

**ВЕСТНИК
ТОМСКОГО
ГОСУДАРСТВЕННОГО
УНИВЕРСИТЕТА**

**УПРАВЛЕНИЕ,
ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА
И ИНФОРМАТИКА**

**TOMSK STATE UNIVERSITY
JOURNAL OF CONTROL AND COMPUTER SCIENCE**

Научный журнал

2023

№ 62

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

Горцев Александр Михайлович – гл. редактор, проф., д-р техн. наук, НИ ТГУ. Тел. +73822529485
Смагин Валерий Иванович – зам. гл. редактора, проф., д-р техн. наук, НИ ТГУ. Тел. +73822529485
Нежельская Людмила Алексеевна – ответственный секретарь, 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
Вишневецкий Владимир Миронович – проф., д-р техн. наук, Институт проблем управления РАН (г. Москва)
Замятин Александр Владимирович – проф., д-р техн. наук, НИ ТГУ
Кошкин Геннадий Михайлович – проф., д-р физ.-мат. наук, НИ ТГУ
Костюк Юрий Леонидович – проф., д-р техн. наук, НИ ТГУ
Матророва 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
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

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

CONTENTS

УПРАВЛЕНИЕ ДИНАМИЧЕСКИМИ СИСТЕМАМИ

CONTROL OF DYNAMICAL SYSTEMS

Рзаева В.Г.
Необходимые условия оптимальности первого и второго порядков в одной задаче оптимального управления, описываемой системой гиперболических интегро-дифференциальных уравнений типа Вольтерра 4

Rzayeva V.G.
Necessary conditions for optimality of first and second orders in one optimal control problem described by system of hyperbolic integro-differential equations of Volterra type 4

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

MATHEMATICAL MODELING

Бураченко А.В., Пазников А.А., Державин Д.П.
Распределенная очередь без использования блокировок в модели удаленного доступа к памяти 13
Дмитриев В.М., Ганджа Т.В., Кочергин М.И.
Матрично-топологический анализ компонентных цепей 25

Burachenko A.V., Paznikov A.A., Derzhavin D.P.
Lock-free distributed queue in remote memory access model 13
Dmitriev V.M., Gandzha T.V., Kochergin, M.I.
Matrix topological analysis of component circuits 25

ОБРАБОТКА ИНФОРМАЦИИ

DATA PROCESSING

Горцев А.М., Веткина А.В.
Применение метода максимального правдоподобия для оценки параметра равномерного распределения длительности непродлевающегося мертвого времени в рекуррентном альтернирующем полусинхронном потоке событий 36

Gortsev A.M., Vetkina A.V.
An application of the maximum likelihood estimation for the parameter of uniform distribution of the duration of unextendable dead time in recurrent alternating semi-synchronous flow 36

Краковский Ю.М., Куклина О.К.
Бинарное прогнозирование динамических показателей на основе методов машинного обучения 50

Krakovsky Y.M., Kuklina O.K.
The binary forecasting of dynamic indicators based on machine learning methods 50

Лиситин Д.В., Гаврилов К.В.
Максиминная задача оценивания параметров в условиях байесовского точечного засорения 56

Lisitsin D.V., Gavrilov K.V.
Maximin problem of parameter estimation in conditions of point Bayesian contamination 56

Нежелская Л.А., Степаненко И.Д.
Применение метода моментов для оценки длительности мертвого времени в рекуррентном обобщенном полусинхронном потоке событий с продлевающимся мертвым временем 65

Nezhel'skaya L.A., Stepanenko I.D.
An application of the method of moments to estimation of the dead time duration in recurrent generalized semi-synchronous flow of events with prolonged dead time 65

Рудько И.М.
Алгоритм обнаружения звукорядов в спектре шумоизлучения морских судов 76

Rudko I.M.
Algorithm for detecting harmonic series in the noise emission spectrum of sea vessels 76

Smagin V.I., Kim K.S.
Robust extrapolation for systems with unknown input and interval uncertainty in system and observations 85

Smagin V.I., Kim K.S.
Robust extrapolation for systems with unknown input and interval uncertainty in system and observations 85

Удод В.А.
Вычисление коэффициента корреляции случайных величин с использованием дополнительной информации о форме их взаимосвязи 92

Udod V.A.
Calculation of the correlation coefficient of random variables using additional information about the form of their relationship 92

Цицашвили Г.Ш., Бочарников В.Н., Краснопеев С.М.
Метод устранения мешающего параметра в статистике пуассоновского потока точек 101

Tsitsiashvili G.Sh., Bocharnikov V.N., Krasnopeev S.M.
The method of eliminating the interfering parameter in the statistics of the Poisson flow of points 101

ИНФОРМАТИКА И ПРОГРАММИРОВАНИЕ

INFORMATICS AND PROGRAMMING

Громов М.Л., Прокопенко С.А., Сотников А.П., Шабалдина Н.В.
Оптимизация клиентской компоненты по критерию отсутствия лишних диалогов с сервером 107

Gromov M.L., Prokopenko S.A., Sotnikov A.P., Shabaldina, N.V.
Optimization of the client component by the criterion of the absence of redundant dialogues with the server 107

Каравай М.Ф., Подлазов В.С., Соколов В.В.
Простая масштабируемая коммутируемая управляющая сеть 115

Karavay M.F., Podlazov V.S., Sokolov V.V.
Simple scalable switched control network 115

ДИСКРЕТНЫЕ ФУНКЦИИ И АВТОМАТЫ

DISCRETE FUNCTION AND AUTOMATONS

Останин С.А., Матросова А.Ю., Андреева В.В.
Построение аппроксимирующих схем для синхронных автоматов в рамках технологии троирования 124

Ostanin S.A., Matrosova A.Yu., Andreeva V.V.
Deriving approximate circuits for TMR technique applied to synchronous sequential circuits 124

УПРАВЛЕНИЕ ДИНАМИЧЕСКИМИ СИСТЕМАМИ

CONTROL OF DYNAMICAL SYSTEMS

Научная статья

УДК 517.977.56

doi: 10.17223/19988605/62/1

**Необходимые условия оптимальности первого и второго порядков
в одной задаче оптимального управления, описываемой системой
гиперболических интегро-дифференциальных уравнений типа Вольтерра****Вафа Гюлага кызы Рзаева***Сумгаитский государственный университет, Сумгаит, Азербайджан, vafa.asgerova77@mail.ru*

Аннотация. Рассматривается задача оптимального управления, описываемая системой гиперболических интегро-дифференциальных уравнений типа Вольтерра с терминальным критерием качества при предположении открытости области управления. Вычислены первая и вторая вариации функционала качества. Получены аналог уравнения Эйлера и общее необходимое условие оптимальности второго порядка. Используя необходимое условие оптимальности второго порядка, с помощью специальных вариаций управления доказан аналог условия Лежандра–Клебша и получено необходимое условие оптимальности особых в классическом смысле управлений.

Ключевые слова: гиперболическое интегро-дифференциальное уравнение; краевая задача; уравнение Эйлера; классическая экстремаль; необходимые условия оптимальности; особое управление.

Для цитирования: Рзаева В.Г. Необходимые условия оптимальности первого и второго порядков в одной задаче оптимального управления, описываемой системой гиперболических интегро-дифференциальных уравнений типа Вольтерра // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2023. № 62. С. 4–12. doi: 10.17223/19988605/62/1

Original article

doi: 10.17223/19988605/62/1

**Necessary conditions for optimality of first and second orders
in one optimal control problem described by system
of hyperbolic integro-differential equations of Volterra type****Vafa G. Rzayeva***Sumgait State University, Sumgait, Azerbaijan, vafa.asgerova@mail.ru*

Abstract. In this paper we consider minimization of the terminal functional

$$S(u) = \varphi(z(t_1, x_1))$$

under restrictions

$$u(t, x) \in U \subset R^r, \quad (t, x) \in D,$$

$$z_{tx}(t, x) = f(t, x, z(t, x), u(t, x)) + \int_{t_0}^t \int_{x_0}^x g(t, x, \tau, s, z(\tau, s), u(\tau, s)) ds d\tau,$$

$$z(t_0, x) = a(x), \quad x \in [x_0, x_1],$$

$$z(t, x_0) = b(t), \quad t \in [t_0, t_1],$$

$$a(x_0) = b(t_0).$$

Here $u(t, x)$ – is measurable and bounded r -dimensional vector-function of control actions, and U is given non-empty open and bounded sets, $\varphi(z)$ – is given twice continuously differentiable scalar function, $a(x)$ and $b(t)$ are given and absolutely continuous n -dimensional vector's of the functions, $f(t, x, z, u)$ and $g(t, x, \tau, s, z, u)$ are given n -dimensional vectors of the function having continuous partial derivatives up to the second order.

Each control function delivering the minimum value to functional under constraints is called optimal control. Using the analogue of the increment method, the first and second order variations of the quality functional are calculated. Using them, various necessary conditions of optimality of the first and second orders are obtained. Separately investigated the case of singular control in the classical sense.

Keywords: hyperbolic integro-differential equations; boundary condition; Euler equation; classical extreme; necessary optimality condition; singular control.

For citation: Rzaeva, V.G. (2023) Necessary conditions for optimality of first and second orders in one optimal control problem described by system of hyperbolic integro-differential equations of Volterra type. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 62. pp. 4–12. doi: 10.17223/19988605/62/1

Введение

Многие процессы описываются различного типа уравнениями в частных производных (см., напр.: [1–4]). Поэтому задачи оптимального управления, описываемые уравнениями в частных производных, интенсивно исследуются (см.: [3–9]). Такие задачи оптимального управления называются задачами оптимального управления с распределенными параметрами.

Задачи оптимального управления, описываемые системами гиперболических уравнений второго порядка с краевыми условиями Гурса, начали изучаться А.И. Егоровым [4, 5]. Был установлен ряд необходимых условий оптимальности в форме принципа максимума Л.С. Понтрягина [4–7] и изучены особые управления [8, 9].

Предлагаемая работа посвящена исследованию одной задачи оптимального управления, описываемой системой гиперболических интегро-дифференциальных уравнений типа Вольтерра с краевыми условиями Гурса.

При предположении открытости области управления установлены необходимые условия оптимальности первого и второго порядков, носящие конструктивный характер.

1. Постановка задачи

Пусть $U \subset R^r$ – заданные непустые и ограниченные множества, $D = [t_0, t_1] \times [x_0, x_1]$ – заданный прямоугольник.

Допустим, что управляемый процесс описывается системой гиперболических интегро-дифференциальных уравнений типа Вольтерра

$$z_{tx}(t, x) = f(t, x, z(t, x), u(t, x)) + \int_{t_0}^t \int_{x_0}^x g(t, x, \tau, s, z(\tau, s), u(\tau, s)) ds d\tau \quad (1)$$

с краевыми условиями

$$z(t_0, x) = a(x), \quad x \in [x_0, x_1],$$

$$z(t, x_0) = b(t), \quad t \in [t_0, t_1],$$

$$a(x_0) = b(t_0). \quad (2)$$

Здесь $f(t, x, z, u)$ и $g(t, x, \tau, s, z, u)$ – заданные n -мерные вектор-функции, непрерывные по совокупности переменных вместе с частными производными по (z, u) до второго порядка включительно, $a(x)$ и $b(t)$ – заданные абсолютно непрерывные n -мерные вектор-функции, $u(t, x)$ r -мерная измеримая и ограниченная вектор-функция управляющих воздействий, удовлетворяющая ограничениям

$$u(t, x) \in U \subset R^r, \quad (t, x) \in D. \quad (3)$$

Такую управляющую функцию назовем допустимым управлением.

Будем предполагать, что каждому допустимому управлению $u(t, x)$ соответствует единственное абсолютно непрерывное решение $z(t, x)$ краевой задачи (1), (2).

Пару $(u(t, x), z(t, x))$ назовем допустимым процессом.

Пусть $\varphi(z)$ – заданная дважды непрерывно дифференцируема скалярная функция.

Рассмотрим задачу нахождения минимального значения функционала

$$S(u) = \varphi(z(t_1, x_1)) \quad (4)$$

при ограничениях (1)–(3).

Допустимое управление $u(t, x)$, доставляющее минимальное значение функционалу (4) при ограничениях (1)–(3), назовем оптимальным управлением.

Целью работы является вывод необходимых условий оптимальности в рассматриваемой задаче.

2. Вычисление первой и второй вариаций (в классическом смысле) функционала (4)

Пусть $u(t, x)$ и $\bar{u}(t, x) = u(t, x) + \Delta u(t, x)$ – некоторые допустимые управления. Через $z(t, x)$ и $\bar{z}(t, x) = z(t, x) + \Delta z(t, x)$ обозначим соответствующие им решения краевой задачи (1), (2) и запишем приращение функционала качества

$$\Delta S(u) = S(\bar{u}) - S(u) = \varphi(\bar{z}(t, x)) - \varphi(z(t, x)). \quad (5)$$

Ясно, что приращение $\Delta z(t, x)$ состояния $z(t, x)$ является решением уравнения

$$\begin{aligned} \Delta z_{ix}(t, x) = & f(t, x, \bar{z}(t, x), \bar{u}(t, x)) - f(t, x, z(t, x), u(t, x)) + \\ & + \int_{t_0}^t \int_{x_0}^x [g(t, x, \tau, s, \bar{z}(\tau, s), \bar{u}(\tau, s)) - g(t, x, \tau, s, z(\tau, s), u(\tau, s))] ds d\tau, \end{aligned} \quad (6)$$

с краевыми условиями

$$\begin{aligned} \Delta z(t_0, x) &= 0, \quad x \in [x_0, x_1], \\ \Delta z(t, x_0) &= 0, \quad t \in [t_0, t_1]. \end{aligned} \quad (7)$$

Пусть $\psi(t, x)$ некоторая n -мерная вектор-функция.

Предполагаем, что вектор-функция $\psi(t, x)$ удовлетворяет тем условиям гладкости, которые будут использованы при дальнейших рассуждениях.

Из (6) получим справедливость тождеств

$$\begin{aligned} \int_{t_0}^{t_1} \int_{x_0}^{x_1} \psi'(t, x) \Delta z_{ix}(t, x) dx dt = & \int_{t_0}^{t_1} \int_{x_0}^{x_1} \psi'(t, x) [f(t, x, \bar{z}(t, x), \bar{u}(t, x)) - f(t, x, z(t, x), u(t, x))] dx dt + \\ & + \int_{t_0}^{t_1} \int_{x_0}^{x_1} \psi'(t, x) \left[\int_{t_0}^t \int_{x_0}^x [g(t, x, \tau, s, \bar{z}(\tau, s), \bar{u}(\tau, s)) - g(t, x, \tau, s, z(\tau, s), u(\tau, s))] ds d\tau \right] dx dt. \end{aligned} \quad (8)$$

Здесь и в дальнейшем штрих (') – операция скалярного произведения для векторов, а для матриц – операция транспонирования.

Используя теорему Фубини (двумерный аналог формулы Дирихле (см., напр.: [10])) получаем, что

$$\begin{aligned} & \int_{t_0}^{t_1} \int_{x_0}^{x_1} \Psi'(t, x) \left[\int_{t_0}^t \int_{x_0}^x \left[g(t, x, \tau, s, \bar{z}(\tau, s), \bar{u}(\tau, s)) - g(t, x, \tau, s, z(\tau, s), u(\tau, s)) \right] ds d\tau \right] dx dt = \\ & = \int_{t_0}^{t_1} \int_{x_0}^{x_1} \left[\int_{t_0}^t \int_{x_0}^x \Psi'(\tau, s) \left[g(t, x, \tau, s, \bar{z}(\tau, s), \bar{u}(\tau, s)) - g(t, x, \tau, s, z(\tau, s), u(\tau, s)) \right] ds d\tau \right] dx dt. \end{aligned} \quad (9)$$

Введем обозначения

$$\begin{aligned} \Delta_{\bar{u}} f[t, x] &\equiv f(t, x, z(t, x), \bar{u}(t, x)) - f(t, x, z(t, x), u(t, x)), \\ \Delta_{\bar{u}} f_z[t, x] &\equiv f_z(t, x, z(t, x), \bar{u}(t, x)) - f_z(t, x, z(t, x), u(t, x)), \\ f_z[t, x] &\equiv f_z(t, x, z(t, x), u(t, x)), \end{aligned}$$

$$\Delta_{\bar{u}} g[t, x, \tau, s] \equiv g(t, x, \tau, s, z(\tau, s), \bar{u}(\tau, s)) - g(t, x, \tau, s, z(\tau, s), u(\tau, s)),$$

$$\Delta_{\bar{u}} g_z[t, x, \tau, s] \equiv g_z(t, x, \tau, s, z(\tau, s), \bar{u}(\tau, s)) - g_z(t, x, \tau, s, z(\tau, s), u(\tau, s)),$$

$$H(t, x, z, u, \psi) = -\Psi' f(t, x, z, u) + \int_{t_0}^t \int_{x_0}^x \Psi'(\tau, s) g(t, x, \tau, s, z(\tau, s), u(\tau, s)) ds d\tau,$$

$$\Delta_{\bar{u}} H[t, x] \equiv H(t, x, z(t, x), \bar{u}(t, x), \psi(t, x)) - H(t, x, z(t, x), u(t, x), \psi(t, x)),$$

$$H_z[t, x] \equiv H_z(t, x, z(t, x), u(t, x), \psi(t, x))$$

и, учитывая тождества (8), (9), из формулы приращения (5) получаем

$$\begin{aligned} \Delta S(u) &= \Phi(\bar{z}(t, x)) - \Phi(z(t, x)) + \int_{t_0}^{t_1} \int_{x_0}^{x_1} \Psi'(t, x) \Delta z_{tx}(t, x) dx dt - \\ &- \int_{t_0}^{t_1} \int_{x_0}^{x_1} \left[H(t, x, \bar{z}(t, x), \bar{u}(t, x), \psi(t, x)) - H(t, x, z(t, x), u(t, x), \psi(t, x)) \right] dx dt. \end{aligned} \quad (10)$$

В силу краевых условий (7) можно доказать, что

$$\begin{aligned} \int_{t_0}^{t_1} \int_{x_0}^{x_1} \Psi'(t, x) \Delta z_{tx}(t, x) dx dt &= \Psi'(t_1, x_1) \Delta z(t_1, x_1) - \int_{x_0}^{x_1} \Psi'_x(t_1, x) \Delta z(t_1, x) dx - \int_{t_0}^{t_1} \Psi'_t(t, x_1) \Delta z(t, x_1) dt + \\ &+ \int_{t_0}^{t_1} \int_{x_0}^{x_1} \Psi'_{tx}(t, x) \Delta z(t, x) dx dt. \end{aligned} \quad (11)$$

Учитывая тождество (11) и применяя формулу Тейлора, из (10) получаем, что

$$\begin{aligned} \Delta S(u) &= \Phi'_z(z(t_1, x_1)) \Delta z(t_1, x_1) + \frac{1}{2} \Delta z'(t_1, x_1) \Phi_{zz}(z(t_1, x_1)) \Delta z(t_1, x_1) + o_1(\|\Delta z(t_1, x_1)\|^2) + \\ &+ \Psi'(t_1, x_1) \Delta z(t_1, x_1) - \int_{x_0}^{x_1} \Psi'(t_1, x) \Delta z(t_1, x) dx - \int_{t_0}^{t_1} \Psi'_x(t, x_1) \Delta z(t, x_1) dt + \\ &+ \int_{t_0}^{t_1} \int_{x_0}^{x_1} \Psi'_{tz}(t, x) \Delta z(t, x) dx dt - \int_{t_0}^{t_1} \int_{x_0}^{x_1} \left[H'_z[t, x] \Delta z(t, x) + H'_u[t, x] \Delta u(t, x) \right] dx dt - \\ &- \frac{1}{2} \int_{t_0}^{t_1} \int_{x_0}^{x_1} \left[\Delta z'(t, x) H_{zz}[t, x] \Delta z(t, x) + 2 \Delta u'(t, x) H_{uz}[t, x] \Delta z(t, x) + \Delta u'(t, x) H_{uu}[t, x] \Delta u(t, x) + \right. \\ &\left. + o_2(\|\Delta z(t, x) + \Delta u(t, x)\|^2) \right] dx dt. \end{aligned} \quad (12)$$

Если предполагать, что вектор-функция $\psi(t, x)$ является решением краевой задачи

$$\psi_{tx}(t, x) = H_z[t, x], \quad (13)$$

$$\psi_t(t, x_1) = 0, \quad \psi_x(t_1, x) = 0, \quad (14)$$

$$\psi(t_1, x_1) = -\varphi_z(z(t_1, x_1)),$$

то формула приращения (12) примет вид:

$$\begin{aligned} \Delta S(u) = & - \int_{t_0}^t \int_{x_0}^x H_u[t, x] \Delta u(t, x) dx dt + \frac{1}{2} \left[\Delta z'(t_1, x_1) \varphi_{zz}(z(t_1, x_1)) \Delta z(t_1, x_1) - \right. \\ & \left. - \int_{t_0}^t \int_{x_0}^x \left[\Delta z'(t, x) H_{zz}[t, x] \Delta z(t, x) + 2 \Delta u'(t, x) H_{uz}[t, x] \Delta z(t, x) + \Delta u'(t, x) H_{uu}[t, x] \Delta u(t, x) \right] dx dt \right] + \\ & + o_1 \left(\left\| \Delta z(t_1, x_1) \right\|^2 \right) - \int_{t_0}^t \int_{x_0}^x o_2 \left(\left\| \Delta z(t, x) + \Delta u(t, x) \right\|^2 \right) dx dt. \end{aligned} \quad (15)$$

Пусть $R(t, x, \tau, s) (n \times n)$ – матричная функция, являющаяся решением линеаризованной краевой задачи

$$\begin{aligned} R_{ts}(t, x, \tau, s) &= R(t, x, \tau, s) f_z[\tau, s] + \int_{\tau}^t \int_s^x R(t, x, \alpha, \beta) g_z[\alpha, \beta, \tau, s] d\alpha d\beta, \\ R_t(t, x, \tau, s) &= 0, \quad R_s(t, x, \tau, s) = 0, \\ R(t, x, t, x) &= E, \end{aligned}$$

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

Из краевой задачи (6), (7) следует, что приращение $\Delta z(t, x)$ состояния $z(t, x)$ является решением уравнения

$$\begin{aligned} \Delta z_{tx}(t, x) &= f_z[t, x] \Delta z(t, x) + f_u[t, x] \Delta u(t, x) + \int_{t_0}^t \int_{x_0}^x \left[g_z[t, x, \tau, s] \Delta z(\tau, s) + g_u[t, x, \tau, s] \Delta u(\tau, s) \right] ds d\tau + \\ & + o_3 \left(\left\| \Delta z(t, x) + \Delta u(t, x) \right\| \right) + \int_{t_0}^t \int_{x_0}^x o_4 \left(\left\| \Delta z(\tau, s) + \Delta u(\tau, s) \right\| \right) ds d\tau, \end{aligned} \quad (16)$$

$$\begin{aligned} \Delta z(t_0, x) &= 0, \\ \Delta z(t, x_0) &= 0. \end{aligned} \quad (17)$$

Если интерпретировать уравнение (16) как линейное неоднородное интегро-дифференциальное уравнение относительно $\Delta z(t, x)$, то решение краевой задачи (16), (17) может быть представлено в виде:

$$\begin{aligned} \Delta z(t, x) &= \int_{t_0}^t \int_{x_0}^x R(t, x, \tau, s) f_u[\tau, s] \Delta u(\tau, s) ds d\tau + \\ & + \int_{t_0}^t \int_{x_0}^x \left[\int_{\tau}^t \int_s^x R(t, x, \alpha, \beta) g_u[\alpha, \beta, \tau, s] d\alpha d\beta \right] \Delta u(\tau, s) ds d\tau + \\ & + \int_{t_0}^t \int_{x_0}^x R(t, x, \tau, s) o_3 \left(\left\| \Delta z(\tau, s) + \Delta u(\tau, s) \right\| \right) ds d\tau + \\ & + \int_{t_0}^t \int_{x_0}^x \left[\int_{\tau}^t \int_s^x R(t, x, \alpha, \beta) o_4 \left(\left\| \Delta z(\tau, s) + \Delta u(\tau, s) \right\| \right) d\alpha d\beta \right] ds d\tau. \end{aligned} \quad (18)$$

Пусть ε – малое по абсолютной величине число, а $\delta u(t, x) \in U \subset R^r, (t, x) \in D$ – произвольная r -мерная, измеримая и ограниченная вектор-функция.

Через $\Delta u(t, x; \varepsilon)$ обозначим специальное приращение управления $u(t, x)$ по формуле

$$\Delta u(t, x; \varepsilon) = \varepsilon \delta u(t, x), (t, x) \in D. \quad (19)$$

Пусть $\Delta z(t, x; \varepsilon)$ – специальное приращение состояния $z(t, x)$, отвечающее специальному приращению (19) управления $u(t, x)$.

С помощью представления (18) доказывается справедливость разложения

$$\Delta z(t, x; \varepsilon) = \varepsilon \delta z(t, x) + o(\varepsilon; t, x), \quad (20)$$

где $\delta z(t, x)$ определяется из соотношения

$$\delta z(t, x) = \int_{t_0}^t \int_{x_0}^x R(t, x, \tau, s) f_u[\tau, s] \delta u(\tau, s) ds d\tau + \int_{t_0}^t \int_{x_0}^x \left[\int_{\tau}^t \int_s^x R(t, x, \alpha, \beta) g_u[\alpha, \beta, \tau, s] d\alpha d\beta \right] \delta u(\tau, s) ds d\tau.$$

При введении обозначения

$$Q(t, x, \tau, s) = R(t, x, \tau, s) f_u[\tau, s] + \int_{\tau}^t \int_s^x R(t, x, \alpha, \beta) g_u[\alpha, \beta, \tau, s] d\alpha d\beta$$

последнее представление записывается в виде:

$$\delta z(t, x) = \int_{t_0}^t \int_{x_0}^x Q(t, x, \tau, s) \delta u(\tau, s) ds d\tau. \quad (21)$$

Учитывая формулы (18) и (20), из формулы приращения (15) получаем справедливость разложения

$$S(u + \varepsilon \delta u) - S(u) = -\varepsilon \int_{t_0}^{t_1} \int_{x_0}^{x_1} H'_u[t, x] \delta u(t, x) dx dt + \frac{\varepsilon^2}{2} \int_{t_0}^{t_1} \int_{x_0}^{x_1} \left[\delta z'(t, x) H_{zz}[t, x] \delta z(t, x) + 2\delta u'(t, x) H_{uz}[t, x] \delta z(t, x) + \delta u'(t, x) H_{uu}[t, x] \delta u(t, x) \right] dx dt + o(\varepsilon^2). \quad (22)$$

Из разложения (22) следует, что первая и вторая вариации функционала (4) имеют соответственно вид:

$$\delta^1 S(u; \delta u) = \int_{t_0}^{t_1} \int_{x_0}^{x_1} H'_u[t, x] \delta u(t, x) dx dt, \quad (23)$$

$$\delta^2 S(u; \delta u) = -l'(t_1, x_1) \Phi_{zz}(z(t_1, x_1)) l(t_1, x_1) - \int_{t_0}^{t_1} \int_{x_0}^{x_1} \left[\delta z'(t, x) H_{zz}[t, x] \delta z(t, x) + 2\delta u'(t, x) H_{uz}[t, x] \delta z(t, x) + \delta u'(t, x) H_{uu}[t, x] \delta u(t, x) \right] dx dt. \quad (24)$$

Если $u(t, x)$ – оптимальное управление, то вдоль него первая вариация (23) функционала (4) равна нулю, а вторая вариация (24) неотрицательна.

Следовательно, при сделанных предположениях для оптимальности допустимого процесса $(u(t, x), z(t, x))$ необходимо, чтобы для всех $\delta u(t, x) \in U \subset R^r, (t, x) \in D$ выполнялись соотношения

$$\int_{t_0}^{t_1} \int_{x_0}^{x_1} H'_u[t, x] \delta u(t, x) dx dt = 0, \quad (25)$$

$$\int_{t_0}^{t_1} \int_{x_0}^{x_1} \left[\delta z'(t, x) H_{zz}[t, x] \delta z(t, x) + 2\delta u'(t, x) H_{uz}[t, x] \delta z(t, x) + \delta u'(t, x) H_{uu}[t, x] \delta u(t, x) \right] dx dt \geq 0. \quad (26)$$

Из тождества (25) в силу произвольности $\delta u(t, x)$ следует

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

$$H_u[\theta, \xi] = 0 \quad (27)$$

выполнялось для всех $(\theta, \xi) \in [t_0, t_1] \times [x_0, x_1]$.

Здесь $(\theta, \xi) \in [t_0, t_1] \times [x_0, x_1]$ – произвольная правильная точка (см., напр.: [11]) управления $u(t, x)$.

Условие (27), следуя классическому вариационному исчислению, назовем уравнением Эйлера для рассматриваемой задачи.

Уравнение Эйлера является необходимым условием оптимальности первого порядка.

Каждое допустимое управление, являющееся решением уравнения Эйлера, назовем классической экстремалью. Вообще говоря, число классических экстремалей может быть достаточно большим. Поэтому надо иметь новые необходимые условия оптимальности, «сужающие» множество классических экстремалей.

Таким необходимым условием оптимальности является необходимое условие оптимальности (26). Но необходимое условие оптимальности (26) является неявным условием оптимальности.

Используя неравенство (26), получим необходимые условия оптимальности, непосредственно выраженные через параметры рассматриваемой задачи.

Займемся преобразованием некоторых слагаемых в неравенстве (26).

Используя представление (22), убедимся в справедливости тождеств

$$\begin{aligned} & \delta z'(t_1, x_1) \phi_{zz}(z(t_1, x_1)) \delta z(t_1, x_1) = \\ & = \iint_D \delta u'(\tau, s) f_u[\tau, s] R'(t_1, x_1, \tau, s) \phi_{zz}(z(t_1, x_1)) R(t_1, x_1, \alpha, \beta) f_u[\alpha, \beta] \delta u(\alpha, \beta) ds d\tau d\alpha d\beta, \end{aligned} \quad (28)$$

$$\iint_{t_0, x_0}^{t_1, x_1} \delta u'(t, x) H_{uz}[t, x] \delta z(t, x) dx dt = \iint_D \left[\iint_{t_0, x_0}^{t_1, x_1} \delta u'(\tau, s) H_{uz}[\tau, s] Q(t, x, \tau, s) \delta u(t, x) ds d\tau \right] dx dt. \quad (29)$$

Далее по аналогии с [8] доказывается, что

$$\begin{aligned} & \iint_{t_0, x_0}^{t_1, x_1} \delta z'(t, x) H_{zz}[t, x] \delta z(t, x) dx dt = \\ & = \iint_D \left[\int_{\max(\tau, \alpha)}^{t_1} \int_{\max(s, \beta)}^{x_1} Q'(t, x, \tau, s) H_{zz}[t, x] Q(t, x, \alpha, \beta) dx dt \right] \delta u(\alpha, \beta) ds d\tau d\alpha d\beta. \end{aligned} \quad (30)$$

Введем обозначение

$$\begin{aligned} B(\tau, s, \alpha, \beta) = & \int_{\max(\tau, \alpha)}^{t_1} \int_{\max(s, \beta)}^{x_1} Q'(t, x, \tau, s) H_{zz}[t, x] Q(t, x, \alpha, \beta) dx dt - \\ & - Q'(t_1, x_1, \tau, s) \phi_{zz}(z(t_1, x_1)) Q(t_1, x_1, \alpha, \beta). \end{aligned} \quad (31)$$

Тогда, учитывая тождества (29), (30) и обозначение (31) в неравенстве (26), получим, что

$$\begin{aligned} & \iint_D \delta u'(\tau, s) B(\tau, s, \alpha, \beta) \delta u(\alpha, \beta) d\tau ds d\alpha d\beta + 2 \iint_D \left[\iint_{t, x}^{t_1, x_1} \delta u'(\tau, s) H_{uz}[\tau, s] Q(\tau, s, t, x) ds d\tau \right] \times \\ & \times f_u[t, x] \delta u(t, x) dx dt + \int_D \delta u'(t, x) H_{uu}[t, x] \delta u(t, x) dx dt \leq 0. \end{aligned} \quad (32)$$

Следовательно, доказана

Теорема 2. Для оптимальности классической экстремали $u(t, x)$ в рассматриваемой задаче необходимо, чтобы неравенство (32) выполнялось для всех $\delta u(t, x) \in U \subset R^r, (t, x) \in D$.

Необходимое условие оптимальности (32) является довольно общим условием оптимальности, учитывая произвольность допустимой вариации $\delta u(t, x)$ управления $u(t, x)$. Из него можно получить ряд легко проверяемых необходимых условий оптимальности, в частности аналог условия Лежандра–Клебша.

Пусть $\varepsilon > 0$ – достаточно малое произвольное число, $v \in R^r$ – произвольный вектор.

Специальную вариацию управления $u(t, x)$ определим по формуле

$$\delta u(t, x; \varepsilon) = \begin{cases} v, & (t, x) \in [\theta, \theta + \varepsilon] \times [\xi, \xi + \varepsilon], \\ 0, & (t, x) \in D \setminus [\theta, \theta + \varepsilon] \times [\xi, \xi + \varepsilon]. \end{cases} \quad (33)$$

Учитывая формулу (33), в неравенстве (32) получим, что

$$\varepsilon^2 (v' H_{uu} [\theta, \xi] v + o(\varepsilon^2)) \leq 0.$$

Отсюда в силу произвольности и достаточной малости ε получаем

$$v' H_{uu} [\theta, \xi] v \leq 0. \quad (34)$$

Теорема 3. Для оптимальности классической экстремали $u(t, x)$ необходимо, чтобы неравенство (33) выполнялось для всех $v \in U \subset R^r$ и $(\theta, \xi) \in [t_0, t_1] \times [x_0, x_1]$.

Заметим, что неравенство (34) является аналогом условия Лежандра–Клебша для рассматриваемой задачи.

Рассмотрим случай вырождения аналога условия Лежандра–Клебша.

Определение 1. Классическую экстремаль $u(t, x)$ назовем особым в классическом смысле управлением, если для всех $v \in U \subset R^r$ и $(\theta, \xi) \in [t_0, t_1] \times [x_0, x_1]$

$$v' H_{uu} [\theta, \xi] v = 0.$$

Из введенного определения ясно, что для особых в классическом смысле управлений аналог условия Лежандра–Клебша (34) вырождается.

