,04002	0,05317
0,14	0,01902	0,02145	0,02467	0,02835	0,03460	0,04298	0,05711
0,15	0,02024	0,02286	0,02632	0,03031	0,03700	0,04602	0,06113
0,16	0,02149	0,02430	0,02801	0,03230	0,03944	0,04910	0,06521
0,17	0,02275	0,02575	0,02971	0,03432	0,04191	0,05223	0,06934
0,18	0,02403	0,02722	0,03143	0,03637	0,04442	0,05538	0,07350
0,19	0,02532	0,02871	0,03317	0,03844	0,04694	0,05857	0,07769
0,2	0,02662	0,03021	0,03492	0,04053	0,04948	0,06177	0,08190
0,23	0,03058	0,03475	0,04020	0,04686	0,05716	0,07143	0,09456
0,25	0,03324	0,03781	0,04375	0,05111	0,06230	0,07788	0,10297
0,3	0,03994	0,04547	0,05263	0,06175	0,07512	0,09387	0,12373
0,4	0,05324	0,06063	0,07012	0,08258	0,10001	0,12458	0,16310
0,5	0,06611	0,07522	0,08686	0,10232	0,12331	0,15289	0,19886
0,6	0,07836	0,08904	0,10262	0,12069	0,14472	0,17853	0,23088
0,7	0,08992	0,10202	0,11732	0,13763	0,16424	0,20160	0,25937
0,8	0,10077	0,11415	0,13096	0,15318	0,18196	0,22226	0,28460
0,9	0,11092	0,12544	0,14357	0,16743	0,19802	0,24077	0,30679
1	0,12042	0,13594	0,15522	0,18046	0,21257	0,25733	0,32618

Из табл. 4 мы видим, что для каждого процента цензурирования существует оптимальный параметр ширины окна. В табл. 4 минимальные значения выделены жирным шрифтом. При использовании оптимальной ширины окна оценка становится ближе к функции плотности.

Заключение

Исследования в области статистики цензурированных наблюдений показали, что степенные оценки вероятности безотказной работы испытываемого на надежность физического устройства (или индивидуума) обладают рядом преимуществ по сравнению с их аналогами, построенными на основе множительных оценок. Особенно это заметно у заранее сглаженной оценки дроби кумулятивных интенсивностей. Методом статистического моделирования показано, что такое преимущество сохраняется и для ядерных и гистограммных оценок плотности.

Список источников

1. Байхельт Ф., Франкен П. Надежность и технической обслуживание : математический подход. М. : Радио и связь, 1988. 387 с.
2. Кошкин Г.М. Основы страховой математики. Томск : Том. гос. ун-т, 2002. 112 с.
3. Абдушукуров А.А. Оценки неизвестных распределений по неполным наблюдениям и их свойства. Saarbrücken : LAP Lambert Academic Publishing, 2011. 308 с.

4. Burke M.D., Csörgő S., Horváth L. Strong approximations of some biometric estimates under random censorship // Z. Wahrschein. Verw. Gebiete. 1981. V. 56. P. 87–112.
5. Csörgő S., Horváth L. On random censorship from the right // Acta Scientiarum Mathematicarum. 1982. V. 44. P. 23–24.
6. Kaplan E.L., Meier P. Nonparametric estimation from incomplete observation // J. Am. Stat. Assoc. 1958. V. 53. P. 457–481.
7. Abdushukurov A.A., Sayfullayeva G.S. Asymptotic properties of modified empirical Kac processes under general random censorship model // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2022. № 61. С. 26–36. doi:10.17223/19988605/61/3
8. Abdushukurov A.A., Nonparametric estimation of the distribution function based on relative-risk function // Commun. Stat. Theory Meth. 1998. V. 27. P. 1991–2012.
9. Abdushukurov A.A. On nonparametric estimation of reliability indices by censored samples // Theory Probab. Appl. 1999. V. 43 (1). P. 3–11.
10. Altshuler B. Theory for the measurement for the competing risks in animal experiments // Math. Biosciences. 1970. V. 6. P. 1–11.
11. Breslow N. Discussion on professor Cox's paper // J. Royal. Statist. Society. Ser. A. 1972. V. 34. P. 216–217.
12. Шуленин В.П. Асимптотические свойства и робастность оценок урезанных вариантов стандартного отклонения и среднего абсолютных отклонений // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2021. № 55. С. 91–102. doi:10.17223/19988605/55/11
13. Зенкова Ж.Н. Учет информации об Sa-равноплечной симметрии при обработке цензурированных данных // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2011. № 2 (15). С. 32–45.
14. Nadaraya E.A. On estimating regression // Probab. Theor. Relat. Fields. 1964. V. 61. P. 405–415. doi: 10.1137/1109020
15. Watson G.S. Smooth regression analysis. // Sankhya. Ser. A. 1964. V. 26. P. 359–372.

References

1. Baykheft, F. & Franken, P. (1988) *Nadezhnos' i tekhnicheskoy obsluzhivanie. Matematicheskiy podkhod* [Reliability and technical maintenance. A mathematical approach]. Moscow: Radio i svyaz'.
2. Koshkin, G. M. (2002) *Osnovy strakhovoy matematiki* [Fundamentals of insurance mathematics]. Tomsk: Tomsk State University.
3. Abdushukurov, A.A. (2011) *Otsenki neizvestnykh raspredeleniy po nepolnym nablyudeniym i ikh svoystva* [Estimation of the unknown distribution based on incomplete observations and their properties]. LAP Lambert Academic Publishing.
4. Burke, M.D., Csörgő, S. & Horváth, L. (1981) Strong approximations of some biometric estimates under random censorship. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*. 56. pp. 87–112.
5. Csörgő, S. & Horváth, L. (1982) On random censorship from the right. *Acta Scientiarum Mathematicarum*. 44. pp. 23–24.
6. Kaplan, E.L. & Meier, P. (1958) Nonparametric estimation from incomplete observation. *Journal of the American Statistical Association*. 53. pp. 457–481.
7. Abdushukurov, A.A. & Sayfullayeva, G.S. (2022) Asymptotic properties of modified empirical Kac processes under general random censorship model. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 61. pp. 26–36. DOI: 10.17223/19988605/61/3
8. Abdushukurov, A.A. (1998) Nonparametric estimation of the distribution function based on the relative-risk function. *Communications in Statistics – Theory and Methods*. 27. pp. 1991–2012.
9. Abdushukurov, A.A. (1999). On nonparametric estimation of reliability indices by censored samples. *Theory of Probability and Its Applications*. 43(1). pp. 3–11.
10. Altshuler, B. (1970) Theory for the measurement of competing risks in animal experiments. *Mathematical Biosciences*. 6. pp. 1–11.
11. Breslow, N. (1972) Discussion on Professor Cox's paper. *Journal of the Royal Statistical Society*. A(34). pp. 216–217.
12. Skulenin, V.P. (2021). Asymptotic properties and robustness of trimmed versions estimates of standard deviation and mean absolute deviation. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 55. pp. 91–102. DOI: 10.17223/19988605/55/11
13. Zenkova, Zh.N. (2011) Using an amount of information about ss-equal – shoulder symmetry at the processing of censored data. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 15(2). pp. 32–45.
14. Nadaraya, E.A. (1964) On estimating regression. *Probability Theory and Related Fields*. 61. pp. 405–415. DOI: 10.1137/1109020
15. Watson, G.S. (1964) Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics*. A(26). pp. 359–372.

Информация об авторах:

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

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

Мансуров Дилшод Равилович – кандидат физико-математических наук, доцент Навоийского государственного педагогического института (Навои, Узбекистан). E-mail: mathematicianmd@gmail.com

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

Information about the authors:

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

Bozorov Sukhrob Bakhodirovich (Post-Graduate Student, Gulistan State University, Gulistan, Uzbekistan). E-mail: suxrobbek_8912@mail.ru

Mansurov Dilshod Ravilovich (Candidate of Physical and Mathematical Sciences, Navoi State Pedagogical Institute, Navoi, Uzbekistan). E-mail: mathematicianmd@gmail.com

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

Поступила в редакцию 27.02.2024; принята к публикации 03.06.2024

Received 27.02.2024; accepted for publication 03.06.2024

Научная статья

УДК 519.233.22

doi: 10.17223/19988605/67/7

Об одном способе выбора компромисса в семействе условно оптимальных оценок

Константин Викторович Гаврилов¹, Евгения Леонидовна Веретельникова²

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

¹ aenigma77@mail.ru

² janever@mail.ru

Аннотация. Работа посвящена развитию теории устойчивого оценивания А.М. Шурыгина в части локально устойчивого подхода, основанного на анализе показателя неустойчивости оценки (L_2 -нормы функции влияния). В рамках данного подхода рассмотрено семейство условно оптимальных оценок, которое может определяться как оптимизирующее асимптотическую дисперсию при ограничении на неустойчивость. На практике нередко возникает вопрос обоснованного выбора одной оценки из семейства, обеспечивающей компромисс между определяющими собой выпуклую линейную комбинацию исходных критериев, которые нормируются таким образом, чтобы центр семейства соответствовал решению максиминной задачи относительно аргументов функционала. Показано, что найденная оценка минимизирует произведение критериев.

Ключевые слова: M -оценки; функция влияния; устойчивые оценки; сниженные оценки; условно оптимальная оценка.

Для цитирования: Гаврилов К.В., Веретельникова Е.Л. Об одном способе выбора компромисса в семействе условно оптимальных оценок // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 67. С. 60–68. doi: 10.17223/19988605/67/7

Original article

doi: 10.17223/19988605/67/7

On one way to choose a compromise in a family of conditionally optimal estimators

Konstantin V. Gavrilov¹, Evgeniya L. Veretel'nikova²

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

¹ aenigma77@mail.ru

² janever@mail.ru

Abstract. The work is devoted to the development of stable estimation theory by A. M. Shurygin in terms of a locally stable approach based on the analysis of the indicator of estimation instability (L_2 -norm of the influence function). Within the framework of this approach, a family of conditionally optimal estimators is considered, which can be defined as optimizing the asymptotic dispersion under a constraint on instability. In practice, there may be a question of a reasonable choice of one estimator from a family that provides a compromise between the criteria defining the family. To solve the problem, it is proposed to form a functional, which is a convex linear combination of the initial criteria are normalized in such a way that the center of the family corresponds to the solution of the maximin problem regarding the arguments of the functional. It is shown that the found estimator minimizes the product of the criteria.

Keywords: M -estimators; influence function; stable estimates; redescending estimators; conditionally optimal estimators.

For citation: Gavrilov, K.V., Veretel'nikova, E.L. (2024) On one way to choose a compromise in a family of conditionally optimal estimators. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 67. pp. 60–68. doi: 10.17223/19988605/67/7

Введение

Важным аспектом теории робастного оценивания статистических моделей является обеспечение устойчивости оценок к асимметричному засорению наблюдений. Таким свойством обладают *сниженные оценки*, в которых влияние периферийных наблюдений снижено по сравнению с классическими робастными решениями или равно нулю [1–3]. Однако до недавнего времени большинство известных сниженных оценок было получено, по существу, эвристически. Ситуация изменилась после того, как А.М. Шурыгин [4] предложил модель байесовского точечного засорения, описывающую воздействие асимметричного засорения на набор данных, и локально устойчивый подход, основанный на показателе *неустойчивости* (L_2 -норме функции влияния). С использованием данных концепций стало возможным получать теоретически обоснованные решения, относящиеся к классу сниженных оценок и, как следствие, обеспечивающие устойчивость к асимметричному засорению.

Примером такого решения в рамках второй из перечисленных концепций является семейство *условно оптимальных оценок* [4–8], которое может определяться как решение задачи оптимизации неустойчивости при ограничении на *асимптотическую дисперсию* (или, что то же, оптимизации асимптотической дисперсии при ограничении на величину неустойчивости). Эта задача сводится к оптимизации линейной комбинации критериев. Вместе с тем обратим внимание, что критерии асимптотической дисперсии и неустойчивости, на базе которых построено семейство, вообще говоря, несопоставимы, поскольку имеют отличающиеся размерности – обычно квадрата и куба случайной величины соответственно (в случаях оценивания параметра сдвига или масштаба). И подобная ситуация типична для аддитивных критериев [9]. Общепринятым подходом здесь является введение нормирующих коэффициентов, которые согласуют размерности частных критериев и их вклады в сумму. В нашем случае нормирующие коэффициенты уже присутствуют.

Проблема состоит в том, что исследователь не всегда располагает априорной информацией, позволяющей достаточно объективно присвоить нормирующие коэффициенты каждому из критериев, если по смыслу задачи даны лишь сами критерии, и эта проблема хорошо известна [10]. Соответствующая задача относится к классу некорректных, поэтому стандартные приемы решения задач многокритериальной оптимизации в данном случае не подходят. Но, оказывается, некоторые теоретические выводы можно получить даже в условиях полностью неопределенных нормирующих коэффициентов, о чем и пойдет речь ниже.

В настоящее время известен ряд способов выбора одной оценки из условно оптимального семейства оценок, которые являются в большей или меньшей степени эвристическими. Например, в [4] для этого предложена *компромиссная* оценка, которая минимизирует сумму отношений асимптотической дисперсии и неустойчивости к их минимальным значениям. Также в [8] рассматривается *равнооптимальная* оценка, которая обеспечивает равные значения этих отношений. В данной работе предлагается еще один подход к выбору разумного компромисса в условно оптимальном семействе, основанный на решении оптимизационной задачи и не привязанный к минимальным значениям критериев.

Для этого рассматривается функционал, представляющий собой выпуклую линейную комбинацию нормированных критериев, оптимизация которого приводит к условно оптимальному семейству. При этом нормировка критериев равносильна репараметризации семейства, которая может быть произведена с точностью до произвольного положительного параметра. Произвол в выборе данного параметра можно устранить путем решения максиминной задачи относительно аргументов функционала при наложении дополнительного регуляризирующего условия – чтобы решение указанной задачи соответствовало центру семейства. Построенная таким образом оценка названа *равновесной* – по аналогии с геометрией, где такая выпуклая комбинация двух точек аффинного пространства определяет центр тяжести отрезка, соединяющего эти точки.

В разд. 1 работы приведены базовые теоретические сведения. В разд. 2 описан метод построения равновесной оценки для условно оптимального семейства. Также показано, что равновесная оценка минимизирует произведение критериев асимптотической дисперсии и неустойчивости. В разд. 3 рас-

смотрен пример построения равновесной оценочной функции для оценки параметра сдвига косинусного распределения.

1. Элементы теории устойчивого оценивания

Пусть $x_1, \dots, x_m$ – независимые наблюдения случайной величины ξ , распределенной с плотностью $f(x, \theta)$, где $x \in X \subseteq R$ и параметр $\theta \in \Theta \subseteq R$. M -оценка неизвестного параметра θ может определяться как решение оценочного уравнения [1–4]

$$\sum_{i=1}^m \psi(x_i, \hat{\theta}) = 0,$$

где $\psi(x, \theta)$ – *оценочная функция* параметра θ . Условие асимптотической несмещенности оценки имеет вид [3, 11]:

$$E \psi(\xi, \theta) = \int_X \psi(x, \theta) f(x, \theta) dx = 0, \quad (1)$$

где E – оператор математического ожидания.

Дифференцируя (1) по θ и допуская возможность изменения порядка дифференцирования и интегрирования, можно записать следующие равенства [1, 11]:

$$N(\theta) \stackrel{df}{=} \lim_{t \rightarrow \theta} \frac{\partial}{\partial t} E \psi(\xi, t) = -E \frac{\partial}{\partial \theta} \psi(\xi, \theta) = \int_X \psi(x, \theta) \frac{\partial}{\partial \theta} f(x, \theta) dx. \quad (2)$$

Условия регулярности. Потребуем, чтобы в окрестности истинного значения параметра θ выполнялось условие асимптотической несмещенности (1), были справедливы равенства (2), функция $N(\theta)$ была непрерывной и неравной нулю [1, 8]. Более полный набор условий регулярности, обеспечивающий также асимптотическую нормальность и $\sqrt{m}$ -состоятельность оценок [3, 12, 13], приведен, например, в [14]. Дополнительно будем предполагать, что там, где это необходимо, допустимо внесение операции дифференцирования под знак интеграла, все выписанные ниже характеристики оценочных функций непрерывно зависят от своих параметров, а также $f(x, \theta) \neq \text{const}$.

Нормированная асимптотическая дисперсия M -оценки определяется выражением [1–3]

$$V(\psi) = \frac{1}{N^2(\theta)} \int_X \psi^2(x, \theta) f(x, \theta) dx = E IF^2(\xi, \theta),$$

где $IF(x, \theta) = \psi(x, \theta)/N(\theta)$ – функция влияния [1–3].

Рассмотрим локально устойчивый подход А.М. Шурыгина, основанный на показателе неустойчивости [4, 5]

$$W(\psi) = \frac{1}{N^2(\theta)} \int_X \psi^2(x, \theta) dx = \int_X IF^2(x, \theta) dx.$$

Минимизация функционала $V(\psi)$ приводит к *оценке максимального правдоподобия* (ОМП) [1–3], минимизация $W(\psi)$ – к *оценке максимальной устойчивости* (ОМУ) [4]. Соответствующие значения оптимизируемых функционалов обозначим $V_{\text{ОМП}}$ и $W_{\text{ОМУ}}$. Эффективностью и устойчивостью оценки называются соответственно относительные характеристики

$$\text{eff } \psi = V_{\text{ОМП}}/V(\psi), \quad \text{stb } \psi = W_{\text{ОМУ}}/W(\psi).$$

В рамках данного подхода построено семейство условно оптимальных оценочных функций [4]

$$\psi(x, \theta) = c(\theta, \lambda) \left[\frac{\partial}{\partial \theta} \ln f(x, \theta) + \beta(\theta, \lambda) \right] \frac{f(x, \theta)}{f(x, \theta) + \lambda}, \quad (3)$$

где $\lambda \geq 0$ – параметр, задающий семейство, $c(\theta, \lambda)$ – произвольная непрерывная функция, неравная нулю для всех $\theta \in \Theta$, функция $\beta(\theta, \lambda)$ определяется из условия (1). Значение параметра $\lambda = 0$ в (3)

соответствует ОМП, значение $\lambda = \infty$ – ОМУ. Промежуточные значения параметра λ позволяют получить различные условно оптимальные оценки, имеющие минимальную неустойчивость при ограничении сверху на величину асимптотической дисперсии либо, что то же, минимальную асимптотическую дисперсию при ограничении сверху на величину неустойчивости [4–6].

Для выбора единственного решения из условно оптимального семейства в [4] предложена компромиссная оценка, которой соответствует значение $\lambda = V_{\text{ОМП}}/W_{\text{ОМУ}}$. Другое решение, называемое равнооптимальной оценкой [8], удовлетворяет условию $\text{eff } \psi = \text{stb } \psi$. Ниже предлагается еще одно решение.

2. Определение и свойство равновесной оценки

Условно оптимальное семейство (3) при $\lambda \geq 0$ может определяться как решение оптимизационной задачи [8]

$$V(\psi) + \lambda W(\psi) \rightarrow \min_{\psi}. \quad (4)$$

Произведем нормировку критериев $V(\psi)$ и $W(\psi)$, что равносильно репараметризации семейства. Для этого введем новый параметр

$$k = \lambda / (\lambda + \lambda_0), \quad (5)$$

где $\lambda_0 > 0$ – неопределенный вспомогательный параметр, который задает способ репараметризации семейства. Нетрудно видеть, что независимо от выбора λ_0 значение $k \in [0; 1]$, т.е. значения $1-k$ и k являются коэффициентами выпуклой линейной комбинации, а значение $\lambda = \lambda_0$ соответствует ее центру (значению $k = 1/2$). Также нам понадобится преобразование, обратное к (5): $\lambda = \lambda_0 k / (1-k)$, откуда следует

$$\lambda' = \lambda_0 / (1-k)^2. \quad (6)$$

Штрихом здесь и ниже обозначена производная по k .

Таким образом, нормировку критериев в (4) можно выполнить путем деления выражения в левой части (4) на $\lambda + \lambda_0$. В результате получаем функционал

$$U(\psi, k) = [V(\psi) + \lambda W(\psi)] / (\lambda + \lambda_0) = (1-k)\tilde{V}(\psi) + kW(\psi), \quad (7)$$

где $\tilde{V}(\psi) = V(\psi)/\lambda_0$ – нормированный критерий (критерий $W(\psi)$ не меняется). Очевидно, минимизация функционала (7) по ψ дает то же решение, что и (4).

Далее нужно разрешить неопределенность в выборе вспомогательного параметра λ_0 . Его предлагается выбирать из условия, что значение $k = 1/2$ (центр семейства) является решением максиминной задачи, которая определяет баланс «интересов» между критериями $\tilde{V}(\psi)$ и $W(\psi)$:

$$\min_{\psi} U(\psi, k) \rightarrow \max_k. \quad (8)$$

Прежде чем сформулировать результат решения (8), заметим, что множество значений параметра λ в выражении (3) можно дополнить отрицательными значениями, такими что $\lambda < -\max_{x \in X} f(x, \theta)$.

Соответствующее семейство оценочных функций является решением задачи минимизации неустойчивости $W(\psi)$ при ограничении-равенстве вида $V(\psi) = V_0$ на асимптотическую дисперсию [6], где $V_{\text{ОМП}} \leq V_0 < V_{\text{max}}$, V_{max} – предел (конечный или бесконечный) асимптотической дисперсии оценки при $\lambda \rightarrow -\max_{x \in X} f(x, \theta)$. Расширение условно оптимального семейства в область отрицательных значений λ влияет на оптимизационную формулировку (4): в общем случае она должна записываться как

$$|V(\psi) + \lambda W(\psi)| \rightarrow \min_{\psi}.$$

Что касается функционала (7), то его запись не меняется, а расширение семейства сводится к расширению диапазона значений параметра k в область $k > 1$ (до некоторого максимального значения).

Пусть $\tilde{V}$, V и W обозначают значения соответствующих функционалов для членов условно оптимального семейства, которые доставляют минимум по ψ функционалу (7).

Теорема 1. Если решение (8) существует, оно однозначно определяется условием $\lambda_0 = V/W$.

Доказательство. Найдем максимум по k функции $U = (1-k)\tilde{V} + kW$. Для этого воспользуемся теоремой 1 и следствием 2 из нее в [6], согласно которым $\partial(V + \lambda W)/\partial\lambda = W$. Учитывая это свойство и выражение (6), проверим необходимое условие максимума функции U :

$$U' = \lambda' \cdot \frac{\partial}{\partial\lambda} \frac{V + \lambda W}{\lambda + \lambda_0} = \frac{\lambda_0}{(1-k)^2} \left[\frac{W}{\lambda + \lambda_0} - \frac{V + \lambda W}{(\lambda + \lambda_0)^2} \right] = \frac{\lambda_0(\lambda_0 W - V)}{(1-k)^2(\lambda + \lambda_0)^2} = W - \frac{V}{\lambda_0} = 0.$$

Его можно записать в виде $\tilde{V} = W$, что отражает баланс «интересов». Осталось показать, что данному условию соответствует не более чем единственное значение k , и оно доставляет максимум функции U . Для этого должно выполняться $U'' < 0$ во всем диапазоне допустимых значений k . Используя следствие 3 из теоремы 1 в [6], согласно которому $\partial V/\partial\lambda = -\lambda \partial W/\partial\lambda$, можно записать

$$U'' = W' - V'/\lambda_0 = \lambda'(1 + \lambda/\lambda_0) \partial W/\partial\lambda = W'/(1-k).$$

Убедимся, что функция W убывает по любому из параметров k или λ при $0 \leq k < 1$ (или $\lambda \geq 0$) и возрастает при $k > 1$ (или $\lambda < -\max_{x \in X} \{\lambda_0, \max f(x, \theta)\}$).

Опуская для краткости аргументы функций, введем обозначения $s = f + \lambda$, $\varphi = (\partial f/\partial\theta + \beta f)/s$.

Также согласно следствию из теоремы 2 в [8] справедливо представление $N/c = \int_X \varphi^2 s dx$. Здесь

функция N определена в (2), s и β – те же, что и в (3). Продифференцируем последнее выражение:

$$\frac{\partial}{\partial\lambda} \frac{N}{c} = 2 \frac{\partial\beta}{\partial\lambda} \int_X \varphi f dx - \int_X \varphi^2 dx.$$

Обратим внимание, что первое слагаемое здесь равно 0 в силу (1). И далее, учитывая (1):

$$\begin{aligned} \frac{\partial^2}{\partial\lambda^2} \frac{N}{c} &= 2 \frac{\partial^2\beta}{\partial\lambda^2} \int_X \varphi f dx + 2 \left(\frac{\partial\beta}{\partial\lambda} \right)^2 \int_X \frac{f^2}{s} dx - 2 \frac{\partial\beta}{\partial\lambda} \int_X \frac{\varphi f}{s} dx - \\ &- 2 \frac{\partial\beta}{\partial\lambda} \int_X \frac{\varphi f}{s} dx + 2 \int_X \frac{\varphi^2}{s} dx = \int_X \frac{2}{s} \left(\frac{\partial\beta}{\partial\lambda} f - \varphi \right)^2 dx = 2 \int_X \left(\frac{\partial\varphi}{\partial\lambda} \right)^2 s dx. \end{aligned}$$

Воспользуемся теоремой 1 в [6]:

$$\frac{\partial W}{\partial\lambda} = \frac{\partial^2}{\partial\lambda^2} \frac{c}{N} = -\frac{\partial}{\partial\lambda} \left[\left(\frac{c}{N} \right)^2 \frac{\partial}{\partial\lambda} \frac{N}{c} \right] = 2 \left(\frac{c}{N} \right)^3 \left(\frac{\partial}{\partial\lambda} \frac{N}{c} \right)^2 - \left(\frac{c}{N} \right)^2 \frac{\partial^2}{\partial\lambda^2} \frac{N}{c}.$$

Учитывая, что $\frac{\partial}{\partial\lambda} \frac{N}{c} = -\int_X \varphi^2 dx = \int_X \frac{\partial\varphi}{\partial\lambda} \varphi s dx$, приходим к неравенству Коши–Буняковского:

$$\frac{1}{2} \left(\frac{N}{c} \right)^3 \frac{\partial W}{\partial\lambda} = \left(\int_X \frac{\partial\varphi}{\partial\lambda} \varphi s dx \right)^2 - \int_X \left(\frac{\partial\varphi}{\partial\lambda} \right)^2 s dx \int_X \varphi^2 s dx < 0.$$

Неравенство здесь строгое, поскольку $f \neq \text{const}$. Нетрудно видеть, что оно остается в силе при $s \leq 0$ (т.е. при $k > 1$ или $\lambda < -\max_{x \in X} \{\lambda_0, \max f\}$). В этом случае в левой части равенства сомножитель

$N/c < 0$, а $\partial W/\partial\lambda > 0$. При $0 \leq k < 1$ или $\lambda \geq 0$ имеем $N/c > 0$, а $\partial W/\partial\lambda < 0$. В любом случае $U'' < 0$ (включая и точку $k=1$, в которой функция непрерывно доопределяется). Теорема доказана.

Итак, поскольку условию $k=1/2$ соответствует $\lambda = \lambda_0 > 0$, уравнение для определения компромисса принимает вид:

$$\lambda = V/W. \quad (9)$$

Определение. Оценочная функция из условно оптимального семейства, которая удовлетворяет уравнению (9) при $\lambda > 0$, называется *равновесной*.

Теорема 2. Равновесная оценочная функция является решением задачи $V(\psi)W(\psi) \rightarrow \min_{\psi}$.

Доказательство. Заметим, что решение данной задачи принадлежит условно оптимальному семейству. Действительно, его члены доставляют минимум одному из критериев $V(\psi)$ или $W(\psi)$ при фиксированном значении другого [4, 6], одновременно минимизируется и их произведение.

Определим член семейства, соответствующий решению задачи. Вновь воспользуемся равенством $\partial V/\partial \lambda = -\lambda \partial W/\partial \lambda$ и запишем необходимое условие минимума функции VW :

$$\frac{\partial}{\partial \lambda}(VW) = \frac{\partial W}{\partial \lambda}V + \frac{\partial V}{\partial \lambda}W = \frac{\partial W}{\partial \lambda}(V - \lambda W) = 0.$$

При $\lambda \rightarrow 0$ имеем $\partial(VW)/\partial \lambda \sim V \partial W/\partial \lambda < 0$. При достаточно больших λ имеем $\partial W/\partial \lambda < 0$ и $V - \lambda W < 0$, поэтому $\partial(VW)/\partial \lambda > 0$, хотя и $\lim_{\lambda \rightarrow \infty} \partial(VW)/\partial \lambda = 0$, поскольку каждый из критериев стремится к постоянной величине. Таким образом, условие $\partial W/\partial \lambda = 0$, которому соответствует ОМУ ($\lambda = \infty$), не определяет решение задачи, а условие $V - \lambda W = 0$ равносильно уравнению (9) и доставляет минимум функции VW при некотором $\lambda > 0$, т.е. определяет решение. Теорема доказана.

3. Пример оценивания параметра сдвига косинусной модели

Рассмотрим задачу оценивания параметра сдвига θ косинусного распределения [15] с плотностью

$$f(x, \theta) = \frac{1}{l} \cos^2 \frac{\pi(x - \theta)}{2l} = \frac{1}{2l} \left[1 + \cos \frac{\pi(x - \theta)}{l} \right], \quad |x - \theta| \leq l. \quad (10)$$

Оценочные функции условно оптимальных оценок, совпадающие с функциями влияния, в модели (10) имеют вид [15]:

$$\psi(x, \theta) = \frac{l(\kappa + \sqrt{\kappa^2 - 1})}{\pi} \sin \frac{\pi(x - \theta)}{l} \left/ \left[\kappa + \cos \frac{\pi(x - \theta)}{l} \right] \right., \quad |x - \theta| \leq l, \quad (11)$$

где $\kappa = 2\lambda l + 1$ – параметр, используемый для более лаконичной записи (11). Поскольку модель (10) является финитной, для ее робастного оценивания следует применять подход, при котором допускаются наблюдения вне промежутка $|x - \theta| \leq l$, а оценочная функция вне данного промежутка принимается равной нулю [16].

Характеристики асимптотической дисперсии и неустойчивости оценочных функций (11) определяются выражениями [15]

$$V = \frac{l^2}{\pi^2} \left(1 + \sqrt{\frac{\kappa - 1}{\kappa + 1}} \right); \quad W = \frac{2l^3}{\pi^2} \left(1 + \frac{\kappa}{\sqrt{\kappa^2 - 1}} \right). \quad (12)$$

В частности, $V_{\text{ОМП}} = l^2/\pi^2$, $W_{\text{ОМУ}} = 4l^3/\pi^2$. В [15] также показано, что в семействе (11) компромиссной оценке (ОК) [4] соответствуют параметры $\lambda_{\text{ОК}} = V_{\text{ОМП}}/W_{\text{ОМУ}} = 1/(4l)$, $\kappa_{\text{ОК}} = 3/2$ и характеристики $\text{stb } \psi_{\text{ОК}} = 2/(1 + 3/\sqrt{5}) \approx 85,41\%$, $\text{eff } \psi_{\text{ОК}} = (1 + 1/\sqrt{5})^{-1} \approx 69,10\%$; равнооптимальной оценке (ОРО) [8] – значения $\lambda_{\text{ОРО}} = 1/(8l)$, $\kappa_{\text{ОРО}} = 5/4$ и характеристики $\text{stb } \psi_{\text{ОРО}} = \text{eff } \psi_{\text{ОРО}} = 75\%$.

Для нахождения равновесной оценки (ОРВ) решим уравнение (9). Используя (12), а также условие $\lambda > 0$ (или $\kappa > 1$), можно записать

$$\frac{V}{W} = \frac{\sqrt{\kappa^2 - 1} + 1 - \kappa}{2l} = \frac{\sqrt{4\lambda^2 l^2 + 4\lambda l}}{2l} - \lambda = \sqrt{\lambda^2 + \frac{\lambda}{l}} - \lambda = \lambda.$$

Сократим это уравнение на λ и приведем подобные: $\sqrt{1 + 1/(\lambda l)} = 2$. Отсюда получаем $\lambda_{\text{ОРВ}} = 1/(3l)$ и $\kappa_{\text{ОРВ}} = 5/3$. Вычислим некоторые характеристики равновесной оценки в модели (10):

$$W_{\text{OPB}} = \frac{9l^3}{2\pi^2}, \text{ stb } \psi_{\text{OPB}} = \frac{8}{9} \approx 88,89\%, V_{\text{OPB}} = \frac{3l^2}{2\pi^2}, \text{ eff } \psi_{\text{OPB}} = \frac{2}{3} \approx 66,67\%.$$

На рис. 1 показаны графики некоторых функций из семейства (11) при $l = 1$ и $\theta = 0$.