Пусть $u(t, x)$ – особое в классическом смысле оптимальное управление. Тогда, используя формулу (33), из неравенства (32) после некоторых преобразований получаем

$$\begin{aligned} & \varepsilon^4 [v' f''_{uu} [\theta, \xi] B[\theta, \xi; \theta, \xi] f_u [\theta, \xi] v + \\ & + \frac{1}{2} v' H_{uz} [\theta, \xi] f_u [\theta, \xi] v] + o(\varepsilon^2) \leq 0. \end{aligned} \quad (35)$$

Из неравенства (35) в силу произвольности $\varepsilon > 0$ приходим к следующему утверждению.

Теорема 4. Для оптимальности особого в классическом смысле управления $u(t, x)$ необходимо, чтобы неравенство

$$v' [f''_{uu} [\theta, \xi] B[\theta, \xi; \theta, \xi] f_u [\theta, \xi] + \frac{1}{2} H_{uz} [\theta, \xi] f_u [\theta, \xi]] v \leq 0$$

выполнялось для всех $v \in U \subset R^r$ и $(\theta, \xi) \in [t_0, t_1] \times [x_0, x_1]$.

Последнее неравенство является необходимым условием оптимальности для классически особых управлений.

Заключение

В работе, учитывая открытость области управления и применяя один вариант метода приращений, вычислены первая и вторая вариации функционала качества.

Из условия равенства нулю первой вариации функционала качества получен аналог уравнения Эйлера.

Из условия неотрицательности второй вариации функционала качества вдоль оптимального процесса выведен ряд необходимых условий оптимальности второго порядка, носящих конструктивный характер.

Список источников

1. Рачинский В.В. Введение в общую теорию динамики сорбции и хроматографии. М.: Наука, 1964. 134 с.
2. Тихонов А.Н., Самарский А.А. Уравнения математической физики. М.: Наука, 1972. 735 с.
3. Сиразитдинов Т.К. Оптимизация систем с распределенными параметрами. М.: Наука, 1977. 479 с.

4. Егоров А.И. Об оптимальном управлении процессами в некоторых системах с распределенными параметрами // Автоматика и телемеханика. 1964. № 5. С. 613–623.
5. Егоров А.И. Оптимальные процессы в системах с распределенными параметрами и некоторые задачи теории инвариантности // Известия АН СССР. Сер. математическая. 1965. № 6. С. 1205–1260.
6. Ахиев С.С., Ахмедов Т.К. Необходимые условия оптимальности для некоторых задач теории оптимального управления // Доклады АН Азербайджанской ССР. 1972. Т. 28, № 5. С. 12–16.
7. Плотников В.И., Сумин В.И. Оптимизация объектов с распределенными параметрами, описываемых системой Гурса–Дарбу // Журнал вычислительной математики и математической физики. 1972. № 1. С. 61–77.
8. Мансимов К.Б. К оптимальности особых в классическом смысле управлений в системах Гурса–Дарбу // Доклады АН СССР. 1986. Т. 286, № 4. С. 808–812.
9. Мансимов К.Б. Об оптимальности квазисобых управлений в системах Гурса–Дарбу // Дифференциальные уравнения. 1986. № 11. С. 1952–1960.
10. Алексеев В.М., Тихомиров В.М., Фомин С.В. Оптимальное управление. М.: Физматлит, 2005. 429 с.
11. Понтрягин Л.С., Болтянский В.Г., Гамкрелидзе Р.В., Мищенко Е.Ф. Математическая теория оптимальных процессов. М.: Наука, 1984. 384 с.

References

1. Rachinsky, V.V. (1964) *Vvedenie v obshchuyu teoriyu dinamiki sorbtzii i khromatografii* [Introduction to the general theory of sorption dynamics and chromatography]. Moscow: Nauka.
2. Tikhonov, A.N. & Samarskii, A.A. (1972) *Urvneniya matematicheskoy fiziki* [Equations of Mathematical Physics]. Moscow: Nauka.
3. Sirazitdinov, T.K. (1977) *Optimizatsiya sistem s raspredelennymi parametrami* [Optimization of systems with distributed parameters]. Moscow: Nauka.
4. Egorov, A.I. (1964) Ob optimal'nom upravlenii protsessami v nekotorykh sistemakh s raspredelennymi parametrami [On optimal process control in some systems with distributed parameters]. *Avtomatika i telemehhanika*. 5. pp. 613–623.
5. Egorov, A.I. (1965) Optimal'nye protsessy v sistemakh s raspredelennymi parametrami i nekotorye zadachi teorii invariantnosti [Optimal processes in systems with distributed parameters and some problems of invariance theory]. *Izv. USSR Academy of Sciences, Ser. math.* 6. pp. 1205–1260.
6. Akhiev, S.S. & Akhmedov, T.K. (1972) Neobkhodimye usloviya optimal'nosti dlya nekotorykh zadach teorii optimal'nogo upravleniya [Necessary optimality conditions for some problems in optimal control theory]. *Dokl. AN Azerbaijan. SSR*. 28(5). pp. 12–16.
7. Plotnikov, V.I & Sumin, V.I. (1972) Optimizatsiya ob"ektov s raspredelennymi parametrami, opisyyaemykh sistemoy Gursa–Darbu [Optimization of objects with distributed parameters, described by the Goursat-Darboux system]. *Zhurnal. Calc. mat. and mat. physics*. 1. pp. 61–77.
8. Mansimov, K.B. (1986) K optimal'nosti osobykh v klassicheskom smysle upravleniy v sistemakh Gursa–Darbu [On the optimality of singular controls in the classical sense in Goursat-Darboux systems]. *DAN SSSR*. 286(4). pp. 808–812.
9. Mansimov, K.B. (1986) Ob optimal'nosti kvaziosobykh upravleniy v sistemakh Gursa–Darbu [On the optimality of quasi-singular controls in Goursat-Darboux systems]. *Diferentsialnye Urvneniya*. 11. pp. 1952–1960.
10. Alekseev, V.M., Tikhomirov V.M., & Fomin, S.V. (2005) *Optimal'noe upravlenie* [Optimal control]. Moscow: Fizmatlit.
11. 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.

Информация об авторе:

Рзаева Вафа Гюлага кызы – старший преподаватель Сумгаитского государственного университета (Сумгаит, Азербайджан).
E-mail: vafa.asgerova77@mail.ru

Автор заявляет об отсутствии конфликта интересов.

Information about the author:

Rzayeva Vafa G. (Senior Lecturer of the Sumgayit State University, Sumgayit, Azerbaijan). E-mail: vafa.asgerova77@mail.ru

The author declares no conflicts of interests.

Поступила в редакцию 08.07.2022; принята к публикации 01.03.2023

Received 08.07.2022; accepted for publication 01.03.2023

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

MATHEMATICAL MODELING

Научная статья

УДК 004.272

doi: 10.17223/19988605/62/2

**Распределенная очередь без использования блокировок
в модели удаленного доступа к памяти****Александр Викторович Бураченко¹, Алексей Александрович Пазников²,
Денис Павлович Державин³**^{1, 2, 3} Государственный электротехнический университет «ЛЭТИ», Санкт-Петербург, Россия¹ ss47305@gmail.com² apaznikov@gmail.com³ derzhavinden002@gmail.com

Аннотация. При разработке программного обеспечения для распределенных вычислительных систем в стандарте MPI наравне с моделью передачи сообщений (message-passing) используется модель удаленного доступа к памяти (remote memory access, MPI RMA, RMA). Модель во многих случаях позволяет повысить эффективность и упростить разработку параллельных программ. В рамках RMA имеют место задачи синхронизации параллельных процессов и потоков при обеспечении доступа к разделяемым (распределенным) структурам данных. В системах с общей памятью для аналогичной задачи активно используется неблокирующая синхронизация (non-blocking), гарантирующая прогресс выполнения операций (lock-free, wait-free, obstruction-free). При таком подходе задержка выполнения операций одним процессом не останавливает выполнения остальных процессов. Мы предполагаем, что такой подход может быть эффективным и при построении распределенных структур данных в модели RMA. Нами рассматривается идея построения неблокируемых распределенных структур данных в RMA на примере очереди, описаны построенные алгоритмы для выполнения основных операций, исследуется эффективность структуры данных, приведено экспериментальное сравнение с блокируемыми аналогами.

Ключевые слова: распределенная очередь; неблокирующие структуры данных; удаленный доступ к памяти; MPI; MPI RMA; lock-free; one-sided communications.

Благодарности: Исследование выполнено за счет гранта Российского научного фонда № 22-21-00686, <https://rscf.ru/project/22-21-00686/>

Для цитирования: Бураченко А.В., Пазников А.А., Державин Д.П. Распределенная очередь без использования блокировок в модели удаленного доступа к памяти // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2023. № 62. С. 13–24. doi: 10.17223/19988605/62/2

Original article

doi: 10.17223/19988605/62/2

Lock-free distributed queue in remote memory access model**Alexander V. Burachenko¹, Alexei A. Paznikov², Denis P. Derzhavin³**^{1, 2, 3} Electrotechnical University "LETI", St. Petersburg, Russian Federation¹ ss47305@gmail.com

² apaznikov@gmail.com

³ derzhavinden002@gmail.com

Abstract. In parallel programming for distributed-memory systems in MPI standard, remote memory access model (one-sided communications, MPI RMA, RMA) is used along with the message-passing. This model in many cases leverages the performance and simplifies parallel programming. Here arises the problem of synchronization of multiple parallel processes and threads accessing shared (concurrent, distributed) data structures. In shared-memory machines, non-blocking synchronization (lock-free, wait-free, obstruction-free) is widely used to solve the similar problem. In non-blocking synchronization, delays in execution of one process (thread) do not suspend execution of other threads. We suppose that this approach could also be effective in designing distributed data structures (in the RMA model particularly). In this article, we discuss the idea of building non-blocking distributed data structures in RMA model on the example of a queue, describe the designed algorithms of operations, investigate the efficiency, and provide an experimental comparison with lock-based counterparts.

Keywords: distributed queue; non-blocking concurrent data structures; remote memory access; MPI; MPI RMA; one-sided communications.

Acknowledgments: This research was supported by Russian Science Foundation (RSF) project 22-21-00686, <https://rscf.ru/en/project/22-21-00686/>

For citation: Burachenko, A.V., Paznikov, A.A., Derzhavin, D.P. (2023) Lock-free distributed queue in remote memory access model. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 62. pp. 13–24. doi: 10.17223/19988605/62/2

Введение

Под распределенной вычислительной системой (ВС) будем понимать композицию элементарных машин (ЭМ), взаимодействующих через коммуникационную сеть. ЭМ могут быть процессорными ядрами, процессорами и узлами. При разработке программ для распределенных ВС наравне с моделью передачи сообщений (message-passing) широко применяются альтернативные модели программирования, например модель удаленного доступа к памяти (Remote Memory Access, MPI RMA, RMA) и модель разделенного глобального адресного пространства (Partitioned Global Address Space, PGAS).

Опишем более детально модель RMA, которая является одной из перспективных и реализована в стандарте MPI в виде подсистемы односторонних коммуникаций (one-sided communications) [1, 2]. В RMA процессы напрямую обращаются к памяти других процессов вместо отправки и получения сообщений. В отличие от PGAS (UPC, CAF, Chapel, X10), RMA тесно интегрирована с библиотеками MPI и может быть использована наравне с моделью передачи сообщений, а также реализует доступ к низкоуровневым примитивам коммуникации и синхронизации. Применение RMA часто позволяет сократить время выполнения программ по сравнению с моделями передачи сообщений или PGAS [3]. Оптимизация достигается, в частности, благодаря поддержке современными коммуникационными сетями (Infiniband, PERCS, Gemini, Aries и др.) технологии RDMA [4], реализующей обращение к удаленным сегментам памяти без участия центрального процессора.

Основными операциями в модели RMA являются неблокируемые MPI_Put (запись в удаленную память) и MPI_Get (чтение) и набор атомарных операций: MPI_Accumulate, MPI_Get_accumulate, MPI_Compare_and_swap и др. RMA-вызовы должны находиться внутри областей (эпох, epochs), в рамках которых выполняется синхронизация. RMA предоставляет активный и пассивный методы синхронизации. В настоящей работе применяется пассивный метод синхронизации (passive target synchronization) [2]: процесс открывает эпоху, затем выполняет RMA-операции для доступа к зарегистрированным сегментам памяти (окнам, window) других процессов. Таким образом, RMA-операции выполняются в одностороннем порядке, без явного участия других процессов, и сопровождаются меньшими накладными расходами, чем при активной синхронизации [5].

В рамках моделей RMA и PGAS имеет место задача обеспечения масштабируемого доступа к структурам данных, которые распределены между узлами ВС. Данная задача не является новой в параллельном программировании. Так, для систем с общей памятью имеет место необходимость в син-

хронизации потоков, обращающихся к разделяемым (concurrent, thread-safe) структурам данных. Такие структуры должны обеспечивать корректный (как правило линейризуемый, linearizable) доступ параллельных потоков в произвольные моменты времени [4, 6, 7]. Эффективность синхронизации существенно влияет на время выполнения программ и обуславливает гарантии прогресса выполнения.

Существующие методы синхронизации можно разделить на два класса: с применением блокировок (locks) и без блокировок (non-blocking). Блокировки обладают простой семантикой и часто достаточно эффективны, однако неблокирующий подход обеспечивает гарантии выполнения и позволяет избежать тупиковых ситуаций (deadlocks), инверсий приоритетов (priority inversion) и некоторых других проблем блокировок. Неблокирующим называется такой алгоритм, в котором задержка выполнения одного потока не останавливает прогресс выполнения программы в целом (при соблюдении ряда условий) [4]. Значительная часть работ в области разделяемых структур данных направлена на создание средств синхронизации для ВС с общей памятью. К ним относятся алгоритмы блокировки потоков [4, 9] (TTS, Backoff, CLH, MCS, Oyama, Flat Combining, RCL и др.). Хотя некоторые методы (Hierarchical Locks, Cohorting и др.) учитывают иерархические уровни системы, они неприменимы в ВС с распределенной памятью. Неблокируемые структуры [4, 6–9] также разработаны для многоядерных ВС и неприменимы в распределенных ВС. Структуры, оптимизированные для NUMA [10], также неприменимы в распределенных средах.

Существует, однако, ряд работ по реализации распределенных неблокирующих структур: FastQueue, CircularQueue [11] и Active Message Queue [12]. Их общая проблема – централизация данных в памяти одного процесса. В результате подхода остальные процессы, работающие со структурой данных, постоянно обращаются к центральному процессу. Это является узким местом и может быть причиной снижения производительности очереди, особенно на большом числе процессов. В [13] рассматривается реализация распределенного неблокирующего стека на основе алгоритма Elimination-Backoff Stack. Экспериментальные оценки указывают на неплохую производительность на небольших подсистемах. Основным недостатком структуры также является ее централизованность.

В данной работе исследуется возможность построения неблокируемых (lock-free) структур данных для распределенных ВС на примере очереди. Предлагается подход на основе децентрализации данных между процессами параллельной программы. Мы предполагаем, что такой подход может быть эффективен при построении и других распределенных структур данных.

1. Неблокирующая распределенная очередь с децентрализацией хранения данных

1.1. Модель и структура очереди

Очередь – коллекция объектов, реализующая дисциплину FIFO («первым вошел – первым вышел»). Основные операции: добавление (insert, enqueue) элемента в последнюю позицию (хвост, tail, T) и извлечение (remove, dequeue) элемента из первой позиции (голова, head, H). Элементы распределенной очереди находятся в памяти процессов, выполняющихся на распределенной ВС.

Пусть имеется ВС из N ЭМ. На каждой ЭМ $i = 1, 2, \dots, N$ имеется локальная оперативная память m_i , причем $m_i \cap m_j = \emptyset, \forall i, j \in \{1, 2, \dots, N\}$. Считаем, что на каждой ЭМ i запущен процесс p_i .

Обозначим RMA-окно w , необходимое для выполнения операций. Процессы $p_i, i = 1, 2, \dots, N$, обладают дескриптором окна w и публикуют в нем часть своей памяти m_i (назовем её m_i^*), таким образом предоставляя доступ к m_i^* остальным процессам.

Для хранения элементов очереди на каждом процессе выделена память под пул элементов (в текущей реализации – массив). Пусть e_i – пул элементов очереди в памяти m_i^* процесса p_i . Размер массива e_i фиксирован и равен K . При добавлении новых элементов процесс p_i использует ячейки из массива $e_i: e_{ij}, \forall i, j: i \in \{1, 2, \dots, N\}, j \in \{1, 2, \dots, N\}$. Тогда голова очереди H – это элемент e_{ij} , являющийся первым в очереди, а хвост очереди T – это элемент e_{ij} , являющийся последним в очереди.

Для определения местоположения элементов очереди введем понятие *ссылка на элемент* e_{ij} – пара чисел (i, j) , которая однозначно идентифицирует элемент j распределенной очереди, который находится в памяти процесса i . Пустую ссылку обозначим (\emptyset, \emptyset) . Каждый элемент e_{ij} характеризуется:

- Временной меткой добавления в очередь – τ_{ij} .
- Пользовательскими данными – v_{ij} .
- Состоянием c_{ij} . Принимает одно из трех значений: F, A, D. F – элемент e_{ij} свободен для использования при следующем добавлении в очередь, A – элемент занят, находится в очереди, D – удален из очереди, но еще не доступен для повторного использования. Состояние элемента циклически меняется: из F в A при добавлении, из A в D при удалении, из D в F при освобождении.

– Ссылкой на следующий элемент – n_{ij} . Если за e_{ij} следует e_{gf} , то $n_{ij} = (g, f)$. При добавлении e_{ij} в очередь $n_{ij} = (\emptyset, \emptyset)$.

– Ссылкой на себя l_{ij} . Для e_{ij} , $l_{ij} = (i, j)$. Данная ссылка необходима для возможности обновления информации об элементе.

Также каждый процесс p_i имеет в своей памяти m_i^* ссылки на голову h_i и хвост t_i очереди, $\forall i \in \{1, 2, \dots, N\}$, которые служат для определения текущего положения актуальных головы H и хвоста T .

Процесс p_1 дополнительно обладает ссылкой s – элемент для достижения консенсуса при добавлении первого элемента. Когда несколько процессов одновременно добавляют новый элемент в очередь, они принимают решение, чей элемент будет добавлен. Для этого применяется операция «сравнение с обменом» (Compare-And-Swap, CAS) с текущего хвоста T на новый хвост n_{ij} . Но сразу после инициализации очередь пуста: она не обладает ни головой, ни хвостом. Поэтому необходима область памяти, в которой можно было бы решить, чей элемент будет первым добавлен в очередь.

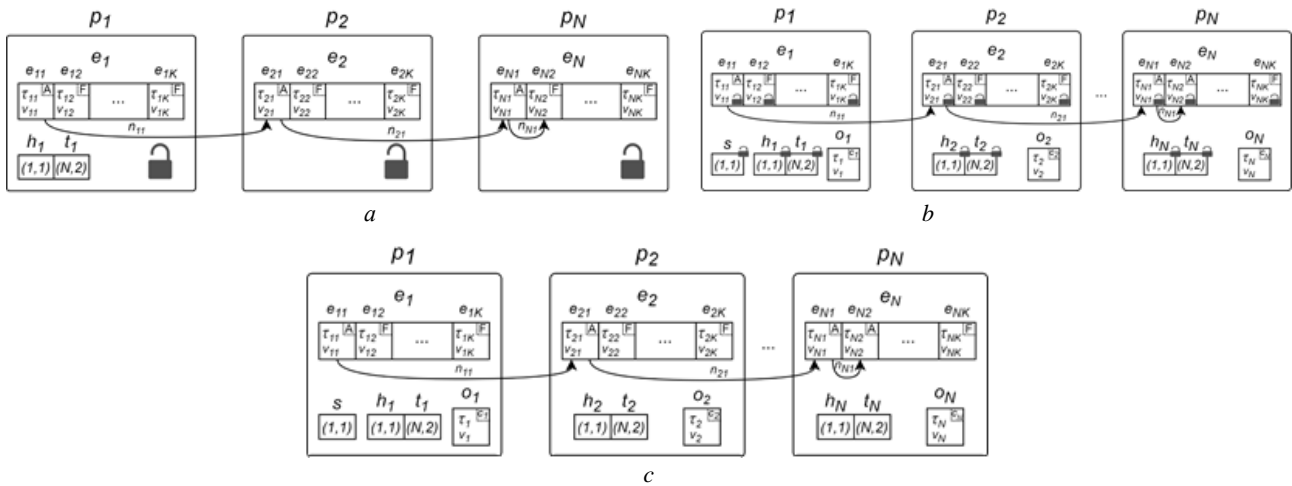


Рис. 1. Структура распределенной очереди. *a* – блокирующая, с централизованным хранением информации о голове / хвосте, *b* – блокирующая, децентрализованная, *c* – неблокирующая, децентрализованная
 Fig. 1. The structure of a distributed queue. *a* – lock-based, with centralized head/tail information storage, *b* – lock-based, decentralized, *c* – non-blocking, decentralized

В рамках централизованного подхода информация о положении головы и хвоста находится в памяти главного процесса p_1 (рис. 1, *a*). Процессы p_i при добавлении или удалении элемента обновляют ссылки t_1 или h_1 соответственно. Такой подход порождает узкое место (bottleneck), так как p_i постоянно обращаются к p_1 , что ведет к снижению пропускной способности (throughput) структуры данных.

1.2. Децентрализованный метод организации очереди

Предлагается подход с децентрализованным хранением информации о положении H и T (рис. 1, *в*). Каждый процесс p_i обладает локальной ссылкой h_i и t_i . При поиске актуальных головы или хвоста

процесс p_i обращается к ссылке h_i или t_i соответственно. После вставки или удаления процесс p_i обновляет ссылки h_j и t_j ($j = 1, 2, \dots, N$) для всех остальных процессов согласно разработанному алгоритму оповещения, в рамках которого p_i сначала обновляет свои ссылки h_i и t_i , затем ссылки на остальных процессах: h_j и $t_j \forall j \in \{1, 2, \dots, N\}, j \neq i$. Алгоритм работает по трем методам: «только голова», «только хвост» и «сначала хвост, затем голова». В зависимости от метода алгоритм обновляет либо ссылки h_j , либо t_j , либо t_j затем h_j для каждого процесса. Подробнее алгоритм рассмотрен в разд. 1.5.

При удалении элементов они переходят в состояние D, т.е. считаются удаленными, но еще не доступны для повторного использования. Для освобождения элементов e_{ij} , т.е. перевода их в состояние F, разработан алгоритм очистки пула e_i (см. разд. 1.6). Алгоритм использует дополнительный элемент o_i . Этим элементом процессы оперируют при чтении элементов очереди: в него копируется содержимое полей элемента, с которым работает процесс p_i в данный момент. o_i находится в памяти m_i^* и обладает теми же атрибутами, что и e_{ij} (для o_i обозначения полей: $\tau_i, v_i, c_i, n_i, l_i$), но не входит в состав пула e_i . Так как важна точность временных меток τ_{ij} и их согласованность, часы процессов синхронизируются.

1.3. Операция добавления элемента в очередь

Опишем операцию добавления (enqueue) (рис. 2). Под *atomic_get* здесь и далее понимается атомарная операция удаленного чтения на основе MPI_Get_accumulate. Операция CAS (MPI_Compare_and_swap) позволяет разрешать консенсус при изменении данных в общем участке памяти для любого числа процессов, модифицирующих очередь. Операции *bcast_t* и *bcast_th* – это вызовы алгоритма оповещения по схемам «только хвост» и «сначала хвост, затем голова».

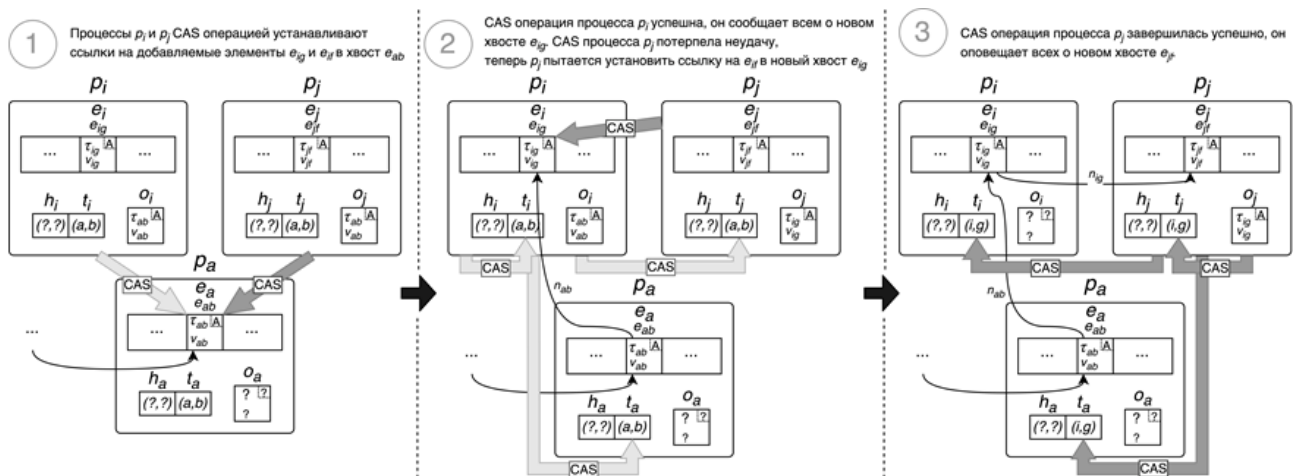


Рис. 2. Схема добавления элемента в очередь
Fig. 2. Scheme of adding an item to the queue

Алгоритм добавления в очередь процессом p_i (табл. 1, а):

1. Инициализация добавляемого элемента e_{ij} . Выбирается свободная ячейка e_{ij} в массиве e_i , в ячейку записываются данные ($v_{ij} = val, c_{ij} = A, n_{ij} = (\emptyset, \emptyset)$) (строка 1).
2. Если массив e_i переполнен, запускается алгоритм очистки. Если после отработки алгоритма очистки свободной ячейки нет, завершить выполнение алгоритма ошибкой (строки 1, 2).
3. Если ссылка на хвост $t_i = (\emptyset, \emptyset)$, получить ссылку s (строки 4, 5). Иначе – перейти к шагу 7.
4. Если $s = (\emptyset, \emptyset)$, установить временную метку τ_{ij} в e_{ij} , установить в s ссылку (i, j) CAS-операцией (строки 6–8). Иначе – перейти к шагу 6.
5. Если CAS завершилась успешно (точка линеаризации), оповестить все процессы о новых хвосте и голове и завершить алгоритм (строки 9, 10). Иначе – обновить s (строка 12) и перейти к шагу 6.
6. Записать s в t_i (строка 14).

7. Получить элемент по ссылке t_i и записать его в o_i (строка 16).
8. Если статус $c_i = A$ (строка 17), перейти к шагу 9. Иначе – перейти к шагу 12.
9. Если $n_i = (\emptyset, \emptyset)$, установить временную метку τ_{ij} в e_{ij} , установить в n_i ссылку (i, j) CAS-операцией (строки 18–20). Иначе – перейти к шагу 11.
10. Если CAS завершилась успешно (точка линеаризации), оповестить процессы о новом хвосте e_{ij} (строка 21) и завершить алгоритм (строка 22). Иначе – обновить o_i и перейти к шагу 11 (строка 24).
11. Получить элемент по ссылке n_i , записать его в o_i и перейти к шагу 8 (строка 26).
12. Если статус $c_i = D$, перейти к шагу 14 (строка 29). Иначе – перейти к шагу 3 (строка 41).
13. Если $n_i = (\emptyset, \emptyset)$, установить метку τ_{ij} в e_{ij} , установить в n_i ссылку (i, j) CAS-операцией (строки 30–32). Иначе – получить элемент по ссылке n_i , записать его в o_i и перейти к шагу 8 (строки 37, 38).
14. Если CAS-операция завершилась успешно (точка линеаризации), оповестить все процессы о новом хвосте и новой голове e_{ij} и завершить выполнение алгоритма (строки 33, 34).

Таблица 1

Алгоритмы выполнения операций для распределенной неблокирующей очереди: a – операция enqueue добавления элемента в очередь, b – операция dequeue извлечения элемента из очереди

a	b
<p>Входные данные: val – добавляемые данные win – окно для выполнения RMA-операций</p> <pre> 1 if init_elem(val, &e_{ij}) == false then 2 return BUFFER_FULL 3 end if 4 if t_i == (-1, -1) then 5 s = atomic_get(s, win) 6 if s == (-1, -1) then 7 e_{ij}.τ_{ij} = get_ts() 8 if cas(s, (-1, -1), (i, j), win) then 9 bcast_th(e_{ij}, win) 10 return SUCCESS 11 end if 12 s = atomic_get(s, win) 13 end if 14 t_i = s 15 end if 16 o_i = atomic_get(t_i, win) 17 if o_i.c_i == A then 18 if o_i.n_i == (-1, -1) then 19 e_{ij}.τ_{ij} = get_ts() 20 if cas(o_i.n_i, (-1, -1), (i, j), win) then 21 bcast_t(e_{ij}, win) 22 return SUCCESS 23 end if 24 o_i = atomic_get(o_i.l_i, win) 25 end if 26 o_i = atomic_get(o_i.n_i, win) 27 goto 17 28 end if 29 if o_i.c_i == D then 30 if o_i.n_i == (-1, -1) then 31 e_{ij}.τ_{ij} = get_ts() 32 if cas(o_i.n_i, (-1, -1), (i, j), win) then 33 bcast_th(e_{ij}, win) 34 return SUCCESS 35 end if 36 else 37 o_i = atomic_get(o_i.n_i, win) 38 goto 17 39 end if 40 end if 41 goto 4</pre>	<p>Входные данные: win – окно для выполнения RMA-операций</p> <pre> 1 if h_i == (-1, -1) then 2 s = atomic_get(s, win) 3 if s == (-1, -1) then 4 return QUEUE_EMPTY 5 end if 6 h_i = s 7 end if 8 o_i = atomic_get(h_i, win) 9 if o_i.c_i == A then 10 if cas(o_i.c_i, A, D, win) then 11 val = o_i.v_i 12 o_i = atomic_get(o_i.l_i, win) 13 if o_i.n_i != (-1, -1) then 14 bcast_h(o_i.n_i, win) 15 end if 16 return SUCCESS 17 end if 18 o_i = atomic_get(o_i.l_i, win) 19 end if 20 if o_i.c_i == D then 21 if o_i.n_i != (-1, -1) then 22 o_i = atomic_get(o_i.n_i, win) 23 goto 9 24 else 25 return QUEUE_EMPTY 26 end if 27 end if 28 goto 1</pre>

1.4. Операция удаления элементов

Опишем алгоритм удаления (dequeue) элементов. Операция *bcast_h* – это вызов алгоритма оповещения с методом «только голова». Алгоритм удаления из очереди процессом p_i представлен на рис. 3, в табл. 1, б).

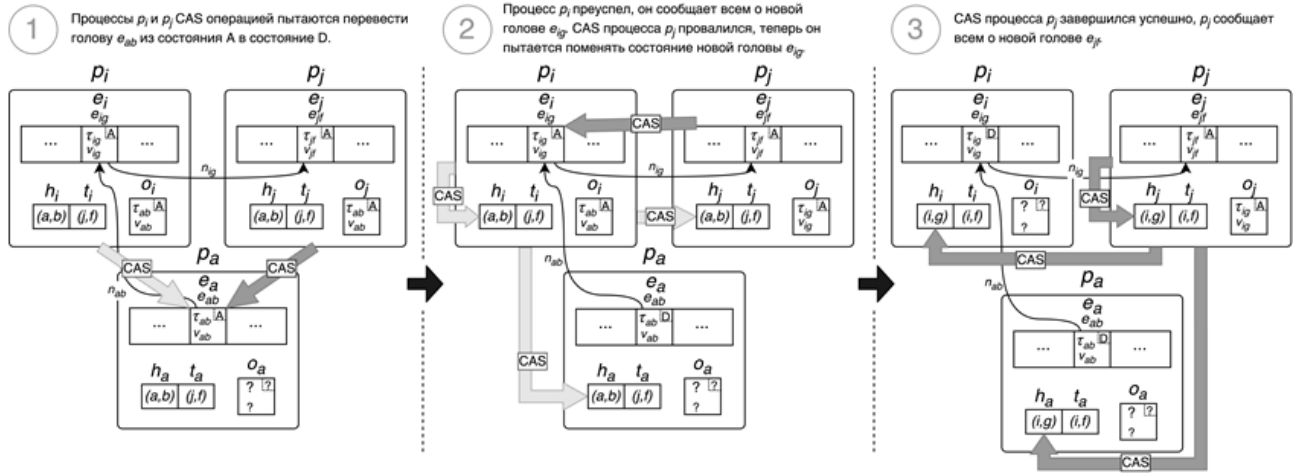


Рис. 3. Схема удаления элемента из очереди
Fig. 3. Scheme of removing an item out of the queue

1. Если ссылка на голову $h_i = (\emptyset, \emptyset)$, получить ссылку s (строки 1, 2). Иначе – перейти к шагу 4.
2. Если $s = (\emptyset, \emptyset)$, значит очередь пуста, завершить выполнение алгоритма (строки 3, 4).
3. Записать s в h_i (строка 6).
4. Получить элемент по ссылке h_i и записать его в o_i (строка 8).
5. Если статус $c_i = A$, изменить c_i на D операцией CAS (строки 9, 10). Иначе – перейти к шагу 8.
6. Если CAS-операция завершилась успешно (точка линейаризации), перейти к шагу 7. Иначе – обновить o_i и перейти к шагу 8 (строки 18–20).
7. Если $n_i \neq (\emptyset, \emptyset)$, оповестить все процессы о новой голове (элемент по ссылке n_i) (строки 12–16). Завершить выполнение алгоритма.
8. Если статус $c_i = D$, перейти к шагу 9 (строка 20). Иначе – перейти к шагу 1 (строка 28).
9. Если $n_i \neq (\emptyset, \emptyset)$, получить элемент по ссылке n_i , записать его в o_i и перейти к шагу 5 (строки 21–23). Иначе очередь пуста – завершить выполнение алгоритма (строка 25).

1.5. Алгоритм оповещения (bcast)

Каждый процесс p_i обладает ссылками на голову h_i и хвост t_i . При изменении головы или хвоста необходимо актуализировать ссылки $h_j, t_j, \forall j \neq i$, всех процессов – для этого используется алгоритм оповещения *bcast*. Алгоритм выполняется в соответствии с тремя схемами (*bcast_method*):

1. Только голова. Обновляются ссылки на хвост t_j . Вызывается после того, как процесс добавил элемент в очередь.
2. Только хвост. Обновляются ссылки на голову h_j . Вызывается после того, как процесс удалил элемент.
3. Сначала хвост, затем голова. Обновляются ссылки t_j и h_j . Вызывается после того, как процесс добавил элемент в очередь вслед за элементом, находящимся в состоянии D.

Алгоритм *bcast* возвращает результат выполнения: SUCCESS – ссылка успешно обновлена, FAIL – ссылка не обновлена (означает, что ее актуализировал другой процесс). Если метод возвращает код FAIL, это признак остановки алгоритма оповещения, так как распространяемый элемент уже не является актуальным хвостом или головой, а актуальную информацию раздает другой процесс.

Ниже описан шаблон алгоритма оповещения для процесса p_i (табл. 2, a). На вход передается тиражируемый элемент e_{ij} , метод оповещения $bcast_method$ и окно для выполнения RMA-операций win .

1. Сформировать массив b рангов процессов, работающих с очередью (строка 1).
2. Обновить h_i и t_i согласно указанному методу $bcast_method$ (строка 2). Если $bcast_method$ вернул код FAIL, завершить выполнение алгоритма (строки 2, 3). Иначе – удалить ранг i из b (строка 5).
3. Если массив b пуст, завершить выполнение алгоритма (строки 7, 8). Иначе – перейти к шагу 4.
4. Случайно выбрать процесс p_a из оставшихся в b (строка 10). Обновить ссылки h_a и t_a согласно указанному методу $bcast_method$ (строка 11). Если $bcast_method$ вернул код FAIL – завершить выполнение алгоритма (строки 11, 12). Иначе – удалить ранг a из b (строка 14) и перейти к шагу 3.

Таблица 2

Алгоритмы оповещения для распределенной неблокирующей очереди: a – шаблон алгоритма оповещения, b – метод «только голова», c – метод «только хвост», d – метод «сначала хвост, затем голова»

a	b
<p>Входные данные: e_{ij} – тиражируемый элемент $bcast_method$ – метод оповещения win – окно для выполнения RMA-операций</p> <pre> 1 b = get_all_ranks() 2 if bcast_method(i, e_{ij}, win) == FAIL then 3 return 4 end if 5 exclude_rank(i, &b) 6 while true 7 if is_empty(b) then 8 return 9 end if 10 a = get_next_rank_rand(b) 11 if bcast_method(a, e_{ij}, win) == FAIL then 12 return 13 end if 14 exclude_rank(a, &b) 15 end </pre>	<p>Входные данные: a – ранг оповещаемого процесса e_{ij} – тиражируемый элемент win – окно для выполнения RMA-операций</p> <pre> 1 do 2 h_a = atomic_get(a, win) 3 if h_a != (-1, -1) then 4 o_i = atomic_get(h_a, win) 5 if e_{ij}.τ_{ij} < o_i.τ_i != (-1, -1) then 6 return FAIL 7 end if 8 end if 9 while cas(h_a, h_a, (i, j), win) == false 10 return SUCCESS </pre>
c	d
<p>Входные данные: a – ранг оповещаемого процесса, e_{ij} – тиражируемый элемент win – окно для выполнения RMA-операций</p> <pre> 1 do 2 t_a = atomic_get(a, win) 3 if t_a != (-1, -1) then 4 o_i = atomic_get(t_a, win) 5 if e_{ij}.τ_{ij} < o_i.τ_i != (-1, -1) then 6 return FAIL 7 end if 8 end if 9 while cas(t_a, t_a, (i, j), win) == false 10 return SUCCESS </pre>	<p>Входные данные: a – ранг оповещаемого процесса, e_{ij} – тиражируемый элемент $check_head$ – флаг требования обновить ссылку на голову $check_tail$ – флаг требования обновить ссылку на хвост, win – RMA-окно</p> <pre> 1 if check_tail then 2 do 3 t_a = atomic_get(a, win) 4 if t_a != (-1, -1) then 5 o_i = atomic_get(t_a, win) 6 if e_{ij}.τ_{ij} < o_i.τ_i != (-1, -1) then 7 check_tail = false 8 break 9 end if 10 end if 11 while cas(t_a, t_a, (i, j), win) == false 12 end if 13 if check_head then 14 do 15 h_a = atomic_get(a, win) 16 if h_a != (-1, -1) then 17 o_i = atomic_get(h_a, win) 18 if e_{ij}.τ_{ij} < o_i.τ_i != (-1, -1) then 19 check_tail = false 20 break 21 end if 22 end if 23 while cas(h_a, h_a, (i, j), win) == false 24 end if 25 if check_tail == false and check_head = false then 26 return FAIL 27 end if 28 return SUCCESS </pre>

Рассмотрим подробнее алгоритмы методов обновления в памяти j -го процесса ссылок на голову и / или хвост. Метод «только голова». Исходные данные: номер процесса a , тиражируемый элемент e_{ij} и RMA-окно win . Алгоритм тогда будет выглядеть (табл. 2, b):

1. Получить ссылку на голову h_a из памяти процесса a (строка 2).
2. Если $h_a \neq (\emptyset, \emptyset)$, получить элемент по ссылке h_a и записать его в o_i (строки 3, 4). Иначе – к шагу 4.
3. Если метка $\tau_{ij} < \tau_j$, вернуть FAIL и завершить алгоритм (строки 5, 6). Иначе – переход к шагу 4.
4. С помощью CAS установить ссылку (i, j) в h_a (строка 9). Если операция завершилась успешно (точка линеаризации), вернуть SUCCESS и завершить алгоритм (строка 10). Иначе перейти к шагу 1.

Метод «только хвост». Исходные данные те же. Алгоритм (табл. 2, c):

1. Получить ссылку на хвост t_a из памяти процесса a (строка 2).
2. Если $t_a \neq (\emptyset, \emptyset)$, получить элемент по ссылке t_a и записать его в o_i (строки 3, 4). Иначе – к шагу 4.
3. Если метка $\tau_{ij} < \tau_j$, вернуть код FAIL и завершить алгоритм (строки 5, 6). Иначе – к шагу 4.
4. С помощью CAS установить ссылку (i, j) в t_a (строка 9). Если операция завершилась успешно (точка линеаризации), вернуть SUCCESS и завершить алгоритм (строка 10). Иначе – переход к шагу 1.

Метод «сначала хвост, затем голова». Данные: номер процесса a , элемент e_{ij} , флаги $check_head$ (обновлять ли ссылку на голову) и $check_tail$ (обновлять ли ссылку на хвост). Параметры $check_head$ и $check_tail$ нужны, так как в какой-то момент e_{ij} может перестать быть актуальным хвостом (головой), и нужно иметь возможность продолжить алгоритм, не проверяя ссылки. Алгоритм (табл. 2, d):

1. Если $check_tail = true$ перейти к пункту 2 (строка 1). Иначе – перейти к пункту 6.
2. Получить ссылку на хвост t_a из памяти процесса a (строка 3).
3. Если $t_a \neq (\emptyset, \emptyset)$, получить элемент по ссылке t_a и записать в o_i (строки 4, 5). Иначе – к шагу 5.
4. Если метка $\tau_{ij} < \tau_j$, записать $false$ в $check_tail$ и перейти к шагу 6 (строки 6–8). Иначе – к шагу 5.
5. CAS-операцией установить ссылку (i, j) в t_a (строка 11). Если операция завершилась успешно (точка линеаризации), перейти к пункту 6. Иначе – перейти к шагу 2.
6. Если $check_head = true$, перейти к шагу 7 (строка 13). Иначе – перейти к шагу 11.
7. Получить ссылку на голову h_a из памяти процесса a (строка 15).
8. Если $h_a \neq (\emptyset, \emptyset)$, получить элемент по ссылке h_a и записать в o_i (строки 16–17). Иначе – к шагу 10.
9. Если метка $\tau_{ij} < \tau_j$, записать $false$ в $check_head$ и перейти к шагу 11 (строки 18–20). Иначе – к шагу 10.
10. CAS-операцией установить ссылку (i, j) в h_a (строка 23). Если операция завершилась успешно (точка линеаризации), перейти к шагу 11. Иначе – перейти к шагу 7.
11. Если $check_tail = false$ и $check_head = false$, вернуть код FAIL (строки 25–26). Иначе вернуть код SUCCESS и завершить выполнение алгоритма (строка 28).

1.6. Алгоритм очистки (cleaning)

Элементы e_{ij} переходят из состояния F в A при добавлении, затем из A в D при удалении. В какой-то момент в массиве e_i не остается доступных элементов (в состоянии F), и последующие операции добавления завершаются ошибкой. Для избегания ситуации применяется *алгоритм очистки*: он переводит элементы из D в F и вызывается, когда при поиске свободной ячейки в пуле e_i не остается свободных элементов при добавлении в очередь. Алгоритм состоит из двух частей: 1) определение минимальной временной метки τ_{min} , элементы с которой сейчас используются процессами; 2) перевод всех элементов в массиве e_i в состояние D с меткой меньше, чем τ_{min} . Опишем алгоритм (для p_i):

1. Сформировать массив b рангов процессов, работающих с очередью. Получить элемент по ссылке t_i и записать его в o_i . Записать τ_i в τ_{min} . Получить элемент по ссылке h_i и записать его в o_i . Если $\tau_i < \tau_{min}$, то записать τ_i в τ_{min} . Удалить ранг i из массива b .
2. Если массив b пуст, перейти к пункту 4. Иначе – перейти к пункту 3.
3. Случайно выбрать процесс p_a из оставшихся в массиве b . Получить элемент по ссылке h_a и записать его в o_i . Если $\tau_i < \tau_{min}$, то записать τ_i в τ_{min} . Получить элемент по ссылке t_a и записать его в o_i .

Если $\tau_i < \tau_{min}$, то записать τ_i в τ_{min} . Получить элемент o_a и записать его в o_i . Если $\tau_i < \tau_{min}$, то записать τ_i в τ_{min} . Удалить ранг a из массива b . Перейти к пункту 2.

4. Пройти по массиву e_i . Для каждого e_{ij} : если $c_{ij} = D$ и $\tau_{ij} < \tau_{min}$, записать F в c_{ij} . Завершить выполнение алгоритма.

2. Проведение экспериментов

Экспериментальное исследование проводилось на кластере Информационно-вычислительного центра Новосибирского государственного университета. В экспериментах использовались 6 узлов по два 6-ядерных процессора Intel Xeon X5670 на каждом (суммарно 72 ядра). MPI: Open MPI 4.1.0.

Был разработан синтетический тест, выполняющий $n = 10\,000$ операций вставки / удаления (тип операции выбирается равновероятно). Число процессов N варьировало от 2 до 72. Измерялась пропускная способность $b = n \times p / t$, где t – время проведения эксперимента, p – количество процессов.

Реализованная распределенная неблокирующая очередь сравнивалась с двумя очередями на основе блокировок. Первая – блокирующая централизованная очередь (см. рис. 1, *a*): информация о местоположении головы / хвоста находится в памяти процесса 0. Блокировка устанавливается на процесс. Перед чтением / изменением информации о голове / хвосте или элемента очереди из памяти некоторого процесса i читающий / изменяющий данные процесс захватывает блокировку процесса i ; после окончания работы – освобождает. Вторая – блокирующая очередь с децентрализованным хранением положения головы / хвоста (см. рис. 1, *b*). Она аналогична неблокирующей, за исключением использования блокировок. В отличие от централизованной очереди, здесь блокируется не процесс, а отдельные элементы. Перед работой с элементом или ссылкой на голову / хвост процесс захватывает блокировку на этом элементе, после работы освобождает.

Пропускная способность разработанной очереди значительно превосходит пропускную способность блокирующих аналогов (рис. 4). Подход lock-free избавляет от накладных расходов на захват и освобождение блокировки и позволяет потоку в случае неудачи операции сразу повторить попытку, не затрачивая время на ожидание. Блокирующая децентрализованная очередь производительнее централизованной: децентрализация позволяет распределить нагрузку на процессы и повышает пропускную способность. Все рассматриваемые очереди масштабируются в умеренной степени. С ростом числа процессов пропускная способность структур данных падает, тем не менее в случае неблокируемой очереди остается на достаточно высоком уровне.

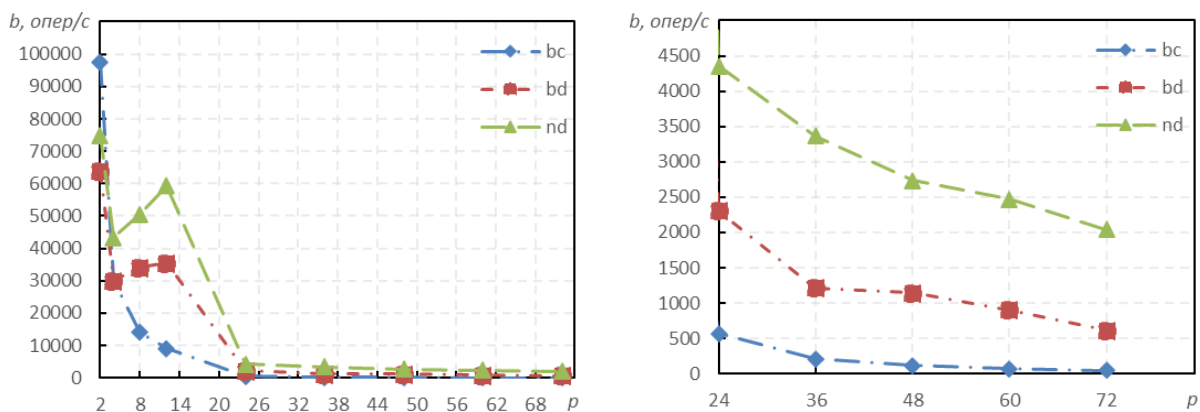


Рис. 4. Пропускная способность очередей (*bc* – блокирующая централизованная, *bd* – блокирующая децентрализованная, *nd* – неблокирующая децентрализованная)

Fig. 4. Throughput of queues (*bc* – blocking centralized, *bd* – blocking decentralized, *nd* – non-blocking decentralized)

Для оценки времени выполнения операций добавления и удаления для неблокирующей очереди выполнялось измерение времени реализации операций и их ключевых этапов. Каждый процесс самостоятельно выполняет измерения с последующей отправкой результатов процессу 0, который вычисляет средние значения. Перечень измерений:

- 1) общее время операции добавления (enq_all);
- 2) время поиска хвоста при добавлении (enq_hop);
- 3) общее время операции удаления (deq_all);
- 4) время поиска головы при удалении (deq_hop);
- 5) общее время алгоритма оповещения (bcast_all).

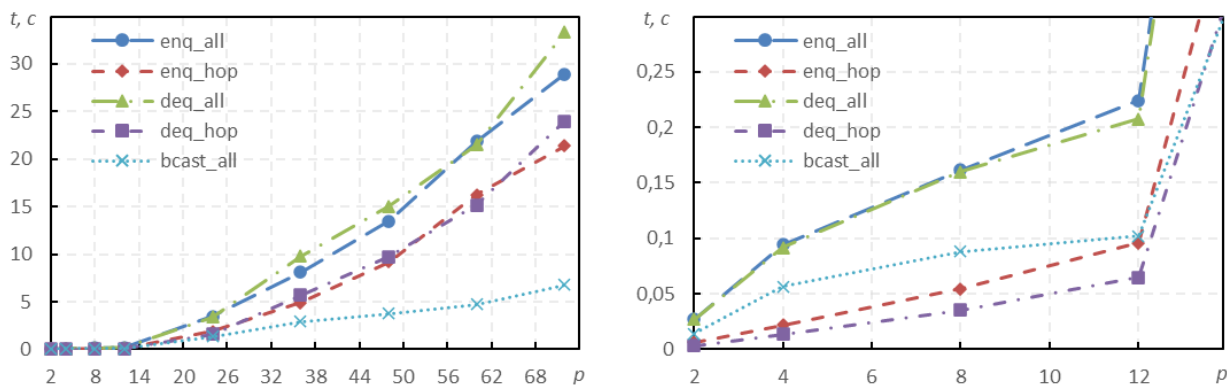


Рис. 5. Время операций добавления и удаления из неблокирующей очереди
Fig. 5. Time of operations of inserting and removing for the non-blocking queue

Большая часть времени уходит на поиск головы / хвоста (рис. 5). Объясняется это тем, что с ростом числа процессов увеличивается число операций с очередью в единицу времени. Отсюда растет вероятность для процесса p_i получить по ссылке t_i элемент, который уже не является хвостом. Тогда процессу приходится переходами по ссылкам на следующие элементы добираться до хвоста. Чем больше операций добавления в очередь выполняется в единицу времени, тем больше времени тратится на поиск хвоста. Та же проблема и с операцией удаления. Как и следовало ожидать, в пределах одного узла (до 12 процессов) операции выполняются достаточно быстро, а с увеличением числа узлов время растет. Существенное влияние оказывают накладные расходы на передачу данных между узлами.

Заключение

В данной статье разработан алгоритм распределенной неблокирующей (lock-free) очереди с децентрализацией хранения информации о положении головы и хвоста в модели MPI RMA. Очередь характеризуется значительно большей пропускной способностью по сравнению с блокирующими аналогами. Оптимизация достигается за счет неблокируемого подхода к синхронизации и уменьшения вероятности появления узких мест при выполнении операций благодаря децентрализации. Созданная структура данных умеренно масштабируется. Выполнен анализ эффективности предложенного алгоритма.

Список источников

1. Liu J., Wu J., Panda D.K. High performance RDMA-based MPI implementation over InfiniBand // International Journal of Parallel Programming. 2004. V. 32. P. 167–198.
2. Hoefler T., Dinan J., Thakur R., Barrett B., Balaji P., Gropp W., Underwood K. Remote memory access programming in MPI-3 // ACM Transactions on Parallel Computing. 2015. V. 2 (2). Art. 9. 26 p.
3. Gerstenberger R., Besta M., Hoefler T. Enabling Highly-Scalable Remote Memory Access Programming with MPI-3 One Sided // Scientific Programming. 2014. V. 2, is. 2. P. 75–91.
4. Herlihy M., Shavit N. The art of multiprocessor programming. Amsterdam et al. : Morgan Kaufmann, 2012. 537 p.
5. Schuchart J., Niethammer C., Gracia J., Bosilca G. Quo Vadis MPI RMA? Towards a More Efficient Use of MPI One-Sided Communication // arXiv: 2111.08142. 2021.
6. Mark M., Shavit N. Concurrent Data Structures. Chapman and Hall/CRC Press, 2004. 32 p.
7. Shavit N. Data structures in the multicore age // Communications of the ACM. 2011. V. 54. P. 76–84.
8. Пазников А.А. Оптимизация делегирования выполнения критических секций на выделенных процессорных ядрах // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2017. № 38. С. 52–58.

9. Аненков А.Д., Пазников А.А. Алгоритмы оптимизации масштабируемого потокобезопасного пула на основе распределяющих деревьев для многоядерных вычислительных систем // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2017. № 39. С. 73–84.
10. Calciu I., Gottschlich J., Herlihy M. Using elimination and delegation to implement a scalable NUMA-friendly stack // 5th {USENIX} Workshop on Hot Topics in Parallelism (HotPar 13). 2013. P. 1–7.
11. Brock B., Buluc A. BCL: A cross-platform distributed data structures library // ICPP. 2019. P. 1–10.
12. Schuchart J., Bouteiller A., Bosilca G. Using MPI-3 RMA for active messages // ExaMPI. 2019. P. 47–56.
13. Diep T.D., Furlinger K. Nonblocking data structures for distributed-memory machines: stacks as an example // 2021 29th Euromicro Int. Conf on Parallel, Distributed and Network-Based Processing (PDP). IEEE, 2021. С. 9–17.

References

1. Liu, J., Wu, J. & Panda, D.K. (2004) High performance RDMA-based MPI implementation over InfiniBand. *International Journal of Parallel Programming*. 32. pp. 167–198.
2. Hoefler, T., Dinan, J., Thakur, R., Barrett, B., Balaji, P., Gropp, W. & Underwood, K. (2015) Remote memory access programming in MPI-3. *ACM Transactions on Parallel Computing*. 2(2). p. 9.
3. Gerstenberger, R., Besta, M. & Hoefler, T. (2004) Enabling Highly-Scalable Remote Memory Access Programming with MPI-3 One Sided. *Scientific Programming*. 2(2). pp. 75–91.
4. Herlihy, M. & Shavit, N. (2012) *The art of multiprocessor programming*. Amsterdam et al.: Morgan Kaufmann.
5. Schuchart, J., Niethammer C., Gracia J. & Bosilca G. (2021) Quo Vadis MPI RMA? Towards a More Efficient Use of MPI One-Sided Communication. *arXiv*. 2111.08142.
6. Mark, M. & Shavit, N. (2004) *Concurrent Data Structures*. Chapman and Hall/CRC Press.
7. Shavit, N. (2001) Data structures in the multicore age. *Communications of the ACM*. 54. pp. 76–84.
8. Pазников А.А. (2017) Optimization method of remote core locking. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 38. pp. 52–58. DOI: 10.17223/19988605/38/8
9. Anenkov, A.D. & Paznikov, A.A. (2017) Algorithms of optimization of scalable thread-safe pool based on diffracting trees for multi-core computing systems. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 39. pp. 73–84. DOI: 10.17223/19988605/39/10
10. Calciu, I., Gottschlich, J. & Herlihy, M. (2013) Using elimination and delegation to implement a scalable NUMA-friendly stack. *5th {USENIX} Workshop on Hot Topics in Parallelism (HotPar 13)*. pp. 1–7.
11. Brock, B. & Buluc, A. (2019) BCL A cross-platform distributed data structures library. *ICPP*. pp. 1–10.
12. Schuchart, J., Bouteiller, A. & Bosilca, G. (2019) Using MPI-3 RMA for active messages: *ExaMPI*. pp. 47–56.
13. Diep, T.D. & Furlinger, K. (2021) Nonblocking data structures for distributed-memory machines: stacks as an example. *29th Euromicro Int. Conf. on Parallel, Distributed and Network-Based Processing (PDP)*. IEEE. pp. 9–17.

Информация об авторах:

Бураченко Александр Викторович – магистрант Санкт-Петербургского государственного электротехнического университета «ЛЭТИ» (Санкт-Петербург, Россия). E-mail: ss47305@gmail.com

Пазников Алексей Александрович – кандидат технических наук, доцент Санкт-Петербургского государственного электротехнического университета «ЛЭТИ» (Санкт-Петербург, Россия). E-mail: apaznikov@gmail.com

Державин Денис Павлович – магистрант Санкт-Петербургского государственного электротехнического университета «ЛЭТИ» (Санкт-Петербург, Россия). E-mail: derzhavinden002@gmail.com

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

Information about the authors:

Burachenko Alexander V. (Saint Petersburg Electrotechnical University “LETI”, St. Petersburg, Russian Federation). E-mail: ss47305@gmail.com

Paznikov Alexei A. (Candidate of Technical Sciences, Associate Professor, Saint Petersburg Electrotechnical University “LETI”, St. Petersburg, Russian Federation). E-mail: apaznikov@gmail.com

Derzhavin Denis P. (Saint Petersburg Electrotechnical University “LETI”, St. Petersburg, Russian Federation). E-mail: derzhavinden002@gmail.com

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

Поступила в редакцию 19.08.2022; принята к публикации 01.03.2023

Received 19.08.2022; accepted for publication 01.03.2023

Научная статья
УДК 004.94
doi: 10.17223/19988605/62/3

Матрично-топологический анализ компонентных цепей

Вячеслав Михайлович Дмитриев¹, Тарас Викторович Ганджа²,
Максим Игоревич Кочергин³

^{1, 2, 3} Томский государственный университет систем управления и радиоэлектроники, Томск, Россия

¹ dmitriewvm@gmail.com

² gtv@main.tusur.ru

³ maksim.i.kochergin@tusur.ru

Аннотация. Предложен теоретико-множественный формализм метода компонентных цепей, позволяющий строить матрично-топологическое описание объекта моделирования. Введены основные понятия, определяющие компонент, его модель, а также состав и содержание связей компонента с системой. Определена роль матричных моделей в методах компонентного моделирования. Даны канонические формы для моделей компонентов и отражена их иерархия. Введенные формальные понятия компонента и компонентной цепи, а также законов сохранения позволяют построить ее матрично-топологическую модель.

Ключевые слова: математическое моделирование; компонентные цепи; компонент; топология цепи.

Для цитирования: Дмитриев В.М., Ганджа Т.В., Кочергин М.И. Матрично-топологический анализ компонентных цепей // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2023. № 62. С. 25–35. doi: 10.17223/19988605/62/3

Original article
doi: 10.17223/19988605/62/3

Matrix topological analysis of component circuits

Vyacheslav M. Dmitriev¹, Taras V. Gandzha², Maxim I. Kochergin³

^{1, 2, 3} Tomsk State University of Control Systems and Radioelectronics, Tomsk, Russian Federation

¹ dmitriewvm@gmail.com

² gtv@main.tusur.ru

³ maksim.i.kochergin@tusur.ru

Abstract. A set theoretic formalism of the component circuits method is proposed, which allows to construct a matrix-topological description of the simulation object. The paper introduces basic concepts defining the component, its model, as well as the composition and content of the component's links with the system. The role of matrix models in the methods of component modeling is determined. Canonical forms for component models are given and their hierarchy is reflected. Introduced formal concepts of a component and a component circuits, as well as conservation laws allow building its matrix topological model.

Keywords: mathematical modeling; component circuits; component; circuit topology.

For citation: Dmitriev, V.M., Gandzha, T.V., Kochergin, M.I. (2023) Matrix topological analysis of component circuits. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 62. pp. 25–35. doi: 10.17223/19988605/62/3

Введение

За последние десятилетия создано довольно много систем имитационного и компьютерного моделирования, перечень которых приведен на официальном сайте Национального общества имитационного моделирования [1]. Рассмотрим некоторые из них.

Система Rand Model Designer [2] (ранее MvStudium [3]) позволяет строить модели многокомпонентных гибридных (дискретно-непрерывных) и непрерывных систем и проводить на их основе активные вычислительные эксперименты. Математические модели структурных блоков задаются в символьном виде. Система AnyLogic поддерживает такие подходы к разработке моделей, как процессно-ориентированный (дискретно-событийный), агентный и системно-динамический, а также разные их комбинации. Указанная среда применяется довольно широко в бизнес-моделировании компаний и в образовании [4]. Simulink [5] – это система, входящая в математический пакет Matlab, осуществляющая блочное моделирование с ориентированными и неориентированными связями у блоков. Simulink имеет обширную библиотеку моделей блоков из различных областей и программное обеспечение для анализа системно-событийной динамики, предназначенное для построения имитационных моделей, анализа и оптимизации динамических систем [6]. Для моделирования схем аналоговой, цифровой и силовой электроники может применяться среда Multisim [7]. Программный комплекс SimInTech [8] (ранее MBТУ – Моделирование в технических устройствах [9]) имеет акцент на моделирование объектов, представленных структурными схемами САУ. Он используется в качестве инструмента для проектирования систем следящих приводов, автоматического управления и роботов-манипуляторов, а также ядерных и тепловых энергетических установок и пр. Одна из ранних отечественных систем – среда многоуровневого моделирования СМ МАРС [10], которая предназначена для автоматизации натуральных и вычислительных экспериментов над сложными управляемыми системами. В большинстве своем названные системы базируются на подходах компонентного моделирования [11].

В данной работе предлагается полный формализм метода компонентных цепей (МКЦ), лежащий в основе СМ МАРС, который ранее не рассматривался. Он приведен для цепей с однородными потоками в связях и установленного заранее класса модели объекта.

1. Математическая модель фирмы

Для разработки общих алгоритмов в МКЦ необходимо формализовать основные понятия к вводимым определениям, согласуя их с требованиями конструктивности, универсальности и программной ориентированности.

Пусть компонент M с n связями (выводами) B_0, \dots, B_{n-1} входит в систему, описываемую Q физическими переменными V^1, V^2, \dots, V^Q . Каждой связи $B_i, i = 0, 1, \dots, n - 1$, соответствует индексное множество $I^i = \{I_1^i, I_2^i, \dots, I_{q_i}^i\}, q_i \leq Q$, задающее типы переменных, действующее на связи B_i . Моделью компонента M будем называть систему уравнений (соотношений), связывающих переменные, определяемые индексным множеством $\bigcup_{i=0}^{n-1} I^i$, то есть переменные $V[I^0], V[I^1], \dots, V[I^m]$.

2. Матричные модели в методах компонентного моделирования

Систему уравнений, описывающую поведение компонента, будем называть моделью компонент, или M -моделью. Эта модель обычно расписывается в виде ряда формул и подлежит либо ручной обработке, либо обработке с применением специальных генераторов модели [12]. Будем расписывать M -модели в форме:

$$\text{ММ [}<\text{название модели}>]: \{F_1, F_2, \dots, F_Q\}, \quad (1)$$

где $\{F_1, F_2, \dots, F_Q\}$ – уравнения модели.

При реализации алгоритмов численного анализа, когда модель M получается в результате выполнения некоторого вычислительного процесса, необходимо ввести понятие вычислительной модели (ВМ), которая задается в форме:

$$\text{ВМ [}<\text{название модели}>]: \{k, I^0, I^1, \dots, I^{n-1}, B_1, B_2, \dots, B_k\}, \quad (2)$$

где k – число связей компонента M ; I^0, I^1, \dots, I^{n-1} – индексы переменных, действующие на связях компонента; B_1, B_2, \dots, B_k – вычислительные блоки, соответствующие уравнениям F_1, F_2, \dots, F_Q . Кроме внешних параметров (переменных) компонент характеризуется и внутренними параметрами (конструктивные, технологические, геометрические и др.).

Величина k , задающая число уравнений в M -модели (1) или число вычислительных блоков в B -модели (2) в общем случае не определяется однозначно. Она зависит от типа уравнений и числа переменных в связях.

Для построения вычислительных алгоритмов анализа цепей необходимо зафиксировать основные канонические формы, которые может принимать M -модель компонента.

3. Канонические формы для моделей компонентов и их иерархия

Модель компонента M с внешними параметрами $V_\alpha, \alpha \in \bigcup_{i=0}^{n-1} I^i$ есть система уравнений, которая может принимать следующие канонические формы.

А. Система линейных алгебраических уравнений

$$\sum_{\alpha} a_{\alpha} V_{\alpha} = C_V, \quad (3)$$

где a_{α} и C_V – переменные коэффициенты, определяемые внутренней структурой компонентов.

Б. Система линейных алгебраических уравнений

$$\sum_{\alpha} b_{\alpha} V_{\alpha} = C_V, \quad (4)$$

где b_{α} – постоянные коэффициенты; C_V – переменные правые части, определяемые входными воздействиями.

В. Система уравнений

$$V_{\alpha} = C_V, \quad (5)$$

где C_V – переменные значения V_{α} .

Г. Система уравнений

$$\sum_{\alpha} b_{\alpha} V_{\alpha} = C_C, \quad (6)$$

где b_{α} и C_C – постоянные коэффициенты.

Д. Система уравнений

$$V_{\alpha} = C_C, \quad (7)$$

где C_C – постоянные значения V_{α} .

Линейные модели (3)–(7) задаются непосредственно в качестве моделей компонентов с привлечением аналитических преобразований [13]. Коэффициенты (3)–(7) можно также получить в результате работы Б-блоков у В-моделей.

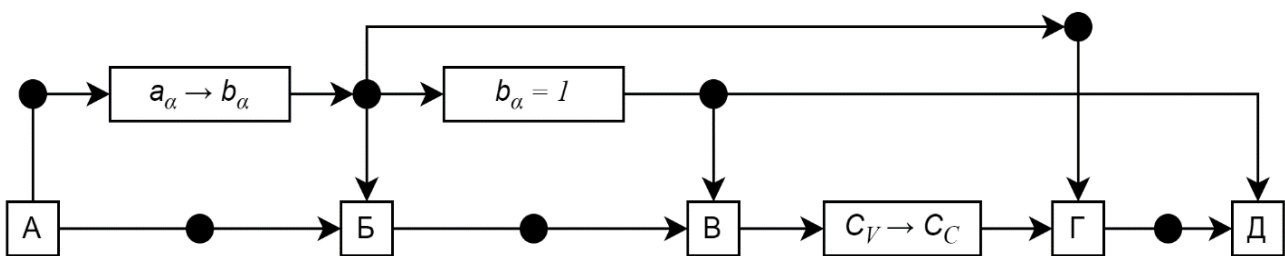


Рис. 1. Иерархическая структура моделей
Fig. 1. Hierarchical structure of models

Очевидно, что модели А, Б, В, Г, Д связаны в иерархическую структуру (рис. 1).

4. Примеры канонических форм моделей

Приведем примеры канонических форм моделей компонентов [14]:

– динамическая нелинейная модель диода (А-форма)

$$C_b \left(\frac{du_1}{dt} - \frac{du_2}{dt} \right) = i_1 - Si \left(e^{\frac{u_1 - u_2}{V_i}} - 1 \right) - \frac{1}{Rb} (u_1 - u_2 - Rmi_1);$$

– модель емкости (Б-форма)

$$\frac{d(u_1 - u_2)}{dt} = \frac{1}{C} i_1;$$

– модель синусоидального источника (В-форма)

$$u_1 - u_2 = A \sin(\omega t + \varphi);$$

– модель резистора (Г-форма)

$$u_1 - u_2 - Ri_1 = 0;$$

– модель постоянного источника (Д-форма)

$$u_1 - u_2 = E.$$

5. Формальное понятие компонента и компонентной цепи

На основании понятия компонента и его модели вводится понятие компонентной цепи (КЦ). Компонентной цепью будем называть произвольную совокупность компонентов, связи которых, именуемые ветвями цепи B_0, \dots, B_{n-1} , объединены в точках, называемых узлами N_0, \dots, N_{n-1} . Каждому узлу цепи соответствуют законы сохранения энергии.

Очевидно, что в зависимости от природы цепи узловой закон сохранения может принимать самые различные формы, однако для построения вычислительных алгоритмов форма законов сохранения должна быть зафиксирована.

Решением модели цепи, состоящей из компонентов M_1, \dots, M_i , будем называть совокупность переменных, действующих на ветвях цепи (8)

$$V = \{ \bigcup (\bigcup V_{M_j}^{I_{M_j}^i}) \}, \quad (8)$$

где $I_{M_j}^i$ – индексы переменных, действующих на i -й связи компонента M_j , В решении V переменные $V[\alpha]$ и $V[\beta]$ будем называть одноименными, если $\alpha = \beta$.

Пусть узел N_i , ($i = 1, \dots, k-1$) объединяет ветви, на которых действуют переменные $V_1[\alpha_1], V_2[\alpha_1], \dots, V_{i_1}[\alpha_1], V_1[\alpha_2], V_2[\alpha_2], \dots, V_{i_2}[\alpha_2], \dots, V_1[\alpha_q], V_2[\alpha_q], \dots, V_{i_q}[\alpha_q]$.