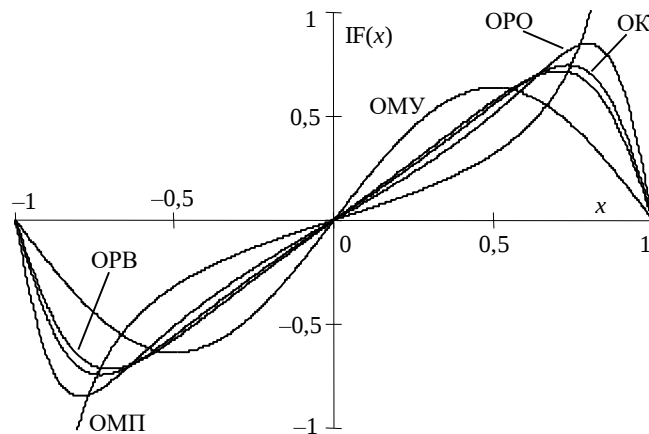


Рис. 1. Функции влияния некоторых условно оптимальных оценок
Fig. 1. Influence functions of some conditionally optimal estimators

Наиболее близкими между собой как визуально, так и по характеристикам оказались функции влияния ОК и ОПВ. Причем похожи и их определения: различие состоит только в том, что в выражении для $\lambda_{\text{ОК}}$ фигурируют минимальные значения показателей V и W , а в уравнении (9) – равновесные. Сходство оценок, с одной стороны, подтверждает разумность соответствующих эвристик, с другой – говорит о том, что компромиссная оценка [4], которая сравнительно просто определяется, в некоторых случаях может использоваться в качестве аппроксимации равновесной оценки.

Заключение

В работе предложен способ обоснованного выбора одной оценки из семейства условно оптимальных оценок [4], которая названа равновесной. Подобные методы выбора компромисса бывают востребованы в тех случаях, когда у исследователя отсутствует априорная информация, позволяющая объективно согласовать критерии, образующие семейство, особенно если критерии имеют несовпадающие размерности. Между тем ввиду некорректности задачи (наличия неустраняемой неопределенности [10]) все подобные методы оказываются в большей или меньшей степени эвристическими. Предложенный подход, в отличие от ранее известных, основан на оптимизационном принципе, который не привязан к минимальным значениям критериев (т.е. более универсальный) и позволяет естественным образом регуляризовать задачу. Благодаря этому полученное решение сбалансированно учитывает «интересы» каждого из критериев.

Также показано, что равновесная оценка минимизирует произведение критериев асимптотической дисперсии и неустойчивости.

Изложенные в работе теоретические построения проиллюстрированы на примере задачи оценивания параметра сдвига косинусного распределения [15]. Для данной модели построена равновесная оценка, вычислены ее характеристики, а также приведены графики функций влияния равновесной и некоторых других оценок из условно оптимального семейства.

В частности, рассмотренный пример показал, что предложенная в [4] компромиссная оценка по своим характеристикам и графику функции влияния оказалась довольно близкой к равновесной оценке. Поскольку компромиссная оценка определяется технически проще, она при необходимости может использоваться вместо равновесной оценки.

Список источников

1. Хьюбер П. Робастность в статистике. М. : Мир, 1984. 303 с.
2. Хампель Ф., Рончетти Э., Рауссеу П., Штаэль В. Робастность в статистике: подход на основе функций влияния. М. : Мир, 1989. 512 с.
3. Shulenin V.P. Robust methods of mathematical statistics. Tomsk : Scientific Technology Publishing House, 2020. 260 p.
4. Шурыгин А.М. Прикладная стохастика: робастность, оценивание, прогноз. М. : Финансы и статистика, 2000. 224 с.
5. Shevlyakov G., Morgenthaler S., Shurygin A. Redescending M -estimators // J. Statist. Plann. Inference. 2008. V. 138 (10). P. 2906–2917.
6. Лисицин Д.В., Гаврилов К.В. О свойствах условно оптимальных оценок // Научный вестник НГТУ. 2015. № 1 (58). С. 76–93.
7. Shevlyakov G.L., Oja H. Robust correlation: theory and applications. Chichester, West Sussex : John Wiley & Sons, 2016. 319 p.
8. Лисицин Д.В., Гаврилов К.В. Об устойчивом оценивании параметров модели при асимметричном засорении данных // Научный вестник НГТУ. 2008. № 1 (32). С. 33–40.
9. Горбунов В.М., Синокова Е.А. Практикум по дисциплине «Теория принятия решений» : учеб. пособие. Томск : Изд-во Том. политехн. ун-та, 2014. 126 с.
10. Подиновский В.В., Потапов М.А. Метод взвешенной суммы критериев в анализе многокритериальных решений: pro et contra // Бизнес-информатика. 2013. № 3 (25). С. 41–48.
11. Смоляк С.А., Титаренко Б.П. Устойчивые методы оценивания: статистическая обработка неоднородных совокупностей. М. : Статистика, 1980. 210 с.
12. Van der Vaart A.W. Asymptotic statistics. Cambridge : Cambridge University Press, 1998. 443 p.
13. DasGupta A. Asymptotic theory of statistics and probability. New York : Springer, 2008. 722 p.
14. Лисицин Д.В., Гаврилов К.В. Максиминная задача оценивания параметров в условиях байесовского точечного засорения // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2023. № 1 (62). С. 56–64.
15. Лисицин Д.В., Гаврилов К.В. Робастное оценивание финитной модели // Сборник научных трудов НГТУ. 2004. № 2 (36). С. 47–56.
16. Лисицин Д.В., Гаврилов К.В. Оценивание параметров финитной модели, устойчивое к нарушению финитности // Сибирский журнал индустриальной математики. 2013. Т. 16, № 2 (54). С. 109–121.

References

1. Huber, P. (1984) *Robastnost' v statistike* [Robust statistics]. Translated from English. Moscow: Mir.
2. Hampel, F., Ronchetti, E., Rousseeuw, P. & Stahel, W. (1989) *Robastnost' v statistike: podkhod na osnove funktsiy vliyaniya* [Robust statistics: The approach based on influence functions]. Translated from English. Moscow: Mir.
3. Shulenin, V.P. (2020) *Robust methods of mathematical statistics*. Tomsk: Scientific Technology Publishing House.
4. Shurygin, A.M. (2000) *Prikladnaya stokhastika: robastnost', otsenivanie, prognoz* [Applied stochastic: Robustness, estimation, forecast]. Moscow: Finansy i statistika.
5. Shevlyakov, G., Morgenthaler, S. & Shurygin, A. (2008) Redescending M -estimators. *Journal of Statistical Planning*. 138(10). pp. 2906–2917.
6. Lisitsin, D.V. & Gavrilov, K.V. (2015) O svoystvakh uslovno optimal'nykh otsenok [About properties of conditionally optimal estimates]. *Nauchnyy vestnik NGTU*. 58(1). pp. 76–93.
7. Shevlyakov, G.L. & Oja, H. (2016) *Robust correlation: theory and applications*. Chichester, West Sussex: John Wiley & Sons.
8. Lisitsin, D.V. & Gavrilov, K.V. (2008) Ob ustoychivom otsenivanii parametrov modeli pri asimmetrichnom zasoreнии dannyykh [On stable estimation of models parameters in presence of asymmetric data contamination]. *Nauchnyy vestnik NGTU*. 32(1). pp. 33–40.
9. Gorbunov, V.M. & Sinyukova, E.A. (2014) *Praktikum po distsipline "Teoriya prinyatiya resheniy"* [Workshop on the discipline "Decision making theory"]. Tomsk: Tomsk Polytechnic University.
10. Podinovski, V.V. & Potapov, M.A. (2013) Metod vzveshennoy summy kriteriev v analize mnogokriterial'nykh resheniy: pro et contra [Weighted sum method in the analysis of multicriterial decisions: pro et contra]. *Biznes-informatika – Business Informatics*. 25(3). pp. 41–48.
11. Smolyak, S.A. & Titarenko, B.P. (1980) *Ustoychivye metody otsenivaniya: statisticheskaya obrabotka neodnorodnykh sovkupnostey* [Stable Estimation Methods: Statistical Processing of Heterogeneous Aggregates]. Moscow: Statistika.
12. Van der Vaart, A.W. (1998) *Asymptotic statistics*. Cambridge: Cambridge University Press.
13. DasGupta, A. (2008) *Asymptotic Theory of Statistics and Probability*. New York: Springer.
14. Lisitsin, D.V. & Gavrilov, K.V. (2023) Maximin problem of parameter estimation in conditions of point Bayesian contamination. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 62(1). pp. 56–64. DOI: 10.17223/19988605/62/6
15. Lisitsin, D.V. & Gavrilov, K.V. (2004) Robastnoe otsenivanie finitnoy modeli [Robust estimation of a finite model]. *Sbornik nauchnykh trudov NGTU*. 36(2). pp. 47–56.

16. Lisitsin, D.V. & Gavrilov, K.V. (2013) Estimation of the parameters of a compactly-supported model stable under the violation of compact supportedness. *Sibirskiy zhurnal industrial'noy matematiki – Journal of Applied and Industrial Mathematics*. 16(2). pp. 109–121.

Информация об авторах:

Гаврилов Константин Викторович – кандидат технических наук, доцент кафедры автоматики Новосибирского государственного технического университета (Новосибирск, Россия). E-mail: aenigma77@mail.ru

Веретельникова Евгения Леонидовна – доцент, кандидат технических наук, доцент кафедры автоматики Новосибирского государственного технического университета (Новосибирск, Россия). E-mail: janever@mail.ru

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

Information about the authors:

Gavrilov Konstantin V. (Candidate of Technical Sciences, Associate Professor of the department of Automatics, Novosibirsk State Technical University, Novosibirsk, Russian Federation). E-mail: aenigma77@mail.ru

Veretel'nikova Evgeniya L. (Candidate of Technical Sciences, Associate Professor, Associate Professor of the department of Automatics, Novosibirsk State Technical University, Novosibirsk, Russian Federation). E-mail: janever@mail.ru

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

Поступила в редакцию 20.12.2023; принята к публикации 03.06.2024

Received 20.12.2023; accepted for publication 03.06.2024

Научная статья

УДК 004.94

doi: 10.17223/19988605/67/8

Вероятностный подход при контроле состояния сложной технической аппаратуры

Юлия Валентиновна Доронина¹, Александр Владимирович Скатков²

^{1,2} Севастопольский государственный университет, Севастополь, Россия

¹ apkSev@yandex.ru

² vm1945@mail.ru

Аннотация. Предложен вероятностный подход каскадно-иерархического типа для контроля состояния сложной технической аппаратуры при производстве с учетом видов контролируемых улов (программный, аппаратный и механический) и типов контроля: техническим средством и соответствующим уровнем квалификации персонала на основе полимоделирования. В отличие от детерминированного процесса контроля, когда последовательные испытания проводятся до получения некоторого исхода, предложенный подход позволил сформировать нестационарные правила оценивания вероятностей исходов процесса динамично, относительно меняющихся условий испытаний и среды. Применение меры Кульбака–Лейблера для оценивания дивергенции функций распределения вероятностей дефектов сложной технической аппаратуры, обладающей высокой чувствительностью, позволило сформировать схему двухконтурной системы поддержки принятия решений по динамическому планированию испытаний сложной технической аппаратуры, что в конечном счете повысило степень обоснованности принимаемых решений.

Ключевые слова: сложная техническая аппаратура; контроль состояния; вероятность обнаружения дефекта; полимоделирование; мера Кульбака–Лейблера.

Для цитирования: Доронина Ю.В., Скатков А.В. Вероятностный подход при контроле состояния сложной технической аппаратуры // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 67. С. 69–82. doi: 10.17223/19988605/67/8

Original article

doi: 10.17223/19988605/67/8

A probabilistic approach to monitoring the condition of complex technical equipment

Yulia V. Doronina¹, Alexander V. Skatkov²

^{1,2} Sevastopol State University, Sevastopol, Russia

¹ apkSev@yandex.ru

² vm1945@mail.ru

Abstract. A probabilistic cascade-hierarchical approach is proposed for monitoring the condition of complex technical equipment in production, taking into account the types of controlled catches (software, hardware and mechanical) and types of control: technical means and the appropriate level of qualification of personnel based on polymodeling. In contrast to the deterministic control process, when sequential tests are carried out until a certain outcome is obtained, the proposed approach made it possible to form non-stationary rules for estimating the probabilities of process outcomes dynamically, relative to changing test conditions and the environment. The use of the Kullback-Leibler measure to assess the divergence of the probability distribution functions of defects in complex technical equipment, having high sensitivity, allowed us to form a scheme of a two-circuit decision support system for dynamic test planning of complex technical equipment, which ultimately increased the degree of validity of the decisions made.

Keywords: complex technical equipment (CTE); condition monitoring; probability of defect detection; polymodeling; Kullback-Leibler measure.

For citation: Doronina, Yu.V., Skatkov, A.V. (2024) A probabilistic approach to monitoring the condition of complex technical equipment. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 67. pp. 69–82. doi: 10.17223/19988605/67/8

Введение

Планирование испытаний при производстве сложной технической аппаратуры (СТА) представляет собой нетривиальную задачу в современных условиях, характеризующихся требованиями максимального сокращения сроков освоения и дальнейшего развития новых изделий, повышения состояния СТА различных видов в рамках координации усилий по их созданию, сокращения стоимости испытаний, что порождает требование минимизации натурных испытаний, и др.

Обобщенными особенностями СТА являются:

- сложная структурная организация, в большинстве случаев иерархического типа;
- наличие в составе подсистем управления (особенно в случае автоматического управления) или подсистем приема управляющих сигналов от внешнего источника управления, что порождает внутреннюю видовую структурную сложность;
- наличие программируемых микропроцессорных элементов, которые требуют принципиально иных подходов к тестированию, отличных от аппаратной части, и др.

Таким образом, объекты испытаний – СТА – будем считать неоднородными, конструктивно разнообразными, выделяя при этом три основные составляющие: программную (П), аппаратную (А) и механическую (М) части. Эти особенности накладывают дополнительные требования и создают проблемы планирования, особенно в области летальных аппаратов ответственного применения.

В целом испытания – процесс определения параметров сложной аппаратуры по установленной методике с целью оценки их соответствия требованиям. Согласно ГОСТ 21194–87 наиболее распространенными при испытаниях являются следующие виды контроля: по способу контроля; по отношению к производственному процессу; по способу определения степени пригодности изделия к использованию; по степени охвата; по характеру воздействия контроля на производственный процесс; по периодичности контроля; по степени участия человека. Операции контроля в процессе производства: выходной контроль; контроль за соблюдением технологического процесса; анализ брака и дефектов; техническая приемка и испытания изделий [1–3].

При формировании планов испытаний СТА крайне важно учитывать интервалы стационарности их характеристик, что позволит получать объективные оценки состояния СТА. Например, часто не учитывается то, что процесс испытаний длительный и реализуется на разных этапах подготовки СТА к сдаче в эксплуатацию, вследствие чего меняющиеся показатели изделий, а также меняющиеся условия испытаний вносят существенные погрешности при оценивании состояния испытуемых СТА.

Источниками неопределенности при проведении и планировании испытаний СТА являются:

- 1) внешние физические условия проведения испытаний (например, погодные условия, скорость направление ветра и т.п.);
- 2) внешние причины, связанные с субъектом организации и проведения испытаний (знания, умения, особенности личности испытателя);
- 3) задачи испытаний, т.е. требуемый процент успешных испытаний на каждом их этапе, что может быть обусловлено нестационарностью задачи исследования характеристик СТА;
- 4) другие, определяемые видом СТА.

Представляется перспективным с целью компенсации этих проблем использовать информационные подходы и адаптивные методы, которые позволяют накапливать информацию и принимать более обоснованные решения, исходя из динамики проведения испытаний СТА.

Очевидно, что испытание прекращается по трем возможным причинам:

- 1 – достигнут допустимый уровень знаний об объекте (энтропия близка к нулю);
- 2 – об объекте в процессе испытаний не получено новой информации (исходная энтропия СТА не изменилась и испытания не привели к пониманию состояния СТА);

3 – качество СТА ухудшилась (оценки энтропии СТА увеличились);

4 – в случае цензурированной выборки, элементами которой являются полные наработки и наработки до цензурирования (неполные наработки) или только значения наработки до цензурирования [4].

Задача проведения испытаний СТА в общем виде представляет собой процесс, на вход которого поступают требования из множества $z = \{\text{найти вероятность успешной реализации функции системы в } i\text{-м испытании; найти вероятность успешной реализации функции системы в комплексе испытаний; обеспечить проведение испытаний в течение заданного времени } T; \text{ учесть неопределенность при проведении испытаний в виде некоторого параметра } \xi; \text{ сформировать траекторию испытаний СТА с изменяющимися параметрами процесса и т.п.}\}$.

С учетом требований к испытаниям формируется множество результатов, которые позволяют осуществить поддержку принятия решений лицом, принимающим решения (ЛПР), например, по изменению длительности испытаний. Примеры наиболее типичных ситуаций, в которых затруднено принятие решений, предложены в следующих формулировках.

Ситуация 1: при проведении 6 испытаний СТА из 10 запланированных все дали положительный результат, после чего принято решение о прекращении дальнейших испытаний.

Ситуация 2: при проведении 6 испытаний СТА из 10 запланированных 5 дали отрицательный результат, после чего принято решение о прекращении испытательного процесса для дополнительной доработки СТА.

Ситуация 3: проведено 8 испытаний из 10 запланированных, 4 дали положительный, 4 – отрицательный результат, целесообразно продолжить испытания.

Таким образом, возникают вопросы: Какова степень обоснованности принимаемых решений? Каковы оценки вероятностей ошибок 1-го и 2-го рода для случаев, отмеченных в примерах? Какова динамика вероятностей отрицательного результата испытаний СТА в зависимости от объема контролируемой партии изделий? Какое количество информации извлечено из этих типовых ситуаций?

Следовательно, важными представляются учет динамичности процесса испытания СТА, учет неопределенности условий проведения испытаний, что позволит в ряде случаев оптимизировать процесс испытаний СТА, например, на основе сокращения сроков сдачи СТА заказчику, либо, напротив, обосновать необходимость увеличения длительности испытаний или их этапов.

Проблемам испытаний изделий СТА посвящено много публикаций. Так, в работах [1, 2] приводится классификация испытаний по назначению (ГОСТ 16504.81), дифференцируются исследовательские, контрольные, сравнительные, определительные, специальные. В работе [4] показано, что минимум дисперсии прогнозной оценки обеспечивается за счет оптимального распределения испытаний в области критической нагрузки объекта, что характерно для критических состояний, однако не имитирует стационарный режим функционирования объектов. Исследования, касающиеся оценивания состояния изделий СТА, приводятся, например, в [5–7]. В [5] – на основе марковской модели с непрерывным временем и одновременного применения экспериментального и расчетно-экспериментального метода оценки показателей надежности и других параметров, при этом уравнения для расчетно-экспериментального метода получены на основе описания процесса эксплуатации объекта матричной моделью. В источниках [6, 7] приводятся оценки, позволяющие оценивать среднюю наработку до отказа, заданную конечной величиной при безотказных испытаниях, проводимых по биномиальному плану, оценка построена в предположении, что наработка до отказа имеет экспоненциальное распределение. Расчет значений потока отказов в [7] сделан для каждого объекта при условии, что статистических данных достаточно для выполнения расчета (зафиксировано 3 и более отказа); по каждому показателю надежности предложено вычислять точечную оценку и ее доверительные интервалы.

Задача определения моментов стохастических различий функций распределения решалась многими авторами на основе методов, представленных, например, в публикациях [8–13]. Многие авторы сводили эту задачу к оценке квалитетических свойств модели [13], предложено выполнять сопоставление модели объекту по целостному свойству через измерение расстояния между соответствующими точками в n -мерном пространстве элементарных свойств объекта.

В связи с преодолением ряда отмеченных трудностей рекомендуется в некоторых случаях применение метода Вальда, который рассматривается многими авторами как основа анализа процесса последовательных испытаний; он позволяет снизить риск принятия неправильных решений, что особенно важно в ситуациях, когда последствия ошибок могут быть значительными, в методе не предусматривается заранее фиксированного объема выборки n и решение о прекращении испытаний выносится в зависимости от проведенных испытаний [6]. Однако метод последовательных испытаний Вальда недостаточно конструктивен в связи с тем, что вероятность исхода испытаний в большинстве случаев непостоянна.

Идеи полимоделирования (мультимоделирования) рассматривались в [14] и эффективно применяются для исследования сложных систем, когда одной модели оказывается недостаточно и требуется несколько иногда разнотипных моделей. Их сопряжение, топологии рассматриваются в [15], где авторами предложена методология построения полимодельных комплексов в рамках каскадно-иерархических структур для моделирования ресурсных характеристик сложных систем при исследовании динамики деградационных отказов и пополнения ресурсов. Предложенный подход позволил в параметрическом модельном пространстве строить планы исследований сложных систем и на основе принципа перманентности оценочного функционала сохранять константность представимости системных характеристик при фазовом укрупнении пространства моделирования, что для данного исследования оказалось полезным при построении двухконтурной системы поддержки принятия решения при планировании испытаний по оценке состояний СТА.

В целом в исследованиях рассматриваемой области можно выделить следующие особенности: методы в большинстве случаев имеют высокую общность и, следовательно, малоприменимы к конкретным видам изделий либо, напротив, ориентированы сугубо на конкретные виды СТА. Авторы настоящего исследования считают целесообразным выделить два основных направления планирования испытаний изделий СТА:

- 1) применение аналитического моделирования с использованием естественного для классических случаев решения данной задачи – оценки энтропии по результатам испытаний сложной аппаратуры – биномиального распределения вероятности в методе Вальда;
- 2) применение моделирования для оценки энтропии по результатам испытаний сложной аппаратуры на основе меры Кульбака–Лейблера.

Исследованию этих направлений при решении задачи повышения эффективности проведения испытаний СТА, планирования этого процесса, ситуационного управления и соответствующего контроля состояния СТА посвящена настоящая статья, в том числе для создания современных систем поддержки принятия решений в этой области на основе систем искусственного интеллекта.

Цель статьи – предложить вероятностную схему контроля состояния сложной технической аппаратуры, учитывающую возможную неопределенность условий проведения испытаний и нестационарность их параметров для поддержки принятия решений ЛПР по дальнейшему проведению испытаний.

1. Математическая модель оценки результатов контроля сложной аппаратуры на основе применения биномиального распределения вероятности

При построении модели рассматривается три состояния СТА: исправное, неисправное состояние (наличие одного или нескольких дефектов – в дальнейшем будет использоваться определение дефекта как комплексного обозначения неисправности СТА) и неопределяемое, которое может быть сформулировано по результатам кластеризации множества состояний СТА.

Рассмотрим процесс испытаний СТА системой контроля состояния (СТК), в результате которого состояние СТА может быть определено как дефектное или исправное с вероятностью p_1 или $(1 - p_1)$ соответственно; p_2 – вероятность обнаружения дефекта СТА (вероятность не обнаружения, соответственно, $(1 - p_2)$). Тогда по результатам текущего испытания мы можем получить один из 4 исходов:

- 1) $p_1 \cdot p_2$ – дефект есть, и он обнаружен;

2) $p_1 \cdot (1 - p_2)$ – дефект есть, и это не обнаружено (ситуация возможна в сложном случае скрытых дефектов);

3) $(1 - p_1) \cdot p_2$ – дефект отсутствует, и это подтверждено результатами контроля;

4) $(1 - p_1) \cdot (1 - p_2)$ – дефект отсутствует, и это не подтверждено (возможно при оценивании качества складываемых изделий; в данном исследовании не рассматривается).

Обозначим априорные вероятности дефектов: p_{Π} – в программной (П), p_A – в аппаратной (А) и p_M – в механической (М) частях СТА; их сумма определяет априорную вероятность дефектов всех видов СТА.

Рассмотрим гипотезы: H_1 – гипотеза, определяющая наличие дефекта при испытании, H_2 – гипотеза, определяющая отсутствие дефекта при испытании. Рассмотрим сложное событие V_l в некоторый фиксированный момент времени, которое может быть представлено в логическом виде и которое отражает выполнение испытаний как полная сумма событий: $V_l = V(S_1) \vee V(S_2) \vee V(S_3)$, $l \in L$, где $V(S_1), V(S_2), V(S_3)$ – события обнаружения дефектов в программной, аппаратной или механической частях СТА соответственно.

Обозначим $V(S_1)|H_1, V(S_1)|H_2, V(S_2)|H_1, V(S_2)|H_2, V(S_3)|H_1, V(S_3)|H_2$ условные вероятности результата испытаний из-за событий $V(S_1), V(S_2), V(S_3)$ соответственно и их вероятности: $P(V(S_1)|H_1), P(V(S_1)|H_2), P(V(S_2)|H_1), P(V(S_2)|H_2), P(V(S_3)|H_1), P(V(S_3)|H_2)$; обозначим $P(H_1), P(H_2)$ – вероятности событий, соответствующих гипотезам H_1, H_2 .

Используя формулу полной вероятности, получим

$$\begin{aligned} P(V(S_1)) &= \sum_i^N P(H_i) \cdot P(V(S_1)|H_i) = P(H_1) \cdot P(V(S_1)|H_1) + P(H_2) \cdot P(V(S_1)|H_2), \\ P(V(S_2)) &= \sum_i^N P(H_i) P(V(S_2)|H_i) = P(H_1) \cdot P(V(S_2)|H_1) + P(H_2) \cdot P(V(S_2)|H_2), \\ P(V(S_3)) &= \sum_i^N P(H_i) P(V(S_3)|H_i) = P(H_1) \cdot P(V(S_3)|H_1) + P(H_2) \cdot P(V(S_3)|H_2), \quad l \in \{L\}, i = \overline{1, N}. \end{aligned} \quad (1)$$

Оценка энтропии состояния СТА определяется на основе (1):

$$H_{V(S_1)} = -P(V(S_1)) \cdot \ln P(V(S_1)); H_{V(S_2)} = -P(V(S_2)) \cdot \ln P(V(S_2)); H_{V(S_3)} = -P(V(S_3)) \cdot \ln P(V(S_3)). \quad (2)$$

Семейство равенств (2) является энтропийной моделью состояний СТА при контроле. На ее основе возможны постановки задач анализа, управления и оптимизации процесса проведения контроля и испытаний. На рис. 1 приведены графики энтропии состояния СТА в зависимости от числа испытаний при контроле (N). С ростом вероятности наличия дефекта в СТА (p_1) энтропия состояния заметно растет (см. рис. 1, б).

Работа системы по оценке состояния СТА поддерживается техническими средствами Tc и персоналом Kp , тогда результат контроля состояния СТА определится декартовым произведением $KK_{ij} = Tc_i \times Kp_j$. При идентификации дефектов q типами технических средств и h типами различных уровней квалификации персонала, получим матрицу вероятностей:

$$KK = \|kk_{ij}\|_{h,q} = \begin{matrix} & \begin{matrix} Tc_1 & \dots & Tc_i & \dots & Tc_q \end{matrix} \\ \begin{matrix} Kp_1 \\ \dots \\ Kp_j \\ \dots \\ Kp_h \end{matrix} & \begin{bmatrix} KK_{11} & \dots & KK_{i1} & \dots & KK_{1q} \\ \dots & \dots & \dots & \dots & \dots \\ KK_{j1} & \dots & KK_{ij} & \dots & KK_{jq} \\ \dots & \dots & \dots & \dots & \dots \\ KK_{h1} & \dots & KK_{hj} & \dots & KK_{hq} \end{bmatrix} \end{matrix}. \quad (3)$$

Запись KK_{12} означает, что контроль состояния выявил дефект техническим средством типа 1 и квалификацией персонала типа 2. На основе (3) формируется три матрицы, связанных с выявлением дефектов в различных частях СТА (А, П, М): $KK^{\Pi} = \|kk_{ij}^{\Pi}\|$, $KK^A = \|kk_{ij}^A\|$, $KK^M = \|kk_{ij}^M\|$, размерности m и n вариативны относительно вида СТА.

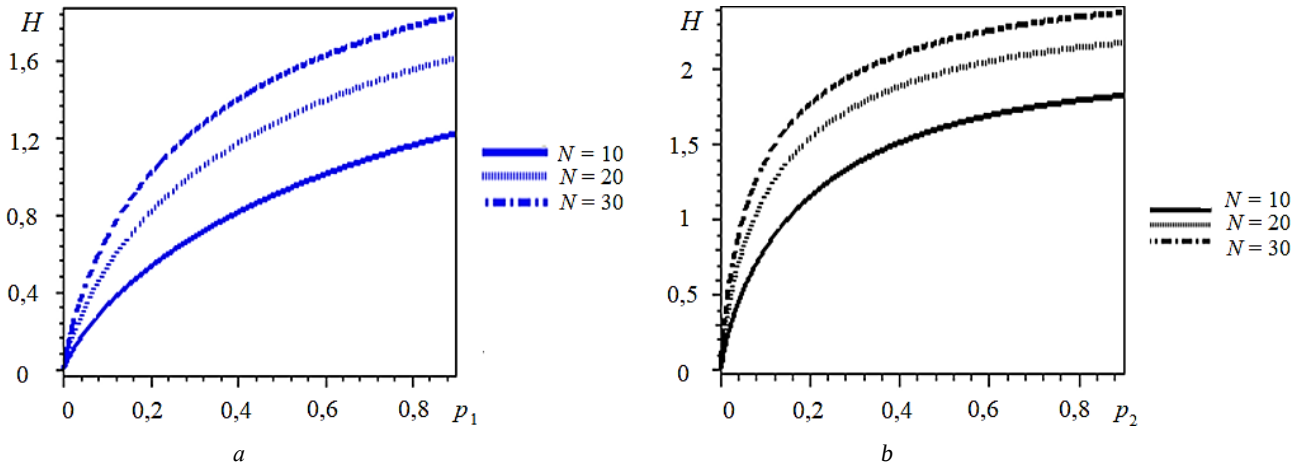


Рис. 1. Энтропия состояний сложной технической системы в зависимости от вероятности обнаружения дефекта при числе испытаний (N) при условии, что вероятность наличия дефекта: $p_1 = 0,1$ (a); $p_1 = 0,4$ (b)

Fig. 1. Entropy of states of complex technical equipment depending on the probability of detecting a defect with a number of tests (N), provided that the probability of a defect: a) $p_1 = 0,1$ b) $p_1 = 0,4$

Платежные матрицы затрат по уменьшению энтропии контроля состояния СТА имеют вид: $ZKK^\Pi = \|zkk_{ij}^\Pi\|$, $ZKK^A = \|zkk_{ij}^A\|$, $ZKK^M = \|zkk_{ij}^M\|$, где zkk_{ij}^Π , zkk_{ij}^A , zkk_{ij}^M – затраты при проведении контроля СТА конкретного вида i -м типом технического средства и j -м типом уровня квалификации персонала.

Следовательно, вероятность обнаружения дефекта СТА p_2 представляет собой полную вероятность по событиям, соответствующим гипотезам H_1 , H_2 с учетом KK^Π , KK^A , KK^M и / или ZKK^Π , ZKK^A , ZKK^M . Учитывая формулы полной вероятности (1) и матрицу вероятностей (2), найдем вероятности того, что событие наличия дефекта в процессе испытаний каждой СТА проявится k раз. Это может быть определено на основе биномиального закона распределения случайной величины (СВ):

$$P(\xi = k) = C_d^k (p_1 \cdot \sum_{Tc_i=1}^q \sum_{Kp_j=1}^h (KK^\Pi \cdot KK^A \cdot KK^M))^k \left(1 - (p_1 \cdot \sum_{Tc_i=1}^q \sum_{Kp_j=1}^h (KK^\Pi \cdot KK^A \cdot KK^M)) \right)^{d-k}, \quad (4)$$

где p_1 и d – параметры биномиального распределения.

Таким образом, согласно (4) стохастической схеме контроля состояния СТА соответствует модель, в которой возможно как управлять параметрами типов контроля (техническими средствами, требуемой квалификацией персонала) и видами дефектов в различных частях СТА, так и решать оптимизационные задачи по управлению требуемыми затратами по каждому из типов локализации дефектов zkk_{ij}^Π , zkk_{ij}^A , zkk_{ij}^M , в том числе, например, на основе минимаксного критерия Вальда.

Результаты численного моделирования вероятностей исходов вариантов контроля состояния СТА на основе биномиального распределения с учетом различных d (число испытаний), вероятности наличия дефекта $p_1 = 0,4$ и вариаций вероятностей обнаружения дефекта СТА (p_2), выполненного с помощью инструмента Maple, приведены на рис.2.