Полагаем, что цепи допускают закон сохранения, заданный в виде равенства одноименных переменных (9) или (10):

$$V_1[\alpha_j] = V_2[\alpha_j] = \dots = V_{i_j}[\alpha_j], \quad j = 1, \dots, q, \quad (9)$$

$$\sum_{k=1}^{i_j} a_k^j V_k[\alpha_j] = 0, \quad j = 1, \dots, q. \quad (10)$$

Можно потребовать, чтобы закон сохранения принял вид:

$$\sum_{k=1}^{i_j} V_k[\alpha_j] = 0, \quad j = 1, \dots, q, \quad (11)$$

т.е. $a_k^j = 1$. Однако для некоторых классов цепей представление вида (11) недопустимо, так как связи компонентов (ветви цепи) могут иметь естественную или общепринятую классификацию.

Очевидно, что при узловых законах вида (9) переменные ветвей цепи можно заменить на узловые переменные. Линейная форма (10) соответствует закону сохранения для переменной $V[\alpha_j]$.

Для разработки методик моделирования и анализа необходимо формальное описание компонентной цепи. Полагаем, что компонентам цепи присвоены номера $1, \dots, l$ (компоненты имеют идентификаторы M_1, M_2, \dots, M_l). Связи компонента $M_i, i = 1, \dots, l$, нумеруются в произвольном и затем фиксируемом порядке. Существует три способа задания компонентной цепи: списковый, графический и аналитический. Рассмотрим здесь первый способ.

Будем полагать, что цепь задана, если имеется список, строка которого имеет вид:

$$M_i, \text{sign}_i^j B_i^j, N_i^j, \quad (12)$$

где $i=1, \dots, n_j, j=1, \dots, l$.

Такое представление именуем далее списковым, или С-цепью.

Знак sign_i^j , где $i=1, \dots, n_j, j=1, \dots, l$, задает ориентацию ветви B_i^j . Очевидно, что при формировании узлового закона сохранения (10) a_k^j соответствуют закону сохранения при определенной ветви. В силу однородности (10) способ установления знака безразличен.

Очевидно, что в (12) $B_i^j, N_i^j, j=1, \dots, l; i=1, \dots, n_j$ задают соответственно ветвь, образованную i -й связью компонента M_j , и узел, в который ведет эта ветвь. Перечисление ветвей в (12) осуществляется в соответствии с фиксированной нумерацией связей компонентов.

Для интерпретации результатов анализа цепей, представление решения V в виде (8) обычно оказывается неудобным. Воспользуемся тем, что ветви цепи имеют сплошную нумерацию B_1, B_2, \dots, B_m , и сгруппируем одноименные переменные:

$$V = V_1[1], \dots, V_m[1], V_1[q]. \quad (13)$$

Очевидно, что цепь является однородной, если выполняется условие

$$I_1^1 = \dots = I_1^{n_1} = \dots = I_l^1 = \dots = I_l^{n_l}. \quad (14)$$

Итак, цепь называется однородной, если на каждой ее ветви действует один и тот же набор физических переменных. В противном случае цепь называется неоднородной. Решение неоднородной цепи представляется в виде (8) либо в виде (13). При этом ряд элементов в (13) в таком случае будет отсутствовать.

6. Структура матрично-топологической модели цепи

Моделью цепи будем называть совокупность моделей ее компонентов M_1, M_2, \dots, M_l , внешним параметрам которых поставлены в соответствие элементы решения V согласно С-цепи (12) и узловых законов сохранения (9) и (10).

Для практического построения модели цепи необходимо задать ее класс. Полагаем, что класс цепи задан, если каждому типу физических переменных поставлен в соответствие узловой закон сохранения, т.е. задано множество пар $\Pi_i = (V^i, P^i), i=1, \dots, q, P_i \in \{P_1, P_2\}$. Через P_1 обозначен закон равенства одноименных переменных (9), через P_2 – закон сохранения (10). Очевидно, что всякая цепь может входить в качестве компонента или подцепи в другую цепь.

Рассмотрим матричное представление канонических форм после применения к ним процедур алгебраизации и линеаризации, предназначенное для алгоритмов численного анализа. Матрицей A канонической формы модели цепи будем называть матрицу

$$A = \begin{vmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,q} & a_{1,q+1} \\ a_{2,1} & a_{2,2} & & a_{2,q} & a_{2,q+1} \\ \vdots & & \ddots & & \vdots \\ a_{p,1} & a_{p,1} & & a_{p,q} & a_{p,q+1} \end{vmatrix}. \quad (15)$$

Полагаем, что канонические формы А, Б, В, Г, Д задаются матрицей $[a_{ij}], i = 1, \dots, q + 1; j = 1, \dots, p$ размерности $(q + 1) \times p$. Столбец $(q + 1)$ соответствует вектору правых частей со своими знаками. Другая форма уравнения (15) имеет вид:

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,q} \\ a_{2,1} & a_{2,2} & & a_{2,q} \\ \vdots & & \ddots & \\ a_{p,1} & a_{p,1} & & a_{p,q} \end{pmatrix} \cdot V_{\text{вн}} = \begin{pmatrix} a_{1,q+1} \\ a_{2,q+1} \\ \vdots \\ a_{p,q+1} \end{pmatrix}, \quad (16)$$

где $V_{\text{вн}}$ – вектор переменных цепи V^α , $\alpha \in \bigcup I^i (i=0, \dots, n-1)$ моделируемого компонента. Вектор $V_{\text{вн}}$ получается из V^α упорядочением в соответствии с фиксированной нумерацией связей и естественным порядком индексов переменных.

Представим (16) в виде:

$$A_1 \cdot V_{\text{вн}} = A_2. \quad (17)$$

Очевидно, что $A = [A_1, A_2]$. К размерности матрицы A предъявляются требования существования и единственности решения цепи, включающих в себя моделируемые компоненты. Рассмотрим основные структурные блоки модели цепи. Узловые законы связей P_1 и P_2 также представляются в матричной форме:

$$PV = 0; \quad P = [P_1 \quad P_2]^t. \quad (18)$$

Матрица P_1 соответствует узловому закону равенства одноименных переменных P_1 , матрица P_2 – закону сохранения P_2 . Элементы матрицы P , очевидно, удовлетворяют условию $P_i^j \in \{0, 1, -1\}$. Для построения (18) формула (9) переписывается в виде:

$$V_1[\alpha_j] - V_2[\alpha_j] = 0, \dots, V_{i,j-1}[\alpha_j] = V_{i,j}[\alpha_j] = 0 \quad j = 1, \dots, q. \quad (19)$$

Матрица P_1 расписывается из коэффициентов (19) для каждого узла цепи. Исходной информацией для построения P_1 служит С-цепь и $\bigcup_{i=1}^q \Pi_i$. При практической реализации метода компонентных цепей матрица P_1 не строится. Совокупность переменных с $\Pi_i = (V^i, P_i)$, относящихся к ветвям, которые ведут в один узел, заменяется согласно системе равенств (9) узловой переменной. В этом случае решение V представляется в виде:

$$V = \bigcup_{i=1}^q V_i, \begin{cases} \Pi_i = (V^i, P_i) & \text{если } \Pi_i = (V^i, P_i), \\ V_i = (V_0^i, \dots, V_{k-1}^i) & \text{если } \Pi_i = (V^i, P_2). \end{cases} \quad (20)$$

Матрица P_2 строится на основе (10). Для каждого из узлов расписывается строка P_2 путем циклического просмотра С-цепи. Ненулевые коэффициенты P_2 рассылаются следующим образом:

$$P_{i,j} = \begin{cases} 1, & \text{sign}_p^q = +1, \quad q = 1, \dots, l; \quad p = 1, \dots, n_i, \\ -1, & \text{sign}_p^q = -1, \quad i = N_q^p, \quad j = B_q^{jp}. \end{cases} \quad (21)$$

С-цепь просматривается для каждой переменной V_i , для которой $\Pi_i = (V^i, P_2)$.

В целом модель компонентной цепи включает в себя матрицы P_1 , P_2 и A -матрицы для каждого компонента цепи.

Матричное уравнение (22) в котором Φ -матрица включает в себя матрицы P_1 и P_2 , а также матрицу A_1 (17) для каждого компонента цепи, будем называть основным уравнением метода компонентных цепей

$$\Phi \cdot V = W. \quad (22)$$

Элементы матрицы A_1 рассылаются в Φ -матрицу в соответствии с С-цепью (12) и установленным порядком переменных в V (13). Вектор W содержит нули в позициях, соответствующих строкам P_1 и P_2 . Элементы W , соответствующие матрицам A_1 , равны элементам матриц A_2 (17). Уравнение (23) и является моделью компонентной цепи с заданными моделями компонентов и узловыми законами сохранения

$$\begin{pmatrix} A_1 \\ P_1 \quad P_2 \end{pmatrix} \cdot V = W. \quad (23)$$

7. Графическая интерпретация компонентной цепи

Будем полагать, что компонентная цепь (КЦ) задана, если имеется список следующего вида:

$$\begin{aligned}
 1: & K_1, \text{sign}_1^1 B_1^1, \dots, \text{sign}_{n_1}^1 B_{n_1}^1, N_1^1, \dots, N_{n_1}^1, \\
 & \vdots \\
 l: & K_l, \text{sign}_1^l B_1^l, \dots, \text{sign}_{n_l}^l B_{n_l}^l, N_l^l, \dots, N_{n_l}^l.
 \end{aligned}
 \tag{24}$$

Знак sign_i^j $j=1, \dots, l; i=1, \dots, n_j$ задает ориентацию соответствующей ветви B_i^j . Полагаем, что связь ведет в компонент, если $\text{sign}_i^j = +1$ (положительное направление ветви), и от компонента, если $\text{sign}_i^j = -1$ (отрицательное направление) (рис. 2).

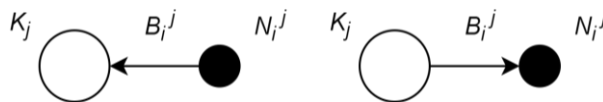


Рис. 2. Варианты ориентации связей
Fig. 2. Link orientation options

Перейдем к графической интерпретации понятия КЦ в виде структурного (бихроматического) графа.

G-графом будем называть граф $G = \{K, N, B, Q\}$, где

- 1) $K = \{K_1, K_2, \dots, K_{l-1}\}$ – множество вершин компонентов графа;
- 2) $N = \{N_1, N_2, \dots, N_{k-1}\}$ – множество вершин – узлов графа;
- 3) $B = \{B_1, B_2, \dots, B_{m-1}\}$ – множество, элементы которого образованы связями компонентов.

Функция Q ставит в соответствие каждой связи $B_j \in B$ упорядоченную пару вершин $(K_i, N_i) \neq (N_i, K_i)$. Ориентированные таким образом связи образуют множество дуг G-графа. Он, в свою очередь, состоит из подграфов. Обозначим их как g – графы компонента. Тогда граф $G = \{\bigcup g_j, N, Q'\}$ – ориентированный граф, где $\bigcup g_j$ – множество, элементы которого образованы объединением g -графов; $N = \{N_1, N_2, \dots, N_{k-1}\}$ – множество вершин – узлов графа; Q' – функция, ставящая в соответствие каждому g -графу упорядоченный набор вершин-узлов.

Для изображения G-графа введем следующие графические обозначения: K и N – множества белых (компоненты) и черных (узлы) вершин. На рис. 3 приведена простая цепь, содержащая компоненты K_1, K_2, K_3, K_4 с узлами N_0, \dots, N_3 с различным числом связей (см. рис. 3, a) и соответствующий цепи G-граф (см. рис. 3, b).

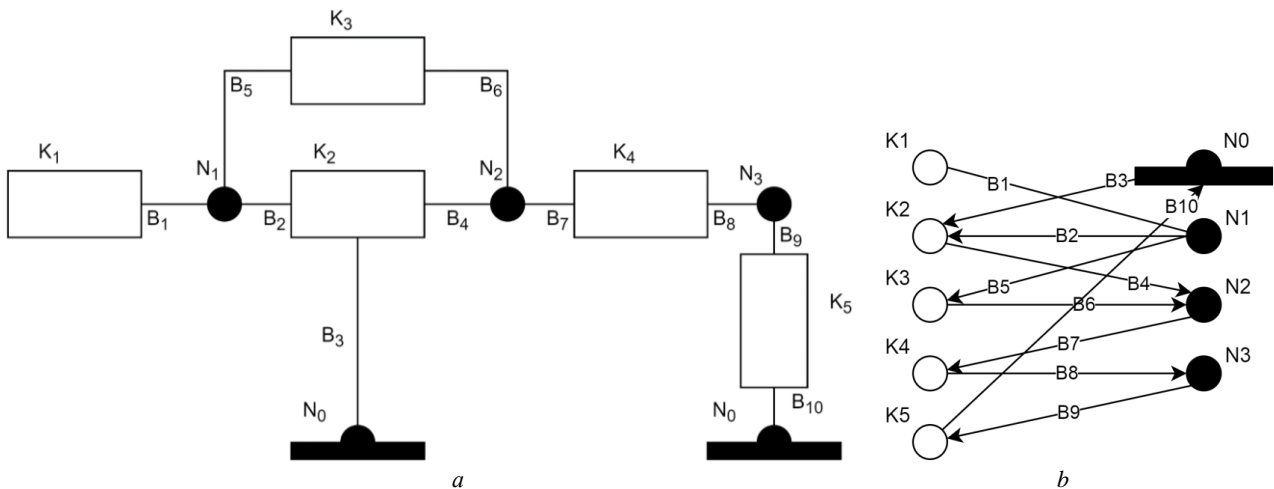


Рис. 3. Компонентная цепь: a – непосредственно компонентная цепь, b – бихроматический граф цепи
Fig. 3. Component chain. a) is directly component chain, b) is the bichromatic graph of the chain

Граф $G = \{K_n, N, B, Q\}$ – структурный (см. рис. 3) и имеет следующие свойства:

- 1) не имеет петель, т.е. две связи одного компонента не присоединяются к одному и тому же узлу;
- 2) все дуги компонента инцидентны аналогичному множеству вершин-узлов и не соединены более чем с одним узлом;

3) множество дуг, образованное только парами вершин-узлов, – пустое множество:

$$B_1 = \{B_1^1, B_2^1, \dots, B_{m-1}^1\} = \emptyset, \text{ где } B_i^1 = (N_j, N_k) \text{ для любых } j \text{ и } k.$$

4) множество дуг, образованное только парами вершин-компонентов, также пустое множество:

$$B_2 = \{B_1^2, B_2^2, \dots, B_{m-1}^2\} = \emptyset, \text{ где } B_i^2 = (K_j, K_k) \text{ для любых } j \text{ и } k.$$

5) множество вершин-узлов, не имеющее инцидентных ветвей, – пустое множество.

В соответствии с перечисленными свойствами G -граф является бихроматическим [15], поскольку элементы любой пары его соседних вершин имеют разный цвет (являются элементами различных множеств K и N). Рассмотрим основные свойства и соотношения в G -графе. Отношение инцидентности здесь записывается следующим образом: $B_i \eta^+ K_j$ или $B_i \eta^+ N_j$; это означает, что вершина $K_j \in K$ или $N_j \in N$ является началом дуги B_j . Соответственно запись $B_i \eta^- K_j$ или $B_i \eta^- N_j$ означает, что вершина $K_j \in K$ или $N_j \in N$ является концом дуги B_j .

Введем матрицы инцидентности $G = (G_1, G_2)$ и матрицу смежности S размерности $|K| \times |N|$.

Матрица смежности будет иметь вид:

$$S = \begin{matrix} & N_1 & N_2 & N_3 & N_4 & N_5 & \\ & \left| \begin{array}{ccccc} 0 & 1 & 0 & 0 & \\ 1 & 1 & 1 & 0 & \\ 0 & 1 & 1 & 0 & \\ 0 & 0 & 1 & 1 & \\ 1 & 0 & 0 & 1 & \end{array} \right| & \begin{array}{l} K_1 \\ K_2 \\ K_3 \\ K_4 \\ K_5 \end{array} \end{matrix}.$$

Расширенная матрица смежности S_p строится в координатах $|K + N| \times |N + K|$ и имеет вид:

$$S_p = \begin{vmatrix} S & 0 \\ 0 & S^T \end{vmatrix},$$

где S^T – транспонированная матрица смежности. Она имеет две нулевые подматрицы.

Если элементы множеств K, B и N заданы с помощью упорядоченного списка, то отношение инцидентности η^+, η^- можно описать матрицей инцидентности $G = g_{i,j}$. Размерность этой матрицы $|B| \times (|K| + |N|)$, где $|B|, |K|$ и $|N|$ – размерности соответствующих множеств. Матрицу G можно представить в виде в виде двух матриц $G = (G_1, G_2)$, где G_1 имеет размерность $|B| \times |K|$, а G_2 – размерность $|B| \times |N|$. Матрицы G_1, G_2 для G -матрицы (см. рис. 3) имеют следующий вид:

$$G_1 = \begin{matrix} & K_1 & K_2 & K_3 & K_4 & K_5 & \\ & \left| \begin{array}{ccccc} 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right| & \begin{array}{l} B_1 \\ B_2 \\ B_3 \\ B_4 \\ B_5 \\ B_6 \\ B_7 \\ B_8 \\ B_9 \\ B_{10} \end{array} \end{matrix}, \quad G_2 = \begin{matrix} & N_0 & N_1 & N_2 & N_3 & \\ & \left| \begin{array}{cccc} 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \end{array} \right| & \begin{array}{l} B_1 \\ B_2 \\ B_3 \\ B_4 \\ B_5 \\ B_6 \\ B_7 \\ B_8 \\ B_9 \\ B_{10} \end{array} \end{matrix}.$$

Оперирова приведенными матрицами смежности и инцидентности, интерпретируем свойства графа цепи (см. рис. 3) для определения возможных топологических ошибок при формировании компонентной цепи.

Введем основные топологические обозначения: $\bigcup B_i$ – совокупность ветвей цепи C_k ; $\bigcup N_j$ – совокупность узлов цепи C_k ; $\bigcup_{j \in K_j} N_j$ – совокупность узлов компонента; $\bigcup_{j \in K_j} B_j$ – совокупность связей компонента.

В этих обозначениях можно представить различные типы топологических ошибок (рис. 4):

- а) висячий узел $\bigcup_{i \in N_j} B_i = 1$ (ранг узла; см. рис. 4, а);
- б) неприсоединенный компонент $\bigcup_{i \in N_j} B_i = 0$ (см. рис. 4, б);
- в) узел без ветвей $\bigcup_{i \in N_j} B_i = 0$ (см. рис. 4, в);
- г) петля из связей $(B_i^k, B_j^k) \in N_j$ (см. рис. 4, д).

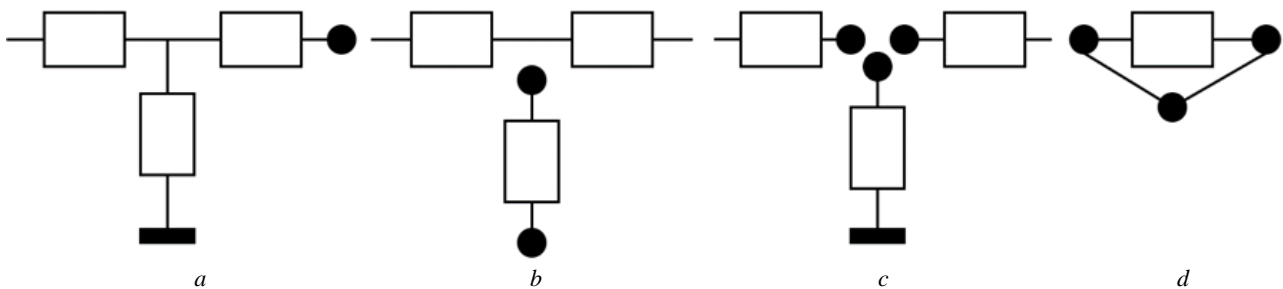


Рис. 4 Типы топологических ошибок
Fig. 4. Types of topological errors

Эти ошибки соответствуют следующим вариантам, применительно к матрицам:

Вариант 1. В матрице G_2 появляется столбец с одним ненулевым элементом.

Вариант 2. В матрице S (смежности) появляется нулевая строка.

Вариант 3. В матрице G_2 появляется столбец только с нулевыми элементами.

Вариант 4. В матрице G_2 появляется столбец с двумя ненулевыми элементами для ветвей одного и того же компонента.

Для данных вариантов могут быть составлены соответствующие алгоритмы контроля топологических ошибок в компонентной цепи.

Заключение

Предложенный в работе теоретико-множественный формализм метода компонентных цепей позволяет строить матрично-топологическое описание объекта моделирования. Введенные понятия определяют компонент, его модель, а также состав и содержание связей компонента с системой. Даны канонические формы для моделей компонентов (модель емкости – Б-форма, модель синусоидального источника – В-форма, модель резистора – Г-форма, модель постоянного источника – Д-форма), позволяющие строить модели систем различной природы, и отражена их иерархия. Введенные формальные понятия компонента и компонентной цепи, а также законов сохранения позволяют построить матрично-топологическую модель компонентной цепи. Приведенная графическая интерпретация компонентной цепи позволяет установить однозначное соответствие между непосредственно компонентной цепью (как моделью некоторой системы) и ее бихроматическим графом, состоящим из узлов и компонентов, соединенных связями. Приведенные четыре случая топологических ошибок в компонентной цепи позволяют выявить их автоматически посредством анализа матриц смежности и инцидентности графа цепи.

Список источников

1. Программные продукты. URL: <http://simulation.su/static/ru-soft.html> (дата обращения: 09.09.2022).
2. Isakov A.A., Kolesov Y.B., Senichenkov Y.B. A new tool for visual modeling-Rand Model Designer 7 // IFAC-Papers On Line. 2015. V. 28 (1). P. 661–662.
3. Senichenkov Yu.B., Kolesov Yu.B. Physical modeling in MvStudium // Differential Equations and Control Processes. 2011. V. 2. P. 34–40.
4. Yakimov I.M., Trusfus M.V., Mokshin V.V., Kirpichnikov A.P. AnyLogic, extendsim and simulink overview comparison of structural and simulation modelling systems // RPC 2018 Proc. of the 3rd Russian-Pacific Conference on Computer Technology and Applications. 2018. Art. 8482152.
5. Deng H., Zhu L., Wang J., He C., Liu J. Hydraulic system simulation of hybrid hydraulic module based on Simulink // Proc. of SPIE – The International Society for Optical Engineering. 2022. V. 12259. Art. 122593M. 9 p.
6. Дьяконов В.П. Simulink 5/6/7. М. : Изд-во ДМК, 2008. 781 с.
7. Peng Z. Comparison and discussion of the functions of logisim and multisim // Journal of Physics: Conference Series. 2021. V. 2030 (1). Art. 012055.
8. Колотыркин И.П., Петухов В.Н. Создание моделей сложных технических систем в среде SimInTech // Компьютерное моделирование в железнодорожном транспорте: динамика, прочность, износ : ст. тез. конф. Брянск : Брянск. гос. техн. ун-т, 2018. С. 45–47.
9. Шабаетв Е.А. Анализ и синтез систем автоматического регулирования на основе программного комплекса «МВТУ» // Научная молодежь: пути и перспективы развития агроинженерной науки: студенческий сб. науч. тр. Зерноград : Азово-Черноморская инж. акад., 2005. С. 13–24.
10. Дмитриев В.М., Шутенков А.В., Зайченко Т.Н., Ганджа Т.В. MARS – среда моделирования технических устройств и систем. Томск : В-спектр, 2011. 277 с.
11. Бенькович Е.С., Колесов Ю.Б., Сениченков Ю.Б. Практическое моделирование динамических систем. СПб. : БХВ – Петербург, 2002. 707 с.
12. Арайс Е.А., Дмитриев В.М. Моделирование неоднородных цепей и систем на ЭВМ. М. : Радио и связь, 1982. 159 с.
13. Арайс Е.А., Сибиряков Г.В. Авто-Аналитик. Новосибирск : Изд-во Высш. и сред. спец. образования РСФСР. 1973. 284 с.
14. Дмитриев В.М., Зайченко Т.Н., Гарганеев А.Г., Шурыгин Ю.А. Автоматизация функционального проектирования электромеханических систем и устройств преобразовательной техники. Томск : Изд-во Том. ун-та, 2000. 292 с.
15. Оре О. Графы и их применение. М. : Мир, 1965. 174 с.

References

1. National Society for Simulation. (n.d.) *Programmnye produkty* [Software]. [Online] Available from: <http://simulation.su/static/ru-soft.html> (Accessed: 9th September 2022).
2. Isakov A.A., Kolesov, Y.B. & Senichenkov, Y.B. (2015) A new tool for visual modeling-Rand Model Designer 7. *IFAC-Papers On Line*. 28(1). pp. 661–662.
3. Senichenkov, Yu.B. & Kolesov, Yu.B. (2011) Physical modeling in MvStudium. *Differential Equations and Control Processes*. 2. pp. 34–40.
4. Yakimov, I.M., Trusfus, M.V., Mokshin, V.V. & Kirpichnikov, A.P. (2018) Any logic, extendsim and simulink overview comparison of structural and simulation modelling systems. *RPC 2018 – Proceedings of the 3rd Russian-Pacific Conference on Computer Technology and Applications*. Art. 8482152. DOI: 10.1109/RPC.2018.8482152
5. Deng, H., Zhu, L., Wang, J., He, C. & Liu, J. (2022) Hydraulic system simulation of hybrid hydraulic module based on Simulink. *Proceedings of SPIE – The International Society for Optical Engineering*. 12259. Art. 122593M. DOI: 10.1117/12.2638851
6. Dyakonov, V.P. (2008) *Simulink 5/6/7*. Moscow: DMK.
7. Peng, Z. (2021) Comparison and discussion of the functions of logisim and multisim. *Journal of Physics: Conference Series*. 2030(1). Art. 012055. DOI: 10.1088/1742-6596/2030/1/012055
8. Kolotyrykin, I.P. & Petukhov, V.N. (2018) Sozdanie modeley slozhnykh tekhnicheskikh sistem v srede SimInTech [Creation of models of complex technical systems in the SimInTech environment]. *Komp'yuternoe mo-delirovanie v zheleznodorozhnom transporte: dinamika, prochnost', iznos* [Computer Modeling in Railway Transport: Dynamics, Strength, Wear]. Proc. of the Conference. Bryansk: Bryansk State Technival University. pp. 45–47.
9. Shabaev, E.A. (2005) Analiz i sintez sistem avtomaticheskogo regulirovaniya na osnove programmno kompleksa “MVTU” [Analysis and synthesis of automatic control systems based on the software complex “MiTD”]. In: *Nauchnaya molodezh': puti i perspektivy razvitiya agroinzhenernoy nauki* [Scientific youth: ways and prospects for the development of agroengineering science]. Zernograd Azovo-Chernomorsk Engineering Academy. pp. 13–24.
10. Dmitriev, V.M., Shutenkov, A.V., Zaychenko, T.N. & Gandzha, T.V. (2011) *MARS – sreda modelirovaniya tekhnicheskikh ustroystv i sistem* [MARS – environment for modeling technical devices and systems]. Tomsk: V-Spektr.
11. Benkovich, E.S., Kolesov, Yu.B. & Senichenkov, Yu.B. (2002) *Prakticheskoe modelirovanie dinamicheskikh sistem* [Practical modeling of dynamic systems]. St. Petersburg: BKhV – Peterburg.

12. Arays, E.A. & Dmitriev, V.M. (1982) *Modelirovanie neodnorodnykh tsepey i sistem na EVM* [Modeling of heterogeneous circuits and systems on a computer]. Moscow: Radio i svyaz'.
13. Arays, E.A. & Sibiryakov, G.V. (1973) *Avto-Analitik* [Auto-Analyst]. Novosibirsk: Higher and Secondary Vocational Education of the RSFSR.
14. Dmitriev, V.M., Zaychenko, T.N., Garganeev, A.G. & Shurygin, Yu.A. (2000) *Avtomatizatsiya funktsional'nogo proektirovaniya elektromekhanicheskikh sistem i ustroystv preobrazovatel'noy tekhniki* [Automation of functional design of electromechanical systems and equipment conversion devices]. Tomsk: Tomsk State University.
15. Ore, O. (1965) *Grafy i ikh primeneniye* [Graphs and Their Application]. Moscow: Mir.

Информация об авторах:

Дмитриев Вячеслав Михайлович – профессор, доктор технических наук, профессор кафедры компьютерных систем в управлении и проектировании Томского государственного университета систем управления и радиоэлектроники (Томск, Россия). E-mail: dmitriewvm@gmail.com

Ганджа Тарас Викторович – доцент, доктор технических наук, профессор кафедры компьютерных систем в управлении и проектировании Томского государственного университета систем управления и радиоэлектроники (Томск, Россия). E-mail: gtv@main.tusur.ru

Кочергин Максим Игоревич – кандидат технических наук, доцент кафедры компьютерных систем в управлении и проектировании Томского государственного университета систем управления и радиоэлектроники (Томск, Россия). E-mail: maksim.i.kochergin@tusur.ru

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

Information about the authors:

Dmitriev Vyacheslav M. (Professor, Doctor of Technical Sciences, Professor of the Department of Computer Systems in Control and Design of Tomsk State University of Control Systems and Radioelectronics, Tomsk, Russian Federation). E-mail: dmitriewvm@gmail.com

Gandzha Taras V. (Associate Professor, Doctor of Technical Sciences, Professor of the Department of Computer Systems in Control and Design, Tomsk State University of Control Systems and Radioelectronics, Tomsk, Russian Federation). E-mail: gtv@main.tusur.ru

Kochergin Maxim I. (Candidate of Technical Sciences, Associate Professor, Department of Computer Systems in Control and Design, Tomsk State University of Control Systems and Radioelectronics, Tomsk, Russian Federation). E-mail: maksim.i.kochergin@tusur.ru

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

Поступила в редакцию 22.09.2022; принята к публикации 01.03.2023

Received 22.09.2022; accepted for publication 01.03.2023

ОБРАБОТКА ИНФОРМАЦИИ

DATA PROCESSING

Научная статья

УДК 519.2

doi: 10.17223/19988605/62/4

Применение метода максимального правдоподобия для оценки параметра равномерного распределения длительности непродлевающегося мертвого времени в рекуррентном альтернирующем полусинхронном потоке событийАлександр Михайлович Горцев¹, Анна Васильевна Веткина²^{1, 2} Томский государственный университет, Томск, Россия¹ a-gortsev@mail.ru² anyavetkina@gmail.com

Аннотация. Исследуется полусинхронный поток событий, относящийся к классу дважды стохастических потоков событий. Поток функционирует в условиях непродлевающегося случайного мертвого времени, распределенного по равномерному закону на отрезке $[0, T^*]$. Рассматривается частный случай функционирования полусинхронного потока событий – альтернирующий поток, который является рекуррентным в общем и особом случаях соотношения его параметров. Производится оценивание параметра T^* равномерного распределения длительности непродлевающегося случайного мертвого времени методом максимального правдоподобия. Приводятся результаты статистических экспериментов.

Ключевые слова: альтернирующий полусинхронный поток событий; непродлевающееся случайное мертвое время; оценка параметра; метод максимального правдоподобия.

Для цитирования: Горцев А.М., Веткина А.В. Применение метода максимального правдоподобия для оценки параметра равномерного распределения длительности непродлевающегося мертвого времени в рекуррентном альтернирующем полусинхронном потоке событий // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2023. № 62. С. 36–49. doi: 10.17223/19988605/62/4

Original article

doi: 10.17223/19988605/62/4

An application of the maximum likelihood estimation for the parameter of uniform distribution of the duration of unextendable dead time in recurrent alternating semi-synchronous flowAlexander M. Gortsev¹, Anna V. Vetkina²^{1, 2} Tomsk State University, Tomsk, Russian Federation¹ a-gortsev@mail.ru² anyavetkina@gmail.com

Abstract. The paper describes alternating recurrent semi-synchronous events flow that is a common mathematical model of information flows of messages operating in telecommunication and information-computing networks, and that belongs to the class of doubly stochastic event flows. General and special cases are considered. Operation of the

flow is considered with random unextendable dead time that has uniform distribution on the interval $[0, T^*]$. Parameter T^* of the dead time is estimated using the maximum likelihood estimation. Results of statistical experiments are presented

Keywords: recurrent semi-synchronous events flow; unextendable random dead time; uniform distribution; estimation of the parameter; maximum likelihood estimation.

For citation: Gortsev, A.M., Vetkina, A.V. (2023) An application of the maximum likelihood estimation for the parameter of uniform distribution of the duration of unextendable dead time in recurrent alternating semi-synchronous flow. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 62. pp. 36–49. doi: 10.17223/19988605/62/4

Введение

Информационные потоки сообщений, функционирующие в современных телекоммуникационных системах и сетях различной конфигурации, наиболее адекватно и точно описываются дважды стохастическими потоками событий [1, 2] – такими потоками, у которых случайными являются и моменты наступления событий, и их интенсивность. В общем случае дважды стохастические потоки событий являются коррелированными потоками [2].

Дважды стохастические потоки делятся на два класса: первый класс составляют потоки, сопровождающий процесс (интенсивность) которых есть непрерывный случайный процесс [3, 4]; второй – потоки, сопровождающий процесс (интенсивность) которых есть кусочно-постоянный случайный процесс с конечным (произвольным) числом состояний [5, 6]. В зависимости от способа перехода интенсивности потока из состояния в состояние выделяется три типа дважды стохастических потоков: 1) синхронные (потоки, у которых состояние сопровождающего процесса меняется в случайные моменты времени, являющиеся моментами наступления событий) [7–10]; 2) асинхронные (потоки, у которых переход из состояния в состояние сопровождающего процесса происходит в случайные моменты времени и не зависит от моментов наступления событий) [11–14]; 3) полусинхронные (потоки, у которых одна часть состояний сопровождающего процесса меняется в моменты наступления событий потока, другая часть состояний сопровождающего процесса меняется в произвольные моменты времени, не связанные с моментами наступления событий потока) [15–18].

В реальных информационных системах часто встречаются ситуации, когда не все события потоков можно наблюдать. Причиной ненаблюдаемости, как правило, служит мертвое время регистрирующих приборов [19], порождаемое зарегистрированным событием, так что другие события, наступившие в этот период, становятся недоступными для наблюдения. Регистрирующие приборы при этом делятся на два вида: с непродлевающимся мертвым временем и продлевающимся. Кроме того, период ненаблюдаемости может быть как детерминированной величиной, одинаковой для всех событий, так и случайной с тем или иным законом распределения. Величина и характер мертвого времени у регистрирующих устройств зависят от многих факторов. При этом регистрирующие приборы обладают значением длительности мертвого времени, ограниченным сверху некоторой величиной. Поэтому, принимая мертвое время случайным [19], вполне естественно рассматривать его распределение как равномерное на некотором отрезке.

Для того чтобы выявить потери событий, возникающие из-за наличия искажающего фактора (непродлевающегося мертвого времени), требуется оценить значение длительности мертвого времени. В настоящей статье оценивается параметр длительности случайного мертвого времени, распределенного по равномерному закону, в полусинхронном дважды стохастическом потоке событий с интенсивностью, являющейся кусочно-постоянным случайным процессом с двумя состояниями. На параметр потока накладывается условие, так что исходный коррелированный поток вырождается в рекуррентный. Производится оценивание для двух случаев функционирования такого потока, общего и особого, с помощью метода максимального правдоподобия. Посредством имитационной модели наблюдаемого потока реализуются статистические эксперименты для получения численных результатов оценивания.

1. Математическая модель наблюдаемого потока. Постановка задачи

Рассматривается стационарный режим функционирования полусинхронного дважды стохастического потока событий, сопровождающий процесс (интенсивность) которого есть кусочно-постоянный стационарный случайный процесс $\lambda(t)$ с двумя состояниями S_1 и S_2 . Будем говорить, что имеет место первое состояние процесса (потока) S_1 , если $\lambda(t) = \lambda_1$, и, наоборот, имеет место второе состояние процесса (потока) S_2 , если $\lambda(t) = \lambda_2$ ($\lambda_1 > \lambda_2 \geq 0$). Если имеет место первое состояние процесса S_1 , то в течение временного интервала, когда $\lambda(t) = \lambda_1$, поступает пуассоновский поток событий с интенсивностью λ_1 . Если имеет место второе состояние процесса S_2 , то в течение временного интервала, когда $\lambda(t) = \lambda_2$, поступает пуассоновский поток событий с интенсивностью λ_2 . Переход из состояния S_1 процесса $\lambda(t)$ в состояние S_2 возможен только в момент наступления события (свойство синхронности потока), при этом этот переход осуществляется с вероятностью p (с вероятностью $1 - p$ процесс $\lambda(t)$ остается в состоянии S_1). Переход из состояния S_2 процесса $\lambda(t)$ в состояние S_1 может осуществляться в произвольный момент времени, не связанный с моментом наступления события (свойство асинхронности потока). При этом длительность пребывания процесса $\lambda(t)$ во втором состоянии есть случайная величина, распределенная по экспоненциальному закону $F(t) = 1 - e^{-\alpha_2 t}$, $t \geq 0$, где α_2 – интенсивность смены состояния S_2 на S_1 . Так как переход из второго состояния в первое не привязан к моменту наступления события во втором состоянии, то поток называется полусинхронным дважды стохастическим потоком событий. В сделанных предположениях $\lambda(t)$ – скрытый марковский процесс ($\lambda(t)$ – принципиально ненаблюдаемый процесс; наблюдаемыми являются только моменты наступления событий потока).

После каждого зарегистрированного события в момент времени t_k наступает период мертвого времени случайной длительности, который порождается этим событием, так что другие события исходного потока, наступившие в течение этого периода мертвого времени, недоступны наблюдению и не вызывают его продления (непродлеваемое мертвое время). Принимается, что случайная длительность мертвого времени распределена по равномерному закону с плотностью вероятности $p(T) = 1/T^*$, где T – значение длительности мертвого времени, $0 \leq T \leq T^*$.

Исследуется частный случай функционирования полусинхронного потока событий, когда $\lambda_2 = 0$, т.е. рассматривается альтернирующий поток, у которого отсутствуют события во втором состоянии. При выполнении такого ограничения исходный полусинхронный поток, действующий в условиях детерминированного мертвого времени, становится рекуррентным потоком:

$$p(\tau_1, \tau_2 | T) = p(\tau_1 | T) p(\tau_2 | T), \quad \tau_1 \geq T, \tau_2 \geq T,$$

где $p(\tau | T)$ – плотность вероятности значений длительности интервала между соседними событиями в наблюдаемом потоке, $\tau = \tau_1, \tau_2$; $p(\tau_1, \tau_2 | T)$ – совместная плотность вероятности [20. С. 254].

Вместе с тем рассматриваются общий и особый случаи соотношения параметров данного потока: когда $\lambda_1 - \alpha_2 \neq 0$ и когда $\lambda_1 - \alpha_2 = 0$ соответственно.

Возможный вариант возникающей ситуации приведен на рис. 1, где S_1 и S_2 – состояния случайного процесса $\lambda(t)$; временная ось $(0, t)$ – ось моментов наступления наблюдаемых событий в моменты времени t_1, t_2, \dots ; временная ось $(0, t^{(1)})$ – ось наступления событий в моменты времени $t_1^{(1)}, t_2^{(1)}, \dots$ в первом состоянии (S_1) процесса $\lambda(t)$, на которой также указаны значения длительностей $T_1^{(1)}, T_2^{(1)}, \dots$ мертвых времен, порождаемых наблюдаемыми событиями потока; на временной оси $(0, t^{(2)})$ нет событий потока, так как $\lambda_2 = 0$; белыми кружками обозначены наблюдаемые события, черными – ненаблюдаемые, штриховкой – периоды мертвого времени; траектория процесса $\lambda(t)$ привязана к временной оси $(0, t^{(1)})$.

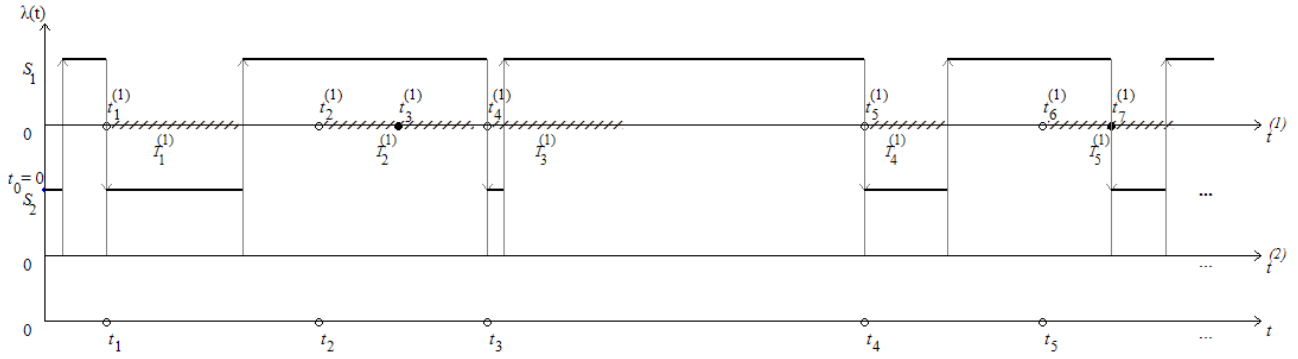


Рис. 1. Формирование наблюдаемого альтернирующего потока событий
Fig. 1. Formation of the observed altering events flow

Цели данной работы:

1. На основании выборки моментов наступления событий t_1, t_2, \dots, t_n наблюдаемых потоков в общем и особом случаях на временном интервале $(0, T_m)$, где T_m – время наблюдения за потоком ($t_n < T_m$), оценить параметр равномерного распределения длительности непродлевающегося случайного мертвого времени T^* методом максимального правдоподобия.

2. Исследовать оценку \hat{T}^* для общего и особого случаев рассматриваемого потока. Для этого провести статистические эксперименты, устанавливающие стационарный режим и определяющие свойства полученных оценок.

2. Приближенная МП-оценка параметра T^*

Введем $\tau_k = t_{k+1} - t_k$, $k = 1, 2, \dots$, – значение длительности k -го интервала между соседними событиями наблюдаемого потока ($\tau_k \geq 0$). В силу того, что рассматривается стационарный режим функционирования наблюдаемого потока, плотность вероятности значений длительности k -го интервала есть $p(\tau_k) = p(\tau)$, $\tau \geq 0$, для любого k , т.е. момент наступления события есть $\tau = 0$.

Из работы [20. С. 217] имеем, что плотность вероятности значений длительности интервала между соседними событиями в рекуррентном альтернирующем полусинхронном потоке, функционирующем в условиях непродлевающегося мертвого времени фиксированной длительности T , в общем случае имеет вид:

$$p(\tau|T) = \begin{cases} 0, & 0 \leq \tau < T, \\ \gamma(T)\lambda_1 e^{-\lambda_1(\tau-T)} + (1-\gamma(T))\alpha_2 e^{-\alpha_2(\tau-T)}, & \tau \geq T, \end{cases}$$

$$\text{где } \gamma(T) = \frac{\alpha_2(\lambda_1 - \alpha_2 - \lambda_1 p)}{(\lambda_1 p + \alpha_2)(\lambda_1 - \alpha_2)} \left[1 + \frac{p(\lambda_1 - \lambda_2)}{\alpha_2 e^{(\lambda_1 p + \alpha_2)T}} \right], \quad \lambda_1 - \alpha_2 \neq 0.$$

Для альтернирующего полусинхронного потока событий с непродлевающимся мертвым временем фиксированной длительности T в особом случае $\lambda_1 - \alpha_2 = 0$ справедливы формулы [20. С. 220]

$$p(\tau|T) = \begin{cases} 0, & 0 \leq \tau < T, \\ \left[\lambda_1 - \lambda_1 \pi_2(T)(1 - \lambda_1(\tau - T)) \right] e^{-\lambda_1(\tau - T)}, & \tau \geq T, \end{cases}$$

$$\text{где } \pi_2(T) = \frac{p}{1+p} \left[1 + p e^{-\lambda_1(1+p)T} \right].$$

Подчеркнем, что внесение непродлевающегося случайного мертвого времени в математическую модель полусинхронного потока событий может только изменить (в меньшую или большую сторону) корреляцию в потоке по сравнению с ситуацией отсутствия мертвого времени ($T^* = 0$) либо с ситуацией наличия детерминированного мертвого времени ($T > 0$), но не устранить ее полностью.

Тогда искомая плотность вероятности $p(\tau)$ примет вид:

$$p(\tau) = \int_{(T)} p(\tau, T) dT = \int_{(T)} p(T) p(\tau | T) dT,$$

где $p(\tau, T)$ – совместная плотность вероятности значений τ и T ; условная плотность вероятности $p(\tau | T)$ определяется выписанными выше выражениями для общего и особого случаев; равномерная плотность $p(T)$ определена в разд. 1; (T) – область интегрирования значений случайной величины – длительности непродлевающегося случайного мертвого времени. Отметим, что внесение непродлевающегося случайного мертвого времени в математическую модель для случая рекуррентного альтернирующего потока, когда $\lambda_2 = 0$, оставляет наблюдаемый поток в классе рекуррентных потоков, так как последнее является следствием значений параметров потока или их соотношений.

Область значений случайной величины мертвого времени представляет собой объединение двух областей, когда $0 \leq \tau < T^*$ и когда $\tau \geq T^*$, поэтому общее выражение для плотности $p(\tau)$ имеет следующий вид:

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

Подставляя выражение $p(\tau, T)$ для общего случая в последнюю формулу, находим

$$p_1(\tau) = \frac{1}{T^*} \left\{ 1 - e^{-\lambda_1 \tau} - e^{-\alpha_2 \tau} + e^{-(\lambda_1 p + \alpha_2) \tau} \right\}, \quad 0 \leq \tau < T^*, \quad (1)$$

$$p_2(\tau) = \frac{1}{T^*} \left\{ e^{-\lambda_1 \tau} \left[-1 + C_1 e^{\lambda_1 T^*} + C_2 e^{(\lambda_1 - \lambda_1 p - \alpha_2) T^*} \right] + e^{-\alpha_2 \tau} \left[-1 + C_1 e^{-\lambda_1 p T^*} + C_2 e^{\alpha_2 T^*} \right] \right\}, \quad \tau \geq T^*, \quad (2)$$

где $C_1 = \frac{\alpha_2 (\lambda_1 - \alpha_2 - \lambda_1 p)}{(\lambda_1 p + \alpha_2)(\lambda_1 - \alpha_2)}$, $C_2 = \frac{\lambda_1^2 p}{(\lambda_1 p + \alpha_2)(\lambda_1 - \alpha_2)}$, $0 < p < 1$, $C_1 + C_2 = 1$, $\lambda_1 \neq \alpha_2$.

Подставляя выражение $p(\tau, T)$ для особого случая в общую формулу для $p(\tau)$, получаем

$$p_1(\tau) = \frac{1}{T^*} \left\{ 1 - 2e^{-\lambda_1 \tau} + e^{-\lambda_1 (1+p) \tau} \right\}, \quad 0 \leq \tau < T^*, \quad (3)$$

$$p_2(\tau) = \frac{1}{(1+p)T^*} \left\{ \left(-2 + e^{\lambda_1 T^*} + e^{-\lambda_1 p T^*} \right) (1+p) + \lambda_1 p (\tau - T^*) \left(e^{\lambda_1 T^*} - e^{-\lambda_1 p T^*} \right) \right\} e^{-\lambda_1 \tau}, \quad \tau \geq T^*. \quad (4)$$

В статье авторов [21] для оценивания параметра T^* использован метод моментов [22]. В настоящей работе для случая рекуррентного альтернирующего полусинхронного потока событий для нахождения оценки параметра T^* воспользуемся методом максимального правдоподобия [22], поскольку длительности интервалов между соседними событиями потока являются взаимно независимыми случайными величинами. Подчеркнем, что аналогичный подход использован в работе [23], где рассмотрена подобная задача для рекуррентного обобщенного асинхронного потока.

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

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

где $p(\tau_k | T^*)$ – плотность вероятности, определенная формулами (1), (2) для общего случая и формулами (3), (4) для особого случая альтернирующего потока, в которых $\tau = \tau_k$ ($\tau^{(k)}$ – измерение); T^* – переменная величина ($T^* > 0$).

Выражение (5) перепишем в следующем виде:

$$L(T^* | \tau_1, \tau_2, \dots, \tau_n) = L_1(T^* | \tau_1, \tau_2, \dots, \tau_n) L_2(T^* | \tau_1, \tau_2, \dots, \tau_n), \quad (6)$$

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

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

$$L(T^* | \tau^{(1)}, \tau^{(2)}, \dots, \tau^{(n)}) = \prod_{k=1}^n p_2(\tau^{(k)} | T^*), \quad 0 < T^* < \tau^{(1)};$$

$$L(T^* | \tau^{(1)}, \tau^{(2)}, \dots, \tau^{(n)}) = p_1(\tau^{(1)} | T^*) \prod_{k=2}^n p_2(\tau^{(k)} | T^*), \quad \tau^{(1)} \leq T^* < \tau^{(2)};$$

$$\dots$$

$$L(T^* | \tau^{(1)}, \tau^{(2)}, \dots, \tau^{(n)}) = \prod_{k=1}^{n-1} p_1(\tau^{(k)} | T^*) p_2(\tau^{(n)} | T^*), \quad \tau^{(n-1)} \leq T^* < \tau^{(n)};$$

$$L(T^* | \tau^{(1)}, \tau^{(2)}, \dots, \tau^{(n)}) = \prod_{k=1}^n p_1(\tau^{(k)} | T^*), \quad \tau^{(n)} \leq T^* < \infty. \quad (7)$$

Теорема. Функция $p(\tau^{(k)} | T^*)$ переменной T^* ($T^* > 0$) как в общем, так и в особом случае достигает своего глобального максимума в точке $T^* = \tau^{(k)}$.

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

Рассмотрим теперь функцию $p_2(\tau^{(k)} | T^*)$ – плотность вероятности, определенную формулой (2), в которой $\tau = \tau^{(k)}$ ($\tau^{(k)}$ – измерение), T^* – переменная величина, $0 \leq T^* \leq \tau^{(k)}$, $k = \overline{1, n}$. При этом $p_1(\tau^{(k)} | T^* = \tau^{(k)}) = p_2(\tau^{(k)} | T^* = \tau^{(k)})$, $k = \overline{1, n}$. Производная по T^* функции $p_2(\tau^{(k)} | T^*)$ выпишется в виде:

$$p_2'(\tau^{(k)} | T^*) = (1/T^*)^2 \psi(T^*), \quad 0 \leq T^* \leq \tau^{(k)}, \quad k = \overline{1, n}, \quad (8)$$

$$\psi(T^*) = e^{-\lambda_1 \tau^{(k)}} + e^{-\alpha_2 \tau^{(k)}} + \frac{1}{(\lambda_1 p + \alpha_2)(\lambda_1 - \alpha_2)} \left\{ e^{-\lambda_1(\tau^{(k)} - T^*)} \left[\alpha_2(\lambda_1 - \lambda_1 p - \alpha_2)(\lambda_1 T^* - 1) + \right. \right.$$

$$\left. \left. + \lambda_1^2 p(T^*(\lambda_1 - \lambda_1 p - \alpha_2) - 1) e^{-(\lambda_1 p + \alpha_2) T^*} \right] + e^{-\alpha_2(\tau^{(k)} - T^*)} \left[\lambda_1^2 p(\alpha_2 T^* - 1) - \right. \right.$$

$$\left. \left. - \alpha_2(\lambda_1 - \lambda_1 p - \alpha_2)(\lambda_1 p T^* + 1) e^{-(\lambda_1 p + \alpha_2) T^*} \right] \right\}. \quad (9)$$

Знак производной (8) определяется знаком функции (9), как функции переменной T^* ($0 \leq T^* \leq \tau^{(k)}$, $k = \overline{1, n}$). Производная функции $\psi(T^*)$ по переменной T^* выпишется в виде:

$$\psi'(T^*) = \frac{T^*}{(\lambda_1 p + \alpha_2)(\lambda_1 - \alpha_2)} \left[\alpha_2 + p(\lambda_1 - \lambda_1 p - \alpha_2) e^{-(\lambda_1 p + \alpha_2) T^*} \right] \times$$

$$\times \left[(\lambda_1 - \lambda_1 p - \alpha_2) e^{-\lambda_1(\tau^{(k)} - T^*)} + p \alpha_2 e^{-\alpha_2(\tau^{(k)} - T^*)} \right]. \quad (10)$$

В (10) возможны четыре ситуации соотношения параметров рассматриваемого потока:

- 1) $\lambda_1 - \alpha_2 > 0$, $\lambda_1 - \alpha_2 - \lambda_1 p > 0$,

- 2) $\lambda_1 - \alpha_2 > 0, \lambda_1 - \alpha_2 - \lambda_1 p < 0,$
- 3) $\lambda_1 - \alpha_2 > 0, \lambda_1 - \alpha_2 - \lambda_1 p = 0,$
- 4) $\lambda_1 - \alpha_2 < 0, \lambda_1 - \alpha_2 - \lambda_1 p < 0.$

Нетрудно показать, что для всех этих ситуаций имеет место $p_2'(\tau^{(k)} | T^*) > 0$ для $0 \leq T^* \leq \tau^{(k)}$, $k = \overline{1, n}$. Последнее означает, что $p_2(\tau^{(k)} | T^*)$ – возрастающая функция переменной T^* ($0 \leq T^* \leq \tau^{(k)}$, $k = \overline{1, n}$) и достигает своего глобального максимума в точке $T^* = \tau^{(k)}$. Объединяя два утверждения относительно глобальных максимумов функций $p_1(\tau^{(k)} | T^*)$ и $p_2(\tau^{(k)} | T^*)$, получаем утверждение теоремы для общего случая.

II. Особый случай. Аналогично общему случаю имеем $dp_1(\tau^{(k)} | T^*) / dT^* < 0$ для $T^* \geq \tau^{(k)}$, где $p_1(\tau^{(k)} | T^*)$ определена в (3). То есть функция $p_1(\tau^{(k)} | T^*)$ является убывающей по переменной T^* и достигает своего глобального максимума в точке $T^* = \tau^{(k)}$, $k = \overline{1, n}$.

Рассмотрим теперь функцию $p_2(\tau^{(k)} | T^*)$ – плотность вероятности, определенную формулой (4), в которой $\tau = \tau^{(k)}$ ($\tau^{(k)}$ – измерение), T^* – переменная величина, $0 \leq T^* \leq \tau^{(k)}$, $k = \overline{1, n}$. При этом $p_1(\tau^{(k)} | T^* = \tau^{(k)}) = p_2(\tau^{(k)} | T^* = \tau^{(k)})$, $k = \overline{1, n}$. Производная по T^* функции $p_2(\tau^{(k)} | T^*)$ выпишется в виде (8), где

$$\psi(T^*) = \frac{1}{1+p} e^{-\lambda_1 \tau^{(k)}} \left\{ 2(1+p) - e^{\lambda_1 T^*} \left(p + [1 + \lambda_1 T^*] [1 + \lambda_1 p(\tau^{(k)} - T^*)] \right) - e^{-\lambda_1 p T^*} \left(1 + p [1 + \lambda_1 p T^*] [1 - \lambda_1(\tau^{(k)} - T^*)] \right) \right\}. \quad (11)$$

Знак производной (8) определяется знаком функции (11) как функции переменной T^* ($0 \leq T^* \leq \tau^{(k)}$, $k = \overline{1, n}$). Производная функции $\psi(T^*)$ по переменной T^* выпишется в виде:

$$\psi'(T^*) = T^* \frac{\lambda_1^2}{1+p} [1 - p + p \lambda_1(\tau^{(k)} - T^*)] [1 - p^2 e^{-\lambda_1(1+p)T^*}] e^{-\lambda_1(\tau^{(k)} - T^*)}. \quad (12)$$

Из (12) вытекает, что $p_2'(\tau^{(k)} | T^*) > 0$ для $0 \leq T^* \leq \tau^{(k)}$, $k = \overline{1, n}$. Последнее означает, что $p_2(\tau^{(k)} | T^*)$ – возрастающая функция переменной T^* ($0 \leq T^* \leq \tau^{(k)}$, $k = \overline{1, n}$) и достигает своего глобального максимума в точке $T^* = \tau^{(k)}$. Объединяя два утверждения относительно глобальных максимумов функций $p_1(\tau^{(k)} | T^*)$ и $p_2(\tau^{(k)} | T^*)$, получаем утверждение теоремы для особого случая, что завершает доказательство теоремы.

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

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

Так как функция $p(\tau^{(k)} | T^*)$ ($T^* > 0$) в точке $T^* = \tau^{(k)}$, $k = \overline{1, n}$, достигает глобального максимума, то будем считать точку $T^* = \tau^{(k)}$, $k = \overline{1, n}$, точкой, подозрительной на локальный максимум функции правдоподобия (7). Тогда алгоритм нахождения значения приближенной МП-оценки \hat{T}^* параметра T^* будет выглядеть следующим образом:

- 1) вычисляются значения функции правдоподобия (7) в точках $T^* = \tau^{(k)}$, $k = \overline{1, n}$;
- 2) находится максимальное значение функции (7) на множестве этих точек;
- 3) в качестве значения приближенной МП-оценки параметра T^* выбирается \hat{T}^* , обеспечивающее максимальное значение функции (7) на предыдущем шаге алгоритма.

3. Результаты статистических экспериментов для наблюдаемого потока

Для установления стационарного режима и определения свойств найденных оценок проведены статистические эксперименты.

Первый статистический эксперимент (установление стационарного режима). Получено 100 реализаций ($N = 100$) имитационной модели наблюдаемого потока для общего и особого случаев при $T^* = 1$ ед. времени и $T_m = 50, 100, \dots, 2\,000$ ед. времени. Для каждого набора параметров получено 100 оценок параметра T^* методом максимального правдоподобия. Каждое i -е решение есть значение оценки \hat{T}_i^* , $i = 1, \dots, 100$, параметра T^* . Заданный набор параметров для рекуррентного альтернирующего полусинхронного потока в общем случае: $\lambda_1 = 2$, $\lambda_2 = 0$, $\alpha_2 = 0,2$, $p = 0,6$; в особом случае: $\lambda_1 = 2$, $\lambda_2 = 0$, $\alpha_2 = 2$, $p = 0,6$.

На основании полученных данных вычислялись выборочное среднее искомым оценок $\hat{M}(\hat{T}^*) = \frac{1}{N} \sum_{i=1}^N \hat{T}_i^*$ и их выборочная вариация $\hat{V}(\hat{T}^*) = \frac{1}{N} \sum_{i=1}^N (\hat{T}_i^* - T^*)^2$, где T^* – известное из имитационной модели значение параметра.

В табл. 1 приведены результаты для общего случая наблюдаемого потока при $T^* = 1$ ед. времени. В первой строке таблицы указано время моделирования T_m (время наблюдения за потоком) ($T_m = 50, 100, \dots, 2\,000$ ед. времени); во второй и третьей строках указано выборочное среднее $\hat{M}(\hat{T}^*)$ и выборочная вариация $\hat{V}(\hat{T}^*)$.

Таблица 1

Численные результаты первого статистического эксперимента для $T^* = 1$, общий случай

T_m	50	100	150	200	250	300	350	400	450	500	
$\hat{M}(\hat{T}^*)$	0,9822	0,8826	0,9919	1,0062	0,9764	0,9277	1,0603	0,9963	0,9825	1,0234	
$\hat{V}(\hat{T}^*)$	1,0369	0,5177	0,2402	0,1225	0,1547	0,1230	0,1150	0,0541	0,0531	0,0614	
	550	600	650	700	750	800	850	900	950	1 000	1 050
	1,0114	1,0061	0,9927	1,0032	1,0150	1,0210	1,0094	1,0370	0,9909	1,0141	0,9866
	0,0766	0,0661	0,0315	0,0379	0,0367	0,0313	0,0339	0,0328	0,0228	0,0191	0,0119
	1 100	1 150	1 200	1 250	1 300	1 350	1 400	1 450	1 500	1 550	1 600
	0,9931	1,0098	1,0145	0,9722	0,9848	0,9967	1,0063	0,9887	1,0087	0,9971	1,0103
	0,0187	0,0250	0,0172	0,0162	0,0204	0,0188	0,0184	0,0133	0,0116	0,0111	0,0100
	1 650	1 700	1 750	1 800	1 850	1 900	1 950	2 000			
	0,9962	1,0039	1,0047	0,9970	0,9860	1,0011	0,9882	1,0078			
	0,0117	0,0117	0,0140	0,0104	0,0150	0,0118	0,0102	0,0125			

Для наглядности на рис. 2, 3 приведены графики зависимостей $\hat{M}(\hat{T}^*)$ и $\hat{V}(\hat{T}^*)$ от значения времени моделирования T_m , построенные по данным табл. 1.

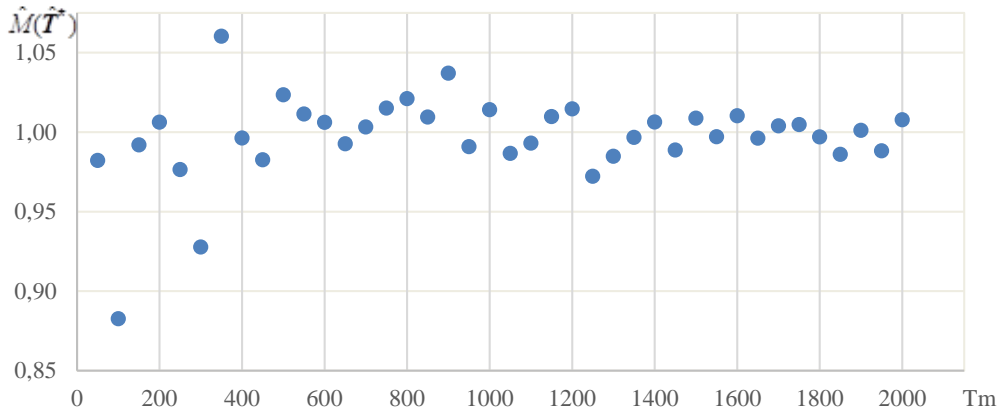


Рис. 2. График зависимости $\hat{M}(\hat{T}^*)$ от T_m при $T^* = 1$ в общем случае

Fig. 2. Plot of $\hat{M}(\hat{T}^*)$ versus T_m with $T^* = 1$ in general case

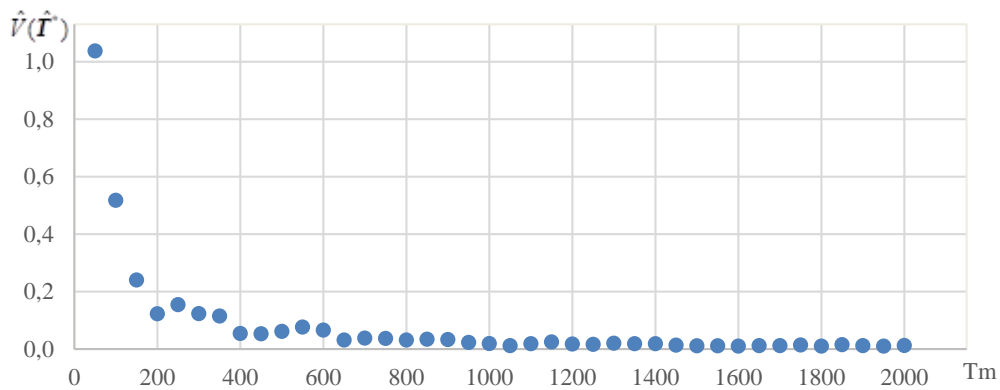


Рис. 3. График зависимости $\hat{V}(\hat{T}^*)$ от T_m при $T^* = 1$ в общем случае

Fig. 3. Plot of $\hat{V}(\hat{T}^*)$ versus T_m with $T^* = 1$ in general case

Из анализа результатов первого статистического эксперимента следует, что для рекуррентного полусинхронного потока событий в общем случае стационарный режим функционирования потока устанавливается при $T_m \geq 1600$ ед. времени, так как выборочное среднее $\hat{M}(\hat{T}^*)$ устанавливается вблизи постоянного значения, когда $T_m \geq 1600$ ед. времени.

Результаты первого статистического эксперимента для особого случая приведены в табл. 2 для $T^* = 1$, а также на рис. 4, 5.

Таблица 2

Численные результаты первого статистического эксперимента для $T^* = 1$, особый случай

T_m	50	100	150	200	250	300	350	400	450	500	
$\hat{M}(\hat{T}^*)$	0,9784	1,0115	0,9836	1,0096	1,0041	0,9964	0,9990	0,9969	0,9973	0,9915	
$\hat{V}(\hat{T}^*)$	0,0437	0,0185	0,0154	0,0095	0,0069	0,0062	0,0055	0,0043	0,0039	0,0044	
	550	600	650	700	750	800	850	900	950	1 000	1 050
	1,0026	1,0066	0,9964	0,9990	1,0002	0,9981	0,9973	1,0000	1,0020	0,9936	1,0006
	0,0030	0,0027	0,0026	0,0029	0,0024	0,0022	0,0021	0,0016	0,0021	0,0018	0,0017
	1 100	1 150	1 200	1 250	1 300	1 350	1 400	1 450	1 500	1 550	1 600
	0,9985	1,0019	0,9990	1,0005	1,0006	0,9944	1,0025	1,0002	0,9968	1,0011	0,9980
	0,0016	0,0020	0,0014	0,0014	0,0014	0,0010	0,0015	0,0013	0,0011	0,0010	0,0010
	1 650	1 700	1 750	1 800	1 850	1 900	1 950	2 000			
	0,9988	1,0000	0,9996	0,9993	0,9961	1,0001	1,0001	1,0017			
	0,0011	0,0011	0,0011	0,0008	0,0010	0,0009	0,0009	0,0008			

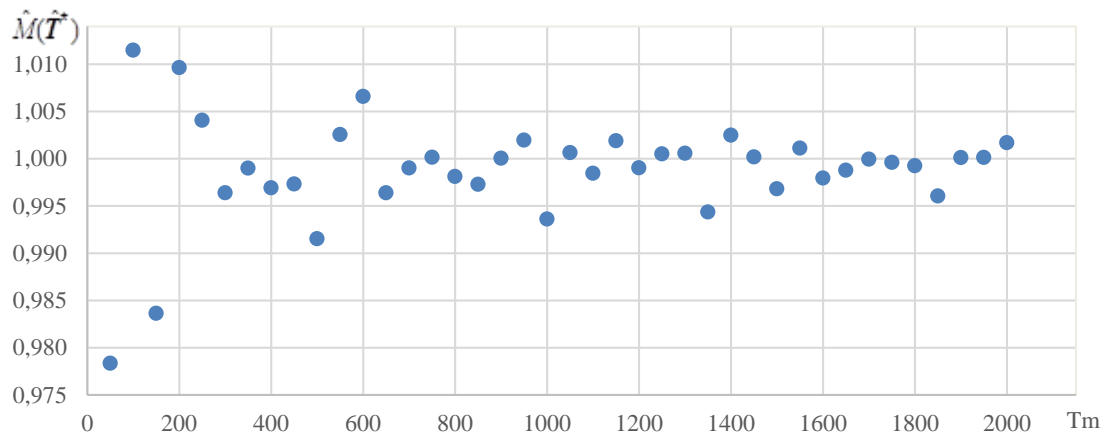


Рис. 4. График зависимости $\hat{M}(\hat{T}^*)$ от T_m при $T^* = 1$ в особом случае

Fig. 4. Plot of $\hat{M}(\hat{T}^*)$ versus T_m with $T^* = 1$ in special case

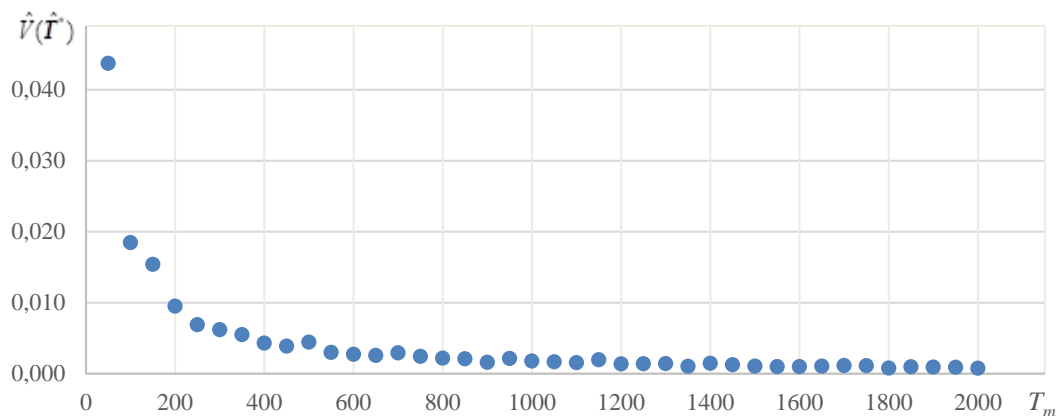


Рис. 5. График зависимости $\hat{V}(\hat{T}^*)$ от T_m при $T^* = 1$ в особом случае

Fig. 5. Plot of $\hat{V}(\hat{T}^*)$ versus T_m with $T^* = 1$ in special case

Аналогично для особого случая рассматриваемого потока следует, что стационарный режим функционирования потока устанавливается при $T_m \geq 1300$ ед. времени.