Показанные на рис. 2 результаты моделирования апостериорных вероятностей дефектов СТА с учетом вероятностей обнаружения p_2 позволяют видеть сглаживание формы графиков. При плане испытаний $d = 10$ вероятности того, что событие наличия дефекта в процессе испытаний каждой СТА проявится k раз более, сконцентрированы около математического ожидания и имеют малые k относительно планов испытаний $d = 15$ и $d = 30$. Наблюдается зависимость: чем больше выборка (объем плана испытаний), тем больше информации получит ЛПР, энтропия событий (в данном случае относительно обнаружения дефектов изделий) снижается, однако растет стоимость испытаний.

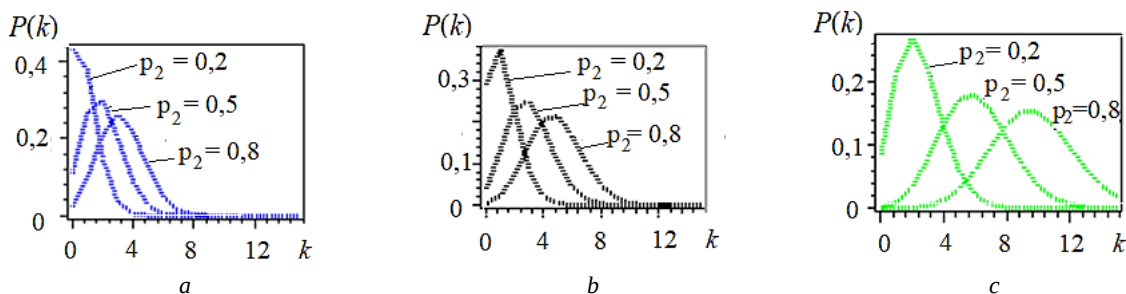


Рис. 2. Распределение вероятностей для вариантов контроля состояния СТА для $p_1 = 0,4$, $p_2 = 0,2; 0,5; 0,8$:

a – при $d = 10$, b – при $d = 15$, c – при $d = 30$

Fig. 2. Probability distribution for quality control options CTE for $p_1 = 0,4$, $p_2 = 0,2; 0,5; 0,8$:

a) at $d = 10$, b) at $d = 15$, c) at $d = 30$

Таким образом, остается проблема определения оптимального плана испытаний для снижения их стоимости при определенном уровне апостериорной энтропии процесса испытаний. Кроме того, рассмотренные особенности процесса испытаний и учет уровней персонала, типов средств контроля испытаний приводят к необходимости построения полимодельного комплекса исследования процесса испытаний СТА. Схема полимоделирования в рамках предложенного метода представлена на рис. 3.

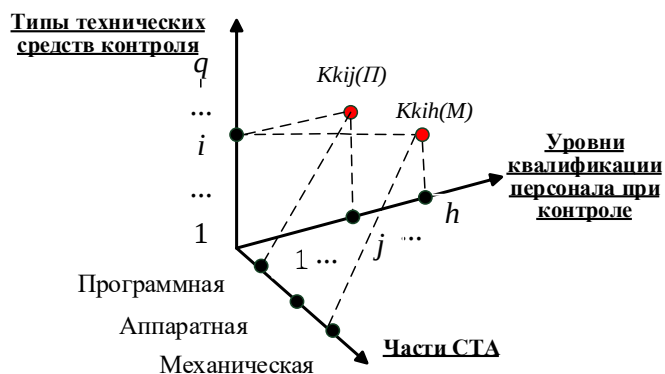


Рис. 3. Схема полимоделирования в стохастическом методе контроля состояния СТА

Fig. 3. Polymodeling scheme in the probabilistic state control CTE method

Можно видеть, что полимодельный комплекс определяется тремя факторами: типами технических средств контроля состояния, уровнями квалификации персонала при контроле состояния и видами локализации дефектов в СТА. Таким образом, схема полимоделирования имеет каскадно-иерархическую топологию [14].

Рассмотренная модель испытаний СТА построена в условиях некоторых допущений, поскольку вероятность события неисправности в результатах испытаний зависит от многих факторов, в том числе определяется непостоянством условий проведения экспериментов, старением или износом аппаратуры, в том числе испытательных стендов и т.п. Исследование причин появления этих факторов не является задачей настоящей статьи. Поскольку модель на основе биномиального распределения не позволяет в достаточной степени поддержать процесс принятия решений, далее рассматривается применение более тонких механизмов поддержки принятия решений по планированию испытаний СТА на основе меры Кульбака–Лейблера.

2. Математическая модель оценки результатов контроля состояния сложной аппаратуры на основе применения меры Кульбака–Лейблера

При построении модели принят тот факт, что вероятность исхода процесса контроля состояния СТА p не предполагается постоянной; возможный диапазон ее изменения: $p \pm \delta, \delta \leq \tilde{\delta}$, где δ – текущее отклонение, а $\tilde{\delta}$ – допустимая его граница. Например, среднее значение вероятности обнаружения

дефектов i -го типа в процессе контроля равно 0,3, при росте вероятности дефекта до 0,4 наблюдается изменение состояния СТА (по одному или группе параметров). Как было рассмотрено ранее, причины важности применения высокочувствительных методов при проведении контроля состояния СТА связаны с многими факторами, такими как типы контроля, виды контролируемых узлов СТА и др.

В простейшем случае рассматриваются два этапа испытаний: начальный и заключительный, которые характеризуются естественным снижением вероятности обнаружения дефектов γ вследствие завершения процесса доработки СТА. Отметим, что в более сложных случаях таких этапов может быть больше. Предположим, что случайная величина (СВ), определяющая изменчивость p , распределена по нормальному закону и для 10 натурных испытаний двух образцов разных уровней сложности (I тип сложности; II тип сложности) получены следующие результаты в виде частот появления вероятностей γp , изображенные на рис. 4.

Введенные градации типов сложности СТА (I – конструктивно монотонная, сложная, II – конструктивно немонотонная, сложная) позволяют иллюстрировать истоки заметных различий в результатах оценок состояния СТА (см. рис. 4, а, б). Под конструктивной монотонностью СТА в статье понимается наличие подавляющего числа однотипных конструктивных элементов.

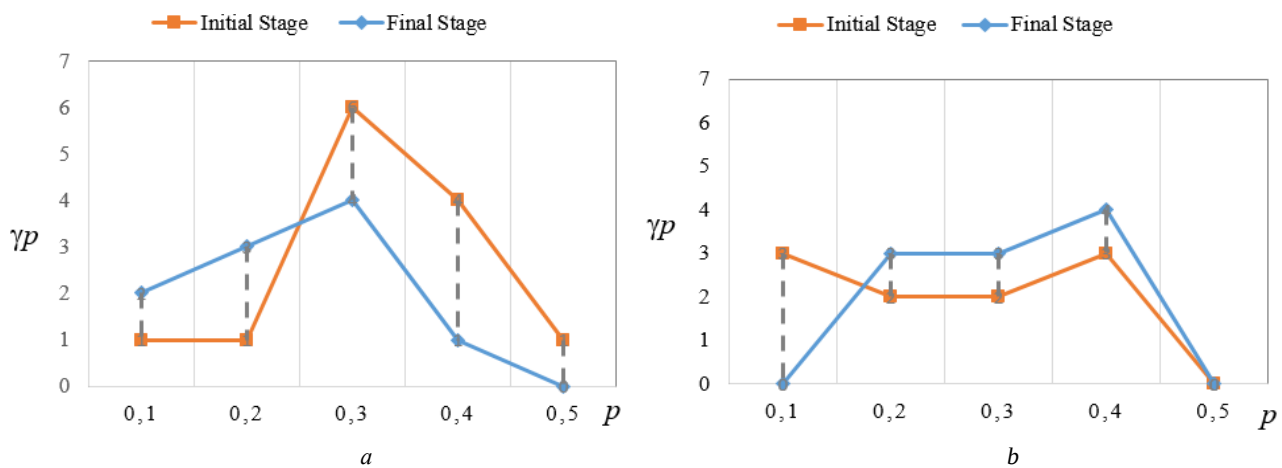


Рис. 4. Частоты обнаруженных дефектов при анализе состояния сложной технической аппаратуры в серии из 10 натурных экспериментов (средняя вероятность обнаружения дефектов $p_2 = 0,31$) на начальном и заключительном этапах испытаний для двух образцов СТА: а – I тип; б – II тип
Fig. 4. The frequency of detected defects in the analysis of the state of complex technical equipment in a series of 10 field experiments (average probability of the probability of detecting defects $p_2 = 0,31$) at the initial and final stages of testing for two samples CTA: а) I type; б) II type

Как видно из результатов обработки натурных испытаний при контроле состояния СТА, приведенных на рис. 4, различия в числе полученных значений вероятностей наличия дефектов на начальном и заключительном этапах производства СТА весьма заметны и составляют для одного типа от 1 до 3 (см. рис. 4, а), для СТА другого типа график различий имеет иной вид, несмотря на то что средние вероятности обнаружения дефектов остались прежними, частоты дефектов не уменьшились, но не превышают значения в первой группе испытаний, а частоты на начальном этапе несколько ниже. В этой связи представляется важным понимать состояние СТА на разных временных отрезках (интервалах) производства на основе определения различия процессов начального и заключительного этапов процесса создания СТА.

Представим набор возможных состояний СТА при проведении r -го испытания четверкой:

$$Q_r \langle M, p_r, \alpha, \xi_r \rangle, \quad (5)$$

где $r \in \overline{1, m}$ (m – количество испытаний), M – математическое ожидание вероятностей события наличия дефектов СТА, p_r – апостериорные вероятности дефектов СТА в r -м испытании, α – коэффициент значимости важного конструктивного параметра (параметров) СТА, ξ_r – коэффициент влияния случайного фактора при проведении k -го испытания.

Тогда состояние СТА после серии из m испытаний с учетом (5) представимо в виде кортежа:

$$Q_{m,r} \langle M_m, M(p_n), \alpha, M(\xi_n), m, N \rangle, \quad (6)$$

где $M(p_r)$ – математическое ожидание вероятностей наличия дефектов СТА в серии из m испытаний и партии N изделий, $M(\xi_n)$ – математическое ожидание наличия случайного фактора в серии из m испытаний и партии N изделий.

В первом приближении ξ_r и $M(\xi_n)$ будем считать некоторыми константами, а конструктивные параметры СТА постоянными в серии испытаний, тогда можно видеть, что состояние СТА определяется $M(p_n)$ или p_r .

В рассмотренных на рис. 3 результатах натурных испытаний речь идет о двух этапах: начальном и заключительном, – иллюстрирующих явные изменения состояний СТА, однако моменты этих изменений лежат в области неопределенности, и именно поэтому выбирались явно выраженные состояния объекта относительно жизненного цикла их производства. Предлагаемое в данной статье использование меры Кульбака–Лейблера дает возможность получать количественные оценки возможных границ состояния СТА за счет оценки расхождения между функциями распределения (ФР) отрицательных результатов при испытаниях с любым законом распределения СВ в рамках близлежащих этапов производства СТА. Основными преимуществами использования меры Кульбака–Лейблера в задаче определения расстояний между ФР СТА являются чувствительность к изменениям состояния системы, малая вычислительная трудоемкость, адаптивность к внешним воздействиям.

Информационная мера Кульбака–Лейблера (расстояние Кульбака–Лейблера; ИМКЛ) распределений p_l относительно q_l рассчитывается следующим образом: $\sum_{l=1}^K (p_l - q_l) \log \frac{p_l}{q_l}$ [11].

На основе ИМКЛ и с учетом того, что в результате оценивания состояния СТА вероятность обнаружения дефектов СТА в k -м испытании изменяется в пределах заданной (требуемой) точности ε , имеем $|P(r) - P(r_0)| < \varepsilon$; выражение ИМКЛ в задаче идентификации состояний СТА запишется в виде:

$$D(P_0(m, \sigma_0, \alpha, \xi) \| P_r(m, \sigma_r, \alpha, \xi_r)) = \sum_{j=1}^R (P_i(M, \sigma_0, \alpha, \xi) - P_i(M, \sigma_j, \alpha, \xi_r)) \log \left(\frac{P_0(M, \sigma_0, \alpha, \xi)}{P_i(M, \sigma_j, \alpha, \xi_k)} \right) < \varepsilon, \quad (7)$$

где $D(P_0(M, \sigma_0, \alpha, \xi) \| P_r(M, \sigma_r, \alpha, \xi_r))$ – дивергенция Кульбака–Лейблера (ДКЛ) как мера расстояния распределений вероятностей P_r относительно P_0 в r -м испытании из R ; σ_0, σ_j – априорное и апостериорное среднее квадратическое отклонение математического ожидания СВ дефекта СТА соответственно; α_0, α – априорный и апостериорный случайные параметры СТА соответственно (в общем случае – вектор).

Значение ε выбирается по согласованию с заказчиком, так же как и σ_0, σ_j , что будет влиять на чувствительность метода оценивания перехода объекта в новое состояние $Q_s, s \in \{S\}$ – множество наблюдаемых состояний СТА. Учитывая предложенные градации типов сложности СТА (I – конструктивно монотонная, сложная, II – конструктивно немонотонная, сложная), введем коэффициент сложности β_I, β_{II} , что отразится в выражении (7):

$$D_I = \beta_I \sum_{j=1}^K (P_0(m_{\beta_I}, \sigma_0, \beta_I \alpha, \xi) - P_i(M_{\beta_I}, \sigma_j, \beta_I \alpha, \xi_r)) \log \left(\frac{P_0(M_{\beta_I}, \sigma_0, \beta_I \alpha, \xi)}{P_i(M_{\beta_I}, \sigma_j, \beta_I \alpha, \xi_r)} \right) < \varepsilon_{\beta_I},$$

$$D_{II} = \beta_{II} \sum_{j=1}^K (P_0(M_{\beta_{II}}, \sigma_0, \beta_{II} \alpha, \xi) - P_i(M_{\beta_{II}}, \sigma_j, \beta_{II} \alpha, \xi_r)) \log \left(\frac{P_0(M_{\beta_{II}}, \sigma_0, \beta_{II} \alpha, \xi)}{P_i(M_{\beta_{II}}, \sigma_j, \beta_{II} \alpha, \xi_r)} \right) < \varepsilon_{\beta_{II}}, \quad (8)$$

где D_I, D_{II} – ДКЛ распределений вероятностей P_k относительно P_0 в k -м испытании из N I и II типов соответственно, а $M_{\beta_I}, M_{\beta_{II}}, \varepsilon_{\beta_I}, \varepsilon_{\beta_{II}}$ – уточненные математические ожидания и требуемые точности ДКЛ каждого типа. Эти уточнения распространяются на СТА, включенные в выбранные I и II типы на основе экспертных решений ЛПР.

Целью моделирования ДКЛ на основе (8) является получение количественных оценок возможного изменения состояния СТА за счет оценки расхождения ФР вероятностей дефектов СТА при оценке состояний с ограничением только возможностями генераторов (псевдогенераторов) случайных чисел, что в дальнейшем позволит принимать обоснованные решения по управлению процессом контроля качества в частности и испытаний СТА в целом.

Для получения числовых оценок различия процессов начального и заключительного этапов производства СТА воспользуемся формулой Кульбака–Лейблера. В табл. 1 приведены результаты имитационного моделирования различия ФР вероятностей дефектов СТА при контроле состояния, а на рис. 5, а – графическое представление дивергенции Кульбака–Лейблера при различных стандартных отклонениях σ : 0,01; 0,02; 0,03. При увеличении числа испытаний до 30 получены результаты, отражающие в том числе область предполагаемого изменения состояний СТА (табл. 2, рис. 5, б).

Таблица 1

Результаты имитационного моделирования

m	0,30	0,301	0,302	0,303	0,304	0,305	0,306	0,307	0,308	0,309	0,31	0,41	0,42	0,43	0,44
$\sigma = 0,03$	0,07922	0,10486	0,08239	0,08653	0,05233	0,05653	0,08957	0,06443	0,08007	0,06757	0,16048	0,20687	0,26639	0,27749	0,27522
$\sigma = 0,05$	0,21129	0,13775	0,12111	0,06826	0,09925	0,05799	0,11183	0,10376	0,11551	0,0806	0,31385	0,33443	0,35724	0,41064	0,43261
$\sigma = 0,01$	0,00385	0,00519	0,00851	0,00513	0,00341	0,00182	0,00299	0,00356	0,00253	0,00266	0,08455	0,11195	0,12346	0,15411	0,18155

Таблица 2

Результаты имитационного моделирования

M	0,301	0,302	0,303	0,304	0,305	0,306	0,307	0,30S	0,309	0,31	0,311	0,312	0,315	0,41
$m = 20$	0,00972	0,01302	0,0103	0,00757	0,01241	0,00707	0,01038	0,00801	0,01179	0,00955	0,01008	0,01253	0,0115	0,02698
$m = 30$	0,01412	0,00972	0,01284	0,01086	0,0138	0,01349	0,00798	0,01244	0,00711	0,01324	0,00998	0,0081	0,01426	0,02707

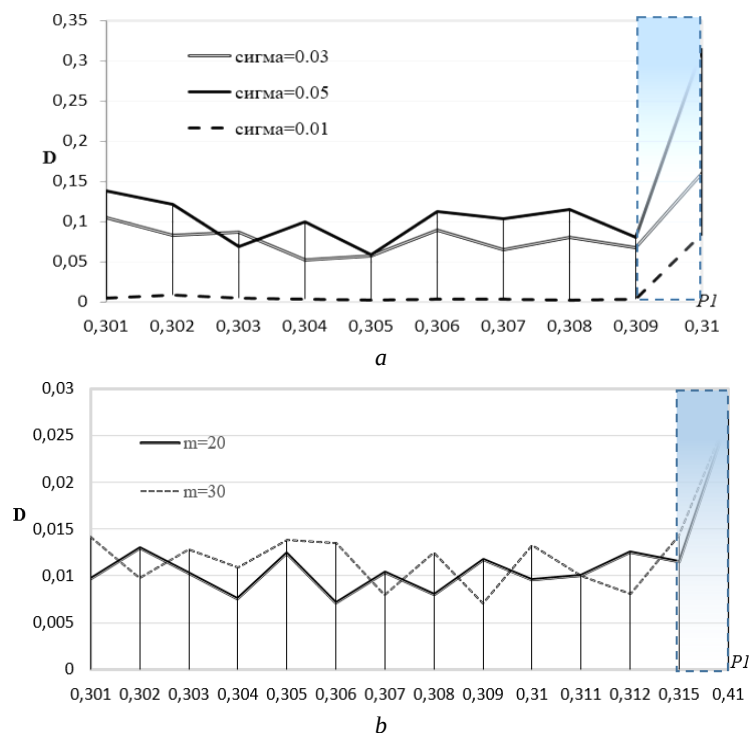


Рис. 5. Результаты имитационного моделирования зависимости дивергенции функций распределения вероятностей дефектов при анализе состояния сложной технической аппаратуры:

а – для $\sigma = 0,01; 0,02; 0,03$; б – в зависимости от числа испытаний m

Fig. 5. The results of simulation modeling of the dependence of the divergence of distribution functions the probabilities of defects in the analysis of the state of complex technical equipment:

а) for $\sigma = 0,01; 0,02; 0,03$; б) depending on the number of tests m

Результаты имитационного моделирования меры различия функций распределения вероятностей дефектов СТА, приведенные на рис. 5, а, показывают, что при изменении стандартного отклонения σ от 0,01 до 0,03 наблюдается рост дивергенции Кульбака–Лейблера начиная с $M = 0,309$, что не противоре-

чит начальной гипотезе об изменении состояния СТА при переходе к вероятности неисправности от 0,3 к 0,31. На рис. 5, а область предполагаемого изменения состояния СТА выделена затемнением.

Результаты имитационного моделирования меры различия функций распределения вероятностей дефектов сложной технической аппаратуры, приведенные на рис. 5, б, показывают, что рост дивергенции Кульбака–Лейблера сдвигается при увеличении объема испытаний, для $m = 10$ – начиная с $M = 0,309$, к $M = 0,14$ (при $m = 20$) и $M = 0,15$ (при $m=30$). Область предполагаемого изменения состояния СТА выделена затемнением. Эти результаты также не противоречат начальной гипотезе об изменении состояния СТА при переходе к вероятности неисправности от 0,3 к 0,4, но и дают возможность определять изменение состояния СТА с высокой чувствительностью.

Таким образом, планирование испытаний СТА реализуется на основе рассчитываемых границ их состояний, и в зависимости от требуемой точности может быть обосновано необходимое количество испытаний m , т.е. получена возможность управления m за счет ε и σ . Безусловно, в более сложных случаях управляющими параметрами являются α и учет неопределенности в виде ξ .

3. Структура системы поддержки принятия решения при планировании испытаний по оценке состояний сложной технической аппаратуры

Предложенный вероятностный подход при контроле состояния СТА, использующий описанную имитационную модель с учетом применения ДКЛ, позволил построить двухконтурную структурную схему системы поддержки принятия решения (СППР) по определению границ изменения состояний при планировании испытаний сложной технической аппаратуры со следующими параметрами: количество испытаний m , стандартное отклонение σ , математическое ожидание дефектов СТА, отклонение ε и др., что позволит формировать высокочувствительный анализ динамики состояний СТА при контроле (рис. 6).

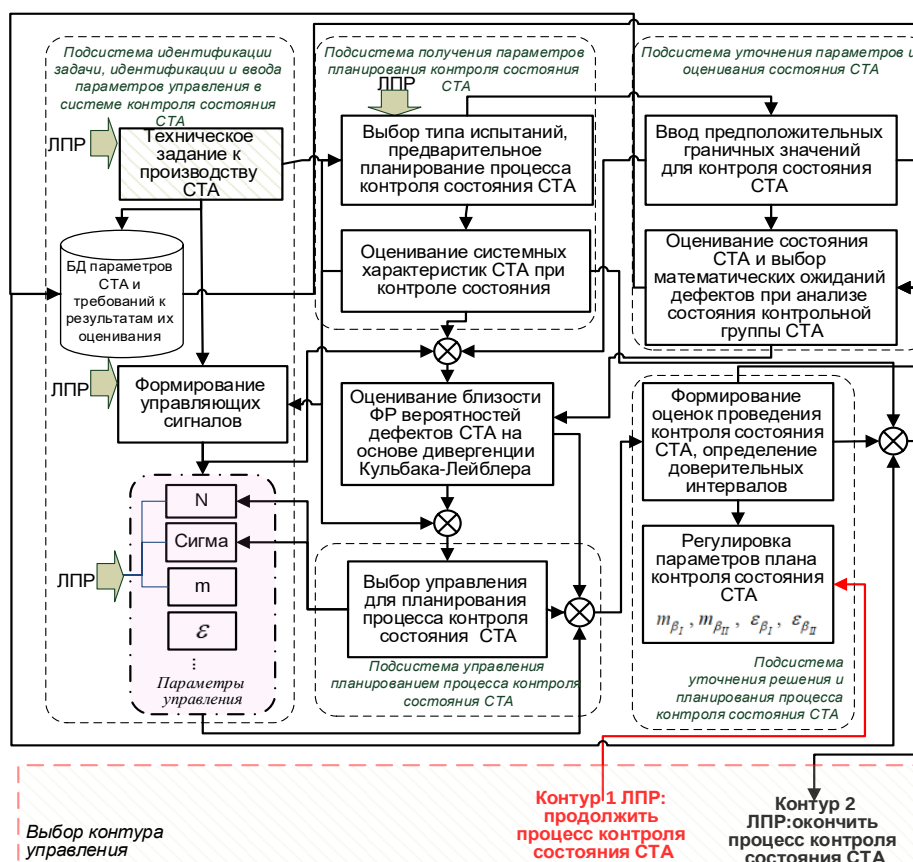


Рис. 6. Структура двухконтурной системы поддержки принятия решения при планировании процесса контроля состояния сложной технической аппаратуры
Fig. 7. The structure of a two-circuit decision support system for planning the process of monitoring the condition of complex technical equipment

Структура предложенной СППР содержит несколько подсистем в рамках двух контуров управления: идентификации задачи, идентификации и ввода параметров управления, получения параметров динамического планирования процесса контроля состояния СТА, уточнения параметров процесса и непосредственно оценивания состояния СТА, управления планированием процесса контроля состояния СТА, уточнения решения при динамическом планировании. В контуре 1 ЛПП предлагается продолжить процесс контроля, а в контуре 2 – окончить процесс контроля в связи с получением приемлемой оценки энтропии состояния СТА. Предполагается разработка уникальных, свойственных отдельным типам СТА, и универсальных алгоритмов анализа состояния на основе анализа различия ФР вероятности наличия дефектов СТА при контроле состояния.

В системе формируются оценки проведения контроля (или мониторинга) состояния СТА на различных этапах производства СТА с высокой чувствительностью благодаря применению меры Кульбака–Лейблера, а также предусмотрен выбор контуров управления ЛПП окончания или продолжения процесса, исходя из анализа промежуточных результатов контроля состояния СТА, что позволяет повысить обоснованность принимаемых решений.

Заключение

Контроль состояния в процессе производства СТА представляет собой нетривиальную задачу в связи с нестационарностью условий проведения испытаний. В статье предложен полимодельный вероятностный подход каскадно-иерархического типа для контроля состояния СТА с учетом видов контролируемых улов (программный, аппаратный и механический) и типов контроля: на основе технических средств и исходя из соответствующего уровня квалификации персонала, что привносит дополнительную факторную нагрузку на модель.

В отличие от детерминированного процесса контроля, когда последовательные испытания проводятся до получения или не получения ожидаемого исхода, предложенный подход позволил сформировать нестационарные правила оценивания вероятностей исходов процесса в системе контроля состояния СТА. Применение меры Кульбака–Лейблера как меры дивергенции функций распределения вероятностей дефектов сложной технической аппаратуры, обладающей высокой чувствительностью, позволило сформировать схему двухконтурной системы поддержки принятия решений по динамическому планированию испытаний сложной технической аппаратуры и повысить степень обоснованности принимаемых решений ЛПП. В том числе в рамках СППР реализуется накопление обучающей информации, и это способствует повышению квалификации персонала. Предложенный подход способствует повышению эффективности проведения испытаний СТА, планирования этого процесса, ситуационного управления, в том числе для создания современных систем поддержки принятия решений в этой области на основе систем искусственного интеллекта.

Список источников

1. Афанасьев П.П., Геращенко А.Н., Голубев И.С. и др. Испытания летательных аппаратов (беспилотные летательные аппараты). Калуга : Изд. Захаров С.И. (СерНа), 2016. 528 с.
2. Денисенко В., Ерещенко П., Кильметов Р., Метелкин Е., Халявко А., Шальнев К. Испытания электронной аппаратуры: быстро и эффективно // Компоненты и технологии. 2004. Т. 39. С. 216–220.
3. Акимов В.Н., Иванов Д.Н., Парафесь С.Г. Методика исследования динамического поведения беспилотного летательного аппарата в наземных условиях с учетом данных летного эксперимента // Научный вестник Московского государственного технического университета гражданской авиации. 2019. Т. 22, № 3. С. 16–24. doi: 10.26467/2079-0619-2019-22-3-16-24
4. Гишваров А.С. Моделирование ускоренных испытаний технических систем на надежность и ресурс // Вестник Уфимского государственного авиационного технического университета. 2007. Т. 9, № 1. С. 26–40.
5. Зеленцов Б.П. Определение показателей надежности электронного оборудования по данным эксплуатационных испытаний // Вестник СибГУТИ. 2018. № 4 (44). С. 3–11.
6. Михайлов В.С., Юрков Н.К. Оценки показателей надежности для безотказных испытаний, проводимых по биномиальному плану // Надежность и качество сложных систем. 2018. № 4 (24). С. 29–39. doi: 10.21685/2307-4205-2018-4-3
7. Баранов С.Н., Тележкин А.М. Метрическое обеспечение программных разработок // Труды СПИИРАН. 2014. Вып. 36. С. 5–27. doi: 10.15622/sp.36.1

8. Штоян Д. Качественные свойства и оценки стохастических моделей. М. : Мир, 1979. 268 с.
9. Kumbhakar S.C., Lien G.D., Hardaker J.B. Technical Efficiency in Competing Panel Data Models: A Study of Norwegian Grain Farming // *Applied Economics*. 2011. V. 12. P. 218–232. doi: 10.1007/s11123-012-0303-1
10. Иванов Н.Н. Метод вычисления функции распределения и числовых вероятностных характеристик времени выполнения стохастического сетевого графика // *Проблемы управления*. 2014. № 6. С. 15–21.
11. Скатков А.В., Брюховецкий А.А., Моисеев Д.В. Метод контроля качественного изменения состояний сетевого трафика на основе информационной меры Кульбака–Лейблера // *Системы контроля окружающей среды – 2016 : тез. докладов междунар. науч.-техн. конф., Севастополь, 24–27 октября 2016 г. Севастополь : Ин-т природно-технических систем, 2016. № 5 (25). С. 53–58.*
12. Pervukhina E., Emmenegger J.-F. Adaptive Time Series Filters Obtained by Minimization of the Kullback-Leibler Divergence Criterion // *International Journal of Applied Mathematics*. 2005. V. 17 (1). P. 69–89.
13. Okhtilev M.Yu., Gnidenko A.S., Alferov V.V., Salukhov V.I., Nazarov D.I. Methods and Algorithms of Integrated Modeling of Complex Technical Objects in Dynamically Changing Conditions // *Proc. of the International Scientific Conference MMET NW*. 2018. P. 282–284.
14. Микони С.В., Соколов Б.В. Юсупов Р.М. Квалиметрия моделей и полимодельных комплексов. М. : РАН, 2018. 314 с. doi: 10.31857/S9785907036321000001
15. Доронина Ю.В., Скатков А.В. Каскадно-иерархическое моделирование в задачах анализа динамики ресурсных характеристик сложных систем // *Информационно-управляющие системы*. 2020. № 3. С. 48–58. doi: 10.31799/1684-8853-2020-3-48-58

References

1. Afanasiev, P.P., Gerashchenko, A.N., Golubev, I.S. et al. (2016) *Ispytaniya letatel'nykh apparatov (bespilotnye letatel'nye apparaty)* [Tests of aircraft (unmanned aerial vehicles)]. Kaluga: S.I. Zakharov (SerNa).
2. Denisenko, V., Ereshchenko, P., Kilmetov, R., Metelkin, E., Kholyavko, A. & Shalnev, K. (2004) *Ispytaniya elektronnoy apparatury: bystro i effektivno* [Electronic equipment testing: fast and efficient]. *Komponenty i tekhnologii*. 39. pp. 216–220.
3. Akimov, V.N., Ivanov, D.N. & Parafes, S.G. (2019) Metodika issledovaniya dinamicheskogo povedeniya bespilotnogo letatel'nogo apparata v nazemnykh usloviyakh s uchetom dannykh letnogo eksperimenta [Methodology for studying the dynamic behavior of a pilotless aircraft in ground conditions, taking into account flight experiment data]. *Nauchnyy vestnik Moskovskogo gosudarstvennogo tekhnicheskogo universiteta grazhdanskoy aviatsii*. 22(3). pp. 16–24. DOI: 10.26467/2079-0619-2019-22-3-16-24
4. Gishvarov, A.S. (2007) Modelirovaniye uskorennykh ispytaniy tekhnicheskikh sistem na nadezhnost' i resurs [Modeling accelerated tests of technical systems for reliability and resource]. *Vestnik Ufimskogo gosudarstvennogo aviatsionnogo tekhnicheskogo universiteta*. 9(1). pp. 26–40.
5. Zelentsov, B.P. (2018) Opredeleniye pokazateley nadezhnosti elektronnoy oborudovaniya po dannym ekspluatatsionnykh ispytaniy [Determination of reliability indicators of electronic equipment based on operational test data]. *Vestnik SibGUTI*. 4(44). pp. 3–11.
6. Mikhailov, V.S. & Yurkov, N.K. (2018) Otsenki pokazateley nadezhnosti dlya bezotkaznykh ispytaniy, provodimykh po binomial'nomu planu [Estimates of reliability indicators for trouble-free tests conducted according to a binomial plan]. *Nadezhnost' i kachestvo slozhnykh sistem*. 4(24). pp. 29–39. DOI: 10.21685/2307-4205-2018-4-3
7. Baranov, S.N. & Telyozkin, A.M. (2014) Metricheskoe obespecheniye programmnykh razrabotok [Metrical support of software development]. *Trudy SPIIRAN*. 36. pp. 5–27. DOI: 10.15622/sp.36.1
8. Shtoyan, D. (1979) *Kachestvennyye svoystva i otsenki stokhasticheskikh modeley* [Qualitative properties and estimates of stochastic models]. Moscow: Mir.
9. Kumbhakar, S.C., Lien, G.D. & Hardaker, J.B. (2011) Technical Efficiency in Competing Panel Data Models: A Study of Norwegian Grain Farming. *Applied Economics*. 12. pp. 218–232. DOI: 10.1007/s11123-012-0303-1
10. Ivanov, N.N. (2014) Metod vychisleniya funktsii raspredeleniya i chislovykh veroyatnostnykh kharakteristik vremeni vypolneniya stokhasticheskogo setevogo grafika [Method of calculating the distribution function and numerical probabilistic characteristics of the execution time of a stochastic network graph]. *Problemy upravleniya*. 6. pp. 15–21.
11. Skatkov, A.V., Bryukhovetsky, A.A. & Moiseev, D.V. (2016) Metod kontrolya kachestvennogo izmeneniya sostoyaniy setevogo trafika na osnove informatsionnoy mery Kul'baka–Leyblera [A method for monitoring qualitative changes in network traffic conditions based on the Kulback-Leibler information measure]. *Sistemy kontrolya okruzhayushchey sredy – 2016* [Environmental Monitoring Systems – 2016]. Proc. of the International Conference. Sevastopol, October 24–27, 2016. Sevastopol: Institute of Natural and Technical Systems. pp. 53–58.
12. Pervukhina, E. & Emmenegger, J.-F. (2005) Adaptive Time Series Filters Obtained by Minimization of the Kullback-Leibler Divergence Criterion. *International Journal of Applied Mathematics*. 17(1). pp. 69–89.
13. Okhtilev M.Yu., Gnidenko, A.S., Alferov, V.V., Salukhov, V.I. & Nazarov, D.I. (2018) Methods and Algorithms of Integrated Modeling of Complex Technical Objects in Dynamically Changing Conditions. *Proceedings of the International Scientific Conference MMET NW*. pp. 282–284.
14. Mikoni, S.V., Sokolov, B.V. & Yusupov, R.M. (2018) *Kvalimetriya modeley i polimodel'nykh kompleksov* [Qualimetry of Models and Polymodel Complexes]. Moscow: RAS. DOI: 10.31857/S9785907036321000001

15. Doronina, Yu.V. & Skatkov, A.V. (2020) Kaskadno-ierarkhicheskoe modelirovanie v zadachakh analiza dinamiki resursnykh kha-rakteristik slozhnykh sistem [Cascade – hierarchical modeling in the tasks of analyzing the dynamics of resource characteristics of complex systems]. *Informatsionno-upravlyayushchie sistemy*. 3. pp. 48–58. DOI: 10.31799/1684-8853-2020-3-48-58

Информация об авторах:

Доронина Юлия Валентиновна – доцент, доктор технических наук, профессор кафедры информационных технологий и компьютерных систем Института информационных технологий Севастопольского государственного университета (Севастополь, Россия). E-mail: apkSev@yandex.ru

Скатков Александр Владимирович – профессор, доктор технических наук, профессор Севастопольского государственного университета (Севастополь, Россия). E-mail: vm1945@mail.ru

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

Information about the authors:

Doronina Yulia V. (Associate Professor, Doctor of Technical Sciences, Professor of the Department of Information Technologies and Computer Systems, Institute of Information Technologies, Sevastopol State University, Sevastopol, Russian Federation). E-mail: apkSev@yandex.ru

Skatkov Alexander V. (Professor, Doctor of Technical Sciences, Sevastopol State University, Sevastopol, Russian Federation). E-mail: vm1945@mail.ru

The author declares no conflicts of interests.