Второй статистический эксперимент (исследование влияния параметра T^ на качество оценок).* Второй статистический эксперимент поставлен при фиксированном времени моделирования, равным времени установления стационарного режима, полученном из первого статистического эксперимента. Варьирует параметр $T^* = 1; 2; 3; 4; 5$, остальные параметры приняты такими же, как и в первом эксперименте.

Результаты второго статистического эксперимента для общего случая рассматриваемого потока приведены в табл. 3 ($T_m = 1600$ ед. времени).

Таблица 3

Численные результаты второго статистического эксперимента, общий случай

T^*	1	2	3	4	5
$\hat{M}(\hat{T}^*)$	1,0022	2,0110	3,0442	4,0410	4,9749
$\hat{V}(\hat{T}^*)$	0,0166	0,0402	0,0969	0,1165	0,2249

На рис. 6, 7 приведены графики зависимости выборочного среднего $\hat{M}(\hat{T}^*)$ и выборочной вариации $\hat{V}(\hat{T}^*)$ от параметра $T^* = 1; 2; 3; 4; 5$, построенные по данным табл. 3.

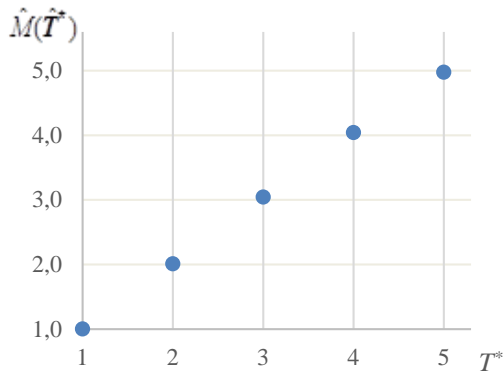


Рис. 6. График зависимости $\hat{M}(T^*)$ от T^* в общем случае

Fig. 6. Plot of $\hat{M}(T^*)$ versus T^* in general case

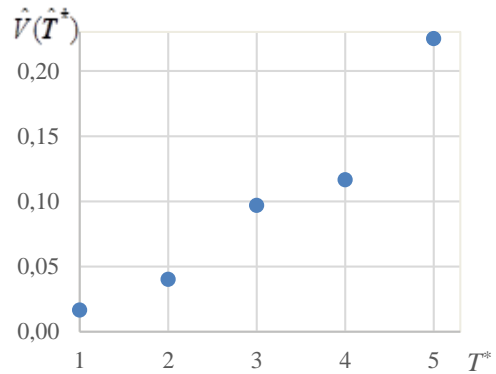


Рис. 7. График зависимости $\hat{V}(T^*)$ от T^* в общем случае

Fig. 7. Plot of $\hat{V}(T^*)$ versus T^* in general case

Результаты второго статистического эксперимента для особого случая приведены в табл. 4 ($T_m = 1300$ ед. времени), а также на рис. 8, 9.

Таблица 4

Численные результаты второго статистического эксперимента, особый случай

T^*	1	2	3	4	5
$\hat{M}(T^*)$	0,9998	2,0032	3,0047	4,0119	4,9929
$\hat{V}(T^*)$	0,0009	0,0020	0,0042	0,0058	0,0081

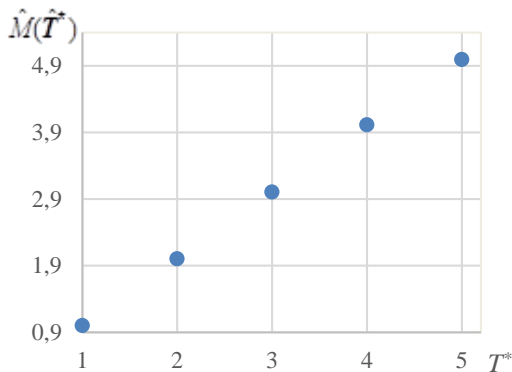


Рис. 8. График зависимости $\hat{M}(T^*)$ от T^* в особом случае

Fig. 12. Plot of $\hat{M}(T^*)$ versus T^* in special case

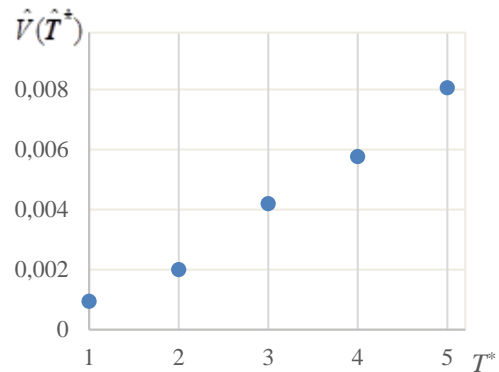


Рис. 9. График зависимости $\hat{V}(T^*)$ от T^* в особом случае

Fig. 13. Plot of $\hat{V}(T^*)$ versus T^* in special case

анный эксперимент демонстрирует тот факт, что при увеличении значений параметра равномерного распределения мертвого времени T^* увеличивается выборочная вариация оценки. Это объясняется тем, что при больших значениях T^* увеличивается число потерянных событий исходного потока и, как следствие, ухудшается качество оценивания. Отклонение значений $\hat{M}(T^*)$ от истинных значений параметра T^* вполне удовлетворительное. Таким образом, получено достаточно высокое качество оценивания параметра T^* методом максимального правдоподобия в смысле критерия выборочной вариации $\hat{V}(T^*)$.

Заключение

В работе рассмотрен рекуррентный альтернирующий полусинхронный дважды стохастический поток событий в общем и особом случаях соотношения его параметров с непродлевающимся случайным мертвым временем, распределенным по равномерному закону.

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

- 1) получен явный вид (7) функции правдоподобия для оценивания параметра T^* равномерного распределения длительности непродлевающегося случайного мертвого времени;
- 2) показано, что точка глобального максимума функции правдоподобия (7) принадлежит отрезку $[\tau^{(1)}, \tau^{(n)}]$;
- 3) описана процедура построения МП-оценки \hat{T}^* ;
- 4) полученные оценки экспериментально исследованы на качество; приведенные результаты численных расчетов показывают, что метод максимального правдоподобия позволяет в достаточной мере эффективно оценивать параметр мертвого времени T^* в смысле малости выборочной вариации оценок.

Список источников

1. Башарин Г.П., Кокотушкин В.А., Наумов В.А. О методе эквивалентных замен расчета фрагментов сетей связи. Ч. 1 // Известия АН СССР. Техническая кибернетика. 1979. № 6. С. 92–99.
2. Вишневецкий В.М., Дудин А.Н., Клименок В.Н. Стохастические системы с коррелированными потоками. Теория и применение в телекоммуникационных сетях. М. : Техносфера, 2018. 564 с.
3. Cox D.R. The analysis of non-Markovian stochastic processes by the inclusion of supplementary variables // Proceedings of the Cambridge Philosophical Society. 1955. V. 51 (3). P. 433–441.
4. Kingman Y.F.C. On doubly stochastic Poisson process // Proceedings of the Cambridge Philosophical Society. 1964. V. 60 (4). P. 923–930.
5. Neuts M.F. A versatile Markovian point process // Journal of Applied Probability. 1979. V. 16. P. 764–779.
6. Lucantoni D.M. New results on the single server queue with a batch markovian arrival process // Communication in Statistics Stochastic Models. 1991. V. 7. P. 1–46.
7. Горцев А.М., Нежелская Л.А. Оценивание длительности мертвого времени и параметров синхронного альтернирующего потока событий // Вестник Томского государственного университета. 2003. № S6. С. 232–239.
8. Gortsev A.M., Nezhel'skaya L.A. Estimation of parameters of synchronously alternating Poisson stream of events by the moment method // Telecommunications and Radio Engineering. 1996. V. 50 (1). P. 56–63.
9. Горцев А.М., Нежелская Л.А. Оценивание параметров синхронного дважды стохастического потока событий методом моментов // Вестник Томского государственного университета. 2002. № S1-1. С. 24–29.
10. Nezhel'skaya L.A. Probability density function for modulated MAP event flows with unextendible dead time // Communication in Computer and Information Science. 2015. V. 564. P. 141–151.
11. Горцев А.М., Нежелская Л.А. Оценивание параметров асинхронного потока с инициированием лишних событий методом моментов // Вестник Томского государственного университета. 2006. № S18. С. 267–273.
12. Леонова М.А., Нежелская Л.А. Вероятность ошибки при оценивании состояний обобщенного асинхронного потока событий // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2012. № 2 (19). С. 88–101.
13. Горцев А.М., Зуевич В.Л. Оптимальная оценка состояний асинхронного дважды стохастического потока событий с произвольным числом состояний // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2010. № 2 (11). С. 44–65.
14. Горцев А.М., Леонова М.А., Нежелская Л.А. Сравнение МП- и ММ-оценок длительности мертвого времени в обобщенном асинхронном потоке событий // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2013. № 4 (25). С. 32–42.
15. Калягин А.А., Нежелская Л.А. Сравнение МП- и ММ-оценок длительности мертвого времени в обобщенном полусинхронном потоке событий // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2015. № 3 (32). С. 23–32.
16. Горцев А.М., Калягин А.А., Нежелская Л.А. Оценка максимального правдоподобия длительности мертвого времени в обобщенном полусинхронном потоке // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2015. № 1 (30). С. 27–37.
17. Нежелская Л.А. Оптимальное оценивание состояний полусинхронного потока событий в условиях его частичной наблюдаемости // Вестник Томского государственного университета. 2000. № 269. С. 95–98.
18. Горцев А.М., Калягин А.А., Нежелская Л.А. Совместная плотность вероятностей длительности интервалов обобщенного полусинхронного потока событий при непродлевающемся мертвом времени // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2014. № 2 (27). С. 19–29.
19. Апанасович В.В., Коляда А.А., Чернявский А.Ф. Статистический анализ случайных потоков в физическом эксперименте. Минск : Университетское, 1988. 256 с.

20. Нежелская Л.А. Оценка состояний и параметров дважды стохастических потоков событий : дис. ... д-ра физ.-мат. наук. Томск, 2016. 341 с.
21. Горцев А.М., Веткина А.В. Оценивание методом моментов параметра равномерного распределения длительности непродлевающегося случайного мертвого времени в рекуррентном полусинхронном потоке событий в общем и особом случаях // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2022. № 61. С. 47–60.
22. Малинковский Ю.В. Теория вероятностей и математическая статистика. Гомель : ГГУ им. Ф. Скорины, 2004. Ч. 2: Математическая статистика. 146 с.
23. Нежелская Л.А., Першина А.А. Оценивание методом максимального правдоподобия параметра распределения случайного мертвого времени в рекуррентном обобщенном асинхронном потоке событий // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2021. № 55. С. 53–64.

References

1. Basharin, G.P., Kokotushkin, V.A. & Naumov, V.A. (1979) O metode ekvivalentnykh zamen rascheta fragmentov setey svyazi. Ch. 1 [On the equivalent substitutions method for computing fragments of communication networks. Part 1]. *Izvestiya AN SSSR. Tekhnicheskaya kibernetika*. 6. pp. 92–99.
2. Vishnevsky, V.M., Dudin, A.N. & Klimenok, V.N. (2018) *Stokhasticheskie sistemy s korrelirovannymi potokami. Teoriya i primeneniye v telekommunikatsionnykh setyakh* [Stochastic systems with correlated flows. Theory and application in telecommunication networks]. Moscow: Tekhnosfera.
3. Cox, D.R. (1955) The analysis of non-Markovian stochastic processes by the inclusion of supplementary variables. *Proceedings of the Cambridge Philosophical Society*. 51(3). pp. 433–441. DOI: 10.1017/S0305004100030437
4. Kingman, Y.F.C. (1964) On doubly stochastic Poisson process. *Proceedings of the Cambridge Philosophical Society*. 60(4). pp. 923–930.
5. Neuts, M.F. (1979) A versatile Markovian point process. *Journal of Applied Probability*. 16. pp. 764–779.
6. Lucantoni, D.M. (1991) New results on the single server queue with a batch markovian arrival process. *Communication in Statistics Stochastic Models*. 7. pp. 1–46.
7. Gortsev, A.M. & Nezhelskaya, L.A. (2003) Otsenivaniye dlitel'nosti mertvogo vremeni i parametrov sinkhronnogo al'terniruyushchego potoka sobyitiy [Dead time period and parameter estimation of synchronous alternating flow of events]. *Vestnik Tomskogo gosudarstvennogo universiteta – Tomsk State University Journal*. 6. pp. 232–239.
8. Gortsev A.M. & Nezhelskaya, L.A. (1996) Estimation of parameters of synchronously alternating Poisson stream of events by the moment method. *Telecommunications and Radio Engineering*. 50(1). pp. 56–63.
9. Gortsev, A.M. & Nezhelskaya, L.A. (2002) Estimation of the parameters of a synchronous doubly stochastic event flow by the method of moments. *Vestnik Tomskogo gosudarstvennogo universiteta – Tomsk State University Journal*. S1-1. pp. 24–29.
10. Nezhelskaya, L.A. (2015) Probability density function for modulated MAP event flows with unextendible dead time. *Communication in Computer and Information Science*. 564. pp. 141–151.
11. Gortsev, A.M. & Nezhelskaya, L.A. (2006) Otsenivaniye parametrov asinkhronnogo potoka s initsirovaniem lishnikh sobyitiy metodom momentov [Estimation of the parameters of an asynchronous flow with the initiation of unnecessary events by the method of moments]. *Vestnik Tomskogo gosudarstvennogo universiteta – Tomsk State University Journal*. S18. pp. 267–273.
12. Leonova, M.A. & Nezhelskaya, L.A. (2012) The probability of wrong decisions in the estimation of states of a generalized asynchronous flow of events. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 2(19). pp. 88–101.
13. Gortsev, A.M. & Zuevich, V.L. (2010) Optimal'naya otsenka sostoyaniy asinkhronnogo dvazhdy stokhasticheskogo potoka sobyitiy s proiz-vol'nym chislom sostoyaniy [Optimal estimation of the states of an asynchronous doubly stochastic flow of events with an arbitrary number of states]. *Vestnik Tomskogo gosudarstvennogo universiteta – Tomsk State University Journal*. 2(11). pp. 44–65.
14. Gortsev, A.M., Leonova, M.A. & Nezhelskaya, L.A. (2013) Sravneniye MP- i MM-otsenok dlitel'nosti mertvogo vremeni v obobshchennom asinkhronnom potoke sobyitiy [Comparison of MP- and MM-estimates of the duration of dead time in a generalized asynchronous flow of events]. *Vestnik Tomskogo gosudarstvennogo universiteta – Tomsk State University Journal*. 4(25). pp. 32–42.
15. Kalyagin, A.A. & Nezhelskaya, L.A. (2015) The comparison of maximum likelihood estimation and method of moments estimation of dead time value in a generalized semysynchronous flow of events. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 3(32). pp. 23–32. DOI: 10.17223/19988605/32/3
16. Gortsev, A.M., Kalyagin, A.A. & Nezhelskaya, L.A. (2015) Maximum likelihood estimation of dead time value at a generalized semysynchronous flow of events. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 1(30). pp. 27–37.
17. Nezhelskaya, L. (2000) Optimal'noe otsenivaniye sostoyaniy polusinkhronnogo potoka sobyitiy v usloviyakh ego chastichnoy nablyudaemosti [Optimal estimation of the states of semi-synchronous event flow under conditions of its partial observability]. *Vestnik Tomskogo gosudarstvennogo universiteta – Tomsk State University Journal*. 269. pp. 95–98.

18. Gortsev, A.M., Kalyagin, A.A. & Nezhelskaya, L.A. (2014) The joint probability density of duration of the intervals in a generalized semisynchronous flow of events with unprolonging dead time. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 2(27). pp. 19–29.
19. Apanasovich, V.V., Kolyada, A.A. & Chernyavsky, A.F. (1988) *Statisticheskiy analiz sluchaynykh potokov v fizicheskom eksperimente* [The statistical analysis of series of random events in physical experiment]. Minsk: Universitetskoe.
20. Nezhelskaya, L.A. (2016) *Otsenka sostoyaniy i parametrov dvazhdy stokhasticheskikh potokov sobyitiy* [Estimation of States and Parameters of Doubly Stochastic Event Flows]. Physics and Mathematics Cand. Diss. Tomsk. p. 341.
21. Gortsev, A.M. & Vetkina, A.V. (2022) Estimation by method of moments of the parameter of the uniform distribution of the duration of unextendable random dead time in recurrent semi-synchronous events flow in general and special case. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 61. pp. 47–60. DOI: 10.17223/19988605/61/5
22. Malinkovsky, Yu.V. (2004) *Teoriya veroyatnostey i matematicheskaya statistika* [Probability Theory and Mathematical Statistics]. Gomel: Francisk Skorina Gomel State University.
23. Nezhelskaya, L.A. & Pershina, A.A. (2021) Estimation of the uniform distribution parameter of unextendable dead time duration in a generalized recurrent asynchronous flow of events by the maximum likelihood method. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 55. pp. 53–64. DOI: 10.17223/19988605/55/7

Информация об авторах:

Горцев Александр Михайлович – профессор, доктор технических наук, заведующий кафедрой прикладной математики Института прикладной математики и компьютерных наук Национального Исследовательского Томского государственного университета. E-mail: a-gortsev@mail.ru

Веткина Анна Васильевна – аспирант кафедры прикладной математики Института прикладной математики и компьютерных наук Национального Исследовательского Томского государственного университета. E-mail: anyavetkina@gmail.com

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

Information about the authors:

Gortsev Alexander M. (Doctor of Technical Sciences, Professor, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: a-gortsev@mail.ru

Vetkina Anna V. (Post-graduate Student, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: anyavetkina@gmail.com

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

Поступила в редакцию 14.09.2022; принята к публикации 01.03.2023

Received 14.09.2022; accepted for publication 01.03.2023

Научная статья

УДК 519.688

doi: 10.17223/19988605/62/5

Бинарное прогнозирование динамических показателей на основе методов машинного обучения

Юрий Мечеславович Краковский¹, Ольга Константиновна Куклина²

¹ Иркутский государственный университет путей сообщения, Иркутск, Россия, 79149267772@yandex.ru

² Читинский институт (филиал) Байкальского государственного университета, Чита, Россия, kuklinaok@bgu-chita.ru

Аннотация. Рассмотрена задача бинарного прогнозирования динамических показателей на основе машинного обучения с приложением к задаче перевозки грузов железнодорожным транспортом. В качестве методов выбраны вероятностная нейронная сеть и логистическая регрессия. Бинарное прогнозирование заключается в оценке прогнозных значений показателя на основе вероятностей принадлежности одному из двух интервалов. Так как при такой процедуре определяется не само будущее значение показателя, а то, в каком интервале оно будет находиться, такое прогнозирование называют бинарным, или интервальным. Программное обеспечение разработано на языке программирования Python с применением сторонних библиотек с открытым исходным кодом. Тестирование созданного программно-алгоритмического обеспечения по реальным исходным данным перевозочного процесса показало высокую точность бинарного прогнозирования и на основе вероятностной нейронной сети, и на основе логистической регрессии.

Ключевые слова: бинарное прогнозирование; вероятностная нейронная сеть; логистическая регрессия; динамические показатели.

Для цитирования: Краковский Ю.М., Куклина О.К. Бинарное прогнозирование динамических показателей на основе методов машинного обучения // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2023. № 62. С. 50–55. doi: 10.17223/19988605/62/5

Original article

doi: 10.17223/19988605/62/5

The binary forecasting of dynamic indicators based on machine learning methods

Yuri M. Krakovsky¹, Olga K. Kuklina²

¹ Irkutsk State Transport University, Irkutsk, Russian Federation, 79149267772@yandex.ru

² Chita Institute of Baikal State University, Chita, Russian Federation, kuklinaok@bgu-chita.ru

Abstract. The problem of binary forecasting of dynamic indicators based on machine learning methods in relation to the problem of cargo transportation by railway transport is considered. The probabilistic neural network and logistic regression were chosen as the methods. The binary forecasting consists on evaluating predictive values of the indicator which is based on the belonging probabilities to one of two intervals. The forecasting is called binary or interval as on this process is calculated interval for the indicator value where it will be, not the predicted value of the indicator. The software is developed using the Python programming language with open source libraries. The software and algorithm test were done on the examples of real values of railway transportation process and shown its high accuracy of binary forecasting both on the probabilistic neural network and logistic regression methods.

Keywords: binary forecasting; probabilistic neural network; logistic regression; dynamic indicators.

For citation: Krakovsky, Y.M., Kuklina, O.K. (2023) The binary forecasting of dynamic indicators based on machine learning methods. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 62. pp. 50–55. doi: 10.17223/19988605/62/5

Введение

Как правило, динамические процессы проявляются в виде ряда последовательно расположенных в хронологическом порядке значений того или иного показателя, характеризующего динамику развития изучаемого явления. В связи с этим важным направлением математического моделирования является прогнозирование показателей динамических процессов. Данному направлению посвящено значительное число работ, но среди них преобладают работы, основанные на регрессионных или авторегрессионных моделях. Для увеличения точности прогноза в условиях неопределенности в последние годы происходит совершенствование методов прогнозирования [1–3], и здесь следует отметить широкое применение ансамблей и методов на основе машинного обучения [4, 5].

Методы прогнозирования показателей можно разделить на две группы: а) точечные, когда определяется будущее значение; б) вероятностные, когда определяется не само значение, а вероятности того, в каком интервале это значение находится. В последнее время наблюдается возрастающий интерес пользователей именно к вероятностным методам прогнозирования [6, 7]. Это можно объяснить тем, что на практике часто не требуется знания самого значения показателя, а необходимо лишь знание о том в каком интервале оно окажется и с какой вероятностью.

Среди методов вероятностного прогнозирования, включая бинарное прогнозирование, можно выделить следующие: вероятностные кластерные методы, логистической регрессии, на основе вероятностных нейронных сетей и др. [8, 9].

В данной работе рассматривается задача прогнозирования динамических показателей, характеризующих перевозочный процесс грузов железнодорожным транспортом. Для России перевозка грузов по железной дороге существенно влияет на ее экономику. Это особенно важно для восточного полигона железнодорожной сети ОАО «РЖД», расположенного в границах четырех дорог: Красноярской, Восточносибирской, Забайкальской и Дальневосточной. Особая роль среди этих дорог принадлежит Дальневосточной, так как она является конечной и взаимодействует через морские порты и погранпереходы с другими странами. В связи с этим для повышения эффективности управления перевозкой грузов прогнозирование динамических показателей перевозочного процесса для этой дороги имеет существенное значение. В условиях неопределенности перевозочного процесса такая задача является актуальной и нетривиальной.

Цель работы – реализация и апробация методов на основе машинного обучения для бинарного прогнозирования динамических показателей с приложением к задаче перевозки грузов железнодорожным транспортом. В качестве методов выбраны вероятностная нейронная сеть (ВНС) и логистическая регрессия (ЛР).

Пусть известен временной ряд некоторого показателя $Q = \{q_t : t \in T\}$. Здесь q_t – значения показателя, заданные в моменты времени t ; время t принимает значения из множества $T = 0, \dots, n-1$; n – количество значений показателя; все значения показателя $q_t > 0$. Обозначим интервал возможных значений показателя в будущем $(c_1; c_2)$, где $c_1 > 0$, $c_2 < \infty$; введем внутреннюю точку c : $c_1 < c < c_2$. Это позволяет разделить этот интервал на два других[^]

$$I_a = (c_1; c], \quad I_b = (c; c_2). \quad (1)$$

Тогда бинарное (интервальное) прогнозирование заключается в оценке прогнозных значений показателя на основе вероятностей принадлежности одному из двух интервалов (1). Так как при такой процедуре определяется не само будущее значение показателя, а то, в каком интервале оно будет находиться, такое прогнозирование названо бинарным[?] или интервальным [1].

Значение внутренней точки c можно определять различными способами. В данной работе она определяется так[^]

$$c = q_{n-1} + \Delta, \quad \Delta = \alpha \cdot \left(\sum_{t=1}^{n-1} |q_t - q_{t-1}| \right) / (n-1), \quad (2)$$

где $\alpha \in [-1; 1]$ – коэффициент, который задается в исходных данных.

В момент времени $t = n - 1$ необходимо определить, в каком из интервалов (1) будет находиться будущее (неизвестное) значение q_{t+p} на основе оценок вероятностей ρ_{t+p}^a и ρ_{t+p}^b , где $p = 1, \dots, r$ – время упреждения; ρ_{t+p}^a – вероятность того, что $q_{t+p} \in I^a$; ρ_{t+p}^b – вероятность того, что $q_{t+p} \in I^b$; $\rho_{t+p}^a + \rho_{t+p}^b = 1$.

Бинарное прогнозирование проводится по правилу:

будущее значение $q_{t+p} \in I^a$, если $\rho_{t+p}^a \geq \rho_{t+p}^b$; будущее значение $q_{t+p} \in I^b$, если $\rho_{t+p}^b > \rho_{t+p}^a$. (3).

1. Математическое описание задачи

1.1. Случай вероятностной нейронной сети

Архитектура ВНМ была предложена в 1988 г. Д. Спехтом [10] для проведения классификации векторов (образов) с неизвестной классификацией. Перед проведением такой классификации ВНМ должна быть обучена на множестве векторов с известной классификацией. Достоинством ВНМ является простота построения и высокая скорость обучения.

Под вектором ВНМ понимается набор значений $V = v_j, j \in J$, где j принимает значения из множества $J = 0, \dots, m - 1$, при этом $m > 0$. Часть векторов (u) – обучающие, другая часть – для тестирования. Для всех векторов ВНМ должно выполняться условие

$$\sum_{j \in J} v_j^2 = 1. \quad (4)$$

Условие (4) обеспечивается специальной обработкой реальных значений.

ВНС включают четыре слоя: входной слой, в котором количество нейронов определяется количеством введенных исходных значений (размерность векторов m); скрытый слой, в котором каждый нейрон имеет m входов; суммирующий слой, где определяются вероятности принадлежности входного вектора к тому или иному классу; выходной слой, на котором проводится сравнение вероятностей и формулируется результат.

На этапе обучения создается матрица весов: $w_{z,j} = v_{z,j}$, где z принимает значения из множества $Z = 0, \dots, u - 1$, $u = n^A + n^B$, где n^A – число векторов при обучении из класса A (прогнозное значение исследуемого показателя попадает в интервал I_a); n^B – число векторов при обучении из класса B , когда прогнозное значение показателя попадает в интервал I_b .

Рассмотрим алгоритм классификации векторов с неизвестной классификацией посредством ВНМ, а также формализуем вид активационной функции нейронов скрытого слоя.

Пусть имеется произвольный вектор $V_h = \{v_{h,j}, j \in J\}$ с неизвестной классификацией, для которого также выполняется условие (4). Этот вектор подается на входы входного слоя ВНМ.

Далее в скрытом слое вычисляется u значений с помощью нелинейной активационной функции

$$H_{h,z} = \exp \left[\frac{-\sum_{j \in J} (v_{h,j} - w_{z,j})^2}{2\sigma^2} \right]. \quad (5)$$

Здесь σ – параметр ВНМ, который задается в исходных данных.

С учетом (5) нейроны суммирующего слоя реализуют вычисления

$$I_h^A = \frac{\sum_{z \in M^A} H_{h,z}}{n^A}, \quad I_h^B = \frac{\sum_{z \in M^B} H_{h,z}}{n^B}, \quad (6)$$

где M^A – множество номеров нейронов скрытого слоя, созданных для векторов класса A ; M^B – множество номеров нейронов скрытого слоя, созданных для векторов класса B . Эти множества определяются на этапе обучения.

Выходной нейрон ВНМ вычисляет бинарное значение

$$\Omega_h = \begin{cases} 1, & \tilde{\rho}_{t+p}^b > \tilde{\rho}_{t+p}^a, \\ 0, & \tilde{\rho}_{t+p}^a \geq \tilde{\rho}_{t+p}^b. \end{cases} \quad (7)$$

В функции (7) приведены оценки вероятностей ρ_{t+p}^a и ρ_{t+p}^b :

$$\tilde{\rho}_{t+p}^a = l_h^A / (l_h^A + l_h^B), \quad \tilde{\rho}_{t+p}^b = l_h^B / (l_h^A + l_h^B). \quad (8)$$

Само бинарное прогнозирование осуществляется по правилу (3): значение 1 в (7) означает, что будущее значение показателя попадет в интервал I_b , а при значении 0 – в интервал I_a . Оценки вероятностей определяются по формуле (8) с учетом (6).

1.2. Случай логистической регрессии

Введем: а) бинарное значение

$$y_{t+p} = \begin{cases} 1, & q_{t+p} > c, \\ 0, & q_{t+p} \leq c, \end{cases} \quad (9)$$

где c – внутренняя точка (2); б) линейную регрессионную функцию

$$s_t = a_0 + \sum_{i=1}^f a_i \cdot q_{t-i+1}, \quad (10)$$

где a_0, \dots, a_f – коэффициенты; f – число регрессоров. Следуя рекомендациям, будем предполагать, что оценка вероятности первого события

$$\tilde{\rho}_{t+p}^b \ y_{t+p} = 1 | s_t = \sigma \ s_t, \quad (11)$$

где $\sigma \ s_t = 1 / (1 + e^{-s_t})$ – логистическая (сигмоидальная) функция; величина s_t определяется согласно выражению (10). Так как y_{t+p} принимает только два возможных значения (9), то оценка вероятности наступления второго события равна

$$\tilde{\rho}_{t+p}^a \ y_{t+p} = 0 | s_t = 1 - \sigma \ s_t. \quad (12)$$

При бинарном прогнозировании искомые вероятности заменяются оценками (11) и (12). Само бинарное прогнозирование осуществляется по правилу (3): значение 1 в (9) означает, что будущее значение показателя попадет в интервал I_b , а при значении 0 – в интервал I_a .

2. Программное обеспечение и обсуждение результатов

Программное обеспечение бинарного прогнозирования разработано на языке программирования Python с применением сторонних библиотек с открытым исходным кодом. Для разработки выбрана среда PyCharm – интеллектуальная Python IDE с полным набором средств для эффективной разработки на языке Python. Пользователю предлагается выбрать подготовленные исходные данные; способ прогнозирования показателей – на основе вероятностной нейронной модели; на основе логистической регрессии (вкладка Load data). На вкладке Plot пользователю доступны инструменты визуализации данных. Результаты интервального прогнозирования доступны на вкладке Result.

Например, построение ВНС выполнено с использованием библиотеки `neupy` и `sklearn`; для обработки и анализа данных – библиотека `pandas`. Для визуализации машинного обучения предложено использовать `yellowbrick.classifier` и библиотеку интерактивной визуализации данных `cufflinks`. По-

строение ВНС выполнено с использованием модуля `neuru.algorithms.rbfnn.pnn` (библиотека для искусственных нейронных сетей и глубокого обучения).

Для визуализации машинного обучения предложено использовать платформу Yellowbrick – проект Python с открытым исходным кодом, который объединяет API-интерфейсы `scikit-learn` и `matplotlib`.

На рис. 1 приведены вкладки разработанного программного обеспечения для бинарного прогнозирования.

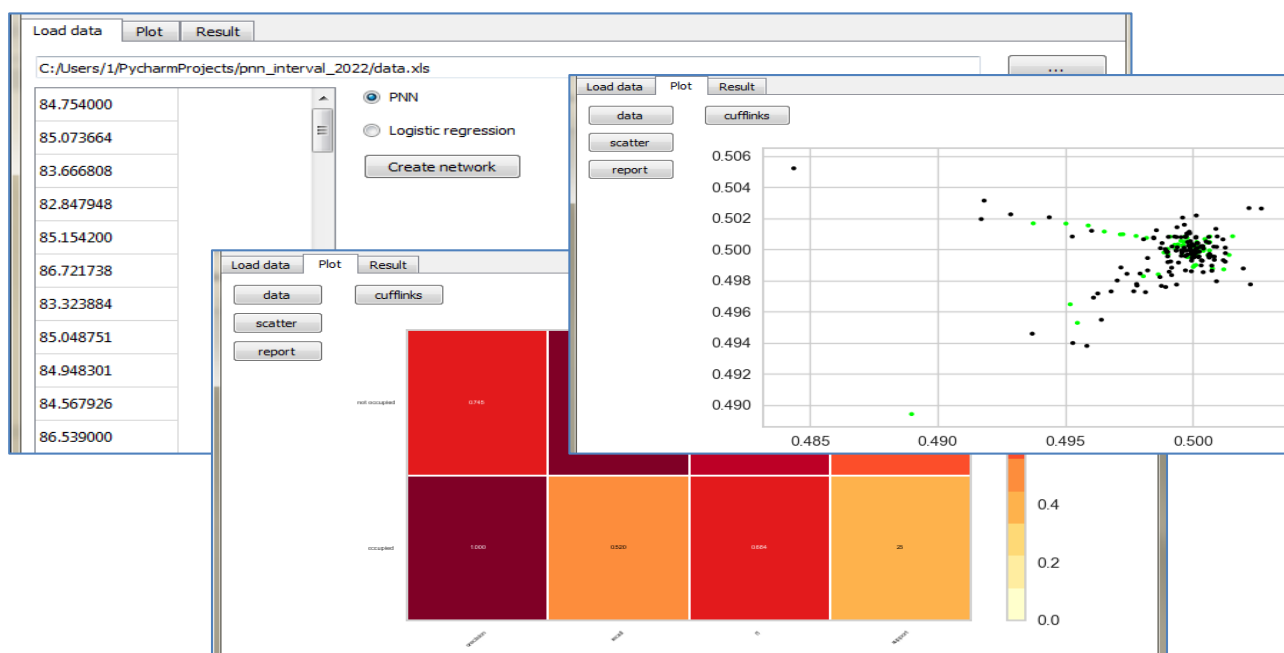


Рис. 1. Вкладки программы для бинарного прогнозирования

Fig. 1. Program tabs of binary forecasting

В результате тестирования созданного программно-алгоритмического обеспечения бинарного прогнозирования базового показателя перевозочного процесса в виде грузооборота на основе ВНС показано, что в 93,3% случаев получены правильные результаты. Применительно к ЛР правильные результаты оказались еще выше и составили 95,1%.

Заключение

Предложено и апробировано программно-алгоритмическое обеспечение бинарного прогнозирования динамических показателей с приложением к задаче перевозки грузов железнодорожным транспортом. В качестве методов машинного обучения выбраны вероятностная нейронная сеть и логистическая регрессия. Тестирование созданного обеспечения по реальным исходным данным перевозочного процесса в виде грузооборота показало его высокую точность и на основе вероятностной нейронной сети, и на основе логистической регрессии.