Поступила в редакцию 21.02.2024; принята к публикации 03.06.2024

Received 21.02.2024; accepted for publication 03.06.2024

Научная статья

УДК 519.2

doi: 10.17223/19988605/67/9

Оптимальное планирование эксперимента на основе винеровской деградационной модели с ковариатами с использованием условной информационной матрицы Фишера

Евгения Алексеевна Осинцева¹, Екатерина Владимировна Чимитова²

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

¹ osinceva.j@gmail.com

² chimitova@corp.nstu.ru

Аннотация. Винеровские деградационные модели с ковариатами широко используются для оценки функции надежности по данным об изменении показателя деградации во времени. В данной статье разработан алгоритм построения оптимального плана эксперимента на надежность, который предусматривает вычисление оптимальных величин нагрузок, количества исследуемых изделий и моментов времени измерения показателя деградации с учетом ограничений на время эксперимента и критического значения показателя деградации. Предложенный алгоритм основан на оптимизации суммы функционалов от информационной матрицы Фишера и условной информационной матрицы Фишера при заданных ограничениях на величину нагрузки, длительность проведения эксперимента и минимальный интервал между моментами времени измерения показателя деградации. Применение разработанного алгоритма рассмотрено на примере данных об исследовании светодиодов (LED).

Ключевые слова: винеровская деградационная модель; ковариаты; информационная матрица Фишера; оптимальное планирование эксперимента.

Для цитирования: Осинцева Е.А., Чимитова Е.В. Оптимальное планирование эксперимента на основе винеровской деградационной модели с ковариатами с использованием условной информационной матрицы Фишера // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 67. С. 83–93. doi: 10.17223/19988605/67/9

Original article

doi: 10.17223/19988605/67/9

Optimal design of experiment based on the Wiener degradation model with covariate using conditional Fisher information matrix

Evgeniya A. Osintseva¹, Ekaterina V. Chimitova²

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

¹ osinceva.j@gmail.com

² chimitova@corp.nstu.ru

Abstract. The Wiener degradation models with covariates are widely used for reliability analysis by degradation index paths. In this paper, we suggest an algorithm for constructing the optimal design based on the Wiener degradation model, which involves determining optimal stress levels, the number of objects under each stress level and the time moments for measuring degradation index, taking into account the limitation on the duration of the experiment and the critical value of the degradation index. The proposed algorithm is based on optimizing the sum of functionals from the Fisher information matrix (FIM) and the conditional Fisher information matrix (CFIM) under given restrictions on the stress levels, the duration of the experiment and the minimal interval between the time points of measuring the degradation index. As an example, the application of the developed algorithm is considered for the LED reliability analysis.

Keywords: Wiener degradation model; covariate; Fisher information matrix; optimal design.

For citation: Osintseva, E.A., Chimitova, E.V. (2024) Optimal design of experiment based on the Wiener degradation model with covariate using conditional Fisher information matrix. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 67. pp. 83–93. doi: 10.17223/19988605/67/9

Введение

В условиях быстрого развития технологий и высоких требований со стороны потребителей срок службы производимых изделий должен быть достаточно долгим. Следовательно, отказы могут не возникнуть в течение ограниченного времени тестирования на надежность. Таким образом, использовать классические методы, основанные только на данных о времени отказа, невозможно. В том случае, когда есть возможность получить данные о деградационном показателе, характеризующем процесс ухудшения работы исследуемых изделий, переходят к анализу надежности на основе деградационных моделей [1]. Несмотря на то, что получение данных о деградации позволяет проводить анализ надежности на более ранних сроках, не дожидаясь выхода из строя всех объектов исследования, этот процесс может занять много времени. Поэтому во многих статьях можно встретить комбинированный подход – получение деградационных данных в процессе проведения ускоренного эксперимента [2].

В качестве вероятностной модели, описывающей распределение времен наступления отказов в зависимости от величины нагрузки на основе данных о деградации, часто используют винеровскую деградационную модель. Особенностью данной модели является то, что приращение деградационного показателя может быть как положительным, так и отрицательным [3–5]. Другие популярные модели, такие как гамма и обратная гауссовская модели, также часто используются в различных приложениях, однако приращение деградационного показателя в этих моделях всегда положительное, что необходимо учитывать при их использовании [6–8].

Ускоренные испытания можно классифицировать по трем типам нагрузок: постоянные нагрузки, ступенчатые нагрузки и прогрессирующие нагрузки [9, 10]. В данной статье рассматривается винеровская деградационная модель с постоянными нагрузками в ходе проведения ускоренного эксперимента. Перед проведением ускоренных испытаний следует задать план эксперимента. Для этого проводят предварительные испытания для построения деградационной модели, после чего можно переходить к подбору оптимальных условий для проведения повторного эксперимента на надежность. Оптимальное планирование эксперимента является распространенным подходом для снижения экономических затрат на проведение самого эксперимента и повышения точности получаемой на его основе модели. Кроме того, для достижения корректных результатов необходимо учитывать различные факторы, такие как свойства материала, граничные условия, геометрические размеры, предельные нагрузки [11].

В литературе представлены различные подходы к построению оптимального плана эксперимента для получения величин оптимальных нагрузок и количества исследуемых объектов с различными ограничениями как на величину допустимых нагрузок, так и на стоимость и продолжительность эксперимента. При этом выбор целевой функции оказывает существенное влияние на результат оптимального планирования ускоренного эксперимента [12]. В [13–15] проводились исследования различных целевых функций, оптимизация которых может приводить к увеличению точности оценок параметров модели или точности прогнозирования срока службы изделий в зависимости от поставленной задачи. Согласно текущим исследованиям, множество целевых функций включает в себя некоторый функционал от информационной матрицы Фишера [13], асимптотическую дисперсию или среднеквадратическую ошибку оценки q -го квантиля распределения наработок до отказа [14, 15] и др. Кроме того, в литературе можно встретить альтернативный подход, основанный на оптимизации функционала от ковариационной матрицы, полученной методами статистического моделирования. Основным минусом данного подхода является то, что процедура моделирования требует большого объема вычислительных ресурсов и времени.

В своих предыдущих работах мы рассматривали подход к оптимальному планированию эксперимента на основе оптимизации некоторого функционала от информационной матрицы Фишера. В [16] акцент был сделан на то, что выбор моментов времени для измерения показателя деградации существенно влияет на точность оценок максимального правдоподобия (ОМП) параметров винеровской деградационной модели. Оптимальное распределение моментов времени измерения зависит от модели, описывающей процесс деградации, а также от условий эксперимента, таких как уровень нагрузок, продолжительность эксперимента и минимальный временной интервал между измерениями показателя деградации. В [17] мы разработали алгоритм построения A - и D -оптимальных планов на основе винеровской деградационной модели, который включает определение оптимальных уровней нагрузок, количества тестируемых объектов и моментов времени измерения показателя деградации. Однако в [17] не был учтен тот факт, что наблюдение за объектом должно быть прекращено, когда значение показателя деградации достигает критического уровня. В [18] был предложен алгоритм, который позволяет получить оптимальные величины нагрузок при условии, что все деградационные процессы ограничены заданным критическим уровнем. Данный алгоритм основан на оптимизации функционала от условной информационной матрицы Фишера.

Основной целью данной работы является разработка общего алгоритма построения D -оптимального плана эксперимента на основе винеровской деградационной модели, который включает определение оптимальных уровней нагрузок, количества тестируемых объектов и моментов времени измерения показателя деградации при условии, что все деградационные процессы ограничены заданным критическим уровнем.

1. Винеровская деградационная модель

Случайный процесс с независимыми приращениями $Z(t)$, характеризующий процесс деградации исследуемых изделий, называется винеровским деградационным процессом, если приращения подчиняются нормальному закону распределения с функцией плотности

$$f(u, \theta_1, \theta_2) = \frac{1}{\theta_2 \sqrt{2\pi}} \exp \left(-\frac{(u - \theta_1)^2}{2\theta_2^2} \right),$$

где $\theta_1 = \mu(\rho(t + \Delta t) - \rho(t))$ – параметр сдвига, $\theta_2 = \sigma \sqrt{\rho(t + \Delta t) - \rho(t)}$ – параметр масштаба, $\sigma > 0$, $\rho(t; \gamma)$ – положительная возрастающая функция тренда, γ – параметр функции тренда, $Z(0) = 0$.

Предполагается, что деградационный процесс наблюдается при некоторой постоянной во времени нагрузке (ковариате) x , диапазон значений которой определяется условиями проведения эксперимента и представляет собой отрезок числовой прямой. В общем случае ковариата x представляет собой векторную величину, поскольку может учитываться влияние нескольких факторов, например температуры, напряжения, давления и пр. Однако в настоящей работе для упрощения математических выражений ковариата рассматривается как скалярная величина.

Влияние ковариаты на изменение деградационного показателя в ходе проведения ускоренного эксперимента можно определить следующим образом:

$$Z_x(t) = Z(t / r(x; \beta)),$$

где $r(x; \beta)$ – положительная возрастающая функция, β – регрессионный параметр. В этом случае математическое ожидание $Z_x(t)$ имеет вид:

$$E(Z_x(t)) = \mu \rho(t / r(x; \beta); \gamma).$$

Время наработки до отказа представляет собой случайную величину

$$\tau = \sup \{t : Z_x(t) \leq z_0\},$$

где z_0 – критический уровень показателя деградации. Таким образом, функцию надежности можно записать в следующем виде:

$$S(t) = P\{\tau > t\} = P\{Z_x(t) < z_0\} = \Phi\left(\frac{z_0 - \mu\rho(t/r(x;\beta);\gamma)}{\sigma\sqrt{\rho(t/r(x;\beta);\gamma)}}\right).$$

Обозначим выборку приращений деградационного показателя, полученную в результате наблюдения деградационных процессов n объектов, следующим образом:

$$\mathbf{X}_n = \left\{ (\Delta Z_{1j}, x_1), (\Delta Z_{2j}, x_2), \dots, (\Delta Z_{nj}, x_n), j = \overline{1, k_i} \right\},$$

где для каждого i -го объекта k_i – количество измерений деградационного показателя, x_i – значение ковариаты, $\Delta Z_{ij} = Z_i(t_j) - Z_i(t_{j-1})$ – значение приращения деградационного показателя в интервале от t_{j-1} до t_j . Моменты времени измерения деградационного показателя:

$$0 = t_0 < t_1 \dots < t_{k_i}, \quad t_{k_i} \leq \sup\{t : Z_{x_i}(t) \leq z_0\}, \quad i = \overline{1, n}.$$

Неизвестные параметры модели оцениваются методом максимального правдоподобия. Логарифмическая функция правдоподобия для винеровской деградационной модели имеет вид:

$$\begin{aligned} \ln L(\mathbf{X}_n) = & -\sum_{i=1}^n k_i (\ln \sqrt{2\pi} + \ln \sigma) - \frac{1}{2} \sum_{i=1}^n \sum_{j=2}^{k_i} \ln(\rho(t_{i,j+1}) - \rho(t_{ij})) - \\ & - \frac{1}{2\sigma^2} \sum_{i=1}^n \sum_{j=1}^{k_i} \frac{[\Delta Z_{ij} - \mu(\rho(t_{i,j+1}) - \rho(t_{ij}))]^2}{(\rho(t_{i,j+1}) - \rho(t_{ij}))}. \end{aligned} \quad (1)$$

2. Информационная матрица Фишера для винеровской деградационной модели

Для получения информационной матрицы Фишера (ИМФ) для винеровской деградационной модели необходимо вычислить математическое ожидание вторых частных производных от функции правдоподобия (1) по параметрам.

Элементы информационной матрицы Фишера можно записать следующим образом:

$$\begin{aligned} I_{11} &= \frac{\partial^2 \ln L(\mathbf{X}_n)}{\partial \sigma^2} = \frac{2}{\sigma^2} \sum_{i=1}^n k_i; \quad I_{23} = \frac{\partial^2 \ln L(\mathbf{X}_n)}{\partial \mu \partial \gamma} = \frac{\mu^2}{\sigma^2} \sum_{i=1}^n \sum_{j=1}^{k_i} \frac{\partial \Delta \rho_{ij}}{\partial \gamma}; \\ I_{12} &= \frac{\partial^2 \ln L(\mathbf{X}_n)}{\partial \sigma \partial \mu} = 0; \quad I_{24} = \frac{\partial^2 \ln L(\mathbf{X}_n)}{\partial \mu \partial \beta} = \frac{\mu^2}{\sigma^2} \sum_{i=1}^n \sum_{j=1}^{k_i} \frac{\partial \Delta \rho_{ij}}{\partial \beta}; \\ I_{13} &= \frac{\partial^2 \ln L(\mathbf{X}_n)}{\partial \sigma \partial \gamma} = \frac{1}{\sigma} \sum_{i=1}^n \sum_{j=1}^{k_i} \frac{1}{\Delta \rho_{ij}} \cdot \frac{\partial \Delta \rho_{ij}}{\partial \gamma}; \quad I_{33} = \frac{\partial^2 \ln L(\mathbf{X}_n)}{\partial \gamma^2} = \sum_{i=1}^n \sum_{j=1}^{k_i} \left(\frac{\partial \Delta \rho_{ij}}{\partial \gamma} \right)^2 \cdot \left(\frac{1}{2\Delta \rho_{ij}^2} + \frac{\mu^2}{\sigma^2} \cdot \frac{1}{\Delta \rho_{ij}} \right); \\ I_{14} &= \frac{\partial^2 \ln L(\mathbf{X}_n)}{\partial \sigma \partial \beta} = \frac{1}{\sigma} \sum_{i=1}^n \sum_{j=1}^{k_i} \frac{1}{\Delta \rho_{ij}} \cdot \frac{\partial \Delta \rho_{ij}}{\partial \beta}; \quad I_{34} = \frac{\partial^2 \ln L(\mathbf{X}_n)}{\partial \gamma \partial \beta} = \sum_{i=1}^n \sum_{j=1}^{k_i} \left(\frac{\partial \Delta \rho_{ij}}{\partial \gamma} \right) \cdot \left(\frac{\partial \Delta \rho_{ij}}{\partial \beta} \right) \cdot \left(\frac{1}{2\Delta \rho_{ij}^2} + \frac{\mu^2}{\sigma^2} \cdot \frac{1}{\Delta \rho_{ij}} \right); \\ I_{22} &= \frac{\partial^2 \ln L(\mathbf{X}_n)}{\partial \mu^2} = \frac{1}{\sigma^2} \sum_{i=1}^n \sum_{j=1}^{k_i} \Delta \rho_{ij}; \quad I_{44} = \frac{\partial^2 \ln L(\mathbf{X}_n)}{\partial \beta^2} = \sum_{i=1}^n \sum_{j=1}^{k_i} \left(\frac{\partial \Delta \rho_{ij}}{\partial \beta} \right)^2 \cdot \left(\frac{1}{2\Delta \rho_{ij}^2} + \frac{\mu^2}{\sigma^2} \cdot \frac{1}{\Delta \rho_{ij}} \right), \end{aligned} \quad (2)$$

где $\Delta \rho_{ij} = \rho(t_{i,j+1}/r(x_i;\beta);\gamma) - \rho(t_{ij}/r(x_i;\beta);\gamma)$.

3. Условная информационная матрица Фишера для винеровской деградационной модели

Для того чтобы учесть факт, что наблюдение за объектом прекращается при достижении критического уровня показателя деградации, при вычислении ИМФ необходимо использовать условную

функцию плотности: $f_{\Delta Z_{ij}}(u | Z_{ik_i} \leq z_0)$, где $Z_{ik_i} = \sum_{j=1}^{k_i} \Delta Z_{ij}$ – значение показателя деградации в момент

времени достижения критического уровня z_0 . Однако ее вычисление предполагает расчет k -кратных интегралов, что представляет собой сложную вычислительную задачу. Таким образом, в данной статье предлагается упростить задачу, используя условную функцию плотности $f_{Z_{ik_i}}(u | Z_{ik_i} \leq z_0)$. Для упрощения выкладок введем обозначение

$$C(z) = \frac{(z - \mu\rho(t_{k_i}))}{\sigma\sqrt{2\rho(t_{k_i})}}.$$

Тогда логарифмическая функция правдоподобия может быть представлена в следующем виде:

$$\ln L(X_n | Z_{k_1} \leq z_0, \dots, Z_{k_n} \leq z_0) = \sum_{i=1}^n (\ln f_{Z_{ik_i}}(Z_{ik_i}) - \ln F_{Z_{ik_i}}(z_0)), \quad (3)$$

$$\text{где } f_{Z_{ik_i}}(z) = \frac{1}{\sigma\sqrt{2\rho(t_{k_i})}} \exp(-C^2(z)), \quad F_{Z_{ik_i}}(z_0) = 0.5 + \frac{0.5}{\sqrt{\pi}} \Gamma(0.5, C^2(z_0)).$$

Для получения элементов условной информационной матрицы Фишера (УИМФ) вычислим математическое ожидание вторых частных производных по параметрам функции правдоподобия (3):

$$\begin{aligned} I(\xi | Z_{k_1} \leq z_0, \dots, Z_{k_n} \leq z_0) &= -E \left(\frac{\partial^2 \ln L(X_n | Z_{k_1} \leq z_0, \dots, Z_{k_n} \leq z_0)}{\partial \theta \partial \theta^T} \right) = -E \left(\frac{\partial^2 \sum_{i=1}^n (\ln f_{Z_{ik_i}}(Z_{ik_i}) - \ln F_{Z_{ik_i}}(z_0))}{\partial \theta \partial \theta^T} \right) = \\ &= -\sum_{i=1}^n E \left(\frac{\partial^2 (\ln f_{Z_{ik_i}}(Z_{ik_i}) - \ln F_{Z_{ik_i}}(z_0))}{\partial \theta \partial \theta^T} \right) = -\sum_{i=1}^n E \left(\frac{\partial^2 \ln f_{Z_{ik_i}}(Z_{ik_i})}{\partial \theta \partial \theta^T} \right) + \sum_{i=1}^n E \left(\frac{\partial^2 \ln F_{Z_{ik_i}}(z_0)}{\partial \theta \partial \theta^T} \right). \end{aligned}$$

$$\text{Элементы первой части УИМФ } I_1(\xi | Z_{k_1} \leq z_0, \dots, Z_{k_n} \leq z_0) = -\sum_{i=1}^n E \left(\frac{\partial^2 \ln f_{Z_{ik_i}}(Z_{ik_i})}{\partial \theta \partial \theta^T} \right):$$

$$\begin{aligned} I_{11} &= \frac{2n}{\sigma^2}; & I_{23} &= I_{32} = \frac{\mu^2}{\sigma^2} \sum_{i=1}^n \frac{\partial \rho_{ik_i}}{\partial \gamma}; \\ I_{12} &= I_{21} = 0; & I_{24} &= I_{42} = \frac{\mu^2}{\sigma^2} \sum_{i=1}^n \frac{\partial \rho_{ik_i}}{\partial \beta}; \\ I_{13} &= I_{31} = \frac{1}{\sigma} \sum_{i=1}^n \frac{1}{\rho_{ik_i}} \cdot \frac{\partial \rho_{ik_i}}{\partial \gamma}; & I_{33} &= \sum_{i=1}^n \left(\frac{\partial \rho_{ik_i}}{\partial \gamma} \right)^2 \cdot \left(\frac{1}{2\rho_{ik_i}^2} + \frac{\mu^2}{\sigma^2} \cdot \frac{1}{\rho_{ik_i}} \right); \\ I_{14} &= I_{41} = \frac{1}{\sigma} \sum_{i=1}^n \frac{1}{\rho_{ik_i}} \cdot \frac{\partial \rho_{ik_i}}{\partial \beta}; & I_{34} &= I_{43} = \sum_{i=1}^n \left(\frac{\partial \rho_{ik_i}}{\partial \gamma} \right) \cdot \left(\frac{\partial \rho_{ik_i}}{\partial \beta} \right) \cdot \left(\frac{1}{2\rho_{ik_i}^2} + \frac{\mu^2}{\sigma^2} \cdot \frac{1}{\rho_{ik_i}} \right); \\ I_{22} &= \frac{1}{\sigma^2} \sum_{i=1}^n \rho_{ik_i}; & I_{44} &= \sum_{i=1}^n \left(\frac{\partial \rho_{ik_i}}{\partial \beta} \right)^2 \cdot \left(\frac{1}{2\rho_{ik_i}^2} + \frac{\mu^2}{\sigma^2} \cdot \frac{1}{\rho_{ik_i}} \right); \end{aligned} \quad (4)$$

где $\rho_{ik_i} = \rho(t_{k_i} / r(x_i; \beta); \gamma)$, $i = \overline{1, n}$.

Элементы второй части УИМФ:

$$I_2(\xi | Z_{k_1} \leq z_0, \dots, Z_{k_n} \leq z_0) = \sum_{i=1}^n E \left(\frac{1}{F_{Z_{ik_i}}(z_0)} \frac{\partial^2 F_{Z_{ik_i}}(z_0)}{\partial \theta \partial \theta^T} - \frac{1}{(F_{Z_{ik_i}}(z_0))^2} \frac{\partial F_{Z_{ik_i}}(z_0)}{\partial \theta} \frac{\partial F_{Z_{ik_i}}(z_0)}{\partial \theta^T} \right),$$

где

$$\frac{\partial F_{Z_{ik_i}}(z_0)}{\partial \theta} = \frac{1}{\sqrt{\pi}} \frac{\partial C(z_0)}{\partial \theta} \exp(-C^2(z_0)),$$

$$\frac{\partial^2 F_{Z_{ik_i}}(z_0)}{\partial \theta \partial \theta^T} = \frac{1}{\sqrt{\pi}} \frac{\partial^2 C(z_0)}{\partial \theta \partial \theta^T} \exp(-C^2(z_0)) - \frac{2C(z_0)}{\sqrt{\pi}} \frac{\partial C(z_0)}{\partial \theta} \frac{\partial C(z_0)}{\partial \theta^T} \exp(-C^2(z_0)).$$

Производные $C(z_0)$ по параметрам винеровской деградационной модели имеют вид:

$$\frac{\partial C(z_0)}{\partial \sigma} = -\frac{(z_0 - \mu \rho_{ik_i})}{\sigma^2 \sqrt{2\rho_{ik_i}}}; \quad \frac{\partial C(z_0)}{\partial \mu} = -\frac{\sqrt{\rho_{ik_i}}}{\sqrt{2}\sigma}; \quad \frac{\partial C(z_0)}{\partial \gamma} = -\frac{\left(\mu + \frac{z_0}{\rho_{ik_i}}\right) \partial \rho_{ik_i}}{2\sigma \sqrt{2\rho_{ik_i}} \partial \gamma}; \quad \frac{\partial C(z_0)}{\partial \beta} = -\frac{\left(\mu + \frac{z_0}{\rho_{ik_i}}\right) \partial \rho_{ik_i}}{2\sigma \sqrt{2\rho_{ik_i}} \partial \beta};$$

$$\frac{\partial^2 C(z_0)}{\partial \sigma^2} = \frac{2(z_0 - \mu \rho_{ik_i})}{\sigma^3 \sqrt{2\rho_{ik_i}}}; \quad \frac{\partial^2 C(z_0)}{\partial \mu \partial \gamma} = -\frac{1}{2\sigma \sqrt{2\rho_{ik_i}}} \frac{\partial \rho_{ik_i}}{\partial \gamma};$$

$$\frac{\partial^2 C(z_0)}{\partial \sigma \partial \mu} = \frac{\sqrt{\rho_{ik_i}}}{\sqrt{2}\sigma^2}; \quad \frac{\partial^2 C(z_0)}{\partial \mu \partial \beta} = -\frac{1}{2\sigma \sqrt{2\rho_{ik_i}}} \frac{\partial \rho_{ik_i}}{\partial \beta};$$

$$\frac{\partial^2 C(z_0)}{\partial \sigma \partial \gamma} = \frac{\left(\mu + \frac{z_0}{\rho_{ik_i}}\right) \partial \rho_{ik_i}}{2\sigma^2 \sqrt{2\rho_{ik_i}} \partial \gamma}; \quad \frac{\partial^2 C(z_0)}{\partial \gamma^2} = \frac{\left(\mu + \frac{z_0}{\rho_{ik_i}}\right)}{\sigma^2 \sqrt{2\rho_{ik_i}}} \left(\frac{1}{2\rho_{ik_i}} \left(\frac{\partial \rho_{ik_i}}{\partial \gamma} \right)^2 - \frac{\partial^2 \rho_{ik_i}}{\partial \gamma^2} \right);$$

$$\frac{\partial^2 C(z_0)}{\partial \sigma \partial \beta} = \frac{\left(\mu + \frac{z_0}{\rho_{ik_i}}\right) \partial \rho_{ik_i}}{2\sigma^2 \sqrt{2\rho_{ik_i}} \partial \beta}; \quad \frac{\partial^2 C(z_0)}{\partial \gamma \partial \beta} = \frac{\left(\mu + \frac{z_0}{\rho_{ik_i}}\right)}{2\sigma \sqrt{2\rho_{ik_i}}} \left(\frac{1}{2\rho_{ik_i}} \frac{\partial \rho_{ik_i}}{\partial \gamma} \frac{\partial \rho_{ik_i}}{\partial \beta} - \frac{\partial^2 \rho_{ik_i}}{\partial \gamma \partial \beta} \right);$$

$$\frac{\partial^2 C(z_0)}{\partial \mu^2} = 0; \quad \frac{\partial^2 C(z_0)}{\partial \beta^2} = \frac{\left(\mu + \frac{z_0}{\rho_{ik_i}}\right)}{2\sigma \sqrt{2\rho_{ik_i}}} \left(\frac{1}{2\rho_{ik_i}} \left(\frac{\partial \rho_{ik_i}}{\partial \beta} \right)^2 - \frac{\partial^2 \rho_{ik_i}}{\partial \beta^2} \right),$$

где $\rho_{ik_i} = \rho(t_{k_i} / r(x_i; \beta); \gamma)$, $i = \overline{1, n}$.

4. Задача построения оптимального плана эксперимента на надежность

Определим план эксперимента как совокупность данных

$$\xi = \left\{ \begin{array}{ccc} x_{(1)} & \dots & x_{(q)} \\ \omega_1 & \dots & \omega_q \\ t_1, \dots, t_{k_1} & \dots & t_1, \dots, t_{k_q} \end{array} \right\},$$

здесь $x_{(1)}, \dots, x_{(q)}$ – опорные точки плана; $\omega_i = \frac{n_i}{n}$, $\omega_i \geq 0$, $\sum_{i=1}^q \omega_i = 1$, где n_i – число объектов, исследуемых при ковариате $x_{(i)}$, $i = \overline{1, q}$, $q \leq n$, $\sum_{i=1}^q n_i = n$, и моменты времени измерения деградационного показателя $t_1 \dots t_{k_i}$. Все тестируемые объекты разделены на q групп в соответствии с величинами нагрузок (опорными точками плана), и $t_{k_i} \leq \sup\{t : \mu\rho(t / r(x_i; \beta); \gamma) \leq z_0\}$, $i = \overline{1, q}$.

Таким образом, задача построения D -оптимального плана эксперимента может быть записана в следующем виде:

$$\begin{cases} M(I(\xi)) = \det(I(\xi)) \rightarrow \max, \\ x_{\min} \leq x_{(i)} \leq x_{\max}, \\ t_0 = 0, \quad t_{k_i} \leq T, \quad t_j - t_{j-1} > \Delta t_{\min}, \quad j = \overline{1, k_i}, \quad i = \overline{1, q}, \end{cases}$$

где $M(\cdot)$ – функционал от информационной матрицы Фишера (ИМФ), $x_{\min}$, $x_{\max}$ – минимальное и максимальное значения нагрузок, определенных условиями задачи, $\Delta t_{\min}$ – минимальный интервал времени между моментами измерения деградационного показателя.

5. Алгоритм прямого поиска оптимального плана эксперимента

Алгоритм прямого поиска оптимального плана эксперимента предполагает решение задачи оптимизации некоторой целевой функции. Для достижения поставленной цели предлагается в качестве целевой функции использовать сумму определителей полученных выражений ИМФ и УИМФ для оптимизации моментов времени измерения деградации $t_1, \dots, t_{k_i}$ и получения оптимальных величин нагрузок $x_{(1)}, \dots, x_{(q)}$ и весов $\omega_1, \dots, \omega_q$.

Таким образом, предлагаемый алгоритм поиска оптимального плана эксперимента можно сформулировать в следующем виде.

Шаг 1. Задать начальный план эксперимента:

$$\xi^0 = \begin{pmatrix} x_{(1)}^0 & \dots & x_{(q)}^0 \\ \omega_1^0 & \dots & \omega_q^0 \\ t_1^0, \dots, t_{k_1}^0 & \dots & t_1^0, \dots, t_{k_q}^0 \end{pmatrix}.$$

Установить $iter = 0$.

Шаг 2. Вычислить ИМФ $I(\xi^0)$ и УИМФ $I(\xi^0 | Z_{k_1} \leq z_0, \dots, Z_{k_q} \leq z_0)$ для начального плана эксперимента.

Шаг 3. Зафиксировать точки плана $\omega_1^0, \dots, \omega_q^0$, $t_1^0, \dots, t_{k_i}^0$, $i = \overline{1, q}$, и решить оптимизационную задачу

$$M(I(\xi | Z_{k_1} \leq z_0, \dots, Z_{k_q} \leq z_0)) + M(I(\xi)) \rightarrow \max_{x_{(1)} \dots x_{(q)}}.$$

Вычислить ИМФ $I(\xi^{iter})$ и УИМФ $I(\xi^{iter} | Z_{k_1} \leq z_0, \dots, Z_{k_q} \leq z_0)$ в соответствии с полученным планом:

$$\xi^{iter} = \begin{pmatrix} x_{(1)}^{iter+1} & \dots & x_{(q)}^{iter+1} \\ \omega_1^{iter} & \dots & \omega_q^{iter} \\ t_1^{iter}, \dots, t_{k_1}^{iter} & \dots & t_1^{iter}, \dots, t_{k_q}^{iter} \end{pmatrix}.$$

Шаг 4. Зафиксировать точки плана $x_{(1)}^{iter+1}, \dots, x_{(q)}^{iter+1}$, $t_1^{iter}, \dots, t_{k_i}^{iter}$, $i = \overline{1, q}$, и решить оптимизационную задачу

$$M(I(\xi | Z_{k_1} \leq z_0, \dots, Z_{k_q} \leq z_0)) + M(I(\xi)) \rightarrow \max_{\omega_1, \dots, \omega_q}.$$

Вычислить ИМФ $I(\xi^{iter})$ и УИМФ $I(\xi^{iter} | Z_{k_1} \leq z_0, \dots, Z_{k_q} \leq z_0)$ в соответствии с полученным планом:

$$\xi^{iter} = \left\{ \begin{array}{ccc} x_{(1)}^{iter+1} & \dots & x_{(q)}^{iter+1} \\ \omega_1^{iter+1} & \dots & \omega_q^{iter+1} \\ t_1^{iter}, \dots, t_{k_1}^{iter} & \dots & t_1^{iter}, \dots, t_{k_q}^{iter} \end{array} \right\}.$$

Шаг 5. Зафиксировать точки плана $x_{(1)}^{iter+1}, \dots, x_{(q)}^{iter+1}$, $\omega_1^{iter+1}, \dots, \omega_q^{iter+1}$ и решить оптимизационную задачу

$$M\left(I\left(\xi \mid Z_{k_1} \leq z_0, \dots, Z_{k_q} \leq z_0\right)\right) + M\left(I(\xi)\right) \rightarrow \max_{t_1, \dots, t_{k_i}}.$$

Вычислить ИМФ $I(\xi^{iter})$ и УИМФ $I(\xi^{iter} \mid Z_{k_1} \leq z_0, \dots, Z_{k_q} \leq z_0)$ в соответствии с полученным планом.

Шаг 6. Проверить условие окончания для получения оптимального плана

$$\xi^{iter+1} = \left\{ \begin{array}{ccc} x_{(1)}^{iter+1} & \dots & x_{(q)}^{iter+1} \\ \omega_1^{iter+1} & \dots & \omega_q^{iter+1} \\ t_1^{iter+1}, \dots, t_{k_1}^{iter+1} & \dots & t_1^{iter+1}, \dots, t_{k_q}^{iter+1} \end{array} \right\}.$$

Если для небольшого положительного числа ε неравенства

$$\left| 100 \cdot \frac{M\left(I\left(\xi^{iter}\right)\right)}{M\left(I\left(\xi^{iter+1}\right)\right)} - 100 \right| \leq \varepsilon \text{ и } \left| 100 \cdot \frac{M\left(I\left(\xi^{iter} \mid Z_{k_1} \leq z_0, \dots, Z_{k_q} \leq z_0\right)\right)}{M\left(I\left(\xi^{iter+1} \mid Z_{k_1} \leq z_0, \dots, Z_{k_q} \leq z_0\right)\right)} - 100 \right| \leq \varepsilon$$

выполняются, то оптимальный план эксперимента найден; в противном случае установить $iter = iter + 1$ и повторить шаги 3–6.

6. Оптимальное планирование эксперимента для проверки надежности светодиодов

Светодиодное освещение имеет множество преимуществ для промышленных и коммерческих предприятий, которые заинтересованы в снижении энергопотребления и затрат.

Вероятностная модель, описывающая функцию надежности, может быть построена по результатам предварительного эксперимента, проведенного в [19]. В течение 250 ч проводилось наблюдение за 24 светодиодами. Нормальный уровень силы тока для светодиода составляет 25 мА. Для проведения ускоренного эксперимента на надежность светодиоды тестировались при двух уровнях электрического тока: 40 мА и 35 мА. Когда интенсивность света уменьшается на 50%, фиксируется отказ светодиода (рис. 1, а).

Винеровская деградиционная модель со степенной функцией тренда, где влияние силы тока описывается степенной функцией от ковариат, оказалась наиболее предпочтительной моделью по результатам исследований в [17]:

$$f(u, \theta_1, \theta_2) = \frac{1}{\sigma \sqrt{2\pi} (t / e^{x\beta})^y} \exp \left(- \frac{\left(u - \mu \left(t / e^{x\beta} \right)^\gamma \right)^2}{2\sigma^2 \left(t / e^{x\beta} \right)^\gamma} \right),$$

с параметрами $(\sigma, \mu, \gamma, \beta) = (0,84; 0,33; 0,48; -1,72)$.

На основе полученной винеровской деградиционной модели построим оптимальный план эксперимента с использованием предлагаемого алгоритма в разделе 5 с учетом заданных ограничений: $25 \leq x_{(i)} \leq 50$, $t_0 = 0$, $t_{k_i} \leq 250$, $t_j - t_{j-1} \geq 5$, $j = 1, k_i$, $i = 1, q$, где $q = 2$.

Рисунок 1, б наглядно иллюстрирует, что в случае использования условной информационной матрицы Фишера оптимальный план эксперимента включает в себя максимально допустимую

нагрузку $x_{(1)} = 50$ и нагрузку $x_{(2)} = 41$, при которой показатель деградации достигает своего критического значения $z_0 = 50$ к моменту завершения эксперимента $T = 250$.

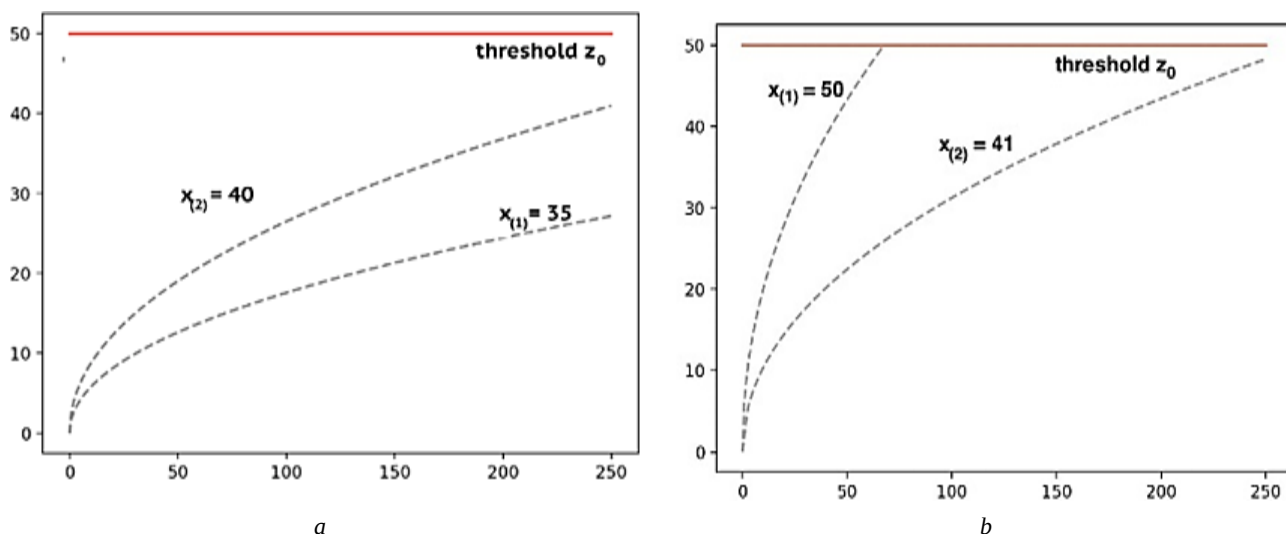


Рис. 1. Функции тренда, соответствующие исходному плану эксперимента (a); функции тренда, соответствующие оптимальному плану эксперимента (b)
Fig. 1. Trend functions for the initial design a); Trend functions for the optimal design b)

В таблице приведены исходный план и полученный D -оптимальный план, а также соответствующие значения определителей ИМФ и УИМФ.

D - оптимальный план эксперимента

	Начальный план	Оптимальный план
План	$\xi^0 = \begin{Bmatrix} 35 & 40 \\ 0.5 & 0.5 \\ 0, 50, \dots, 250 & 0, 50, \dots, 250 \end{Bmatrix}$	$\xi^* = \begin{Bmatrix} 41 & 50 \\ 0.5 & 0.5 \\ 0, 65, 155, 240, 245, 250 & 0, 65 \end{Bmatrix}$
ИМФ	$\det(I(\xi^0)) = 3,9e + 15$	$\det(I(\xi^*)) = 7,8e + 15$
УИМФ	$\det(I(\xi^0)) = 2,6e + 16$	$\det(I(\xi^*)) = 9,2e + 16$

На основании полученных результатов можно сделать вывод, что точность оценок параметров стала выше, так как определитель ИМФ увеличился в 2 раза, УИМФ увеличился в 3 раза.

Заключение

В данной работе предложен алгоритм прямого поиска оптимального плана эксперимента на основе винеровской деградационной модели. Алгоритм позволяет определить оптимальные уровни нагрузок, количество тестируемых объектов и моменты времени измерения деградационного показателя при условии, что все деградационные процессы ограничены заданным критическим уровнем. В качестве примера была рассмотрена задача анализа надежности светодиодов. Следуя предложенному алгоритму, получен D -оптимальный план проверки надежности светодиодов. На основании полученных результатов можно сделать вывод, что точность оценок параметров стала выше, так как функционал от ИМФ и от УИМФ увеличился.

Список источников

1. Nima G., Lin M., Murthy M., Prasad Y., Yong S. A review on degradation models in reliability analysis // Proc. of the 4th World Congress on Engineering Asset Management. Athens, Greece. 2009. P. 379–384. doi: 10.1007/978-0-85729-320-6_42

2. Wang C., Hu Q., Yu D. Wiener processes with random initial degradation values for step-stress accelerated degradation tests data // 2020 Asia-Pacific International Symposium on Advanced Reliability and Maintenance Modeling (APARM). Vancouver, BC, Canada. 2020. P. 1–6. doi: 10.1109/APARM49247.2020.9209514
3. Chimitova E.V., Chetvertakova E.S., Sergeeva S.A., Osinceva E.A. A comparative analysis of the Wiener, Gamma and Inverse gaussian degradation models // Proc. of the intern. workshop “Applied methods of statistical analysis. Nonparametric methods in cybernetics and system analysis”. Krasnoyarsk, 2017. P. 160–167.
4. Meeker W., Hong Y., Escobar L. Degradation models and analyses // Encyclopedia of statistical sciences. John Wiley & Sons, 2011. P. 1–23. doi: 10.1002/0471667196.ess7148
5. Ye Z.-Z., Tsui K.-L., Wang Y., Pecht M. Degradation Data Analysis Using Wiener Processes With Measurement Errors // IEEE Transactions on Reliability. 2013. V. 62 (4). P. 772–780. doi: 10.1109/TR.2013.2284733
6. Chetvertakova E.S., Chimitova E.V. Testing significance of random effects for the gamma degradation model // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2019. № 49. С. 92–100. doi: 10.17223/19988605/49/11
7. Chen X., Ji G., Sun X., Li Z. Inverse Gaussian-based model with measurement errors for degradation analysis // Risk and Reliability. 2019. V. 233, is. 6. doi: 10.1177/1748006X19860682
8. Teng F., Wang H. Analyzing Accelerated Degradation Data via an Inverse Gaussian Degradation Model with Random Parameters // Proc. of Prognostics and System Health Management Conference. Chongqing, China. 2018. P. 1031–1036. doi: 10.1109/PHM-Chongqing.2018.00183
9. Chen Z., Li S., Pan E. Optimal Constant-Stress Accelerated Degradation Test Plans Using Nonlinear Generalized Wiener Process // Hindawi Publishing Corporation. Mathematical Problems in Engineering. 2016. P. 15. doi: 10.1155/2016/9283295
10. Wang X., Wang B., Wu W., Hong Y. et al. Reliability analysis for accelerated degradation data based on the Wiener process with random effects // Quality and Reliability Engineering. 2020. V. 36 (6). P. 1969–1981. doi: 10.1002/qre.2668
11. Chen W.-H., Gao L., Pan J., Qian P., He Q.-C. A novel optimal accelerated degradation test design method considering multiple decision variables // Chinese Journal of Mechanical Engineering. 2018. P. 15. doi: 10.1186/s10033-018-0206-9
12. Ye Z.-S., Chen N., Shen Y. A new class of Wiener process models for degradation analysis // Reliability Engineering & System Safety. 2015. V. 139. P. 58–67. doi: 10.1016/j.ress.2015.02.005
13. Ge Z., Li X. Optimal design for Step-stress ADT based on D-optimality // Reliability & Maintainability Symposium. 2011. P. 1–6. doi: 10.1080/16843703.2016.1189179
14. Lim H., Yum B.J. Optimal design of ADTs based on wiener process models // Journal Applied Statistics. 2011. V. 38, is. 2. P. 309–325. doi: 10.1080/02664760903406488
15. Wang Y., Zhang C. Simulation-based optimal design for ADTs // Maintainability and Safety. 2009. V. 45 (12). P. 1302–1306. doi: 10.1109/ICRMS.2009.5270017
16. Осинцева Е.А., Чимитова Е.В. Информационная матрица Фишера для винеровской деградиационной модели с учетом объясняющих переменных // Обработка информации и математическое моделирование: материалы Рос. науч.-техн. конф. Новосибирск, 2019. С. 92–97.
17. Osintseva, E. Chimitova E. Optimal Design of reliability experiment based on the Wiener degradation model with covariates // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2022. № 59. С. 23–33. doi: 10.17223/19988605/59/3
18. Osintseva, E. Chimitova E. Optimal design of reliability experiment based on the Wiener degradation model under limitation for the degradation index // Proc. of 15th International Conference “Intelligent Systems” (INTELS’22). Moscow, Russia, 14–16 December 2022. doi: 10.3390/engproc2023033064
19. Chaluvadi V. Accelerated life testing of electronic revenue meters : Master Thesis / Clemson University. Clemson, SC, USA. 2008. 130 p.

References

1. Nima, G., Lin, M., Murthy, M., Prasad, Y. & Yong, S. (2009) A review on degradation models in reliability analysis. *Proceedings of the 4th World Congress on Engineering Asset Management*. Athens, Greece. pp. 379–384. DOI: 10.1007/978-0-85729-320-6_42
2. Wang, C., Hu, Q. & Yu, D. (2020) Wiener processes with random initial degradation values for step-stress accelerated degradation tests data. *2020 Asia-Pacific International Symposium on Advanced Reliability and Maintenance Modeling (APARM)*. Vancouver, BC, Canada. pp. 1–6. DOI: 10.1109/APARM49247.2020.9209514
3. Chimitova, E.V., Chetvertakova, E.S., Sergeeva, S.A., Osinceva, E.A. (2017) A comparative analysis of the Wiener, Gamma and Inverse gaussian degradation models. *Proceedings of the Intern. Workshop “Applied methods of statistical analysis. Nonparametric methods in cybernetics and system analysis”*. Krasnoyarsk. pp. 160–167.
4. Meeker, W., Hong, Y. & Escobar, L. (2011) Degradation Models and Analyses. In: *Encyclopedia of Statistical Sciences*. John Wiley & Sons. pp. 1–23. DOI: 10.1002/0471667196.ess7148
5. Ye, Z.-Z., Tsui, K.-L., Wang, Y. & Pecht, M. (2013) Degradation Data Analysis Using Wiener Processes with Measurement Errors. *IEEE Transactions on Reliability*. 62(4). pp. 772–780. DOI: 10.1109/TR.2013.2284733
6. Chetvertakova, E.S. & Chimitova, E.V. (2019) Testing significance of random effects for the gamma degradation model. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 49. pp. 92–100. DOI: 10.17223/19988605/49/11

7. Chen, X., Ji, G., Sun, X. & Li, Z. (2019) Inverse Gaussian-based model with measurement errors for degradation analysis. *Risk and Reliability*. 233(6). DOI: 10.1177/1748006X19860682
8. Teng, F. & Wang, H. (2018) Analyzing Accelerated Degradation Data via an Inverse Gaussian Degradation Model with Random Parameters. *Proceedings of Prognostics and System Health Management Conference*. Chongqing, China. pp. 1031–1036. DOI: 10.1109/PHM-Chongqing.2018.00183
9. Chen, Z., Li, S. & Pan, E. (2016) Optimal Constant-Stress Accelerated Degradation Test Plans Using Nonlinear Generalized Wiener Process. *Mathematical Problems in Engineering*. p. 15. DOI: 10.1155/2016/9283295
10. Wang, X., Wang, B., Wu, W., Hong, Y. et al. (2020) Reliability analysis for accelerated degradation data based on the Wiener process with random effects. *Quality and Reliability Engineering*. 36(6). pp. 1969–1981. DOI: 10.1002/qre.2668
11. Chen, W.-H., Gao, L., Pan, J., Qian, P. & He, Q.-C. (2018) A novel optimal accelerated degradation test design method considering multiple decision variables. *Chinese Journal of Mechanical Engineering*. p. 15. DOI: 10.1186/s10033-018-0206-9
12. Ye, Z.-S., Chen, N. & Shen, Y. (2015) A new class of Wiener process models for degradation analysis. *Reliability Engineering & System Safety*. 139. pp. 58–67. DOI: 10.1016/j.res.2015.02.005
13. Ge, Z. & Li, X. (2011) Optimal design for Step-stress ADT based on D-optimality. *Reliability & Maintainability Symposium*. pp. 1–6. DOI: 10.1080/16843703.2016.1189179
14. Lim, H. & Yum, B.J. (2011) Optimal design of ADTs based on wiener process models. *Journal Applied Statistics*. 38(2). pp. 309–325. DOI: 10.1080/02664760903406488
15. Wang, Y. & Zhang, C. (2009) Simulation-based optimal design for ADTs. *Maintainability and Safety*. 45(12). pp. 1302–1306. DOI: 10.1109/ICRMS.2009.5270017
16. Osintseva, E.A. & Chimitova, E.V. (2019) Fisher's information matrix for the Wiener degradation model with covariate information. *Information Processing and Mathematical Modeling*. Proc. of the Russian Conference. Novosibirsk: [s.n.]. pp. 92–97.
17. Osintseva, E. & Chimitova, E. (2022) Optimal design of reliability experiment based on the Wiener degradation model with covariates. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 59. pp. 23–33. DOI: 10.17223/19988605/59/3
18. Osintseva, E. & Chimitova, E. (2023) Optimal design of reliability experiment based on the Wiener degradation model under limitation for the degradation index. *Proceedings of 15th International Conference "Intelligent Systems" (INTELS'22)*. Moscow, Russia, December 14–16, 2022. DOI: 10.3390/engproc2023033064
19. Chaluvadi, V. (2008) *Accelerated life testing of electronic revenue meters*. Master Thesis. Clemson University. Clemson, SC, USA.

Информация об авторах:

Осинцева Евгения Алексеевна – аспирант кафедры теоретической и прикладной информатики Новосибирского государственного технического университета (Новосибирск, Россия). E-mail: osinceva.j@gmail.com

Чимитова Екатерина Владимировна – доктор технических наук, профессор кафедры теоретической и прикладной информатики Новосибирского государственного технического университета (Новосибирск, Россия). E-mail: chimitova@corp.nstu.ru

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

Information about the authors:

Osintseva Evgeniya A. (Post-Graduate Student of the Department of Theoretical and Applied Informatics of Novosibirsk State Technical University, Novosibirsk, Russian Federation). E-mail: osinceva.j@gmail.com

Chimitova Ekaterina V. (Doctor of Technical Science, Professor of the Department of Theoretical and Applied Informatics of Novosibirsk State Technical University, Novosibirsk, Russian Federation). E-mail: ekaterina.chimitova@gmail.com

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

Поступила в редакцию 07.11.2023; принята к публикации 03.06.2024

Received 07.11.2023; accepted for publication 03.06.2024

ИНФОРМАТИКА И ПРОГРАММИРОВАНИЕ

INFORMATICS AND PROGRAMMING

Научная статья

УДК 519.688

doi: 10.17223/19988605/67/10

Параллельная реализация микромасштабной модели турбулентного движения воздуха и переноса примеси с использованием технологии OpenMP**Дмитрий Викторович Лещинский¹, Евгений Александрович Данилкин²,
Александр Васильевич Старченко³**^{1, 2, 3} *Национальный исследовательский Томский государственный университет, Томск, Россия*^{1, 2, 3} *Институт оптики атмосферы им. академика В.Е. Зуева СО РАН, Томск, Россия*¹ *360flip182@gmail.com*² *ugin@math.tsu.ru*³ *starch@math.tsu.ru*

Аннотация. Представлена разрабатываемая микромасштабная модель турбулентного движения воздуха и переноса пассивной газообразной примеси в уличных каньонах и городских кварталах. Для создания параллельной версии программного обеспечения данной модели предварительно на примере решения одного уравнения конвекции-диффузии были рассмотрены различные технологии параллельного программирования, такие как MPI, OpenMP, OpenACC и CUDA. Для каждой параллельной реализации оценивались ускорение, эффективность и трудоемкость применения рассматриваемой технологии. По результатам вычислительных экспериментов была выбрана наиболее подходящая для распараллеливания всей модели технология параллельного программирования – OpenMP. Результаты применения OpenMP при параллельной реализации микромасштабной модели турбулентного движения воздуха и переноса примеси показали высокую (> 60%) эффективность разработанного параллельного программного обеспечения.

Ключевые слова: технологии параллельного программирования; моделирование турбулентности; численные расчеты; вычислительные эксперименты; ускорение и эффективность.

Благодарности: Работа выполнена при поддержке Министерства науки и высшего образования РФ (№ 075-02-2024-1437)

Для цитирования: Лещинский Д.В., Данилкин Е.А., Старченко А.В. Параллельная реализация микромасштабной модели турбулентного движения воздуха и переноса примеси с использованием технологии OpenMP // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 67. С. 94–105. doi: 10.17223/19988605/67/10

Original article

doi: 10.17223/19988605/67/10

**Parallel implementation of microscale model of turbulent air movement
and pollutant transport using OpenMP technology****Dmitriy V. Leshchinskiy¹, Evgeniy A. Danilkin², Alexander V. Starchenko³**^{1, 2, 3} *National Research Tomsk State University, Tomsk, Russian Federation*^{1, 2, 3} *V.E. Zuev Institute of Atmospheric Optics of SB RAS, Tomsk, Russian Federation*¹ *360flip182@gmail.com*

Abstract. The work presents the developed microscale model of turbulent air movement and the transfer of passive gaseous pollutant in street canyons and city blocks. To parallelize this model, using the example of solving one transport equation, various parallel programming technologies were considered: MPI, OpenMP, OpenACC and CUDA. For each parallel implementation, the acceleration, efficiency and labor intensity of using the technology in question were assessed. Based on the results of computational experiments, the most suitable parallel programming technology – OpenMP for parallelizing the entire model was chosen. The results of using OpenMP in parallel implementation of a microscale model of turbulent air movement and pollutant transfer are presented.

Keywords: parallel programming technologies; turbulence modeling; numerical calculations; computational experiments.

Acknowledgments: This work was supported by the Ministry of Science and Higher Education of Russia (No. 075-02-2024-1437).

For citation: Leshchinskiy, D.V., Danilkin, E.A., Starchenko, A.V. (2024) Parallel implementation of microscale model of turbulent air movement and pollutant transport using OpenMP technology. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 67. pp. 94–105. doi: 10.17223/19988605/67/10

Введение

Современные задачи численного моделирования нередко требуют больших вычислительных затрат. Поэтому наряду с возможностью применения высокопроизводительных вычислительных систем необходимо провести серьезную работу по оптимизации алгоритма для максимально эффективного его использования на системах с общей или распределенной памятью. Множество современных научных работ помимо решения основной задачи включают в себя исследования по максимально эффективному с точки зрения времени подходу для получения искомого решения [1–5]. Самыми популярными и легко сравниваемыми между собой технологиями параллельного программирования являются Open Multi-Processing (OpenMP) и Message Passing Interface (MPI), которые имеют ряд преимуществ и недостатков при их применении.

Параллельное программирование открывает новые возможности эффективного использования современных многоядерных процессоров и гетерогенных вычислительных систем. Однако переход от последовательного программирования к параллельному не лишен проблем [6]. Например, такие опции директив OpenMP, как PRIVATE, THREADPRIVATE, FIRSTPRIVATE, LASTPRIVATE, REDUCTION и COPYPRIVATE, прямо или косвенно используют операции копирования, которые сопровождаются дополнительными накладными расходами на память. Таким образом, применение механизма быстрого копирования больших фрагментов данных будет целесообразнее использовать для крупномасштабных приложений [7].

Если MPI предполагает ручное распределение данных по процессам, и вся ответственность за качество распределения ложится на плечи программиста, то при написании кода с использованием OpenMP распределение данных опционально и не так очевидно на первый взгляд. Эксперименты, представленные в статье [8], говорят о том, что самостоятельное распределение данных не имеет особого смысла, потому как аппаратное обеспечение современных мультипроцессоров хорошо оптимизировано для снижения соотношения задержек при доступе к удаленной и локальной памяти. В [8] рассматривается работа с привычной локальной памятью, а также с удаленной памятью, которая приставляет собой локальную память другого компьютера. Таким образом, любая разумно сбалансированная схема автоматического распределения данных OpenMP будет работать на том же уровне, что и ручная-оптимальная [8], при том что на разработку своей оптимальной схемы распределения необходимы существенные временные затраты.

Область применения OpenMP и MPI также расширяется и на гетерогенные вычислительные системы. Как утверждается в статье [9], в подобных системах, а также в системах, сочетающих центральные процессоры с графическими (GPU) и многоинтегрированными ядрами (MIC), MPI не имеет

большого преимущества по сравнению с OpenMP. В [9] авторы продемонстрировали эффективность параллелизма OpenMP в вычислительной гидродинамике, а также его превосходство над MPI.

Одним из существенных плюсов использования технологии OpenMP является простота ее применения. Для распараллеливания довольно простых циклов настройки любых современных компиляторов по умолчанию не дают сильного расхождения во времени работы параллельной программы. Это связано с тем, что распараллеливание циклов является одной из тем, которая очень интенсивно исследовалась в прошлом и прекрасно реализована в OpenMP. Таким образом, синергия архитектурных факторов и достижений в системном программном обеспечении позволяет простым моделям программирования для систем с общей памятью, таким как OpenMP, сохранять конкурентоспособную позицию по сравнению с моделью передачи сообщений и MPI, сохраняя простоту и переносимость [8].

Цель данной работы – разработка эффективной параллельной реализации микромасштабной модели турбулентного течения с использованием технологии OpenMP. Обоснование выбора данной технологии базируется на сравнении значений ускорения параллельных версий программы численного решения одного уравнения конвекции-диффузии (главной составляющей математической модели), разработанных с помощью технологий MPI, OpenMP, OpenACC (Open Accelerators) и CUDA (Compute Unified Device Architecture). Также при выборе оптимального способа распараллеливания оценивалась трудоемкость процесса применения упомянутых выше технологий. Такой подход к распараллеливанию позволяет оценить трудоемкость процесса параллелизации алгоритма еще на ранних этапах реализации алгоритма и выбрать подходящую технологию параллельного программирования относительно имеющихся вычислительных и временных ресурсов.

1. Постановка задачи

Рассматривается стационарное неизотермическое трехмерное турбулентное движение воздуха в уличном каньоне [10], который представляет собой протяженный участок улицы с параллельно расположенными высокими зданиями с обеих ее сторон (рис. 1). По дороге между зданиями интенсивно движется автотранспорт, выбрасывающий в атмосферу продукты сгорания углеводородного топлива. Высота зданий (H) соизмерима с шириной улицы (W). Примесь считается химически инертной и не осаждается на образующих уличного каньона. Стены и дно уличного каньона могут иметь температуру, отличную от температуры окружающей среды. Это возможно в результате антропогенной деятельности или солнечного нагрева.

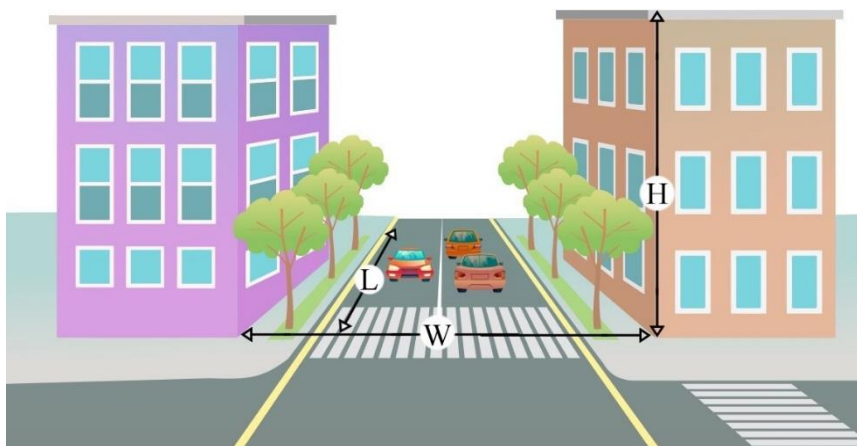


Рис. 1. Иллюстрация физической постановки задачи

Fig. 1. Illustration of the physical statement of the problem

Математическая модель рассматриваемого физического процесса включает в себя осредненные по Рейнольдсу уравнения неразрывности, Навье–Стокса, теплообмена и переноса примеси [10]. Влияние силы плавучести на движение воздуха в уличном каньоне моделируется с помощью приближения Обербека–Буссинеска:

$$\frac{\partial \bar{u}_j}{\partial x_j} = 0; \quad (1)$$

$$\frac{\partial \bar{u}_i \bar{u}_j}{\partial x_j} = -\frac{1}{\rho} \frac{\partial \bar{p}}{\partial x_i} + \frac{\partial}{\partial x_j} \left(\nu \frac{\partial \bar{u}_i}{\partial x_j} \right) - \frac{\partial \bar{u}'_i \bar{u}'_j}{\partial x_j} - \frac{g_i (\bar{T} - T_0)}{T_0}; \quad i = 1, 2, 3; \quad (2)$$

$$\frac{\partial \bar{T} \bar{u}_j}{\partial x_j} = \frac{\partial}{\partial x_j} \left(a \frac{\partial \bar{T}}{\partial x_j} \right) - \frac{\partial \bar{T}' u'_j}{\partial x_j};$$

$$\frac{\partial \bar{C} \bar{u}_j}{\partial x_j} = \frac{\partial}{\partial x_j} \left(D \frac{\partial \bar{C}}{\partial x_j} \right) - \frac{\partial \bar{C}' u'_j}{\partial x_j} + S_c, \quad (4)$$

здесь $\bar{u}_i$, u'_i – осредненная и пульсационная компоненты вектора скорости ($i = 1, 2, 3$), p – давление, $g_i(0, 0, -g)$ – компоненты ускорения свободного падения, $\bar{T}$ – температура, $\bar{C}$ – осредненная концентрация примеси, S_c – функция, описывающая распределение источников примеси, ρ, ν, a, D – плотность, кинематическая вязкость, температуропроводность и коэффициент диффузии, $\bar{u}'_i \bar{u}'_j$ – тензор напряжений Рейнольдса, $\bar{T}' u'_j$, $\bar{C}' u'_j$ – турбулентные потоки тепла и массы, T_0 – температура окружающей среды. По повторяющемуся индексу $j = 1, 2, 3$ производится суммирование.