Список источников

1. Shumway R.H. Time series analysis and its applications with R examples. Springer, 2011. 609 p.
2. Mitrea C.A. A Comparison between neural networks and traditional forecasting methods: a case study // International Journal of Engineering Business Management. 2009. V. 1 (2). P. 19–24.
3. Vernay M., Lafaysse M., Merindol L. Ensemble forecasting of snowpack conditions and avalanche hazard // Cold Regions Science and Technology. 2015. V. 120. P. 251–262.
4. Краковский Ю.М., Курчинский Ю.В., Лузгин А.Н. Интервальное прогнозирование интенсивности кибератак на объекты критической информационной инфраструктуры // Доклады ТУСУР. 2018. Т. 21, № 1. С. 71–79.

5. Wang H., Li G., Wang H. Deep learning based ensemble approach for probabilistic wind power forecasting // *Applied Energy*. 2017. V. 188. P. 56–70.
6. Yoder M., Cering A.S., Navidi W.C. Short-term forecasting of categorical changes in wind power with Markov chain models // *Wind Energy*. 2014. V. 17. P. 1425–1439.
7. Krakovsky Y., Luzgin A. Robust interval forecasting algorithm based on a probabilistic cluster model // *Journal of Statistical Computation and Simulation*. 2018. V. 88 (12). P. 2309–2324.
8. Ivanyo Y.M., Krakovsky Y.M., Luzgin A.N. Interval forecasting of cyber-attacks on industrial control systems // *IOP Conference Series: Materials Science and Engineering (Simulation and Automation of Production Engineering)*. 2018. V. 327. Art. 022044.
9. Munkhdorj B., Yuji S. Cyber attack prediction using social data analysis // *Journal of High Speed Networks*. 2017. V. 23 (2). P. 109–135.
10. Spetch D.F. Probabilistic Neural networks // *Neural Networks*. 1990. V. 3. P. 109–118.

References

1. Shumway, R.H. (2011) *Time Series Analysis and Its Applications with R Examples*. Springer.
2. Mitrea, C.A. (2009) A Comparison between neural networks and traditional forecasting methods: a case study. *International Journal of Engineering Business Management*. 1(2). pp. 19–24.
3. Vernay, M., Lafayssse, M. & Merindol, L. (2015) Ensemble forecasting of snowpack conditions and avalanche hazard. *Cold Regions Science and Technology*. 120. pp. 251–262.
4. Krakovsky, Y.M., Kurchinsky, B.V. & Luzgin, A.N. (2018) Cyber-attack intensity interval forecasting on objects of critical information infrastructure. *Doklady TUSUR – Proceedings of the TUSUR University*. 21(1). pp. 71–79.
5. Wang, H., Li, G. & Wang, H. (2017) Deep learning based ensemble approach for probabilistic wind power forecasting. *Applied Energy*. 188. pp. 56–70.
6. Yoder, M., Cering, A.S. & Navidi, W.C. (2014) Short-term forecasting of categorical changes in wind power with Markov chain models. *Wind Energy*. 17. pp. 1425–1439.
7. Krakovsky, Y. & Luzgin, A. (2018) Robust interval forecasting algorithm based on a probabilistic cluster model. *Journal of Statistical Computation and Simulation*. 88(12). pp. 2309–2324. DOI: 10.1080/00949655.2018.1462809
8. Ivanyo, Y.M., Krakovsky, Y.M. & Luzgin, A.N. (2018) Interval forecasting of cyber-attacks on industrial control systems. *IOP Conference Series: Materials Science and Engineering (Simulation and automation of production engineering)*. 327. Art. 022044. DOI:10.1088/1757-899X/327/2/022044
9. Munkhdorj, B. & Yuji, S. (2017) Cyber attack prediction using social data analysis. *Journal of High Speed Networks*. 23(2). pp. 109–135.
10. Spetch, D.F. (1990) Probabilistic Neural networks. *Neural Networks*. 3. pp. 109–118.

Информация об авторах:

Краковский Юрий Мечеславович – профессор, доктор технических наук, профессор кафедры информационных систем и защиты информации Иркутского государственного университета путей сообщения (Иркутск, Россия). E-mail: 79149267772@yandex.ru

Куклина Ольга Константиновна – старший преподаватель кафедры информационных технологий и высшей математики Читинского института (филиала) Байкальского государственного университета (Чита, Россия). E-mail: kuklinaok@bgu-chita.ru

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

Information about the authors:

Krakovsky Yuri Mecheslavovich (Professor, Doctor of Engineering sciences, Professor of the Information Systems and Information Security Department, Irkutsk State University of Railway Transport, Irkutsk, Russian Federation). E-mail: 79149267772@yandex.ru

Kuklina Olga Konstantinovna (Senior Professor of the Information Technology and Higher Mathematics Department, Chita Institute of Baikal State University, Chita, Russian Federation). E-mail: kuklinaok@bgu-chita.ru

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

Поступила в редакцию 16.08.2022; принята к публикации 01.03.2023

Received 16.08.2022; accepted for publication 01.03.2023

Научная статья

УДК 519.233.22

doi: 10.17223/19988605/62/6

Максиминная задача оценивания параметров в условиях байесовского точечного засорения

Даниил Валерьевич Лисицин¹, Константин Викторович Гаврилов²^{1, 2} Новосибирский государственный технический университет, Новосибирск, Россия¹ lisitsin@ami.nstu.ru² k.gavrilov@corp.nstu.ru

Аннотация. Работа посвящена развитию теории устойчивого оценивания А.М. Шурыгина в части подхода, основанного на модели байесовского точечного засорения. Данный подход удобен для построения и анализа различных устойчивых M -оценок и по сравнению с классическими робастными процедурами предоставляет более широкие возможности. Вариационными методами получено решение максиминной задачи для наиболее широкого множества распределений засоряющей точки, что позволило установить единственность ранее найденного А.М. Шурыгиным решения. Также установлено, что функции, соответствующие решению задачи, определяют седловую точку функционала асимптотического квадратичного отклонения оценки.

Ключевые слова: M -оценки; байесовское точечное засорение; максиминная задача; медианная оценка; седловая точка.

Для цитирования: 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. № 62. С. 56–64. doi: 10.17223/19988605/62/6

Original article

doi: 10.17223/19988605/62/6

Maximin problem of parameter estimation in conditions of point Bayesian contamination

Daniil V. Lisitsin¹, Konstantin V. Gavrilov²^{1, 2} Novosibirsk State Technical University, Novosibirsk, Russian Federation¹ lisitsin@ami.nstu.ru² k.gavrilov@corp.nstu.ru

Abstract. The work is devoted to the development of the theory of stable estimation by A.M. Shurygin in terms of the approach based on the point Bayesian contamination model. This approach is convenient for constructing and analyzing various stable M -estimates and provides more opportunities compared to classical robust procedures. The maximin problem have been solved by variational methods for the widest set of contamination point distributions that allowed us to establish the uniqueness of the solution previously found by A.M. Shurygin. It is also established that the functions corresponding to the solution of the problem determine the saddle point of the functional of the asymptotic quadratic deviation of the estimate.

Keywords: M -estimates; point Bayesian contamination; maximin problem; median estimate; saddle point.

For citation: Lisitsin, D.V., Gavrilov, K.V. (2023) Maximin problem of parameter estimation in conditions of point Bayesian contamination. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 62. pp. 56–64. doi: 10.17223/19988605/62/6

Введение

Современные подходы к решению задачи оценивания параметров статистических моделей [1] позволяют обеспечить устойчивость получаемых решений к возможным отклонениям реальной ситуации от принятых предположений [2–5]. Примеры классических робастных решений для математического ожидания нормального распределения – оценка Хьюбера и выборочная медиана. Обе они являются оценками максимального правдоподобия, первая – для симметричной плотности, нормальной в центре и экспоненциально убывающей на «хвостах», вторая – для распределения Лапласа, также имеющего симметричную экспоненциально убывающую плотность.

Однако на практике «хвосты» могут быть более тяжелыми [4], кроме того, могут проявляться асимметричные искажения модельного распределения. А.М. Шурыгиным [3] была предложена довольно гибкая модель байесовского точечного засорения (БТЗ), которая позволяет решать в том числе и эти проблемы. Модель предполагает наличие серии выборок, каждая из которых имеет засорение в виде распределения, сосредоточенного в одной точке, являющейся фиксированной в пределах одной выборки, но имеющей некоторое распределение по серии выборок. Каждому распределению засоряющей точки соответствует некоторая наилучшая оценочная функция, полученная в результате минимизации функционала асимптотического квадратичного отклонения оценки [3].

Модель БТЗ служит теоретической основой для конструирования широкого спектра оптимальных устойчивых оценок, имеющих важное практическое значение, в том числе робастных решений, которые характеризуются заданием непараметрической («полноразмерной») окрестности модельного распределения [4, 6, 7]. Возможный метод построения оптимальных оценок заключается в выборе наилучшего в некотором смысле распределения засоряющей точки. В частности, *максиминный* подход [3] подразумевает поиск наилучшего распределения засоряющей точки для наилучшей оценочной функции относительно одного и того же функционала качества. В рамках модели БТЗ оказываются оптимальными известные робастные оценки параметра сдвига, например оценки Мешалкина, Эндрюса, Смита, Бернулли, бивес-оценка Тьюки, оценка Хьюбера типа урезанного среднего, обобщенные оценки Шарбонье [3, 8, 9].

А.М. Шурыгин в докторской диссертации [10] показал, что медианная оценка является непараметрическим максиминным решением для наиболее широкого множества распределений засоряющей точки, ограниченного лишь условиями регулярности. Частным случаем медианных оценок (при оценивании параметра сдвига симметричного распределения) выступает выборочная медиана [4, 5], которая при некоторых условиях является наиболее *B*-робастной и наиболее *V*-робастной оценкой, а также минимаксной относительно асимптотического смещения в модели засорения Хьюбера.

Однако предложенное А.М. Шурыгиным доказательство не отвечает на вопрос о единственности решения максиминной задачи. Вместе с тем он пишет о данной задаче, что найти наилучшую плотность распределения вариационными методами трудно из-за сложной зависимости от нее оптимизируемого функционала. Попытка получить медианную оценку путем прямого решения максиминной задачи, предпринятая А.М. Шурыгиным в [3] (теорема I.6.2), оказалась неудачной из-за ошибочного доказательства. Основная ошибка заключалась в том, что не учитывалась неявная зависимость знаменателя оптимизируемого функционала и параметра оценочной функции, обеспечивающего асимптотическую несмещенность оценки, от искомой плотности распределения засоряющей точки. В связи с этим мы восполняем данный пробел. В отличие от доказательства в [10], решение максиминной задачи вариационными методами позволяет установить единственность решения.

1. Элементы теории устойчивого оценивания

Пусть x_1, \dots, x_m – 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 [1, 2, 4] как решение оптимизационной задачи

$$\hat{\theta} = \arg \min_{\theta} \sum_{i=1}^m \rho(x_i, \theta),$$

где $\rho(x, \theta) : X \times \Theta \rightarrow R$ – непрерывная, дифференцируемая почти всюду функция. Условие равенства нулю производной по θ оптимизируемой функции служит альтернативной (хотя и не эквивалентной [1, 2]) формулировкой задачи M -оценивания:

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

называемой *оценочным уравнением*, где функция

$$\psi(x, \theta) = c(\theta) \dot{\rho}(x, \theta) \quad (1)$$

называется *оценочной функцией* для параметра θ , $c(\theta)$ – произвольная непрерывная функция, не равная нулю для всех $\theta \in \Theta$. Множитель $c(\theta)$ в оценочной функции задает семейство *эквивалентных оценочных функций*. Здесь и далее точкой сверху обозначено дифференцирование по оцениваемому параметру.

Оценочная функция также должна удовлетворять условию асимптотической несмещенности, которое принимает вид [1, 2, 11]:

$$\mathbf{E} \psi(\xi, \theta) = \int_X \psi(x, \theta) f(x, \theta) dx = 0, \quad (2)$$

где \mathbf{E} – оператор математического ожидания.

Дифференцируя (2) по θ и допуская возможность изменения порядка дифференцирования и интегрирования, можно записать следующие равенства [4, 11]:

$$N(\theta) = - \lim_{t \rightarrow \theta} \frac{df}{dt} \mathbf{E} \psi(\xi, t) = - \mathbf{E} \dot{\psi}(\xi, \theta) = \int_X \psi(x, \theta) \dot{f}(x, \theta) dx. \quad (3)$$

Условия регулярности. Потребуем, чтобы в некоторой окрестности истинного значения параметра θ :

- 1) выполнялось условие (2);
- 2) были справедливы равенства (3), функция $N(\theta)$ была непрерывной и не равной нулю [4];
- 3) выполнялось условие $\left| \lim_{t \rightarrow \theta} \frac{\partial^2}{\partial t^2} \mathbf{E} \psi(\xi, t) \right| < \infty$;
- 4) выполнялось условие $\mathbf{E} \psi^2(\xi, \theta) < \infty$.

При выписанных условиях M -оценка $\hat{\theta}$ является \sqrt{m} -состоятельной и асимптотически нормальной [1, 2, 12].

Рассмотрим подход Шурыгина, основанный на модели БТЗ [3, 13]. Пусть плотность распределения случайной величины имеет вид:

$$(1 - \alpha) f(x, \theta) + \alpha \delta(x - \nu),$$

где α – доля аномальных наблюдений, $0 \leq \alpha < 1$; δ – функция Дирака; ν – засоряющая точка, такая что в пределах одной выборки $\nu = \text{const}$, а в серии выборок представляет собой случайную величину, распределенную на X с плотностью $s = s(x, \theta)$.

Пусть α – бесконечно малая величина с порядком малости, меньшим $1/2$, т.е. $\alpha = \gamma m^{-\zeta/2}$, где $\gamma > 0$ и $0 < \zeta < 1$. Тогда асимптотическое квадратичное отклонение оценки

$$\gamma^{-2} \lim_{m \rightarrow \infty} m^{\zeta} \mathbf{E}(\hat{\theta} - \theta)^2$$

не зависит от параметров γ, ζ и определяется выражениями [3, 13]:

$$U(\psi, s, f) = \frac{1}{N^2(\theta)} \int_X \psi^2(x, \theta) s(x, \theta) dx = \int_X \mathbf{I}F^2(x; \psi, f) s(x, \theta) dx, \quad (4)$$

где $IF(x; \psi, f) = \psi(x, \theta)/N(\theta)$ – функция влияния Хампеля [5] для M -оценок. В [3] показано, что функционал (4) достигает минимума по ψ на функции

$$\psi_s(x, \theta) = c(\theta) \left[\frac{\partial}{\partial \theta} \ln f(x, \theta) + \beta(\theta) \right] \frac{f(x, \theta)}{s(x, \theta)}, \quad (5)$$

где функция $c = c(\theta)$ имеет тот же смысл, что и в (1), функция $\beta = \beta(\theta)$ определяется из условия (2). Выражение (5) определяет наилучшую оценочную функцию в модели БТЗ.

Для краткости далее часто будем опускать аргументы функций. Например, (5) можно записать в виде: $\psi_s = c(\dot{f} + \beta f)/s$.

2. Максиминная задача

Пусть S – некоторое множество плотностей $s(x, \theta)$, где $x \in X$, $\theta \in \Theta$. Максиминная оценочная функция ψ_* в модели БТЗ получается в результате решения оптимизационной задачи [3]

$$s_*(x, \theta) = \arg \max_{s \in S} \min_{\psi} U(\psi, s, f) = \arg \max_{s \in S} U(\psi_s, s, f) \quad (6)$$

и последующей подстановки (6) в (5). Здесь ψ_s определена выражением (5), функционал U определен в (4). Плотность (6) – это наихудшая плотность распределения засоряющей точки на множестве плотностей S .

Лемма 1. Задача (6) имеет следующую эквивалентную формулировку:

$$\Phi(s, f) = \int_X \frac{(\dot{f} + \beta f)^2}{s} dx \rightarrow \min_{s \in S}, \quad (7)$$

где функция β та же, что и в выражении (5).

Доказательство. Применим теорему 2 в [14], согласно следствию из которой асимптотическое квадратичное отклонение (4) для оценочной функции (5) может быть записано в виде:

$$U(\psi_s, s, f) = \frac{1}{N^2} \int_X \psi_s^2 s dx = c^2 / \int_X \psi_s^2 s dx,$$

Подставив сюда $\psi_s = c(\dot{f} + \beta f)/s$, убеждаемся, что $U(\psi_s, s, f) = 1/\Phi(s, f)$. Таким образом, задачи $U(\psi_s, s, f) \rightarrow \max_{s \in S}$ и $\Phi(s, f) \rightarrow \min_{s \in S}$ эквивалентны. Лемма доказана.

Далее нам понадобится результат, обобщающий лемму Дюбуа–Реймона [15, 16] на случай неограниченного промежутка.

Лемма 2. Пусть $X = (x_1; x_2)$ – ограниченный или неограниченный промежуток, для которого формально допускаются границы вида $x_1 = -\infty$ или $x_2 = +\infty$. Если для непрерывной, ограниченной

функции $b(x): X \rightarrow R$ и всех непрерывных функций $v(x): X \rightarrow R$, таких что $\int_{x_1}^{x_2} v(x) dx = 0$, выполня-

ется $\int_{x_1}^{x_2} b(x)v(x) dx = 0$, то $b(x) = \text{const}$.

Доказательство. Пусть $w(x) = \int_{x_1}^x v(t) dt$ – первообразная к функции $v(x)$. Нетрудно видеть, что

$w(x_1) = w(x_2) = 0$. Учитывая непрерывность и ограниченность функции $b(x)$, можно записать

$$\int_{x_1}^{x_2} b(x)v(x) dx = \int_{x_1}^{x_2} b(x) dw(x) = b(x)w(x)|_{x_1}^{x_2} - \int_{x_1}^{x_2} w(x) db(x) = - \int_{x_1}^{x_2} w(x) db(x) = 0.$$

Последнее равенство, очевидно, выполняется при $b(x) = \text{const}$. Покажем, что это единственно возможный вариант. Допустим, $b(x) \neq \text{const}$. Рассмотрим произвольный промежуток знакопостоянства $(a_1; a_2)$ функции $b'(x)$. Положим, что функция $w(x) > 0$ на $x \in (a_1; a_2)$, а вне этого промежутка $w(x) = 0$. При этом $w(x)$ является непрерывно дифференцируемой, что обеспечивает существование соответствующей $v(x)$. В качестве конкретного примера можно привести функцию

$$w(x) = \begin{cases} 1 + \cos \frac{\pi(2x - a_1 - a_2)}{a_2 - a_1}, & x \in (a_1; a_2); \\ 0, & x \notin (a_1; a_2). \end{cases}$$

Тогда, очевидно,

$$\int_{x_1}^{x_2} b(x)v(x) dx = - \int_{a_1}^{a_2} w(x)b'(x) dx \neq 0.$$

Пришли к противоречию, следовательно $b(x) = \text{const}$. Лемма доказана.

Рассмотрим медианную оценку с оценочной функцией [10]:

$$\Psi_{\text{med}} = c \operatorname{sgn}(\dot{f} + \beta f), \quad (8)$$

где функции c и β имеют тот же смысл, что и в (5). Покажем вариационными методами, что (8) является единственной максиминной оценочной функцией в модели БТЗ для наиболее широкого множества S (ограниченного лишь условиями регулярности).

Теорема. Пусть функции f и \dot{f} непрерывны по x , S – множество непрерывных функций плотности s , таких что $|\psi_s| < \infty$ равномерно на S . Тогда максиминная оценочная функция ψ_* единственная и соответствует медианной оценке с оценочной функцией (8).

Доказательство. Согласно лемме 1 оценочной функции ψ_* соответствует плотность s_* , являющаяся решением задачи (7). Таким образом, требуется найти экстремаль s_* функционала $\Phi(s, f)$ и соответствующую функцию (5), затем показать, что s_* доставляет минимум функционалу.

В соответствии с методом вариаций [16] введем обозначения:

$$s = s_* + t \delta s, \quad (9)$$

$\varphi(t) = \Phi(s, f)$ – функция, в которой все переменные, входящие в (7), кроме t , полагаются фиксированными. Здесь s_* – неизвестная наихудшая плотность (6), δs – допустимая вариация плотности s , т.е. не равная тождественно нулю непрерывная функция, такая что $\int_X \delta s dx = 0$, и $t \in [t_1; t_2]$ – число, такое что $s \geq 0$ для всех $x \in X$. Границы отрезка $[t_1; t_2]$ определяются в зависимости от δs в соответствии с выражениями

$$t_1 = \sup_x \inf \{t : s_* + t \delta s \geq 0\} \leq 0; \quad t_2 = \inf_x \sup \{t : s_* + t \delta s \geq 0\} \geq 0.$$

Поскольку в нулях плотности s (если они имеются) интегрант может иметь бесконечные разрывы, рассмотрим эту ситуацию подробнее. Плотность s имеет нули, во-первых, при $t = t_1$ и $t = t_2$; во-вторых, при всех $t \in [t_1; t_2]$, если нули имеет s_* и хотя бы один из них совпадает с нулем δs . В этих случаях регулярность выражения должна обеспечиваться нулями в соответствующих точках числителя интегранта.

Найдем производную функции $\varphi(t)$, используя условие (2) и учитывая, что функция β , вообще говоря, зависит от t через s [17]:

$$\begin{aligned} \varphi'(t) &= \int_X \left[2 \frac{\partial \beta}{\partial t} \frac{(\dot{f} + \beta f)f}{s} - \frac{(\dot{f} + \beta f)^2}{s^2} \delta s \right] dx = \\ &= \frac{2}{c} \frac{\partial \beta}{\partial t} \int_X \psi f dx - \int_X \frac{(\dot{f} + \beta f)^2}{s^2} \delta s dx = - \int_X \frac{(\dot{f} + \beta f)^2}{s^2} \delta s dx. \end{aligned}$$

Убедимся, что дифференцирование под знаком интеграла здесь возможно. Действительно, условие $|\psi_s| < \infty$ эквивалентно условию $(\dot{f} + \beta f)^2 / s^2 < \infty$, а из него следует условие $(\dot{f} + \beta f)^2 / s < \infty$. Это обеспечивает непрерывность подынтегральных выражений $\varphi(t)$ и $\varphi'(t)$ в точках x , где $s = 0$ (если они имеются). Кроме того, это условие обеспечивает равномерную сходимость интеграла $\varphi'(t)$ в случае неограниченного X . Непрерывность подынтегральных выражений по x также обеспечивается непрерывностью входящих в них функций, по t – видом зависимости (9).

Условие стационарности функционала имеет вид:

$$\delta\Phi(s_*, f) = \varphi'(0) = - \int_X \frac{(\dot{f} + \beta f)^2}{s_*^2} \delta s \, dx = 0.$$

Для данного выражения выполнены условия леммы 2, поэтому

$$(\dot{f} + \beta f)^2 / s_*^2 = \lambda = \text{const}$$

или, учитывая неотрицательность s ,

$$s_* = |\dot{f} + \beta f| / \sqrt{\lambda}. \quad (10)$$

Подставив найденную плотность s_* в (5), получаем выражение (8), где функция c поглощает λ . Таким образом, медианная оценка соответствует экстремали функционала $\Phi(s, f)$. Обратим внимание, что решение (10) удовлетворяет условиям регулярности.

Покажем теперь, что найденная экстремаль s_* является решением задачи (7) на множестве S . Для этого определим знак второй производной функции $\varphi(t)$, продифференцировав $\varphi'(t)$. Внесение операции дифференцирования под знак интеграла здесь произведем формально, поскольку нас интересует лишь знак второй производной. Действительно, положим $\varphi'(t) = \int_X y(x, t) dx$. Если подынте-

гральная функция $y(x, t)$ возрастает по t для каждого фиксированного x , т.е. $\partial y(x, t) / \partial t > 0$, то, очевидно, возрастает и $\varphi'(t)$, т.е. $\varphi''(t) > 0$. При поиске $\varphi''(t)$ не будем отбрасывать нулевое слагаемое в $\varphi'(t)$ и опять воспользуемся условием (2):

$$\begin{aligned} \varphi''(t) &= \frac{d}{dt} \left[\frac{2}{c} \frac{\partial \beta}{\partial t} \int_X \psi f \, dx - \int_X \frac{(\dot{f} + \beta f)^2}{s^2} \delta s \, dx \right] = \\ &= \frac{2}{c} \frac{\partial^2 \beta}{\partial t^2} \int_X \psi f \, dx + 2 \frac{\partial \beta}{\partial t} \int_X \left(\frac{f^2}{s} \frac{\partial \beta}{\partial t} - \frac{\dot{f} + \beta f}{s^2} f \delta s \right) dx + \\ &\quad + 2 \int_X \left[\frac{(\dot{f} + \beta f)^2}{s^3} (\delta s)^2 - \frac{\partial \beta}{\partial t} \frac{\dot{f} + \beta f}{s^2} f \delta s \right] dx = \\ &= 2 \int_X \left[\frac{f^2}{s} \left(\frac{\partial \beta}{\partial t} \right)^2 - 2 \frac{\partial \beta}{\partial t} \frac{\dot{f} + \beta f}{s^2} f \delta s + \frac{(\dot{f} + \beta f)^2}{s^3} (\delta s)^2 \right] dx = \\ &= \int_X \frac{2}{s} \left(f \frac{\partial \beta}{\partial t} - \frac{\dot{f} + \beta f}{s} \delta s \right)^2 dx > 0. \end{aligned}$$

Действительно, $\varphi''(t) = 0$, только если подынтегральное выражение данной функции тождественно равно нулю, но в этом случае $\partial \beta / \partial t = (\dot{f} + \beta f) \delta s / (f s)$, где правая часть равенства есть функция, зависящая от x , а левая – не зависящая, следовательно, это равенство не может являться тождеством.

Запишем разложение $\varphi(t)$ в ряд Маклорена по степеням t с остаточным членом в форме Лагранжа:

$$\varphi(t) = \varphi(0) + t \varphi'(0) + \frac{t^2}{2} \varphi''(\omega t),$$

где ω – некоторое число, $0 < \omega < 1$. Поскольку для всех допустимых вариаций δs имеют место $\varphi'(0) = 0$ и $\varphi''(\omega t) > 0$ для всех $t \in [t_1; t_2]$, экстремальная задача (7) без учета условий регулярности является выпуклой со строго выпуклым оптимизируемым функционалом. Отсюда следует $\Phi(s, f) \geq \Phi(s_*, f)$, причем равенство достигается только при $s = s_*$. То есть единственным решением данной экстремальной задачи [18] является плотность (10). Теорема доказана.

В [6] показано, что медианная оценка (8) является также минимаксным решением в модели БТЗ:

$$\psi_{\text{med}} = \arg \min_{\psi} \max_{s \in S} U(\psi, s, f),$$

где S – множество плотностей, ограниченное лишь условиями регулярности. При этом значению $\max_{s \in S} U(\psi, s, f) = \max_{x \in X} (\psi/N)^2$ соответствует распределение засоряющей точки, сосредоточенное в точках максимума $|\psi/N|$, где функция N определена в (3). Однако, поскольку $|\psi_{\text{med}}| = \text{const}$ почти всюду на $x \in X$, нетрудно видеть (см. также [10]), что значение функционала (4) при $\psi = \psi_{\text{med}}$ не зависит от плотности s , т.е. он достигает своего максимального значения $\max_{x \in X} (\psi_{\text{med}}/N)^2$ для любой непрерывной плотности s . По этой причине из минимаксной формулировки медианной оценки плотность s_* однозначно не определяется, но может быть получена в результате решения максиминной задачи (6). Решению $\psi_* = \psi_{\text{med}}$ соответствует плотность (10).

Таким образом, функционал (4) имеет седловую точку (ψ_*, s_*) , которой соответствует медианная оценка (8), так что удовлетворяется условие

$$U(\psi_*, s, f) = U(\psi_*, s_*, f) \leq U(\psi, s_*, f)$$

для любых допустимых ψ и s .

С точки зрения теории игр седловая точка является точкой равновесия в чистых стратегиях антагонистических игр [19]. Ей соответствует такая ситуация, при которой ни один из игроков не может увеличить свой выигрыш, меняя решение в одностороннем порядке. При этом стратегия обоих игроков является наилучшей реакцией на действия своего оппонента. Положим, что оппонентом исследователя выступает природа, которая воздействует на оценки через плотность s . Выбором $\psi = \psi_*$ исследователь оказывается застрахованным от любой плотности s : какая бы из них ни реализовалась в природе, это не ухудшит качество оценки (правда, и не улучшит). С другой стороны, вместо незаинтересованной природы исследователь может столкнуться с намеренным искажением данных. В этом случае можно считать, что оппонент действует наихудшим для исследователя образом, зная, что тот стремится выбрать наилучшую оценочную функцию. При такой стратегии окажется $s = s_*$, а исследователь выбором $\psi = \psi_*$ получит заведомо лучшую оценку, чем при любом другом выборе.

Заключение

Модель БТЗ является достаточно универсальным инструментом исследователя, позволяющим конструировать различные устойчивые оценочные функции, оптимальные для тех или иных условий. В том числе на основе данной модели можно получить уже известные робастные решения, например выборочную медиану, оптимальную в классических теориях робастности. Выборочная медиана – это частный случай медианной оценки, предложенной А.М. Шурыгиным [3]. В работах [6] и [10] показано, что медианная оценка является соответственно минимаксным и максиминным решением в модели БТЗ при наиболее слабых ограничениях на множество распределений засоряющей точки.

Между тем оставался открытым вопрос единственности решения максиминной задачи. Для ответа на него задача в данной работе была решена вариационными методами, и в результате установлена единственность решения. Показано, что функции, соответствующие решению задачи, определяют седловую точку функционала асимптотического квадратичного отклонения оценки. Дана интерпретация

решения с точки зрения теории игр, где седловые точки играют важную роль. В качестве дополнительного результата получено обобщение леммы Дюбуа–Реймона [15, 16] на случай неограниченного промежутка.

Список источников

1. Боровков А.А. Математическая статистика. СПб. : Лань, 2021. 704 с.
2. Shulenin V.P. Robust methods of mathematical statistics. Tomsk : Scientific Technology Publishing House, 2020. 260 p.
3. Шурыгин А.М. Прикладная стохастика: робастность, оценивание, прогноз. М. : Финансы и статистика, 2000. 224 с.
4. Хьюбер П. Робастность в статистике. М. : Мир, 1984. 303 с.
5. Хампель Ф., Рончетти Э., Рауссеу П., Штаэль В. Робастность в статистике: подход на основе функций влияния. М. : Мир, 1989. 512 с.
6. Лисицин Д.В. Оценивание при байесовском точечном засорении: связь с подходом Хампеля и минимаксная оценка // Сборник научных трудов НГТУ. 2011. Вып. 3 (65). С. 61–66.
7. Shevlyakov G., Morgenthaler S., Shurygin A. Redescending M-estimators // J. Statist. Plann. Inference. 2008. V. 138. P. 2906–2917.
8. Лисицин Д.В., Гаврилов К.В. Оценивание параметров распределения ограниченной случайной величины, робастное к нарушению границ // Научный вестник НГТУ. 2016. № 2 (63). С. 70–89.
9. Lisitsin D.V., Usoltsev A.G. Minimum gamma-divergence estimation for non-homogeneous data with application to ordered probit model // Applied methods of statistical analysis. Statistical computation and simulation. Proceedings of the Int. Workshop. Novosibirsk, 18–20 Sept. 2019. Novosibirsk : NSTU, 2019. P. 227–234.
10. Шурыгин А.М. Асимптотическая теория устойчивого оценивания : дис. ... д-ра техн. наук. М., 2002. 225 с.
11. Смоляк С.А., Титаренко Б.П. Устойчивые методы оценивания: статистическая обработка неоднородных совокупностей. М. : Статистика, 1980. 208 с.
12. Van der Vaart A.W. Asymptotic Statistics. Cambridge : Cambridge University Press, 1998. 443 p.
13. Shurygin A.M. New Approach to Optimization of Stable Estimation // Proceedings of the First US/Japan Conference on the Frontiers of Statistical Modeling: An Informational Approach. Dordrecht: Kluwer, 1994. V. 3. P. 315–340.
14. Лисицин Д.В., Гаврилов К.В. Об устойчивом оценивании параметров модели при асимметричном засорении данных // Научный вестник НГТУ. 2008. № 1 (32). С. 33–40.
15. Du Bois-Reymond P. Erläuterungen zu den Anfangsgründen der Variationsrechnung // Math. Ann. 1979. V. 15. P. 283–314.
16. Габасов Р. и др. Методы оптимизации : пособие. Минск : Четыре четверти, 2011. 472 с.
17. Лисицин Д.В., Гаврилов К.В. О некоторых свойствах M-оценок // Сборник научных трудов НГТУ. 2011. Вып. 2 (64). С. 61–68.
18. Магарил-Ильев Г.Г., Тихомиров В.М. Выпуклый анализ и его приложения. М. : Едиториал УРСС, 2003. 175 с.
19. Петросян Л.А., Зенкевич Н.А., Шевкопляс Е.В. Теория игр. СПб. : БХВ-Петербург, 2012. 432 с.

References

1. Borovkov, A.A. (2021) *Matematicheskaya statistika* [Mathematical Statistics]. St. Petersburg: Lan'.
2. Shulenin, V.P. (2020) *Robust methods of mathematical statistics*. Tomsk: Scientific Technology Publishing House.
3. Shurygin, A.M. (2000) *Prikladnaya stokhastika: robastnost', otsenivanie, prognoz* [Applied stochastics: robustness, estimation, forecast]. Moscow: Finansy i statistika.
4. Huber, P. (1984) *Robastnost' v statistike* [Robust Statistics]. Translated from English. Moscow: Mir.
5. 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.
6. Lisitsin, D.V. (2011) Otsenivanie pri bayesovskom tochechnom zasoreni: svyaz' s podkhodom Khampelya i minimaksnaya otsenka [Estimating in presence of Bayesian dot contamination: Connection with Hampel's approach and minimax estimator]. *Sbornik nauchnykh trudov NGTU*. 3(65). pp. 61–66.
7. Shevlyakov, G., Morgenthaler, S. & Shurygin, A. (2008) Redescending M-estimators. *Journal of Statist. Plann. Inference*. 138. pp. 2906–2917.
8. Lisitsin, D.V. & Gavrilov, K.V. (2016) Otsenivanie parametrov raspredeleniya ogranichennoy sluchaynoy velichiny, robastnoe k narusheniyu granits [Estimation of distribution parameters of a bounded random variable robust to bound disturbance]. *Nauchnyy vestnik NGTU*. 2(63). pp. 70–89.
9. Lisitsin, D.V. & Usoltsev, A.G. (2019) Minimum gamma-divergence estimation for non-homogeneous data with application to ordered probit model. *Applied Methods of Statistical Analysis. Statistical Computation and Simulation*. Proc. of the International Workshop. Novosibirsk, September 18–20, 2019. Novosibirsk: NSTU. pp. 227–234.
10. Shurygin, A.M. (2002) *Asimptoticheskaya teoriya ustoychivogo otsenivaniya* [Asymptotic Theory of Stable Estimation]. Engineering Science. Dr. Diss. Moscow.
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. 443 p.
13. Shurygin, A.M. (1994) New Approach to Optimization of Stable Estimation. In: *Proceedings of the First US/Japan Conference on the Frontiers of Statistical Modeling: An Informational Approach*. Vol. 3. Dordrecht: Kluwer. pp. 315–340.
14. Lisitsin, D.V. & Gavrilov, K.V. (2008) Ob ustoychivom otsenivanii parametrov modeli pri asimmetrichnom zasoreнии dannykh [On stable estimation of models parameters in presence of asymmetric data contamination]. *Nauchnyy vestnik NGTU*. 1(32). pp. 33–40.
15. Du Bois-Reymond, P. (1979) Erläuterungen zu den Anfangsgründen der Variationsrechnung. *Mathematische Annalen*. 15. pp. 283–314.
16. Gabasov, R. et al. (2011) *Metody optimizatsii* [Optimization Methods]. Minsk: Chetyre chetverti.
17. Lisitsin, D.V. & Gavrilov, K.V. (2011) O nekotorykh svoystvakh M-otsenok [About some properties of M-estimates]. *Sbornik nauchnykh trudov NGTU*. 2(64). pp. 61–68.
18. Magaril-Ilyayev, G.G. & Yikhomirov, V.M. (2003) *Vypuklyy analiz i ego prilozheniya* [Convex Analysis and its Applications]. Moscow: Editorial URSS.
19. Petrosyan, L.A., Zenkevich, N.A. & Shevkoplyas, E.V. (2012) *Teoriya igr* [Game Theory]. St. Petersburg: BHV-Peterburg.