Замыкание системы уравнений проводится с использованием градиентно-диффузионной гипотезы Буссинеска [11]:

$$\bar{u}'_i \bar{u}'_j = -\nu_T \left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right) + \frac{2}{3} \delta_{ij} k; \quad i, j = 1, 2, 3; \quad (5)$$

$$\bar{T}' u'_j = -\frac{\nu_T}{Pr_T} \frac{\partial \bar{T}}{\partial x_j}; \quad \bar{C}' u'_j = -\frac{\nu_T}{Sc_T} \frac{\partial \bar{C}}{\partial x_j}; \quad j = 1, 2, 3. \quad (6)$$

Здесь $Sc_T = 0,5$ – турбулентное число Шмидта, $Pr_T = 0,9$ – турбулентное число Прандтля. Для вычисления турбулентной вязкости ν_T используется двухпараметрическая k - ϵ -модель [11], учитывающая влияние плавучести на характеристики турбулентности:

$$\frac{\partial k \bar{u}_j}{\partial x_j} = \frac{\partial}{\partial x_j} \left(\left(\nu + \frac{\nu_T}{\sigma_k} \right) \frac{\partial k}{\partial x_j} \right) + P + G - \epsilon; \quad (7)$$

$$\frac{\partial \epsilon \bar{u}_j}{\partial x_j} = \frac{\partial}{\partial x_j} \left(\left(\nu + \frac{\nu_T}{\sigma_\epsilon} \right) \frac{\partial \epsilon}{\partial x_j} \right) + C_{\epsilon 1} \frac{\epsilon}{k} (P + C_{\epsilon 3} G) - C_{\epsilon 2} \frac{\epsilon^2}{k}; \quad (8)$$

$$\nu_T = \frac{C_\mu k^2}{\epsilon}; \quad P = \nu_T \left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right) \frac{\partial \bar{u}_i}{\partial x_j}; \quad G = \delta_{i3} \frac{g_j}{T_0} \frac{\nu_T}{Pr_T} \frac{\partial \bar{T}}{\partial x_j}, \quad (9)$$

где ν_T – турбулентная вязкость, k – кинетическая энергия турбулентности, ϵ – диссипация турбулентной кинетической энергии. Коэффициенты модели турбулентности $\sigma_k = 1,0$, $\sigma_\epsilon = 1,22$, $C_{\epsilon 1} = 1,44$, $C_{\epsilon 2} = 1,92$, $C_{\epsilon 3} = \tanh(|\bar{u}_3| / \sqrt{\bar{u}_1^2 + \bar{u}_2^2})$, $C_\mu = 0,09$.

В качестве граничных условий для системы уравнений (1)–(9) использовались:

- условия первого рода на входе в расчетную область;
- равенство нулю производных по нормали на верхней и боковых границах, а также на выходе из расчетной области;
- на твердых поверхностях условия прилипания и непротекания для соответствующих компонент скорости, равенство нулю производных по нормали от концентрации примеси и метод пристеночных функций для определения трения и теплового потока на стенках и турбулентных характеристик в пристеночных областях [12].

2. Численный метод решения задачи

Рассматриваемая задача решалась численно на неравномерной структурированной сетке, сгущающейся при приближении к твердым поверхностям. При построении конечно-разностных аналогов дифференциальных уравнений используется метод конечного объема, обеспечивающий точное выполнение закона сохранения массы на разностном уровне. Для учета препятствий используется метод фиктивных областей, в соответствии с которым компоненты скорости через грани конечного объема, принадлежащего препятствию, устанавливаются равными нулю. Аппроксимация конвективных членов уравнений переноса выполнена с использованием схемы MLU Ван Лира [13], а диффузионных слагаемых – с помощью центральных разностей. Согласование полей скорости и давления с целью обеспечения выполнения закона сохранения массы (1) на разностном уровне осуществлялось на основе процедуры SIMPLE [14]. Для источников членов конвективно-диффузионных уравнений (4), (7), (8) применялась линеаризация, обеспечивающая неотрицательность в глобальном итерационном процессе таких величин, как концентрация примеси, энергия турбулентности и скорость ее диссипации. Полученные в результате аппроксимации разностные схемы для каждого уравнения конвекции-диффузии решались численно методом релаксации.

3. Подходы к распараллеливанию алгоритма решения одного уравнения переноса

С целью оценки трудозатратности и эффективности способов распараллеливания всех вычислительных блоков модели было проведено распараллеливание алгоритма решения одного уравнения конвекции-диффузии с источником членом (4). В этом случае рассматривается процесс распространения газообразной инертной примеси от постоянного точечного источника за счет движения воздуха и турбулентного перемешивания в идеализированной области, представляющей собой прямоугольный параллелепипед. Источник примеси находится в центре области исследования на некоторой высоте над землей. Предполагается, что примесь не поглощается почвой и может свободно покидать область исследования. Математическая постановка задачи сводится к решению обобщенного трехмерного дифференциального уравнения конвективно-диффузионного переноса вида (4) с граничными условиями Неймана (производная по нормали к поверхности равна нулю). Компоненты вектора скорости и коэффициенты диффузии постоянны и известны.

Для дискретизации используется декартова равномерная сетка. Аппроксимация дифференциальной задачи выполнена методом конечного объема со вторым порядком аппроксимации по пространству. Используемый для численного решения метод идентичен алгоритму численного решения уравнений переноса в модели, которая описана в разделе 2. Численное решение сравнивается с известным аналитическим решением рассматриваемой упрощенной задачи. Предварительные результаты по исследованию ускорения были получены на примере решения одного уравнения конвекции-диффузии на сетке $256 \times 256 \times 32 = 2\,097\,152$ узлов. Распараллеливание выполнено с использованием следующих технологий параллельного программирования: OpenMP, MPI, OpenACC и CUDA. Расчеты проводились на вычислительной системе со следующими характеристиками: два 12-ядерных процессора Intel(R) Xeon(R) Silver 4214 CPU @ 2.20GHz, 192 Гб общей оперативной памяти и два графических ускорителя NVidia RTX2080 Ti.

3.1. Технология Open Multi-Processing (OpenMP)

Рассмотрим самую простую с точки зрения исполнения технологию параллельного программирования – Open Multi-Processing (OpenMP). Данная технология распараллеливания применяется для многопроцессорных (многоядерных) вычислительных систем с общей оперативной памятью. OpenMP также позволяет запускать исполнительный код в режиме гипертрейдинга – технологии, разработанной компанией Intel, которая обеспечивает более эффективное использование ресурсов процессора, позволяя выполнять несколько потоков на каждом ядре. Технология гипертрейдинга дает возмож-

ность получить дополнительную производительность при более низких затратах на транзисторы и электроэнергию [15]. На рис. 2 приведен фрагмент параллельной части OpenMP-программы.

```
double time = 0;
while (time < time_fin)
{
    #pragma omp parallel for private (/*Список приватных переменных*/)
    for (i = 1; i <= Nx; i++)
        for (j = 1; j <= Ny; j++)
            for (k = 1; k <= Nz; k++)
                C[i][j][k] = ...; //Расчет концентрации

    #pragma omp parallel for private (i, j, k) collapse(3)
    for (i = 0; i <= Nx + 1; i++)
        for (j = 0; j <= Ny + 1; j++)
            for (k = 0; k <= Nz + 1; k++)
                C0[i][j][k] = C[i][j][k];

    time += tau;
}
```

Рис. 2. Фрагмент параллельной части OpenMP-программы для решения уравнения переноса
Fig. 2. Fragment of the parallel part of the OpenMP program for solving the transport equation

Для исполнения разработанного исходного последовательного алгоритма в многопоточном режиме перед вычислительными циклами были добавлены директивы, указывающие на параллельное исполнение следующего за ней блока (см. рис. 2). В каждом таком блоке итерации циклов равномерно распределяются между потоками параллельно исполняемых инструкций при условии независимости вычислений элементов массивов. Параллельное программирование с использованием библиотеки OpenMP можно без преувеличения считать простым с точки зрения трудозатратности.

На рис. 3 представлен график ускорения параллельной OpenMP-программы. Ускорение рассчитывается по формуле

$$S_p = \frac{T_1}{T_p}, \quad (10)$$

где T_1 – время выполнения программы одним потоком, T_p – время выполнения программы на p параллельных потоках.

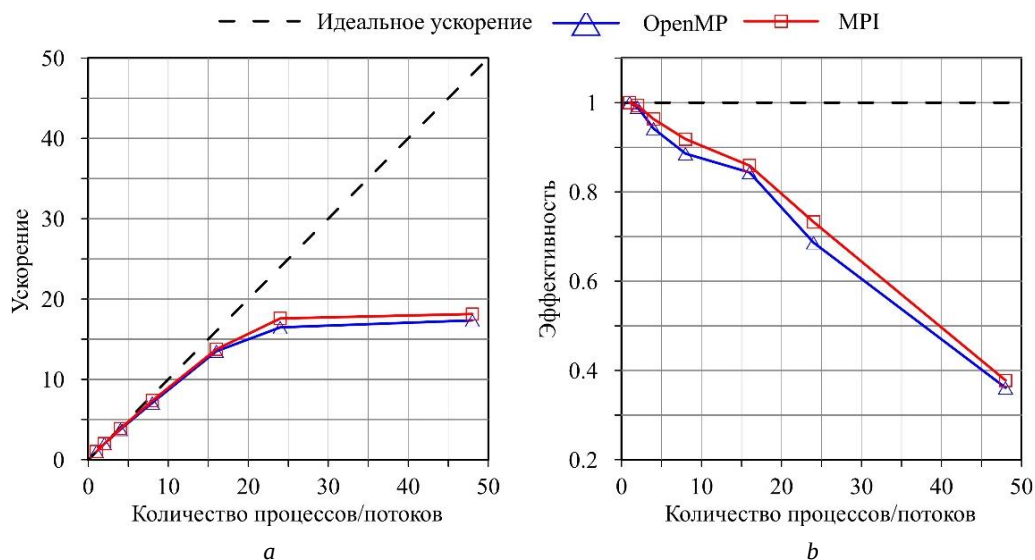


Рис. 3. Ускорение (a) и эффективность (b) работы параллельных MPI- и OpenMP-программ
Fig. 3. Acceleration (a) and efficiency (b) of parallel MPI and OpenMP programs

Результаты показывают, что задействование всех параллельных потоков, доступных на сервере, позволяет ускорить процесс получения решения в 18 раз. Видно, что отклонение от линейного ускорения начинает проявляться при переходе от 16 к 24 потокам. Замедление прироста ускорения объясняется увеличением нагрузки на память, а точнее, уменьшением пропускной способности памяти в расчете на одно активное ядро при задействовании всех доступных физических ядер центрального процессора. Вычислительная система имеет 24 вычислительных ядра, использование гипертрейдинга дает возможность запускать программу на 48 потоках. Однако, как видно из графиков (см. рис. 3), использование гипертрейдинга не дало значительного прироста ускорения, что в целом характерно для программ, активно взаимодействующих с оперативной памятью. Это объясняется тем, что хотя логические ядра процессора имеют свои собственные наборы регистров, они совместно используют почти все физические и аппаратные ресурсы физического ядра [15].

3.2. Технология *Message Passing Interface (MPI)*

Библиотека *Message Passing Interface (MPI)* является средством передачи сообщений между параллельными процессами на системах с распределенной памятью [16]. Использование программистом *MPI* подразумевает самостоятельное распределение вычислительной работы по процессам и организацию обмена данными между параллельными процессами с помощью функций библиотеки передачи сообщений *MPI*. В качестве основного подхода распараллеливания выбрана двумерная (2D) декомпозиция трехмерной сеточной области на подобласти. При разработке параллельной программы используется ряд функций *MPI* для создания двумерной декартовой логической топологии, создания пользовательских типов данных для обеспечения эффективной пересылки между процессами, а также функций *MPI_Isend()* и *MPI_Irecv()* для проведения парных неблокирующих пересылок значений сеточной функции в соседние подобласти двумерной $O_x \times O_y$ -декомпозиции сеточной области.

Ускорение полученного алгоритма в зависимости от числа используемых процессов представлено на рис. 3. Сравнение показывает, что в целом на рассматриваемом вычислительном сервере с общей памятью *MPI*- и *OpenMP*-программы показывают одинаковое ускорение при использовании от 1 до 16 ядер. При дальнейшем увеличении количества задействованных процессов / потоков использование *MPI* позволяет получить большее ускорение работы программы. Важно отметить, что распараллеливание с помощью технологии *MPI* требует значительно больше времени на разработку параллельной программы.

3.3. Технологии создания параллельных программ на графических процессорах

При рассмотрении гибридных вычислительных систем (CPU + GPU) появляется возможность использования вычислительной мощности графических карт. Технология *Open Accelerators (OpenACC)*, как и технология *OpenMP*, с помощью директив компилятора позволяет отмечать участки программы, которые следует выполнять параллельно на графическом процессоре. Предварительно указывается, какие данные должны быть переданы в память GPU, а какие обратно скопированы в общую оперативную память. Результаты расчетов по параллельной программе, созданной с помощью технологии *OpenACC*, показали ускорение в 31 раз по сравнению с последовательной программой.

Compute Unified Device Architecture (CUDA) – технология параллельного программирования, разработанная компанией *NVidia*, позволяющая проводить вычисления на графических процессорах. Распараллеливание работы алгоритма основывалось на принципе двумерной (2D) декомпозиции данных. Использование технологии параллельного программирования *CUDA* для реализации рассматриваемого алгоритма позволяет ускорить получение решения в 38 раз по сравнению с последовательной программой.

Использование технологий параллельного программирования *OpenACC* или *CUDA* и графической карты *NVidia RTX2080Ti* обеспечивает получение достаточно хорошего ускорения в 31 и 38 раз соответственно, однако такое решение сопровождается и высокими временными затратами на парал-

лелизацию алгоритма. Таким образом, при небольшом объеме передаваемых между программными модулями данных или небольшом количестве решаемых уравнений перспективным является применение параллельных технологий программирования на графических процессорах.

4. Параллельная реализация микромасштабной модели турбулентного движения воздуха и переноса примеси

Результаты расчетов решения одного конвективно-диффузионного уравнения показали, что ускорение и эффективность OpenMP-программы практически не уступают аналогичным параметрам MPI-программы на многопроцессорном сервере с общей памятью. Также достаточно хорошее ускорение было получено при использовании технологий OpenACC и CUDA. Параллелизация программ с использованием технологии OpenMP является более простой в сравнении с MPI, OpenACC или CUDA. Исходя из вышеизложенного, для ускорения проведения расчетов по микромасштабной модели турбулентного течения были проведены работы по распараллеливанию программного кода с использованием технологии параллельного программирования OpenMP, ориентированной на многоядерные вычислительные системы с общей оперативной памятью.

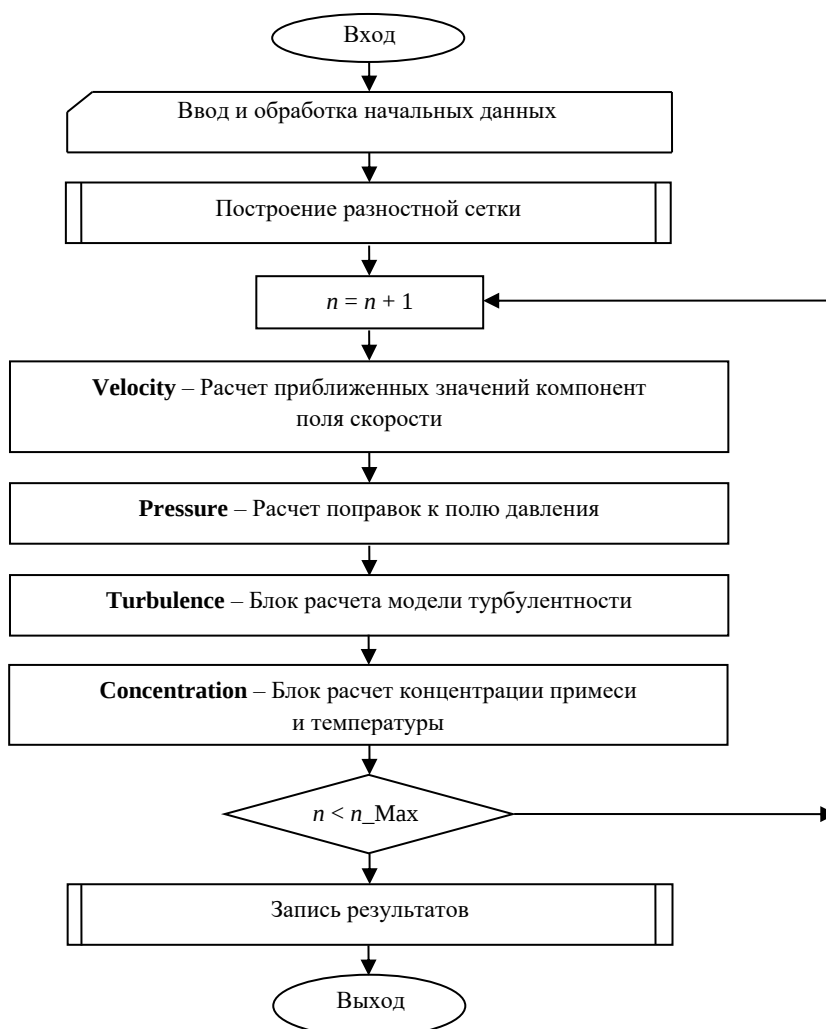


Рис. 4. Блок-схема программы
Fig. 4. Block diagram of the program

При создании параллельной многопоточной программы с использованием технологии OpenMP нужно распараллелить вычислительные блоки, в которых производится перевычисление значений сеточных функций на каждой итерации. Основные по трудоемкости проведения расчетов вычисли-

тельные компоненты программы рассматриваемой микромасштабной модели (рис. 4) – блоки итерационного решения разностного уравнения для поправки давления и разностного уравнения для скалярной сеточной функции (кинетической энергии турбулентности, скорости ее диссипации, температуры, концентрации примеси). Для демонстрации эффективности полученной параллельной версии микромасштабной модели были проведены расчеты на сетке $162 \times 162 \times 47 = 1\,233\,468$ узлов. Расчет длился 1 000 глобальных итераций (~ 14 мин на 24 потоках). На каждой глобальной итерации осуществлялось 120 итераций для решения разностного уравнения для поправки давления с коэффициентом релаксации $\alpha = 1$, 10 итераций для разностных уравнений для скалярных переменных и 6 для компонент скорости с нижней релаксацией ($\alpha = 0,2$). Расчеты проводились на том же вычислительном сервере, что и расчеты по решению одного уравнения конвекции-диффузии (см. разд. 3).

На рис. 5, а представлены графики ускорения вычислительной программы модели турбулентного течения и переноса примеси в уличном каньоне (Total) и ее компонентов (блок расчета компонентов скорости – Velocity, блок расчета поправки давления – Pressure, блок расчета модели турбулентности – Turbulence, блок расчета концентрации и температуры – Concentration). Из графиков видно, что вплоть до использования 24 потоков удастся получить высокие значения ускорения при проведении параллельных вычислений – до 15,5 раз для всей вычислительной модели и до 16,5–17,5 раз для отдельных ее компонентов. Блок с наименьшим ускорением – блок расчета скорости. Ускорение данного блока составило лишь около 13.

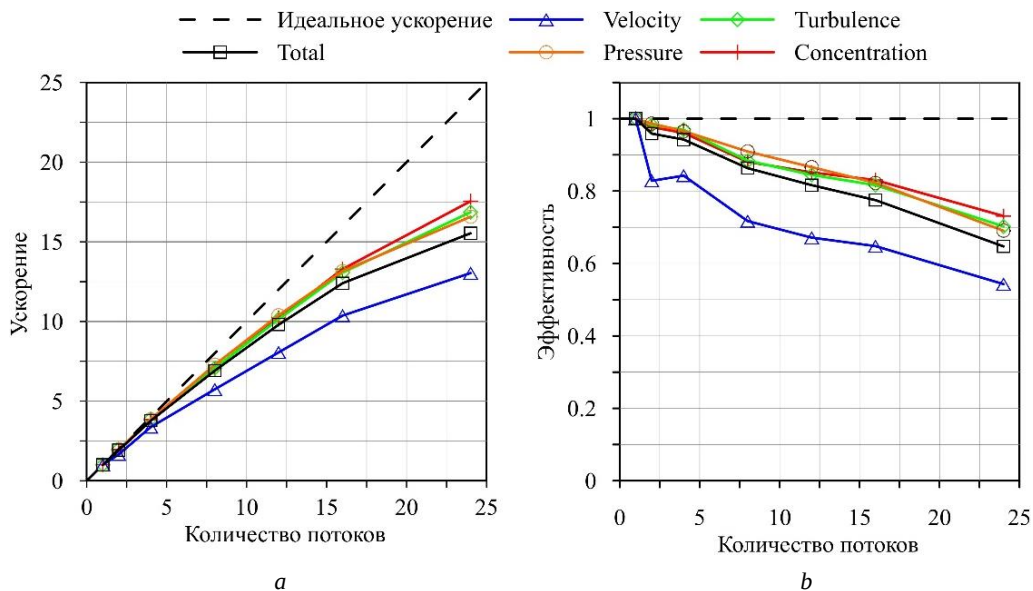


Рис. 5. Ускорение (а) и эффективность (б) параллельной программы и ее компонентов в зависимости от числа используемых потоков

Fig. 5. Acceleration (a) and efficiency (b) of a parallel program and its components depending on the number of threads used

На рис. 5, б приведены графики эффективности вычислительной программы модели в целом и ее компонентов в отдельности. Видно, что повышение эффективности распараллеливания блока расчета компонентов скорости может поднять общий уровень эффективности распараллеливания до 70% на 24 потоках рассматриваемой вычислительной системы, что является неплохим результатом. Также стоит отметить, что использование гипертрейдинга, как и в случае решения одного уравнения переноса, не дает значительного прироста ускорения.

Заключение

Для выбора наиболее перспективной технологии параллельного программирования для параллелизации микромасштабной модели турбулентного движения воздуха и переноса примеси, включающей в себя численное решение неоднородного стационарного трехмерного конвективно-диффузионного

уравнения с помощью метода конечного объема и явно-неявных разностных схем, были проведены вычислительные эксперименты на вычислительном сервере с двумя 12-ядерными процессорами Intel и графическими картами NVIDIA RTX2080 Ti. В качестве технологий параллельного программирования рассматривались международные стандарты программирования OpenMP – для многопроцессорных систем с общей памятью, MPI – для многопроцессорных систем с распределенной памятью, а также OpenACC и CUDA – для гибридных вычислительных систем с графическими процессорами.

Предварительные расчеты по решению одного уравнения конвекции-диффузии, проведенные на структурированных разностных сетках с количеством узлов более 2 млн, показали, что наиболее высокопроизводительную программу для решения рассматриваемой задачи можно получить с помощью технологии MPI. На 24 процессах она дает ускорение вычислений в 17,3 раза по сравнению с расчетом последовательной программы. Более «бюджетные» варианты – использование технологии OpenACC или CUDA для графических процессоров NVIDIA (1 CPU + 1 GPU на одном вычислительном узле) – при тех же условиях дают ускорение вычислений почти в 31 и 38 раз соответственно. Применение технологии OpenMP показало, что ускорение и эффективность OpenMP-программы практически не уступает аналогичным параметрам MPI-программы, что в совокупности с простотой ее использования делает данную технологию наиболее перспективной для параллелизации всей микромасштабной модели.

Результаты параллелизации микромасштабной модели с использованием технологии OpenMP показали, что вплоть до использования 24 потоков удастся получить высокие значения ускорения при проведении параллельных вычислений – до 15,5 раз для всей вычислительной модели и до 16,5–17,5 раз для отдельных ее компонентов. Самый низкий показатель ускорения показал блок вычисления компонент скорости – 13.

Список источников

1. Santos G., Couto E., Paliga A., Cavaleiro G., Avila J. Uma aplicação OpenMP para implementação da estratégia Elemento por Elemento em Análise de Elementos Finitos // Revista de Engenharia Civil IMED. 2023. V. 9. P. 54–67. doi: 10.18256/2358-6508.2022.v9i2.4263
2. Akimova E.N., Sultanov M.A., Misilov V.E., Nurlanuly Y. Parallel Algorithm for Solving the Inverse Two-Dimensional Fractional Diffusion Problem of Identifying the Source Term // Fractal Fract. 2023. V. 7. Art. 801. doi: 10.3390/fractalfract7110801
3. Tang L., Ge M., Schuh H., Wang J., Zhu H., Xu A. Multi-GNSS ultra-rapid orbit determination through epoch-parallel processing // Journal of Geodesy. 2023. V. 97 (11). Art. 99. doi: 10.1007/s00190-023-01787-1
4. Meng F., Shen M. A numerical study on non-spherical droplet impact with solidification in additive manufacturing // Acta Mechanica. 2023. V. 234. P. 1–17. doi: 10.1007/s00707-023-03689-x
5. Amritkar A. Parallel implementation and application of particle scale heat transfer in the Discrete Element Method : PhD Thesis. Blacksburg, VA, 2013. xiv, 123 p. doi: 10.13140/2.1.4214.3526
6. Vinay T.R., Satish E.G., Megha J. Design and Programming for Multicore machines: An Empirical study on time and effort required by programmer // ITM Web Conf. 2023. V. 57. Art. 01016. doi: 10.1051/itmconf/20235701016
7. Marowka A., Liu Z., Chapman B. Openmp-oriented applications for distributed shared memory architectures // Concurr. Comput. Pract. Exper. 2004. V. 16. P. 371–384. doi: 10.1002/cpe.v16:4
8. Nikolopoulos D., Papatheodorou T., Polychronopoulos C., Labarta J. A Transparent Runtime Data Distribution Engine for OpenMP // Scientific Programming. 2002. V. 8 (3). doi: 10.1155/2000/417570
9. Resch M., Sander B., Loebich I. A comparison of OpenMP and MPI for the parallel CFD test case // Proc. of the 1st European Workshop on OpenMP. 2002. P. 71–75.
10. Старченко А.В., Нутерман Р.Б., Данилкин Е.А. Численное моделирование турбулентных течений и переноса примеси в уличных каньонах. Томск : Изд-во Том. ун-та. 2015. 252 с.
11. Henkes R.A.W.M., van der Flugt F.F., Hoogendoorn C.J. Natural Convection Flow in a Square Cavity Calculated with Low-Reynolds-Number Turbulence Models // Int. J. Heat Mass Transfer. 1991. V. 34. P. 1543–1557. doi: 10.1016/0017-9310(91)90258-G
12. Launder B.E., Spalding D.B. The numerical computation of turbulent flows // Computational Methods in Applied Mechanics and Engineering. 1974. V. 3 (2). P. 269–289. doi: 10.1016/0045-7825(74)90029-2
13. Van Leer B. Towards the ultimate conservative difference scheme. II. monotonicity and conservation combined in a second order scheme // Journal of Computational Physics. 1974. V. 14. P. 361–370. doi: 10.1016/0021-9991(74)90019-9
14. Patankar S. Numerical heat transfer and fluid flow. New York : Hemisphere Publ. Corporation, 1980. 214 p. doi: 10.1201/9781482234213

15. Marr D.T., Binns F., Hill D.L. et al. Hyper-Threading Technology Architecture and Microarchitecture // *Intel Technology Journal*. 2002. V. 6, is. 1. P. 4–15.
16. Старченко А.В., Данилкин Е.А., Проханов С.А., Лещинский Д.В. К выбору гибридной схемы распараллеливания для численного решения уравнений пространственной мезомасштабной метеорологической модели атмосферного пограничного слоя // Десятая Сибирская конференция по параллельным и высокопроизводительным вычислениям. Томск : Изд-во НТЛ, 2021. С. 7–17. URL: <http://vital.lib.tsu.ru/vital/access/manager/Repository/koha:000890941>

References

1. Santos, G. & Couto, E., Paliga, A., Cavaleiro, G. & Avila, J. (2023) Uma aplicação OpenMP para implementação da estratégia Elemento por Elemento em Análise de Elementos Finitos. *Revista de Engenharia Civil IMED*. 9. pp. 54–67. DOI: 10.18256/2358-6508.2022.v9i2.4263
2. Akimova, E.N., Sultanov, M., Misilov, V. & Nurlanuly, E. (2023) Parallel Algorithm for Solving the Inverse Two-Dimensional Fractional Diffusion Problem of Identifying the Source Term. *Fractal and Fractional*. 7. Art. 801. DOI: 10.3390/fractalfract7110801
3. Tang, L., Ge, M., Schuh, H., Wang, J., Zhu, H. & Xu, A. (2023) Multi-GNSS ultra-rapid orbit determination through epoch-parallel processing. *Journal of Geodesy*. 97. pp. 1–19. DOI: 10.1007/s00190-023-01787-1
4. Meng, F. & Shen, M. (2023). A numerical study on non-spherical droplet impact with solidification in additive manufacturing. *Acta Mechanica*. 234. pp. 1–17. DOI: 10.1007/s00707-023-03689-x
5. Amritkar, A. (2013) *Parallel implementation and application of particle scale heat transfer in the Discrete Element Method*. PhD Thesis. DOI: 10.13140/2.1.4214.3526.
6. Vinay, T.R., Satish, E.G. & Megha, J. (2023) Design and Programming for Multicore machines: An Empirical study on time and effort required by programmer. *ITM Web of Conferences*. 57. Art. 01016. DOI: 10.1051/itmconf/20235701016
7. Marowka, A., Liu, Z. & Chapman, B. (2004) OpenMP-oriented applications for distributed shared memory architectures. *Concurrency and Computation: Practice and Experience*. 16. pp. 371–384. DOI: 10.1002/cpe.v16:4
8. Nikolopoulos, D., Papatheodorou, Th., Polychronopoulos, C. & Labarta, J. (2002) A Transparent Runtime Data Distribution Engine for OpenMP. *Scientific Programming*. 8(3). DOI: 10.1155/2000/417570
9. Resch, M., Sander, B. & Loebich, I. (2002) A comparison of OpenMP and MPI for the parallel CFD test case. *Proceedings of the 1st European Workshop on OpenMP*. pp. 71–75.
10. Starchenko, A.V., Nuterman, R.B. & Danilkin, E.A. (2015) *Chislennoe modelirovanie turbulentnykh techeniy i perenos primesi v ulichnykh kan'onakh* [Numerical modeling of turbulent flows and pollutant transport in street canyons]. Tomsk: Tomsk State University.
11. Henkes, R.A.W.M., van der Flugt, F.F. & Hoogendoorn, C.J. (1991) Natural Convection Flow in a Square Cavity Calculated with Low-Reynolds-Number Turbulence Models. *International Journal of Heat and Mass Transfer*. 34. pp. 1543–1557. DOI: 10.1016/0017-9310(91)90258-G
12. Launder, B.E. & Spalding, D.B. (1974) The numerical computation of turbulent flows. *Computational Methods in Applied Mechanics and Engineering*. 3(2). pp. 269–289. DOI: 10.1016/0045-7825(74)90029-2
13. van Leer, B. (1974) Towards the Ultimate Conservative Difference Scheme. II. Monotonicity and Conservation Combined in a Second-order Scheme. *Journal of Computational Physics*. 14. pp. 361–370. DOI: 10.1016/0021-9991(74)90019-9
14. Patankar, S. (1980) *Numerical Heat Transfer and Fluid Flow*. New York: Hemisphere Publ. Corporation. DOI: 10.1201/9781482234213
15. Marr, D.T., Binns, F., Hill, D.L. et al. (2002) Hyper-Threading Technology Architecture and Microarchitecture. *Intel Technology Journal*. 6(1). pp. 4–15.
16. Starchenko, A.V., Danilkin, E.A., Prokhanov, S.A. & Leshchinskiy, D.V. (2021) K vyboru gibridnoy skhemy rasparallelivaniya dlya chislennogo resheniya uravneniy prostranstvennoy mezomasshtabnoy meteorologicheskoy modeli atmosfernogo pogrannichnogo sloya [Towards the selection of a hybrid parallelization scheme for the numerical solution of the equations of a spatial mesoscale meteorological model of the atmospheric boundary layer]. In: *Desyataya Sibirskaya konferentsiya po parallel'nym i vysokoproizvoditel'nyim vychisleniyam* [Tenth Siberian Conference on Parallel and High-Performance Computing]. Tomsk: NTL. pp. 7–17. [Online] Available from: <http://vital.lib.tsu.ru/vital/access/manager/Repository/koha:000890941>

Информация об авторах:

Лещинский Дмитрий Викторович – старший преподаватель кафедры вычислительной математики и компьютерного моделирования, младший научный сотрудник Регионального научно-образовательного математического центра Национального исследовательского Томского государственного университета (Томск, Россия); младший научный сотрудник лаборатории прогнозирования состояния атмосферы Института оптики атмосферы им. В.Е. Зуева СО РАН (Томск, Россия). E-mail: 360flip182@gmail.com

Данилкин Евгений Александрович – кандидат физико-математических наук, доцент кафедры вычислительной математики и компьютерного моделирования, старший научный сотрудник Регионального научно-образовательного математического центра Национального исследовательского Томского государственного университета (Томск, Россия); старший научный сотрудник лаборатории прогнозирования состояния атмосферы Института оптики атмосферы им. В.Е. Зуева СО РАН (Томск, Россия). E-mail: ugin@math.tsu.ru

Старченко Александр Васильевич – профессор, доктор физико-математических наук, заведующий кафедрой вычислительной математики и компьютерного моделирования, ведущий научный сотрудник Регионального научно-образовательного математического центра Национального исследовательского Томского государственного университета (Томск, Россия); главный научный сотрудник лаборатории прогнозирования состояния атмосферы Института оптики атмосферы им. В.Е. Зуева СО РАН (Томск, Россия). E-mail: starch@math.tsu.ru

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

Information about the authors:

Leshchinskiy Dmitriy V. (Senior Lecturer of the Department of Computational Mathematics and Computer Modelling, Junior Researcher of the Regional Scientific and Educational Mathematical Center, National Research Tomsk State University, Tomsk, Russian Federation; Junior Staff Scientist, Laboratory for Atmospheric Prediction, V.E. Zuev Institute of Atmospheric Optics of Siberian Branch of the Russian Academy of Science, Tomsk, Russian Federation). E-mail: 360flip182@gmail.com

Danilkin Evgeniy A. (Candidate of Physics and Mathematics, Associate Professor of the Department of Computational Mathematics and Computer Modelling, Senior Researcher of the Regional Scientific and Educational Mathematical Center, National Research Tomsk State University, Tomsk, Russian Federation; Senior Staff Scientist, Laboratory for Atmospheric Prediction, V.E. Zuev Institute of Atmospheric Optics of Siberian Branch of the Russian Academy of Science, Tomsk, Russian Federation). E-mail: ugin@math.tsu.ru

Starchenko Alexander V. (Professor, Doctor of Physical and Mathematical Sciences, Head of Department of Computational Mathematics and Computer Modelling, Leading Researcher of the Regional Scientific and Educational Mathematical Center, National Research Tomsk State University, Tomsk, Russian Federation; Main Staff Scientist, Laboratory for Atmospheric Prediction, V.E. Zuev Institute of Atmospheric Optics of Siberian Branch of the Russian Academy of Science, Tomsk, Russian Federation). E-mail: starch@math.tsu.ru

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

Поступила в редакцию 18.02.2024; принята к публикации 03.06.2024

Received 18.02.2024; accepted for publication 03.06.2024

Original article

UDC 519.876.5

doi: 10.17223/19988605/67/11

Path planning for mobile objects based on modification of the probabilistic roadmap method

Tran Van Truc¹, Anatoly M. Korikov²

^{1,2} Tomsk State University of Control Systems and Radioelectronics, Tomsk, Russian Federation

¹ att82glass@gmail.com

² anaiolii.m.korikov@tusur.ru

Abstract. The article formalizes the task of planning the trajectory of movement of mobile objects, and the minimum length of the trajectory is chosen as a criterion for the optimality of movement. When solving the problem, preference is given to the method of probabilistic route networks and modifications of this method are proposed. The study of the effectiveness of modified algorithms of probabilistic route networks is carried out using examples of solving the problem of automation and optimization of planning the trajectory of a mobile object in an external environment with restrictions on the size of the object. The developed software of the modified algorithms is illustrated by program listings and trajectories of movement of mobile objects.

Keywords: path planning; Probabilistic Roadmap Method; mobile objects; modification of PRM; program structure.

For citation: Tran, V.T., Korikov, A.M. (2024) Path planning for mobile objects based on modification of the probabilistic roadmap method. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 67. pp. 106–115. doi: 10.17223/19988605/67/11

Научная статья

doi: 10.17223/19988605/67/11

Планирование движения мобильных объектов на основе модификации метода вероятностной дорожной карты

Тран Ван Трук¹, Анатолий Михайлович Кориков²

^{1,2} Томский государственный университет систем управления и радиоэлектроники, Томск, Россия

¹ att82glass@gmail.com

² anaiolii.m.korikov@tusur.ru

Аннотация. Формализована задача планирования траектории движения мобильных объектов, в качестве критерия оптимальности движения выбрана минимальная длина траектории. При решении задачи отдано предпочтение методу вероятностных маршрутных сетей и предложены модификации этого метода. Исследование эффективности модифицированных алгоритмов вероятностных маршрутных сетей выполнено на примерах решения задачи автоматизации и оптимизации планирования траектории движения мобильного объекта во внешней среде с ограничениями на габариты объекта. Разработанное программное обеспечение модифицированных алгоритмов иллюстрируется листингами программ и траекториями движения мобильных объектов.

Ключевые слова: планирование движения; вероятностные маршрутные сети; мобильные объекты; модификация алгоритмов; листинги программ.

Для цитирования: Тран В.Т., Кориков А.М. Планирование движения мобильных объектов на основе модификации метода вероятностной дорожной карты // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 67. С. 106–115. doi: 10.17223/19988605/67/11

Introduction

The task of planning the movement of mobile objects is important in many sectors of the economy, health care, and in many fields and applications of technology. Planning the movement of an object (solid body or structure) is the development of the trajectory of its movement from start to finish in the external spatial environment (in the air, on the ground, on water and under water), taking into account obstacles, fuel consumption and other environmental factors [1-3]. When solving this problem, the 3D space including the space without obstacles and the space containing obstacles is called the configuration space of the object [4]. With this concept, the best trajectory is the one with the shortest length and the least number of connecting nodes located in the space without obstacles of the object's configuration space. This trajectory connects the initial point of the object and its ending point, ensuring there is no collision of the object with obstacles in the environment and satisfying the kinematic conditions of the object [4]. From the well-known literature [1–5], it follows that the Probabilistic Roadmap Method (PRM) method has important advantages over other path planning methods. In the PRM method, the process of planning the trajectory of a mobile object (MO) contains two stages: analysis of the external environment and creation of a trajectory of the MO. The PRM method generates movement trajectories on a route network map [4. P. 259–260]. The effectiveness of the route network construction method when applying the PRM method was studied in [6]. The most important characteristic of a route network is connectivity, which can be interrupted in narrow zones of permissible states of the external environment. This problem is usually solved by increasing the number of communication nodes on the map area allowed for MO movement, which leads to an increase in the amount of calculations. Increasing the number of communication nodes complicates the task of finding neighboring vertices and connecting them to each other, even if the planning of the movement of the MO was successful. There may also be problems with increasing the connection radius of neighboring vertices and the number of incident edges. Consequently, the application of the PRM method in practice becomes more difficult. A review of possible ways to solve these problems is contained in [4. P. 259–265]. Application software packages are also known [7–10] that provide solutions to the problems of planning the trajectory of a MO. However, well-known methods and programs show insufficient efficiency in real situations when calculating the trajectories of a MO in space with a complex topology. This article uses the advantages of the PRM method and overcomes the disadvantages of this method based on modifying the PRM and optimizing the trajectory of the MO.

1. Formulation of the problem

Motion trajectory planning, navigation and control of the MO are performed in the external environment. Let us call an acceptable state of the external environment such a state in which direct contact (collision) of the MO with obstacles located in the area of the planned routes of the MO in the external environment is excluded, and which satisfies the specified restrictions on the kinematics and dynamics of the MO. Let us introduce the following sets: X_{PS} is the set of permissible states of the external environment, X_{FS} – the set of forbidden states of the external environment, X_{MO} is the set of MO states. The general formulation of the MO motion planning problem is formulated as follows. For two given states of the MO, $x_{mo}^i \in X_{PS}$ is the initial state and $x_{mo}^f \in X_{FS}$ is the final state, it is required to determine a continuous trajectory of the MO $r(t) : [0,1] \rightarrow X_{PS}$ such that $r(0) = x_{mo}^i$ and $r(1) = x_{mo}^f$. In this formulation, an infinite number of solutions are possible - admissible trajectories (routes) of the MO movement (for some, there may be no even one solution), therefore, we specify the problem of planning the MO movement as an optimization problem with an objective function that allows us to determine the minimum length of the MO trajectory (route). In this formulation of the problem, if the power of the set of states of the external environment that do not conflict with obstacles is large, then this will lead to the creation of a large number of motion trajectories. On the contrary, if the power of the set of states of the external environment that do not conflict with obstacles is small, then this will lead to the fact that the number of created motion trajectories will not be large, or it is even possible that

the motion path of the MO will not be determined. Selecting a sufficiently large set of states of the external environment without colliding with obstacles located in the external environment, planning and optimizing the trajectory of the MO in this external environment is the goal of our research. The optimization problem requires specifying an optimality criterion. The choice of optimality criterion depends on the characteristics of the external environment and the type of MO. When planning the movement of the MO, the following factors must be taken into account: the maximum length of the movement trajectory; obstacles located in the vicinity of the MO movement; dangerous zones in the vicinity of the movement of the MO; consumption of resources (for example, fuel) during the movement of the MO, etc. The best trajectory of the MO movement is the one that has the shortest length, ensures the least consumption of resources (fuel) during the movement of the MO and the greatest safety of its movement. Obviously, such a problem is multiextremal, so let's consider the problem in a simpler formulation: the best trajectory for the MO is the trajectory that has the shortest length. So, the task is set of planning the movement of the MO along the trajectory with the shortest length based on the PRM algorithm and the shortest path search algorithm.

2. Algorithm PRM

The PRM algorithm was proposed in [5] and has been widely applied in creating motion trajectories for unmanned devices such as: unmanned vehicles; unmanned aircraft; robotics, it is also applied to building motion trajectories for robotic manipulators in industry [4]. To create a trajectory for the MO, we first create a route map consisting of two sections: a section without obstacles and a section with obstacles. On the MO route map (Fig. 1), to ensure absolute safety of the MO during its movement, the radius of the area containing obstacles will be increased by an amount depending on the size of the MO and its kinetic characteristics. The white areas in this figure are the areas that will contain a set of states in which there are no collisions between the MO and obstacles, or in other words, these are the areas in which there can be safe movement of the MO. Black areas represent areas containing obstacles and danger zones. The dimensions of these areas have been increased to ensure safe movement of the MO. The route network is formed according to the algorithms described in [4, 5]. To create the trajectory of the MO, a route network is first formed. The creation of a route network is carried out according to the following rule: first, random nodes are created in white areas, then random nodes located next to each other at a given distance d are connected together, forming a route network. The connection of random nodes to each other is carried out according to given rules [4, 5]. After successfully constructing a route network, we connect the starting and ending points of the MO movement trajectory with the route network. After connecting random nodes together and connecting the starting and ending points to the route network, we apply the shortest path algorithm to determine the optimal trajectory of the MO [11, 12].

The quality of the routing network depends on the number of random nodes generated and the radius that allows connections between random nodes. If the number of random nodes is created and the radius allowing connection between them is large, we will have a good routing network, creating a trajectory for the MO will also become easier, but calculating and determining the trajectory for the MO becomes complicated. On the contrary, if the number of random nodes created is small, and the radius allowing connections between random nodes is small, then the created routing network will not be good enough, which will lead to the fact that the search for the desired MO trajectory will be unsuccessful.

Looking at Figures 1 and 2, we can easily see: for Figure 1, after applying the PRM algorithm and the algorithm to find the shortest path with a random number of nodes are 100 nodes and a radius of allowed connection $d = 100\text{m}$ we obtain a moving trajectory consisting of 34 nodes and 33 line segments. It's easy to see that this routing network is not good enough and the moving trajectory is not the best trajectory either. In the routing network, there are still many random nodes that are not connected to each other, the moving trajectory is too long and there are too many navigation points for MO; for Figure 2, applying the method with a random number of nodes are 200 nodes and a connection radius $d = 200\text{m}$, the results show a quite clear difference. A routing network is created that is almost perfect with full connections between random nodes, besides, a more perfect moving trajectory is created with 9 nodes and 8 line segments.

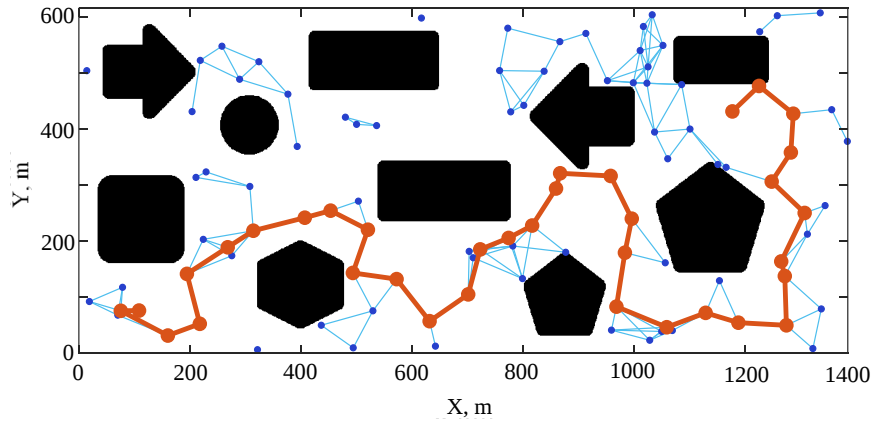


Fig. 1. Route map of a 1400 m x 600 m and shortest path on a network of 100 nodes and $d = 100$ m

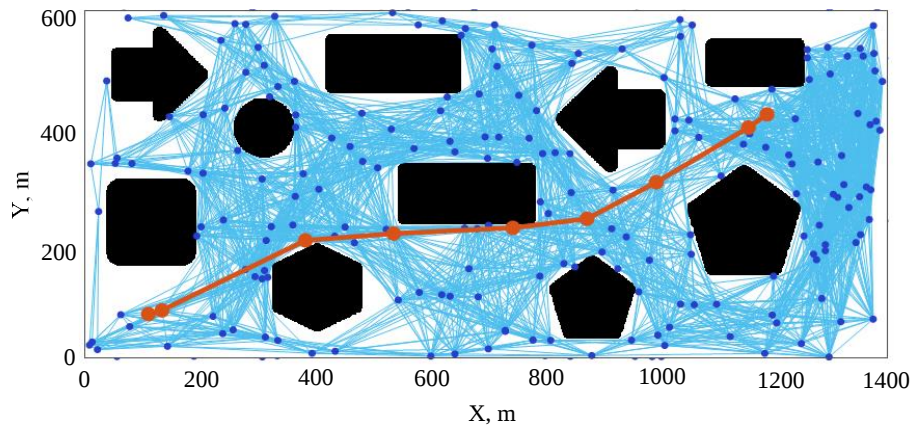


Fig. 2. Shortest path on a network of 200 nodes, $d = 300$ m

So for the PRM algorithm, to create a good moving trajectory we need to create a large number of random nodes and increase the radius that allows connections between them, this also means that a route network is created to be complex, leading to the calculation process of determining the moving trajectory becoming complex and time-consuming, which is not suitable for the navigation characteristics of the MO. In this article we will create a new moving trajectory for the MO based on improving the results of the PRM algorithm and the shortest path finding algorithm.

3. Methods to create trajectory of a mobile object based on PRM algorithm

The MO's moving trajectory is a polyline consisting of many straight lines connecting random nodes together and connected to the start and end points. The best moving trajectory is one, which has the shortest length and the fewest line segments of the trajectory. In this content, we propose 3 methods to create a new flight trajectory for MOs by removing unnecessary nodes after applying the PRM algorithm with the number of randomly generated nodes not being too many and the radius allowing connections between nodes not being too long, this is suitable for low-configuration computers and the navigation characteristics of MOs. Methods include: sequentially connecting nodes in the moving trajectory to the last node, checking and removing unnecessary nodes; check for collisions of 3 consecutive nodes and remove middle nodes; check for collisions by connecting each node in the trajectory to a chain of consecutive nodes, removing unnecessary nodes.

3.1. Create a new trajectory by connecting each node to the last node, checking for collisions, and removing unnecessary nodes

The method of creating a new trajectory for the MO by connecting each node to the last node, checking for collisions, and removing unnecessary nodes is determined as follows:

Input:

path (trajectory): the MO's trajectory is created after using the PRM algorithm and the shortest path finding algorithm
 inroutemap: MO route map

Output:

MO's moving trajectory after being optimized

1. Initialize counter variables and determine the initial length of the MO's moving trajectory

$i = 1;$

$k = 0;$

$l = \text{length}(\text{path})$; (l : total number of nodes on the MO's moving trajectory)

2. Perform sequential connection of nodes on the MO's moving trajectory with the last node ($l-k$) and check for collisions with obstacles, remove redundant points and update the trajectory:

for $i = 1 : l-k$

Check for collision between node i and the last node of the MO's moving trajectory

(Occ = 0: do not collide with obstacles; Occ = 1: collide with obstacles):

Occ = Occupancy(inflatedmap, path($i, :$), path($l-k, :$));

Consider the case where there is no collision with an obstacle:

if Occ == 0

Delete the intermediate nodes between node i and the last node, then update the motion trajectory:

path($i+1:l-k-1, :$) = [];

Reset i to perform the checking process again:

$i = 1;$

Replace the endpoint with the next point adjacent to the endpoint and repeat the process:

$k = k + 1;$

Consider the case of a collision with an obstacle:

elseif Occ == 1

Connect the next point to the last point and continue checking for collisions:

$i = i + 1;$

end

end

3. Update the new movement trajectory of the MO:

newpath = path;

After applying the PRM algorithm and the shortest path finding algorithm, a new motion trajectory is created. However, the number of random nodes and the connection distance between nodes are not large enough, so the created motion trajectory has many unnecessary line segments and is not the trajectory with the best length. To optimize the trajectory, connect the nodes one by one to the last node and check for collisions with obstacles. If collisions with obstacles are absent, redundant nodes are removed from the trajectory.

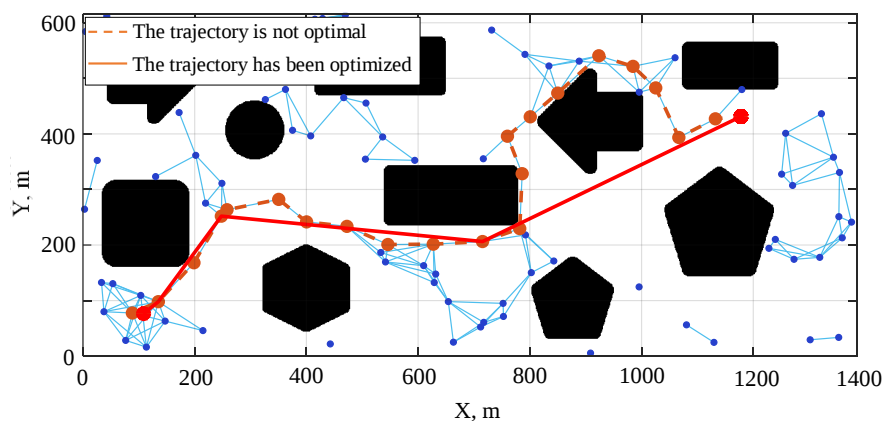


Fig. 3. Create a new trajectory by connecting each node to the last node, checking for collisions, and removing unnecessary nodes

From Figure 3 we see, after applying the PRM algorithm, a trajectory is created with 23 nodes and 22 line segments. Connect each node to the last node in turn and check for collisions. The results show that connecting the 12th node to the last node gives collision test results: no collision with obstacles, so nodes from 13 to node 22 are unnecessary and are removed from the trajectory. Continue to connect the remaining nodes to the 12th node one by one and check for collisions. The results show that when connecting the 5th node to the 12th node, there is no conflict, so nodes 6 to 11 are unnecessary and are removed. Continue connecting the remaining nodes to the 5th node one by one and check that there is no collision at the 3rd node, remove the 4th node. Continue checking the first and third node and remove the 2nd node.

So from Figure 3 we see that after applying this method we obtain a new trajectory consisting of 5 connecting nodes and 4 line segments, this is a quite good trajectory that can be applied to MO navigation.

3.2. Create a new trajectory by checking the collision of 3 consecutive nodes and removing the middle node

The method of creating a new trajectory by checking the collision of 3 consecutive nodes and removing the middle node is determined as follows:

Input:

path (trajectory): the MO's moving trajectory is created after using the PRM algorithm and the shortest path finding algorithm
 inroutemap: MO route map

Output:

MO's moving trajectory after being optimized

1. Initialize the counter variable and determine the initial length of the MO's moving trajectory:

$i = 1;$

$l = \text{length}(\text{path})$; (l: total number of nodes on the MO's moving trajectory)

2. Connect node i and node $i + 2$ on the MO's moving trajectory, check for collisions with obstacles, remove redundant points and update the trajectory:

for $i = 1 : l - 2$

 Check for collision between point i and node $i + 2$ of the MO's moving trajectory

 (Occ = 0: do not collide with obstacles; Occ = 1: collide with obstacles):

 Occ = Occupancy(inflatedmap, path(i,:), path(i+2,:));

 Consider the case where there is no collision with an obstacle:

if Occ == 0

 Delete node $i + 1$ and update the motion trajectory:

 path(i+1,:) = [];

 Reset counter variable i to perform the checking process again:

$i = 1;$

 Consider the case of a collision with an obstacle:

elseif Occ == 1

 Increase i by 1 unit and continue checking for collisions:

$i = i + 1;$

end

end

3. Update the new movement trajectory of the MO:

 newpath = path;

Similar to the first method, after applying the PRM algorithm, we obtain a MO trajectory with 23 nodes and 22 line segments on the trajectory. To create a new, better trajectory, we perform a collision check with obstacles of 3 consecutive nodes on the trajectory and remove the node in the middle. To implement this method, we connect node i and node $i + 2$, then check for collisions with obstacles. If there is no collision with obstacles, we remove node $i + 1$ and update the flight trajectory.

From Figure 4, we first connect the first node with nodes 3, 4 and 5, as a result, there is a collision with the obstacle at node 5, at nodes 3 and 4 there is no collision, so nodes 2, 3 are unnecessary and are removed from the trajectory (connect node 1 and 3, check for collision and delete node 2, update trajectory, continue connecting node 1 and 4 (node 4 as node 3 in the new trajectory) check for collision and delete node 3). Continue connecting node 4 with nodes 6, 7, 8, 9 and 10 and check for collision. As a result, at node 10 there was a collision, at nodes 5, 6, 7, 8 and 9 there was no collision, so nodes 5, 6, 7 and 8 are unnecessary and are removed. Doing the same, we will remove nodes: 10, 11 and 12; nodes 14 and 15; node 17; nodes 19, 20, 21 and 22.

Looking at Figure 4 we see, after applying this method, a new trajectory is created consisting of 7 nodes and 6 line segments, this is also considered a quite good trajectory, this method is simple, the results will be improved if the initial trajectory is simpler.

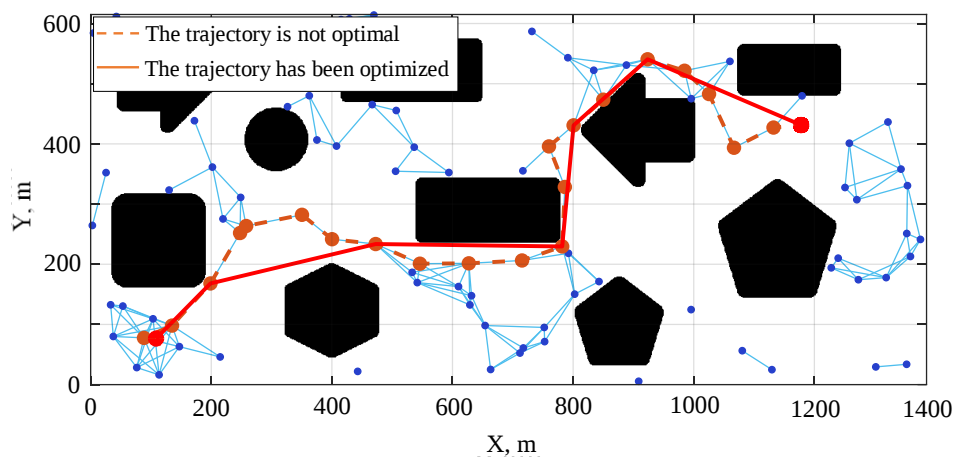


Fig. 4. Create a new trajectory by checking the collision of 3 consecutive nodes and removing the middle node

3.3. Create new trajectories by connecting each node one by one to all other nodes, checking for collisions and removing unnecessary nodes

The method of creating a new trajectory by connecting each node one by one to all other nodes, checking for collisions and removing unnecessary nodes is determined as follows:

Input:

path (trajectory): the MO's moving trajectory is created after using the PRM algorithm and the shortest path finding algorithm

inroutemap: MO route map

Output:

MO's moving trajectory after being optimized

1. Initialize the counter variable and determine the initial length of the MO's moving trajectory:

$i = 1;$

$k = 3;$

$l = \text{length}(\text{path})$; (l : total number of nodes on the MO's moving trajectory)

2. Connect node i and k and check for collisions with obstacles, remove nodes between i and k , and

update the trajectory:

for $i = 1 : l - 2$

for $k = 3 : l$

Check for collision between node i and node k of the MO's moving trajectory ($\text{Occ} = 0$: do not collide with obstacles; $\text{Occ} = 1$: collide with obstacles):

$\text{Occ} = \text{Occupancy}(\text{inflatedmap}, \text{path}(i, :), \text{path}(k, :));$

Consider the case where there is no collision with an obstacle:

if $\text{Occ} == 0$

```

Delete nodes between node i and node k, update the motion trajectory:
path(i+1:k-1, :) = [];
Reset counter variable k to perform the checking process again:
k = 3;
Consider the case of a collision with an obstacle:
elseif Occ == 1
    Increase the counter variable k by 1 and continue checking for collisions:
    k = k + 1;
end
end
Repeat the process of checking collisions between the next node and nodes in the trajectory:
i = i + 1;
k = k + 1;
end
3. Update the new movement trajectory of the MO:
newpath = path;

```

Continue working with the trajectory of 23 nodes and 22 line segments created by applying the PRM algorithm and the shortest path algorithm. To create a new, better trajectory, we connect each node one by one with all remaining nodes, check for collisions and remove unnecessary nodes.

From Figure 5, we first connect the first node with all remaining nodes in the MO's trajectory and check for collisions with obstacles. The collision test results show that there is no collision with obstacles between the first node and the 8th node, so the nodes between the first node and the 8th node are unnecessary and are removed from the trajectory, then update the trajectory. Next, we connect the 8th node with all the remaining nodes in the trajectory and check for collisions. The collision test results show that there is no collision between the 8th node and the 12th node, so the nodes between the 8th node and the 12th node are unnecessary and are removed from the trajectory. Next, connect the 12th node with all remaining nodes in the trajectory, continuing to check for collisions. Collision test results: there is no collision between node 12 and the last node, so the nodes between the 12th node and the last node are unnecessary and are removed.

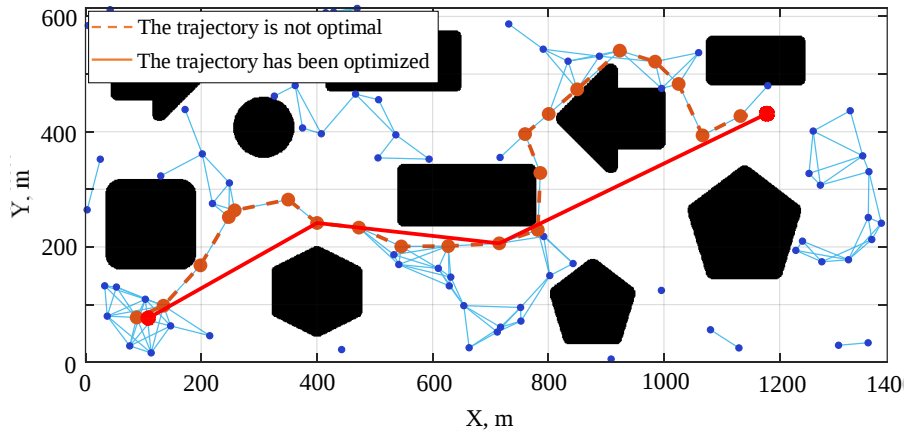


Fig. 5. Create new trajectories by connecting each node one by one to all other nodes, checking for collisions and removing unnecessary nodes

After removing unnecessary nodes and updating the trajectory, we obtain the results as shown in Figure 5. Looking at Figure 5, we see that the newly created trajectory is a polylines consisting of 4 nodes and 3 line segments. This trajectory is much better than the original trajectory, suitable for MO navigation.

3.4. Create a new trajectory using the combined method

In this article, we examined three methods for planning MO movement. From the analysis of Figures 3-5, we see that each method creates a MO trajectory with different lengths, number of nodes and number of line

segments. With this initial trajectory of 23 nodes and 22 line segments, we can have an initial opinion about which method is the best method, but different initial trajectories will produce different results. Therefore, it is impossible to conclude which method is the best method.

To create a new, more optimal trajectory, in this content we add another method, the combined method. Looking at Figure 6, we see that after using the method of creating a new trajectory by checking for collisions with obstacles of 3 consecutive nodes, we obtain a new trajectory which is a blue polylines, including 7 nodes and 6 line segments. Looking at the Figure 6 we see that this is still not the best trajectory, we can remove nodes 3, 4 and 5 or nodes 5 and 6 of the new trajectory. After applying the method of creating a new trajectory by connecting the nodes on the trajectory with the last node in turn for a new trajectory consisting of 7 nodes and 6 line segments, we obtain the final trajectory which is a red trajectory, includes 4 nodes and 3 line segments. This is a trajectory with much less length and line segments.

Depending on the characteristics of each initial trajectory and the desired length after optimization, we can choose 1 of the 3 proposed methods to create a new trajectory for the MO, we can also use a combination of 2 of the 3 methods or a combination of all 3 methods to create a trajectory.

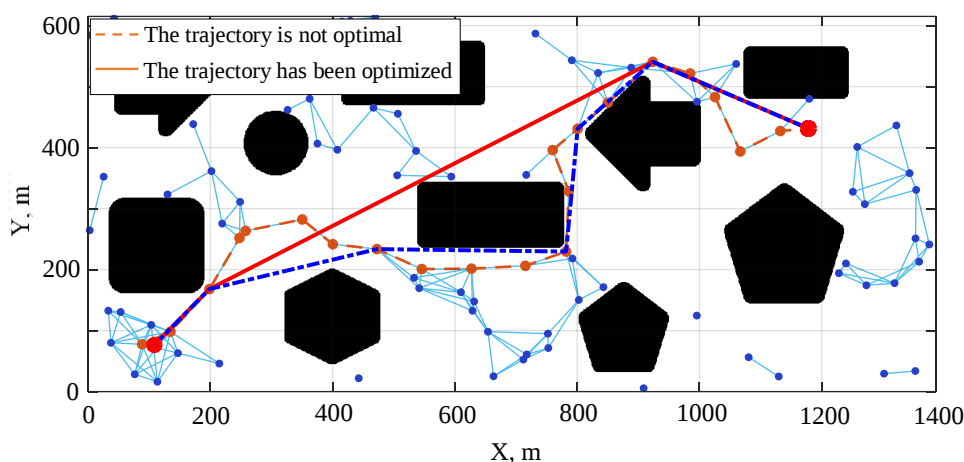


Fig. 6. Create a new trajectory using the combined method

The results of modeling the PRM algorithm and the proposed modifications of the PRM algorithm, presented in Figs 1–6, demonstrate that using modifications of PRM algorithm for MO path planning completely applicable. From the research results of the proposed modified PRM algorithms (Figs. 3–6) follow that the combination of modified PRM algorithms can provide higher efficiency in solving the problem of automation and optimization for MO path planning.

Conclusion

Problems of path planning arise when automating complex technological processes in many sectors of the economy and in particular when planning the path of automated transport. The article formalizes the problem of planning the trajectory of a MO, selects a criterion for the optimality of MO movement, and gives preference to the PRM method. The article proposes three modifications of PRM algorithms. The study of the effectiveness of modified PRM algorithms and combinations of these algorithms was carried out using examples to solve the problem of automation and optimization of trajectory planning for a MO. The simulation results can be recommended for the development of application software for planning routes for various vehicles.

References

1. Choset, H., Lynch, K., Seth, H., Kantor, G., Burgard, W., Kavraki, L.E. & Thrun, S. (2005) *Principles of Robot Motion-Theory. Algorithms and Implementation*. MIT Press.
2. Geraerts, R. & Overmars, M.H. (2007) Creating High-quality Paths for Motion Planning. *International Journal Robotics Research*. 26(8). pp. 845–863. DOI: 10.1177/0278364907079280

3. Karaman, S. & Frazzoli, E. (2011) Sampling-based algorithms for optimal motion planning. *International Journal of Robotics*. 30(7). pp. 846–894.
4. Kazakov, K.A. & Semenov, V.A. (2016) Obzor sovremennykh metodov planirovaniya dvizheniya [Review of modern methods of traffic planning]. *Proceedings of ISP RAS*. 28(4). pp. 241–294.
5. Kavraki, L.E., Svestka, P., Latombe, J.C. & Overmars, M.H. (1996) Probabilistic roadmaps for path planning in high-dimensional configuration spaces. *IEEE Transactions on Robotics and Automation*. 12(4). pp. 566–580. DOI: 10.1109/70.508439
6. Amato, N.M., Bayazit, O.B., Dale, L.K., Jones, C. & Vallejo, D. (2000) Choosing good distance metrics and local planners for probabilistic roadmap methods. *IEEE Transactions on Robotics and Automation*. 16(4). pp. 442–447.
7. Laumond, J.-P. (2006) Kineo CAM: A success story of motion planning algorithms. *IEEE Robotics and Automation Magazine*. 13(2). pp. 90–93. DOI: 10.1109/MRA.2006.1638020
8. Rockel, S., Klimentjew, D., Zhang, L. & Zhang, J. (2014) An hyperreality imagination-based reasoning and evaluation system (HIRES). *Proc. IEEE Int. Conf. on Robotics and Automation*. pp. 5705–5711.
9. Sucan, I., Moll, M. & Kavraki, L.E. (2012) The Open Motion Planning Library. *IEEE Robotics and Automation Magazine*. 19(4). pp. 72–82. DOI: 10.1109/MRA.2012.2205651
10. Porta, J.M., Ros, L., Bohigas, O., Manubens, M., Rosales, C. & Jaillet, L. (2014) The CUIK Suite: Motion Analysis of Closed-chain Multibody Systems. *IEEE Robotics and Automation Magazine*. 21(3). pp. 105–114. DOI: 10.1109/MRA.2013.2287462
11. Chernorutsky, I.G. (2011) *Optimization Methods*. Computer Technologies. St. Petersburg: BHV-Petersburg.
12. Wikipedia. (n.d.) *Dijkstra's algorithm*. [Online] Available from: https://wikipedia.org/wiki/Dijkstra%27s_algorithm (Accessed: 22nd October 2023).

Information about the authors:

Tran Van Truc (Post-Graduate Student of the Department of Automated Control Systems, Tomsk State University of Control Systems and Radioelectronics (TUSUR), Tomsk, Russian Federation). E-mail: att82glass@gmail.com

Korikov Anatoly M. (Doctor of Technical Science, Professor, Department of Automated Control Systems, Tomsk State University of Control Systems and Radioelectronics (TUSUR), Tomsk, Russian Federation). E-mail: anaiolii.m.korikov@tusur.ru

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

Информация об авторах:

Тран Ван Трук – аспирант кафедры «Автоматизированные системы управления» Томского государственного университета систем управления и радиоэлектроники (Томск, Россия). E-mail: att82glass@gmail.com

Кориков Анатолий Михайлович – профессор, доктор технических наук, профессор кафедры «Автоматизированные системы управления» Томского государственного университета систем управления и радиоэлектроники (Томск, Россия). E-mail: anaiolii.m.korikov@tusur.ru

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

Received 15.01.2024; accepted for publication 03.06.2024

Поступила в редакцию 15.01.2024; принята к публикации 03.06.2024

ПРОЕКТИРОВАНИЕ И ДИАГНОСТИКА ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

DESIGNING AND DIAGNOSTICS OF COMPUTER SYSTEMS

Научная статья

УДК 519.7

doi: 10.17223/19988605/67/12

Построение эквивалентных схем в системах троирования

Анжела Юрьевна Матросова¹, Антонина Дмитриевна Пасько², До Там Дык³

^{1, 2, 3} *Национальный исследовательский Томский государственный университет, Томск, Россия*

¹ *mau11@yandex.ru*

² *pasko.tonya@yandex.ru*

³ *kmaddt@gmail.com*

Аннотация. Троирование – один из широко используемых на практике подходов к обеспечению надежности функционирования логических схем. Появившиеся в последние годы возможности одновременного введения в каждую копию и соответствующие линии вредоносных подсхем (Trojan Circuits) делают метод троирования уязвимым к таким действиям. Выходом из этой ситуации является использование вместо трех идентичных схем либо двух аппроксимирующих схем, либо трех эквивалентных схем, отличающихся структурной реализацией друг от друга. Построение эквивалентных схем в отличие от использования аппроксимирующих схем не порождает незащищенной области в системе троирования. Предлагается подход к синтезу эквивалентных схем на основе искажения безыбыточной системы ДНФ (БСДНФ), описывающей поведение рабочей схемы системы троирования с последующим использованием искаженных систем в качестве задания на синтез схем тем же методом, что и синтез по БСДНФ.

Ключевые слова: комбинационные схемы; безыбыточные системы ДНФ (БСДНФ); эквивалентные схемы; троирование.

Для цитирования: Матросова А.Ю., Пасько А.Д., До Там Дык. Построение эквивалентных схем в системах троирования // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2024. № 67. С. 116–123. doi: 10.17223/19988605/67/12

Original article

doi: 10.17223/19988605/67/12

Deriving equivalent circuits for TMR technique

Anzhela Yu. Matrosova¹, Antonina D. Pasko², Do Tam Duc³

^{1, 2, 3} *National Research Tomsk State University, Tomsk, Russian Federation,*

¹ *mau11@yandex.ru*

² *pasko.tonya@yandex.ru*

³ *kmaddt@gmail.com*

Abstract. Triple-Modular Redundancy (TMR) technique is one of the widely used approaches to provide reliable functioning of logical circuits. In recent years, it has become possible to insert Trojan Circuits into the same line of each identical circuit of TMR, such an opportunity makes the TMR technique vulnerable. The way out of this situation is to use instead of three identical circuits either two approximate circuits or three equivalent circuits that differ in structural implementation from each other. The obtained equivalent circuits, in contrast to the use of approximating

circuits, does not create an unprotected area in the TMR technique. An approach is proposed to obtain equivalent circuits based on distortion of the irredundant system of SoPs, which describes the behavior of the primary circuit of the TMR technique. The resulting distorted systems of SoPs and the original system of SoPs are used for the synthesis of circuits.

Keywords: combinational circuits; irredundant system of SoPs; equivalent circuits; TMR technique.

For citation: Matrosova, A.Yu., Pasko, A.D., Do Tam Duc (2024) Deriving equivalent circuits for TMR technique. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 67. pp. 116–123. doi: 10.17223/19988605/67/12

Введение

Системы троирования являются одним из широко распространенных подходов к обеспечению надежности функционирования логических схем. Предполагается, что одна из схем в системе троирования может быть неисправной. В этом случае в условиях подключения схемы голосования к одноименным выходам трех схем значение выхода системы троирования определяется значением большинства одноименных выходов схем, ее составляющих. Схемы голосования предполагаются исправными. В современных условиях производство логической схемы может выполняться различными фирмами, в том числе и в разных частях света. В связи с этим появляется возможность включения вредоносной подсхемы в каждую из идентичных схем, например в соответствующую линию каждой из трех схем системы троирования. Вредоносная подсхема (Trojan circuit) может изменить в нужный момент значение сигнала на этих линиях с целью искажения работы системы или извлечения конфиденциальной информации из устройства, содержащего рассматриваемую систему. Это значит, что технология троирования оказывается уязвимой в условиях возможности включения вредоносных подсхем. Одним из известных подходов к защите системы троирования является использование вместо трех идентичных схем либо двух аппроксимирующих схем и рабочей схемы [1–3], либо трех эквивалентных схем, отличающихся структурной реализацией друг от друга. Построение эквивалентных схем в отличие от использования аппроксимирующих схем не порождает незащищенной области в системе троирования. Предлагается алгоритм расщепления простых импликант БСДНФ, сохраняющий систему функций, заданную БСДНФ рабочей схемы. С помощью расщепления части конъюнкций формируются системы ДНФ для синтеза эквивалентных схем системы троирования, содержащие конъюнкции меньших рангов, чем БСДНФ. Все три эквивалентные системы ДНФ подаются на вход системы синтеза логических схем автоматизированной системы проектирования с целью получения реализации систем ДНФ в виде логических схем. Экспериментальные результаты показали, что расщепление 10–20% конъюнкций приводит к получению эквивалентных схем различной структуры, незначительно отличающихся по числу вентилях в них от рабочей схемы системы троирования.

1. Построение эквивалентных систем булевых функций

Пусть описание поведения рабочей комбинационной схемы представлено в виде БСДНФ, задающей систему $G(x_1, \dots, x_n)$ из m булевых функций. БСДНФ состоит из простых импликант системы функций. Каждая простая импликанта (K, h) БСДНФ состоит из конъюнкции K и характеристики h , в характеристике перечислены функции системы, для которых эта конъюнкция является допустимой. Будем иметь в виду, что ни одна из функций не может быть исключена из характеристики простой импликанты БСДНФ.

Конъюнкция K допустима, если пересечение соответствующего ей интервала с областью нулевых наборов значений переменных булевой функции, упомянутой в характеристике конъюнкции, пусто. Это значит, что рассматриваемый интервал содержится в области единичных наборов значений переменных этой функции.

Поскольку конъюнкция БСДНФ является простой импликантой системы булевых функций, то из нее нельзя исключить ни одной (любой) литеры в условиях сохранения характеристики этой конъюнкции.

Будем иметь в виду, что выбрасывание переменной x_i из конъюнкции K простой импликанты системы приводит к пересечению с областью нулевых значений не каждой функции из характеристики h этой конъюнкции, а только некоторых из них. Непустое пересечение обязательно хотя бы для одной из функций характеристики h .

Выделим функцию характеристики h , для которой выбрасывание переменной x_i из конъюнкции K не приводит к искажению функции.

Будем называть конъюнкцию K расширяемой по переменной x_i для исследуемой функции.

Таким образом, для каждой переменной x_i конъюнкции K из БСДНФ множество функций ее характеристики h разбивается на два подмножества h_p и h_{np} , расширяемых и не расширяемых по переменной x_i .

Конъюнкция K является простой импликантой функции, принадлежащей характеристике h , если для каждой переменной, присутствующей в конъюнкции, h_p пусто. Следовательно, для конъюнкции K множество функций ее характеристики делится на два подмножества h_p простых импликант и подмножество h_{np} , для которых эта конъюнкция простой импликантой не является. Введем операцию расщепления простой импликанты (K, h) БСДНФ на две конъюнкции.

Операция расщепления простой импликанты системы булевых функций

1. Для каждой функции f_i из h_{np} строим простую импликанту $K(f_i)$, выбрасывая в некотором порядке соответствующие литеры конъюнкции K .
2. Выбираем функцию из h_{np} , для которой $K(f_i)$ характеризуется минимальным рангом.
3. Обозначаем выбранную конъюнкцию (K^*, f_i) . Она является простой импликантой заданной системы булевых функций с единственной функцией в характеристике.
4. Заменяем конъюнкцию (K, h) двумя конъюнкциями (K^*, f_i) и конъюнкцией $(K, h/f_i)$.

Замену конъюнкции (K, h) двумя конъюнкциями (K^*, f_i) , $(K, h/f_i)$ назовем расщеплением простой импликанты БСДНФ, а полученные конъюнкции будем называть продуктами расщепления.

Утверждение 1. Операция расщепления конъюнкции (K, h) не приводит к искажению области единичных значений функций системы $G(x_1, \dots, x_n)$.

Доказательство следует из способа реализации операция расщепления.

Следствие. Выполнение операций расщепления нескольких конъюнкций БСДНФ сохраняет функции системы $G(x_1, \dots, x_n)$, т.е. позволяет получать формулы, представляющие эквивалентные системы ДНФ.

Формулу, полученную в результате расщепления нескольких конъюнкций БСДНФ, будем обозначать СДНФ.

Поясним процедуру расщепления конъюнкции БСДНФ на примере. Пусть табл. 1 представляет систему БСДНФ. В ее левых столбцах записаны троичные векторы, задающие конъюнкции, а в правых столбцах представлены характеристики этих конъюнкций. Единичная компонента указывает номера функций, для которых рассматриваемая конъюнкция входит в область их единичных наборов.

Выберем для расщепления конъюнкцию $K = 00-11-$ с соответствующей ей характеристикой $h = 00110101$. Выбранная конъюнкция и ее характеристика выделены в табл. 1 красным цветом. Матричные задания функций данной характеристики представлены на рис. 1. Синим цветом выделен интервал, представленный троичным вектором $00-11-$, соответствующий конъюнкции K . Из приведенных матриц видно, что рассматриваемая конъюнкция является простой импликантой для функций f_4 , f_8 . Поэтому расширяем до простых импликант конъюнкцию K для функций f_3 , f_6 . Имеем: $---11-$, f_6 и $-011-$, f_3 . Для расщепления выбираем простую импликанту меньшего ранга $---11-$, f_6 . Простые импликанты функций f_3 , f_6 на рис. 1 выделены красным цветом.

Итак, конъюнкция $K = 00-11-$ с характеристикой $h = 00110101$ расщепляется на импликанты $00-11-$ с характеристикой 00110001 и простую импликанту $---11-$ с характеристикой 00000100 .

После выполнения операции расщепления над конъюнкцией БСДНФ проверяем полученную систему на безызбыточность по характеристикам. В простых импликантах $01-11-$, 01100110 ; $10-11-$, 10100110 ; $11-11-$, 11000110 функция f_6 в характеристиках является избыточной, ее следует удалить

из характеристик этих конъюнкций. В результате имеем: 01-11-, 01100010; 10-11-, 10100010; 11-11-, 11000010.

Таблица 1

Безыбыточная система ДНФ

x ₁	x ₂	x ₃	x ₄	x ₅	x ₆	f ₁	f ₂	f ₃	f ₄	f ₅	f ₆	f ₇	f ₈
0	0	1	1	—	—	0	0	1	1	1	1	0	0
0	0	1	—	1	—	0	0	1	1	1	1	0	0
0	0	—	—	1	1	0	0	1	1	1	1	0	0
—	1	1	1	—	—	0	0	1	1	1	0	1	0
1	—	1	1	—	—	0	0	1	1	1	0	1	0
0	0	1	—	—	1	0	0	1	1	1	0	1	0
—	1	1	—	1	—	0	0	1	1	1	0	0	1
1	—	1	—	1	—	0	0	1	1	1	0	0	1
—	1	1	—	—	1	0	0	1	1	0	1	1	0
1	—	1	—	—	1	0	0	1	1	0	1	1	0
0	0	—	1	—	1	0	0	1	1	0	1	0	1
—	1	—	1	—	1	0	0	1	1	0	1	0	1
1	—	—	1	—	1	0	0	1	1	0	1	0	1
0	1	—	1	1	—	0	1	1	0	0	1	1	0
1	0	—	1	1	—	1	0	1	0	0	1	1	0
1	1	—	1	1	—	1	1	0	0	0	1	1	0
0	0	—	1	—	1	0	0	1	1	0	0	1	1
—	1	—	—	1	1	0	0	1	1	0	0	1	1
1	—	—	—	1	1	0	0	1	1	0	0	1	1

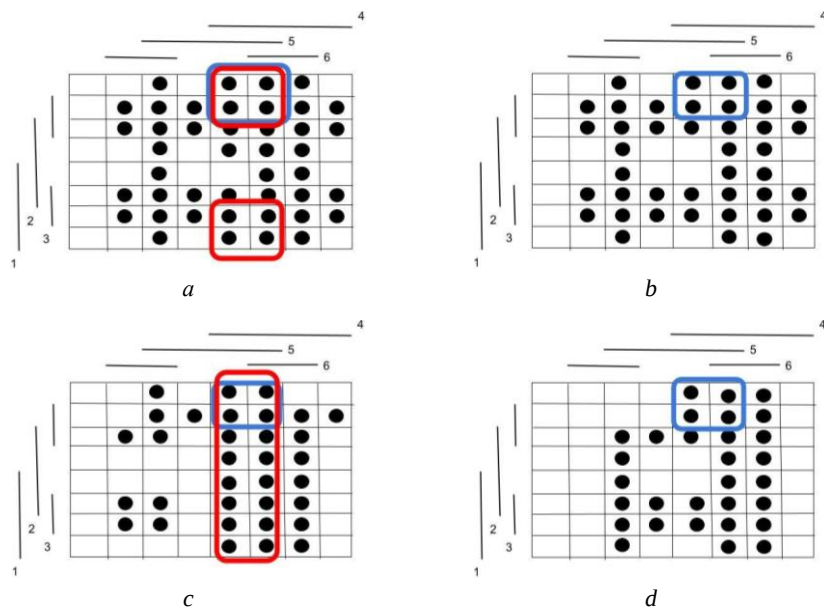


Рис. 1. Матричные представления функций из характеристики $h = 00110101$:

a – функция f_3 , b – функция f_4 , c – функция f_6 , d – функция f_8

Fig. 1. Matrix representations of functions from the characteristic $h = 00110101$:

a is function f_3 , b is function f_4 , c is function f_6 , d is function f_8

Теперь опишем алгоритм построения эквивалентных систем булевых функций:

1. Выбираем импликанту БСДНФ.
2. Проводим ее расщепление.
3. Заменяем порождающую расщепление импликанту БСДНФ парой продуктов расщепления и вносим результат расщепления в первую эквивалентную систему СДНФ₁.
4. Выполняем пункты 1–3 для следующей конъюнкции и вносим результат расщепления во вторую эквивалентную систему СДНФ₂.

5. Включив результаты расщепления в каждую из строящихся эквивалентных систем, добавляем в каждую из систем те конъюнкции из БСДНФ системы $G(x_1, \dots, x_n)$, которые не порождали пары для данной эквивалентной системы.

6. В полученных системах для некоторых конъюнкций характеристики могут содержать избыточные функции, их следует удалить.

В исходной БСДНФ определены только множества наборов единичных значений функций системы, но для проверки того, что конъюнкция после удаления переменной x_i является импликантой рассматриваемой функции, полезно знать множества наборов нулевых значений функций этой системы. Алгоритм проверки основан на применении операций над ROBDD-графами. Итак, имеем конъюнкцию K/x_i и соответствующую ей функцию f_i .

1. Строим для f_i ROBDD-граф, меняем его терминальные вершины местами, получая тем самым ROBDD-граф R_1 для инверсии функции f_i .

2. Строим для проверяемой конъюнкции K/x_i ROBDD-граф R_2 .

3. Вычисляем пересечение ROBDD-графов $R_1 \cap R_2$. Если в результате получим пустое множество, то конъюнкция K/x_i является импликантой функции f_i .

2. Результаты экспериментов над контрольными схемами

Для синтеза схем используется система логического проектирования ABC [4]. ABC – система синтеза и верификации комбинационных и последовательностных схем, разработанная в Беркли, США. Описание полученной синтезированной схемы представлено в формате BLIF.

Для экспериментов использовались наборы бенчмарков LGSynth'91 [5] и IWLS'93 [6]. Среди бенчмарков IWLS'93 рассматривались схемы, уже заданные в формате .pla (табл. 2, 4). Схемы LGSynth91 изначально были описаны в виде конечных автоматов (формат kiss2). Для схем в формате kiss2 сначала производилось кодирование состояний кодами минимальной длины, а затем получались безыбыточные системы ДНФ с помощью программы ESPRESSO [7] (табл. 3, 5).

Ниже представлены сравнительные таблицы полученных схем (см. табл. 2–5), где i – количество входов, o – количество выходов, p – количество конъюнкций БСДНФ, k – количество расщепляемых конъюнкций, S_i – процентное изменение количества элементов эквивалентной системы по отношению к количеству элементов исходной системы. Эксперименты проводились с разным количеством расщепляемых конъюнкций. В табл. 2,3 представлены результаты для систем, полученных в результате расщепления 10% конъюнкций от общего количества конъюнкций исходной (рабочей) системы, а в табл. 4, 5 – для систем, полученных в результате расщепления 20% конъюнкций от общего количества конъюнкций исходной (рабочей) системы.

Таблица 2

Схемы IWLS'93

Схема	i	o	p	k	Кол-во эл-тов (осн. схема)	Кол-во эл-тов (эквив. схема 1)	S_1 %	Кол-во эл-тов (эквив. схема 2)	S_2 %
squar5	5	8	25	3	65	62	–4,62	60	–7,69
bw	5	28	22	2	255	243	–4,71	243	–4,71
inc	7	9	32	3	151	151	0	146	–3,31
rd84	8	4	255	26	275	337	22,55	340	23,64
ex5p	8	63	74	7	1 307	1 382	5,74	1 397	6,89
apex4	9	19	436	44	3 130	3 109	–0,67	3 023	–3,42
sao2	10	4	58	6	169	174	2,96	167	–1,18
ex1010	10	10	284	28	2 177	2 169	–0,37	2 157	–0,92
table3	14	14	175	18	2 146	2 116	–1,4	2 137	–0,42
misex3c	14	14	197	20	701	704	0,43	702	0,14
pdc	16	40	145	15	715	739	3,36	739	3,36
spla	16	46	260	26	1 055	1 112	5,4	1 099	4,17
table5	17	15	158	16	1 959	1 968	0,46	1 993	1,74
duke2	22	29	86	9	678	660	–2,65	644	–5,01
apex1	45	45	206	21	2 607	2 590	–0,65	2 599	–0,31

Таблица 3

Схемы LGSynth'91

Схема	i	o	p	k	Кол-во эл-тов (осн. схема)	Кол-во эл-тов (эквив. схема 1)	S_1 %	Кол-во эл-тов (эквив. схема 2)	S_2 %
dk27	4	5	11	1	30	31	3,33	33	10
dk16	7	8	85	9	369	378	2,44	382	3,52
tbk	11	8	244	24	1 378	1 364	-1,02	1 332	-3,34
s298	11	14	974	97	7 330	7 358	0,38	7 331	0,01
keyb	12	7	90	9	337	337	0	340	0,89
s1	13	11	94	9	634	621	-2,05	612	-3,47
planet	13	25	102	10	979	954	-2,55	980	0,1
planet1	13	25	105	11	969	1 006	3,82	988	1,96
styr	14	15	124	12	936	955	2,03	970	3,63
s1488	14	25	132	13	782	767	-1,92	765	-2,17
s1494	14	25	132	13	744	725	-2,55	732	-1,61
kirkman	16	10	188	19	388	424	9,28	440	13,4
sand	16	14	115	12	926	943	1,84	930	0,43
s820	23	24	90	9	344	331	-3,78	336	-2,33
s832	23	24	90	9	344	337	-2,03	331	-3,78
scf	34	63	165	17	1 558	1 564	0,39	1 548	-0,64

Таблица 4

Схемы IWLS'93

Схема	i	o	p	k	Кол-во эл-тов (осн. схема)	Кол-во эл-тов (эквив. схема 1)	S_1 %	Кол-во эл-тов (эквив. схема 2)	S_2 %
squar5	5	8	25	5	65	60	-7,69	60	-7,69
bw	5	28	22	4	255	248	-2,75	243	-4,71
inc	7	9	32	6	151	139	-7,95	138	-8,61
ex5p	8	63	74	15	1 307	1 349	3,21	1 360	4,06
apex4	9	19	436	87	3 130	3 071	-1,88	3 052	-2,49
sao2	10	4	58	12	169	165	-2,37	173	2,37
ex1010	10	10	284	57	2 177	2 133	-2,02	2 143	-1,56
table3	14	14	175	35	2 146	2 133	-0,61	2 142	-0,19
pdc	16	40	145	29	715	731	2,24	762	6,57
spla	16	46	260	52	1 055	1 122	6,35	1 154	9,38
table5	17	15	158	32	1 959	2 076	5,97	1 992	1,68
duke2	22	29	86	17	678	631	-6,93	655	-3,39
apex1	45	45	206	41	2 607	2 608	0,04	2 602	-0,19

Таблица 5

Схемы LGSynth'91

Схема	i	o	p	k	Кол-во эл-тов (осн. схема)	Кол-во эл-тов (эквив. схема 1)	S_1 %	Кол-во эл-тов (эквив. схема 2)	S_2 %
dk27	4	5	11	2	30	33	10	33	10
dk16	7	8	85	17	369	383	3,79	375	1,63
tbk	11	8	244	49	1 378	1 331	-3,41	1 338	-2,9
s386	11	11	38	8	207	207	0	204	-1,45
sse	11	11	38	8	175	163	-6,86	175	0
cse	11	11	60	12	340	338	-0,59	321	-5,59
s298	11	14	974	195	7 330	7 014	-4,31	6 924	-5,54
keyb	12	7	90	18	337	336	-0,3	341	1,19
tma	12	11	39	8	213	202	-5,16	204	-4,23
s1	13	11	94	19	634	602	-5,05	603	-4,89
pma	13	13	56	11	336	352	4,76	346	2,98
planet	13	25	102	20	979	949	-3,06	943	-3,68
planet1	13	25	105	21	969	988	1,96	988	1,96
styr	14	15	124	25	936	957	2,24	967	3,31
ex1	14	24	70	14	407	423	3,93	431	5,9

Окончание табл. 5

Схема	i	o	p	k	Кол-во эл-тов (осн. схема)	Кол-во эл-тов (эквив. схема 1)	S_1 %	Кол-во эл-тов (эквив. схема 2)	S_2 %
s1488	14	25	132	26	782	763	-2,43	762	-2,56
s1494	14	25	132	26	744	725	-2,55	735	-1,21
kirkman	16	10	188	38	388	415	6,96	436	12,37
sand	16	14	115	23	926	930	0,43	928	0,22
s510	25	13	68	14	300	299	-0,33	286	-4,67
scf	34	63	165	33	1 558	1 523	-2,25	1 526	-2,05

Из полученных данных следует, что в рассмотренных примерах обеспечивается либо приемлемое усложнение эквивалентных схем по сравнению со схемой, построенной по заданной ВСДНФ, при искажении эквивалентных систем за счет появления дополнительных конъюнкций меньших рангов, либо даже незначительное упрощение эквивалентных схем. В случае упрощения схемы доля сокращенных в схеме элементов в процентах по сравнению с числом элементов рабочей схемы представлена со знаком минус.

Заметим, что чем больше количество расщепляемых конъюнкций, тем сильнее эквивалентные схемы отличаются по своей структуре от исходной (рабочей) схемы и, следовательно, тем сложнее вносить в эквивалентные схемы вредоносные подсхемы, обеспечивающие одинаковые с рабочей схемой искажения поведения. Как показали проведенные нами эксперименты, увеличение числа расщепляемых конъюнкций до 30% и более (для рассматриваемых схем) приводит к существенному возрастанию числа элементов в эквивалентных схемах, что нежелательно.

Отметим, что для некоторых ВСДНФ количество конъюнкций, для которых можно выполнить расщепления, мало. Следовательно, нельзя получить эквивалентные схемы с помощью предложенного алгоритма. В этой ситуации в рамках технологии троирования можно использовать аппроксимирующие схемы. Аппроксимирующие схемы реализуют систему булевых функций, которая не является точной копией системы, реализуемой оригинальной (рабочей) схемой, но достаточно близка к ней [8].

Заключение

В данной работе рассмотрена проблема обеспечения надежности функционирования логических схем в условиях возможности внедрения в них вредоносных подсхем. Предложен алгоритм построения эквивалентных систем булевых функций путем расщепления конъюнкций, для того чтобы получить по ним эквивалентные схемы с целью использования их в системах троирования. Проведены эксперименты на нескольких ВСДНФ, которые показали, что в результате расщепления получаются конъюнкции низких рангов по сравнению с конъюнкциями исходной ВСДНФ. Это означает, что некоторые конъюнкции эквивалентных систем существенно отличаются от конъюнкций заданной ВСДНФ. Следствием этого факта является отличие рабочей схемы от эквивалентных схем, затрудняющее введение вредоносных подсхем в эквивалентные схемы. При подходящем выборе расщепляемых конъюнкций обеспечивается приемлемое усложнение эквивалентных схем по сравнению с рабочей схемой. Возникающее при этом изменение структуры эквивалентных схем затрудняет введение в систему троирования вредоносных подсхем.

Список источников

1. Chaudhury M.R., Mohandram K. Approximate Logic Circuits for Low Overhead Non-Intrusive Concurrent Error Detection // DATE '08: Proc. of the conference on Design, automation and test in Europe. New-York : Association for Computing Machinery, 2008. P. 902–908.
2. Sanchez-Clemente A., Entrena L., Garcia-Valderas M., Lopez-Ongil C. Logic Masking for SET Mitigation Using Approximate Logic Circuits // 2012 IEEE 18th International On-Line Testing Symposium (IOLTS). New-York : IEEE, 2012. P. 176–181. doi: 10.1109/IOLTS.2012.6313868
3. Sanchez-Clemente A. Transient Error Mitigation by Means of Approximate Logic Circuits : Tesis Doctoral. Universidad Carlos III de Madrid, 2017.

4. ABC: A System for Sequential Synthesis and Verification. URL: <https://people.eecs.berkeley.edu/~alanmi/abc/> (accessed: 22.01.2023).
5. Yang S. Logic Synthesis and Optimization Benchmarks User Guide: Version 3.0 : Technical Report 1991-IWLS-UG-Saeyang, MCNC, 1991.
6. McElvain K. IWLS'93 Benchmark Set: Version 4.0 : Distributed as a part of IWLS'93 benchmark set, 1993.
7. ESPRESSO: Logic Minimization Software. URL: <https://ptolemy.berkeley.edu/projects/embedded/pubs/downloads/espresso/index.htm> (accessed: 20.11.2023).
8. Останин С.А., Матросова А.Ю., Андреева В.В. Построение аппроксимирующих схем для синхронных автоматов в рамках технологии троирования // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2023. № 62. С. 124–131. doi: 10.17223/19988605/62/14

References

1. Chaudhury, M.R. & Mohandram, K. (2008) Approximate Logic Circuits for Low Overhead Non-Intrusive Concurrent Error Detection. *DATE '08. Proc. of the conference on design, automation and test in Europe*. New-York: Association for Computing Machinery. pp. 902–908.
2. Sanchez Clemente, A., Entrena, L., Garsea Valderas, M. & Lopez-Ongil, C. (2012) Logic Masking for SET Mitigation Using Approximate Logic Circuits. *2012 IEEE 18th International On-Line Testing Symposium (IOLTS)*. New-York: IEEE. pp. 176–181. DOI: 10.1109/IOLTS.2012.6313868
3. Sanchez Clemente, A.J. (2017) *Transient Error Mitigation by Means of Approximate Logic Circuits*. Doctoral Thesis. Universidad Carlos III de Madrid.
4. Berkley. (n.d.) ABC: A System for Sequential Synthesis and Verification. [Online] Available from: <https://people.eecs.berkeley.edu/~alanmi/abc/> (Accessed: 22nd January 2023).
5. Yang, S. (1991) *Logic Synthesis and Optimization Benchmarks User Guide Version 3.0*. Technical Report 1991-IWLS-UG-Saeyang, MCNC.
6. McElvain, K. (1993) *IWLS'93 Benchmark Set: Version 4.0, Distributed as a part of IWLS'93 benchmark set*.
7. Berkley. (n.d.) ESPRESSO: Logic Minimization Software. [Online] Available from: <https://ptolemy.berkeley.edu/projects/embedded/pubs/downloads/espresso/index.htm> (Accessed: 20th November 2023).
8. Ostanin, S.A., Matrosova, A.Yu. & Andreeva, V.V. (2023) Deriving approximate circuits for TMR technique applied to synchronous sequential circuits. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 62. pp. 124–131. DOI: 10.17223/19988605/62/14

Информация об авторах:

Матросова Анжела Юрьевна – доктор технических наук, профессор кафедры компьютерной безопасности Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: mau11@yandex.ru

Пасько Антонина Дмитриевна – студент кафедры компьютерной безопасности Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: pasko.tonya@yandex.ru

До Там Дык – студент кафедры компьютерной безопасности Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: kmaddt@gmail.com

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

Information about the authors:

Matrosova Anzhela Yu. (Doctor of Technical Sciences, Professor, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: mau11@yandex.ru

Pasko Antonina D. (Student, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: pasko.tonya@yandex.ru

Do Tam Duc (Student, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: kmaddt@gmail.com

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

Received 08.02.2024; accepted for publication 03.06.2024

Поступила в редакцию 08.02.2024; принята к публикации 03.06.2024

Научный журнал

**ВЕСТНИК
ТОМСКОГО
ГОСУДАРСТВЕННОГО
УНИВЕРСИТЕТА**

**УПРАВЛЕНИЕ,
ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА
И ИНФОРМАТИКА**

**TOMSK STATE UNIVERSITY
JOURNAL OF CONTROL AND COMPUTER SCIENCE**

2024. № 67

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 17.06.2024 г. Формат 60х84^{1/8}.
Гарнитура Times. Усл. печ. л. 14,4.
Тираж 250 экз. Заказ № 5944. Цена свободная.

Дата выхода в свет 20.06.2024 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