Информация об авторах:

Лиситин Даниил Валерьевич – профессор, доктор технических наук, профессор кафедры теоретической и прикладной информатики Новосибирского государственного технического университета (Новосибирск, Россия). E-mail: lisitsin@ami.nstu.ru
Гаврилов Константин Викторович – кандидат технических наук, доцент кафедры автоматизации Новосибирского государственного технического университета (Новосибирск, Россия). E-mail: k.gavrilov@corp.nstu.ru

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

Information about the authors:

Lisitsin Daniil V. (Doctor of Technical Sciences, Professor, Professor of the department of Theoretical and Applied Computer Science, Novosibirsk State Technical University, Novosibirsk, Russian Federation). E-mail: lisitsin@ami.nstu.ru
Gavrilov Konstantin V. (Candidate of Technical Sciences, Associate Professor of the department of Automatics, Novosibirsk State Technical University, Novosibirsk, Russian Federation). E-mail: k.gavrilov@corp.nstu.ru

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

Поступила в редакцию 08.06.2022; принята к публикации 01.03.2023

Received 08.06.2022; accepted for publication 01.03.2023

Научная статья

УДК 519.2

doi: 10.17223/19988605/62/7

Применение метода моментов для оценки длительности мертвого времени в рекуррентном обобщенном полусинхронном потоке событий с продлевающимся мертвым временем

Людмила Алексеевна Нежелская¹, Илья Денисович Степаненко²

^{1,2} Томский государственный университет, Томск, Россия

¹ ludne@mail.ru

² 97step@mail.ru

Аннотация. Исследуется обобщенный полусинхронный поток событий с двумя состояниями, относящийся к классу дважды стохастических потоков. Его функционирование происходит в условиях продлевающегося мертвого времени фиксированной длительности. Разработан алгоритм оценивания длительности мертвого времени. Приводятся результаты статистических экспериментов.

Ключевые слова: дважды стохастический обобщенный полусинхронный поток событий; продлевающееся мертвое время фиксированной длительности; плотность вероятности; совместная плотность вероятности; условия рекуррентности; рекуррентный поток; преобразование Лапласа; уравнение моментов; оценка длительности мертвого времени.

Для цитирования: Нежелская Л.А., Степаненко И.Д. Применение метода моментов для оценки длительности мертвого времени в рекуррентном обобщенном полусинхронном потоке событий с продлевающимся мертвым временем // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2023. № 62. С. 65–75. doi: 10.17223/19988605/62/7

Original article

doi: 10.17223/19988605/62/7

An application of the method of moments to estimation of the dead time duration in recurrent generalized semi-synchronous flow of events with prolonged dead time

Lyudmila A. Nezhel'skaya¹, Ilya D. Stepanenko²

^{1,2} Tomsk State University, Tomsk, Russian Federation

¹ ludne@mail.ru

² 97step@mail.ru

Abstract. In this article we investigate a generalized semi-synchronous flow of events with two states and which is related to the class of double stochastic flows. The functioning of flow occurs in conditions of the prolonged dead time of fixed duration. An algorithm for estimating the duration of dead time has been developed. The results of statistical experiments are presented.

Keywords: generalized semi-synchronous double stochastic flow of events; prolonged dead time of fixed duration; probability density; joint probability density; flow recurrence conditions; recurrent flow; Laplace transform; equation of moments; estimation of the dead time duration.

For citation: Nezhel'skaya, L.A., Stepanenko, I.D. (2023) An application of the method of moments to estimation of the dead time duration in recurrent generalized semi-synchronous flow of events with prolonged dead time. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 62. pp. 65–75. doi: 10.17223/19988605/62/7

Введение

Теория массового обслуживания (ТМО) является составной частью исследования операций. С развитием информационных и телекоммуникационных технологий появились важные прикладные задачи ТМО – проектирование и создание информационно-вычислительных сетей, спутниковых сетей связи, телекоммуникационных сетей и т.п., что повлекло, в свою очередь, появление задач, связанных с исследованием и оптимизацией протекающих в разного рода сетях процессов [1, 2].

Первым потоком для исследования и применения был простейший поток событий. Однако в настоящее время в той или иной степени данная модель является непригодной для описания реальных информационных потоков сообщений (запросов, заявок, событий) в телекоммуникационных сетях из-за разнородности передаваемых данных и их взаимной коррелированности [2]. Таким образом, требования практики послужили стимулом к рассмотрению дважды стохастических потоков (коррелированных потоков) [2–6], которые являются широко используемой математической моделью реальных потоков сообщений в телекоммуникационных системах и сетях, глобальных компьютерных сетях, спутниковых сетях связи.

Большинством авторов работ по ТМО исследуются математические модели потоков событий, когда все события потока доступны наблюдению. В реальности же зарегистрированное событие может создать период мертвого времени для регистрирующего прибора (период ненаблюдаемости) [7], в течение которого другие события потока становятся недоступными для регистрирующего прибора (теряются).

Можно считать, что мертвое время выступает искажающим фактором при решении задач оценивания как состояний потока, так и его параметров.

В данной работе рассматривается обобщенный полусинхронный поток, функционирующий в условиях продаваемого мертвого времени фиксированной длительности (далее – поток). Получены условия рекуррентности, после чего решается уравнение моментов для нахождения оценки длительности мертвого времени в рекуррентном потоке. Приведены результаты статистических экспериментов по оцениванию длительности мертвого времени, поставленных на имитационной модели рассматриваемого потока.

1. Математическая модель потока. Постановка задачи

Рассматривается дважды стохастический обобщенный полусинхронный поток событий, сопровождающий процесс которого есть кусочно-постоянный принципиально ненаблюдаемый случайный процесс $\lambda(t)$ с двумя состояниями [8]: если $\lambda(t) = \lambda_1$, будем говорить, что имеет место первое состояние (S_1) процесса $\lambda(t)$ (потока), если $\lambda(t) = \lambda_2$, то второе состояние (S_2) процесса $\lambda(t)$ (потока), где $\lambda_1 > \lambda_2 \geq 0$. В течение временного интервала, когда $\lambda(t) = \lambda_i$, имеет место пуассоновский поток событий с параметром λ_i , $i = 1, 2$. Переход из первого состояния процесса $\lambda(t)$ во второе ($S_1 \rightarrow S_2$) возможен только в момент наступления события, при этом переход осуществляется с вероятностью p ($0 < p \leq 1$), с вероятностью $1 - p$ процесс $\lambda(t)$ остается в первом состоянии. Тогда длительность временного интервала, на котором значение процесса $\lambda(t) = \lambda_1$ – участка стационарности процесса $\lambda(t)$ в первом состоянии – есть случайная величина с экспоненциальной функцией распределения $F_1(t) = 1 - e^{-p\lambda_1 t}$, $t \geq 0$ [6]. При этом длительность пребывания процесса $\lambda(t)$ во втором состоянии распределена по экспоненциальному закону $F_2(t) = 1 - e^{-\alpha_2 t}$, $t \geq 0$. Переход из второго состояния процесса $\lambda(t)$ в первое состояние ($S_2 \rightarrow S_1$) осуществляется в произвольный момент времени, не связанный с моментом наступления события потока. При переходе процесса $\lambda(t)$ из второго состояния в первое инициируется с вероятностью δ ($0 \leq \delta \leq 1$) дополнительное событие в первом состоянии

(т.е. сначала осуществляется переход, а затем инициируется дополнительное событие; отметим, что переход и наступление события происходят мгновенно). Матрицы интенсивностей переходов (инфинитезимальных характеристик процесса $\lambda(t)$) принимают вид [6]:

$$\mathbf{D}_0 = \begin{vmatrix} -\lambda_1 & 0 \\ (1-\delta)\alpha_2 & -(\lambda_2 + \alpha_2) \end{vmatrix}, \quad \mathbf{D}_1 = \begin{vmatrix} (1-p)\lambda_1 & p\lambda_1 \\ \delta\alpha_2 & \lambda_2 \end{vmatrix}.$$

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

Будем рассматривать обобщенный полусинхронный поток при его неполной наблюдаемости, т.е. когда не все события потока доступны наблюдению. Каждое зарегистрированное в потоке событие порождает период ненаблюдаемости фиксированной длительности T (мертвое время), другие события, произошедшие в этот период, недоступны наблюдению (теряются). Хотя события и не наблюдаются в течение периода мертвого времени, каждое из них вызывает продление периода ненаблюдаемости на ту же величину T . Следующее наблюдаемое событие регистрируется после окончания последнего периода ненаблюдаемости и снова порождает период мертвого времени длительности T . Таким образом, общий период ненаблюдаемости является случайной величиной.

На рис. 1 представлена одна из возможных реализаций процесса $\lambda(t)$ и наблюдаемого потока, где 1, 2 – состояния случайного процесса $\lambda(t) = \lambda_i, i = 1, 2$; $\alpha = \alpha_2$; черные кружки – ненаблюдаемые события; ξ – значения периодов ненаблюдаемости потока; t_1, t_2, \dots – моменты наступления событий в наблюдаемом потоке.

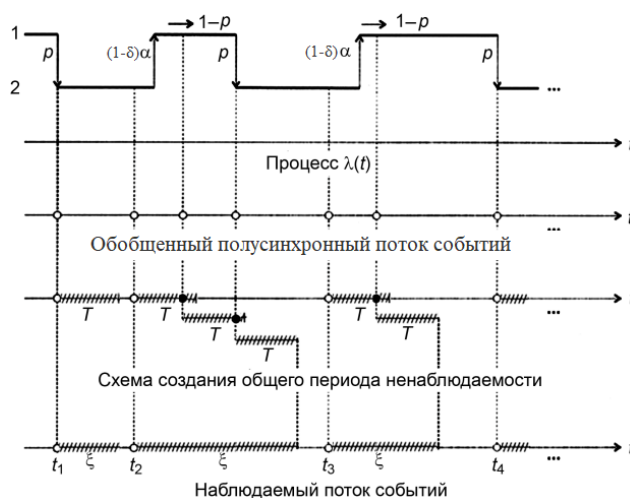


Рис. 1. Реализация обобщенного полусинхронного потока с продлевающимся мертвым временем фиксированной длительности T

Fig. 1. Realization of Generalized semi-synchronous flow with an extended dead time of fixed duration T

Утверждение 1.1. Для обобщенного полусинхронного потока событий сопровождающий случайный процесс $\lambda(t)$ является марковским.

Утверждение 1.2. Моменты наступления событий $t_1, t_2, \dots, t_k, \dots$ в наблюдаемом потоке порождают вложенную цепь Маркова $\{\lambda(t_k)\}$.

Цель работы – вывод уравнения моментов для оценки длительности мертвого времени и проведение ряда статистических экспериментов для установления качества полученных результатов. Настоящая статья является продолжением [9].

2. Плотность вероятности длительности интервала между соседними событиями в потоке

В данном разделе приводится явная формула плотности вероятности длительности интервала между соседними событиями в обобщенном полусинхронном потоке в условиях полной наблюдаемости за потоком [8].

Рассматривается обобщенный полусинхронный поток событий в условиях полной наблюдаемости, $T = 0$, математическое описание которого приведено в разд. 1.

Обозначим $\tau_k = t_{k+1} - t_k$, $k = 1, 2, \dots$, – значение длительности k -го интервала между соседними событиями потока, $p(\tau_k)$ – плотность вероятности длительности k -го интервала между соседними событиями в рассматриваемом потоке. Так как рассматривается стационарный режим функционирования потока, то $p(\tau_k) = p(\tau)$ для всех $k = 1, 2, \dots$, $\tau \geq 0$. Поэтому без ограничения общности момент наступления события t_k можно положить равным нулю, т.е. момент наступления события есть $\tau = 0$.

Теорема 2.1. Плотность вероятности длительности интервала между соседними событиями в обобщенном полусинхронном потоке событий в случае $\lambda_1 - (\lambda_2 + \alpha_2) \neq 0$ имеет вид [8]:

$$p(\tau) = \gamma z_1 e^{-z_1 \tau} + (1 - \gamma) z_2 e^{-z_2 \tau}, \quad \tau \geq 0, \quad \gamma = 1 - \pi_2(0) \frac{\lambda_1 - (\lambda_2 + \delta \alpha_2)}{\lambda_1 - (\lambda_2 + \alpha_2)}, \quad (1)$$

где $z_1 = \lambda_1$, $z_2 = \lambda_2 + \alpha_2$, $\pi_2(0) = \frac{p(\lambda_2 + \alpha_2)}{p(\lambda_2 + \delta \alpha_2) + \alpha_2}$.

3. Совместная плотность вероятности длительностей двух смежных интервалов

В силу стационарности функционирования потока расположение одного интервала (t_k, t_{k+1}) либо двух смежных интервалов (t_k, t_{k+1}) , (t_{k+1}, t_{k+2}) между моментами наступления событий потока $t_1, t_2, \dots, t_k, \dots$ на временной оси может быть произвольным. Поэтому рассмотрим два соседних временных интервала (t_1, t_2) и (t_2, t_3) со значениями длительностей интервалов $\tau_1 = t_2 - t_1$ и $\tau_2 = t_3 - t_2$ соответственно, $\tau_1 \geq 0$, $\tau_2 \geq 0$. Соответствующая совместная плотность вероятности при $T = 0$ есть $p(\tau_1, \tau_2)$, $\tau_1 \geq 0$, $\tau_2 \geq 0$.

Теорема 3.1. Обобщенный полусинхронный поток событий является коррелированным, и совместная плотность вероятности длительностей смежных интервалов для случая $\lambda_1 - (\lambda_2 + \alpha_2) \neq 0$ имеет вид [10]:

$$p(\tau_1, \tau_2) = p(\tau_1)p(\tau_2) + \gamma(1 - \gamma) \frac{(\lambda_2 - p(\lambda_2 + \delta \alpha_2))}{\lambda_2 + \alpha_2} \left[z_1 e^{-z_1 \tau_1} - z_2 e^{-z_2 \tau_1} \right] \left[z_1 e^{-z_1 \tau_2} - z_2 e^{-z_2 \tau_2} \right], \quad (2)$$

где $\tau_1 \geq 0$, $\tau_2 \geq 0$, γ , z_1 , z_2 , $p(\tau_k)$ определены в (1) для $\tau = \tau_k$, $k = 1, 2$.

3.1. Условие рекуррентности потока

Рассмотрим случай, когда поток становится рекуррентным [10]. Анализируя выражение (2), выпишем одно из условий факторизации совместной плотности $\lambda_2 - p(\lambda_2 + \delta \alpha_2) = 0$, $p \neq 1$, и соответствующий этому условию вид одномерной плотности (1):

$$p(\tau) = \gamma z_1 e^{-z_1 \tau} + (1 - \gamma) z_2 e^{-z_2 \tau}, \quad \tau \geq 0, \quad \gamma = 1 - p \frac{\lambda_1 - (\lambda_2 + \delta \alpha_2)}{\lambda_1 - (\lambda_2 + \alpha_2)}, \quad z_1 = \lambda_1, \quad z_2 = \lambda_2 + \alpha_2.$$

При остальных условиях рекуррентности ($\gamma(1 - \gamma) = 0$) поток вырождается в простейший.

4. Оценивание длительности мертвого времени методом моментов

Оценка длительности мертвого времени в рекуррентном обобщенном полусинхронном потоке событий в условиях продевающегося мертвого времени производится методом моментов [11]. Для получения уравнения моментов воспользуемся преобразованием Лапласа.

Пусть ξ – значение длительности общего периода ненаблюдаемости в рекуррентном обобщенном полусинхронном потоке событий. Последовательность моментов наступления событий наблюдаемого потока t_1, t_2, \dots образует вложенную цепь Маркова $\{\lambda(t_k)\}$, и рекуррентность наблюдаемого потока сохранится.

Для нахождения преобразования Лапласа $g_\xi(s)$ плотности вероятности общего периода ненаблюдаемости рекуррентного обобщенного полусинхронного потока событий воспользуемся теоремой о преобразовании Лапласа $g_\xi(s)$ плотности вероятности общего периода ненаблюдаемости для произвольного рекуррентного дважды стохастического потока событий [12].

Теорема 4.1. Преобразование Лапласа $g_\xi(s)$ плотности вероятности значений длительности общего периода ненаблюдаемости в рекуррентном дважды стохастическом потоке событий, функционирующем в условиях продевающегося мертвого времени, имеет вид:

$$g_\xi(s) = \frac{\varphi_0(T)}{e^{sT}} \left[1 - \int_0^T e^{-sx} \tilde{p}(x) dx \right]^{-1}, \quad (3)$$

где $\varphi_0(T) = \int_0^T \tilde{p}(x) dx = 1 - \int_0^T \tilde{p}(x) dx$ – функция Пальма – вероятность того, что на интервале $(0, T)$ событие потока не наступило; $\tilde{p}(x)$ – плотность вероятности длительности интервала между соседними событиями в рекуррентном потоке.

Следствие теоремы 4.1. Математическое ожидание случайной величины ξ определяется в виде:

$$M\xi = -g'_\xi(s)|_{s=0} = T + \frac{1}{\varphi_0(T)} \int_0^T x \tilde{p}(x) dx. \quad (4)$$

В качестве условия рекуррентности в дальнейшем изложении используется только случай $\lambda_2 = p(\lambda_2 + \delta\alpha_2)$, $0 < p < 1$.

Теорема 4.2. Преобразование Лапласа $g_\xi(s)$ плотности вероятности общего периода ненаблюдаемости рекуррентного обобщенного полусинхронного потока событий, функционирующего в условиях продевающегося мертвого времени, имеет вид:

$$g_\xi(s) = \frac{\varphi_0(T)}{e^{sT}} \left[1 - \frac{\gamma z_1 (1 - e^{-(z_1+s)T})}{(z_1+s)} - \frac{(1-\gamma)z_2 (1 - e^{-(z_2+s)T})}{(z_2+s)} \right]^{-1}, \quad (5)$$

где z_1, z_2, γ определены в разд. 3.1; $\varphi_0(T) = \gamma e^{-z_1 T} + (1-\gamma)e^{-z_2 T}$.

Доказательство. Применяя к плотности (1) условие рекуррентности $\lambda_2 = p(\lambda_2 + \delta\alpha_2)$, $0 < p < 1$, переобозначив в (3) $\tilde{p}(x)$ на $p(\tau)$, определенную в разд. 3.1, находим

$$\int_0^T e^{-s\tau} p(\tau) d\tau = \int_0^T e^{-s\tau} (\gamma z_1 e^{-z_1 \tau} + (1-\gamma)z_2 e^{-z_2 \tau}) d\tau = \frac{\gamma z_1}{z_1+s} + \frac{(1-\gamma)z_2}{z_2+s} - \frac{\gamma z_1 e^{-(z_1+s)T}}{z_1+s} - \frac{(1-\gamma)z_2 e^{-(z_2+s)T}}{z_2+s}. \quad (6)$$

Функция Пальма примет вид:

$$\begin{aligned} \varphi_0(T) &= \int_0^T p(\tau) d\tau = 1 - \int_0^T p(\tau) d\tau = 1 - \int_0^T \gamma z_1 e^{-z_1 \tau} d\tau - \int_0^T (1-\gamma)z_2 e^{-z_2 \tau} d\tau = \\ &= 1 - \left[\gamma - \gamma e^{-z_1 T} + (1-\gamma) - (1-\gamma)e^{-z_2 T} \right] = \gamma e^{-z_1 T} + (1-\gamma)e^{-z_2 T}. \end{aligned} \quad (7)$$

Подставляя (6) и (7) в (3), получим (5). Теорема доказана.

Следствие теоремы 4.2. Математическое ожидание общего периода ненаблюдаемости рекуррентного обобщенного полусинхронного потока событий, функционирующего в условиях продлевающегося мертвого времени, имеет вид:

$$M\xi = \frac{z_2\gamma(1 - e^{-z_1T}) + z_1(1 - \gamma)(1 - e^{-z_2T})}{z_1z_2\phi_0(T)}, \quad (8)$$

где $\phi_0(T)$, z_1 , z_2 определены в (5).

Доказательство. Используя следствие теоремы 4.1 и вычислив производную от $g_\xi(s)$, определенную в (5), в точке $s=0$, по формуле (4) получим (8). Следствие теоремы доказано.

Рассмотрим временной интервал (t_k, t_{k+1}) , значение длительности которого есть $\tau_k = t_{k+1} - t_k$. Введем случайную величину η_k , которая имеет смысл длительности интервала между моментом окончания общего периода ненаблюдаемости и моментом наступления следующего события t_{k+1} . Таким образом, получим следующее равенство для значений случайной величины: $\eta_k = \tau_k - \xi_k$. Так как в данной работе рассматривается рекуррентный поток, то $\tau_k, k=1, 2, \dots$, – независимые случайные величины, поэтому равенство можно переписать в виде: $\eta = \tau - \xi$; здесь индекс k опущен, поскольку рассматривается произвольный интервал времени между соседними событиями в наблюдаемом потоке. Случайные величины η и ξ являются зависимыми. Тогда плотность вероятности длительности интервала между соседними событиями в рекуррентном наблюдаемом потоке примет вид:

$$p(\tau) = \int_0^\tau p(\xi)p(\eta|\xi)d\xi = \int_0^\tau p(\xi)p(\tau - \xi|\xi)d\xi. \quad (9)$$

Найдем преобразование Лапласа $g_\tau(s)$ плотности вероятности $p(\tau)$.

Теорема 4.3. Преобразование Лапласа $g_\tau(s)$ плотности вероятности значений длительности интервала между соседними событиями в рекуррентном потоке событий с продлевающимся мертвым временем имеет вид:

$$g_\tau(s) = \frac{z_1}{z_1 + s} g_\xi(s) + sp \frac{(z_1 - z_2 - \alpha_2(\delta - 1))}{(z_1 + s)(z_2 + s)(p\lambda_1 + \alpha_2)} \left[\lambda_1 g_\xi(s) - \lambda_1(1 - p)g_\xi(s + p\lambda_1 + \alpha_2) \right], \quad (10)$$

где z_1, z_2 , определены в (1), $g_\xi(s), g_\xi(s + p\lambda_1 + \alpha_2)$ определены в (5) для аргументов s и $s + p\lambda_1 + \alpha_2$ соответственно.

Доказательство. Пусть момент наступления события в наблюдаемом потоке есть $\tau=0$. Рассмотрим интервал $(0, \tau) = (0, \xi + \eta)$. Зафиксируем ξ . Введем вероятности $p_{ij}(\tau - \xi)$ – условные вероятности того, что на интервале длительности $\eta = \tau - \xi$ не наступит событий наблюдаемого потока и в момент времени $\tau = \xi + \eta$ значение процесса $\lambda(\xi + \eta) = \lambda_j$ при условии, что в момент времени $\tau = \xi$ значение процесса $\lambda(\xi) = \lambda_i, i, j=1, 2$. Соответствующие $p_{ij}(\tau - \xi)$ имеют вид:

$$\begin{aligned} p_{11}(\tau - \xi) &= e^{-\lambda_1(\tau - \xi)}, p_{12}(\tau - \xi) = 0, p_{22}(\tau - \xi) = e^{-(\lambda_2 + \alpha_2)(\tau - \xi)}, \\ p_{21}(\tau - \xi) &= \frac{(1 - \delta)\alpha_2}{\lambda_1 - (\lambda_2 + \alpha_2)} \left[e^{-(\lambda_2 + \alpha_2)(\tau - \xi)} - e^{-\lambda_1(\tau - \xi)} \right], \tau \geq 0. \end{aligned} \quad (11)$$

Имеем $P_1(\tau - \xi) = p_{11}(\tau - \xi) + p_{12}(\tau - \xi)$ – условная вероятность того, что на интервале (ξ, τ) событий наблюдаемого потока не произойдет при условии, что в момент времени $\tau = \xi$ значение процесса $\lambda(\xi) = \lambda_i, i=1, 2$. Тогда условная плотность вероятности значения длительности интервала (ξ, τ) (без наступления событий на этом интервала) по определению есть $p_i(\tau - \xi) = -P_i'(\tau - \xi), i=1, 2$. С учетом (11) имеем

$$p_1(\tau - \xi) = \lambda_1 e^{-\lambda_1(\tau - \xi)}, p_2(\tau - \xi) = (\lambda_2 + \alpha_2) e^{-(\lambda_2 + \alpha_2)(\tau - \xi)}, \xi \leq \tau. \quad (12)$$

Условную плотность $p(\tau - \xi | \xi)$ в формуле (9) представим в виде:

$$p(\tau - \xi | \xi) = \pi_1(\tau = \xi | \xi)p_1(\tau - \xi) + \pi_2(\tau = \xi | \xi)p_2(\tau - \xi), \quad (13)$$

где $\pi_i(\tau = \xi | \xi)$ – условные вероятности того, что в момент времени $\tau = \xi$ значение процесса $\lambda(\tau = \xi) = \lambda_i$, $i = 1, 2$, при условии, что в момент времени $\tau = 0$ событие наступило и наступил общий период ненаблюдаемости длительности ξ .

Используя условие рекуррентности $\lambda_2 = p(\lambda_2 + \delta\alpha_2)$ и считая $\tau = \xi$, выпишем вероятности $\pi_i(\tau = \xi | \xi)$, $i = 1, 2$ в виде [10]:

$$\begin{aligned} \pi_1(\tau = \xi | \xi) &= \frac{\alpha_2(1 - p + p\delta) + \pi_1(\lambda_2 - p(\lambda_2 + \delta\alpha_2))\left[1 - e^{-(p\lambda_1 + \alpha_2)\xi}\right]}{\alpha_2 + p(\lambda_2 + \delta\alpha_2) + (\lambda_2 - p(\lambda_2 + \delta\alpha_2))\left[1 - e^{-(p\lambda_1 + \alpha_2)\xi}\right]}, \\ \pi_2(\tau = \xi | \xi) &= \frac{p(\lambda_2 + \alpha_2) + \pi_2(\lambda_2 - p(\lambda_2 + \delta\alpha_2))\left[1 - e^{-(p\lambda_1 + \alpha_2)\xi}\right]}{\alpha_2 + p(\lambda_2 + \delta\alpha_2) + (\lambda_2 - p(\lambda_2 + \delta\alpha_2))\left[1 - e^{-(p\lambda_1 + \alpha_2)\xi}\right]}, \end{aligned} \quad (14)$$

где $\pi_1 = \frac{\alpha_2}{p\lambda_1 + \alpha_2}$, $\pi_2 = \frac{p\lambda_1}{p\lambda_1 + \alpha_2}$, $\pi_1 + \pi_2 = 1$.

Подставляя (12) и (14) в (13), находим

$$p(\tau - \xi | \xi) = 0, \quad 0 \leq \tau < \xi, \quad p(\tau - \xi | \xi) = \Gamma(\xi)\lambda_1 e^{-\lambda_1(\tau - \xi)} + [1 - \Gamma(\xi)](\lambda_2 + \alpha_2)e^{-(\lambda_2 + \alpha_2)(\tau - \xi)}, \quad \tau \geq \xi, \quad (15)$$

где $\Gamma(\xi) = 1 - \pi_2(\xi) \frac{\lambda_1 - \lambda_2 - \delta\alpha_2}{\lambda_1 - (\lambda_2 + \alpha_2)}$, $\pi_2(\xi) = \pi_2 - [\pi_2 - \pi_2(\tau = \xi | \xi)]e^{-(p\lambda_1 + \alpha_2)\xi}$, $\pi_2(\tau = \xi | \xi) = p$.

Преобразование Лапласа плотности $p(\tau)$ с учетом (9) примет вид:

$$g_\tau(s) = \int_0^\infty e^{-s\tau} p(\tau) d\tau = \int_0^\tau p(\xi) \int_\xi^\infty e^{-s\tau} p(\eta | \xi) d\xi d\tau = g_\xi(s) \int_\xi^\infty e^{-s\eta} p(\eta | \xi) d\eta = g_\xi(s) \int_0^\infty e^{-s(\tau - \xi)} p(\tau - \xi | \xi) d\tau. \quad (16)$$

Подставляя в (16) выражение (15) и проделывая необходимые преобразования, находим

$$g_\tau(s) = \frac{z_1}{z_1 + s} g_\xi(s) + \left(\frac{z_2}{z_2 + s} - \frac{z_1}{z_1 + s} \right) \int_0^\infty e^{-s\xi} (1 - \Gamma(\xi)) p(\xi) d\xi. \quad (17)$$

Подставляя $\Gamma(\xi)$ из (15) в (17), получим (10). Теорема доказана.

Теорема 4.4. Математическое ожидание длительности интервала между соседними событиями в рекуррентном обобщенном полусинхронном потоке, функционирующем в условиях продлевающегося мертвого времени, имеет вид:

$$M\tau = \frac{1}{z_1} + \frac{1}{z_2} - \frac{1}{pz_1 + \alpha_2} + M\xi - p[z_1(1 - p) - \alpha_2] \frac{z_1 - z_2 - \alpha_2(\delta - 1)}{z_1 z_2 (pz_1 + \alpha_2)} g_\xi(pz_1 + \alpha_2), \quad (18)$$

где z_1 , z_2 , $g_\xi(pz_1 + \alpha_2)$ определены в (1) и (5) при $s = pz_1 + \alpha_2$; $M\xi$ определена в (8).

Доказательство. Используя следствие теоремы 4.1, заменив ξ на τ и вычислив производную от $g_\tau(s)$, определенную в (10), в точке $s = 0$, по формуле (4) получим (19). Теорема доказана.

5. Уравнение моментов для оценки длительности мертвого времени

В данном разделе статьи описывается оценивание длительности мертвого времени T методом моментов.

Введем статистику $C_m = \frac{1}{n} \sum_{k=1}^n \tau_k^m$, где $\tau_k = t_{k+1} - t_k$ – значение длительности интервала (t_k, t_{k+1}) .

Зафиксируем параметры потока λ_1 , λ_2 , α_2 , p , δ . Тогда будет достаточно первого момента, чтобы оценить величину T . Уравнение моментов, позволяющее оценить длительность мертвого времени T ,

имеет вид: $M\tau = C_1$, где статистика C_1 при $n \rightarrow \infty$ стремится к начальному моменту $M\tau$ [11]. Используя формулу (18), получим уравнение моментов для оценивания длительности мертвого времени T :

$$\frac{1}{z_1} + \frac{1}{z_2} - \frac{1}{pz_1 + \alpha_2} + \frac{z_2\gamma(1 - e^{-z_1T}) + z_1(1 - \gamma)(1 - e^{-z_2T})}{z_1z_2(\gamma e^{-z_1T} + (1 - \gamma)e^{-z_2T})} - p[z_1(1 - p) - \alpha_2] \frac{z_1 - z_2 - \alpha_2(\delta - 1)}{z_2(pz_1 + \alpha_2)} \times$$

$$\times \frac{(\gamma e^{-(z_1 + pz_1 + \alpha_2)T} + (1 - \gamma)e^{-(z_2 + pz_1 + \alpha_2)T})}{e^{(pz_1 + \alpha_2)T}} \left[1 - \frac{z_1\gamma(1 - e^{-(z_1 + pz_1 + \alpha_2)T})}{(z_1 + pz_1 + \alpha_2)} - \frac{z_2(1 - \gamma)(1 - e^{-(z_2 + pz_1 + \alpha_2)T})}{(z_2 + pz_1 + \alpha_2)} \right]^{-1} = C_1, \quad (19)$$

где z_1, z_2, γ определены в (5); $0 \leq T \leq \tau_{\min}$, $\tau_{\min} = \min \tau_k, k = \overline{1, n}$.

Решение уравнения (19) возможно только численно. Обозначим левую часть данного уравнения, как $M(T)$. Очевидно, что при увеличении T математическое ожидание $M\tau$ возрастает с ростом T . Таким образом, $M(T)$ будет возрастающей функцией при $0 \leq T$. Следовательно, уравнение (19) имеет единственное решение. Тогда оценка \hat{T} длительности мертвого времени является состоятельной [11]. Корень уравнения (19) на интервале $0 \leq T \leq \tau_{\min}$ есть значение оценки \hat{T} . Если на интервале $0 \leq T \leq \tau_{\min}$ уравнение (19) решения не имеет, то возможны два варианта:

- 1) если $M(0) \geq C_1$, то в качестве значения оценки \hat{T} естественно выбрать $\hat{T} = 0$;
- 2) если $M(\tau_{\min}) < C_1$, то $\hat{T} = \tau_{\min}$.

6. Численные результаты оценивания длительности мертвого времени

Статистика C_1 является результатом работы имитационной модели обобщенного полусинхронного потока событий, функционирующего в условиях продлевающегося мертвого времени фиксированной длительности T . Имитационная модель реализована на языке программирования C# [13].

Алгоритм оценивания:

- 1) рассчитывается статистика C_1 ;
- 2) методом Ньютона рассчитывается корень уравнения (19);
- 3) N раз повторяя шаги 1 и 2, получаем значения оценок $\hat{T}_1, \hat{T}_2, \dots, \hat{T}_N$;
- 4) рассчитывается выборочное среднее и выборочная вариация оценки длительности мертвого времени T по формулам

$$\hat{M}(\hat{\mathbf{T}}) = \frac{1}{N} \sum_{i=1}^N \hat{T}_i, \quad \hat{V}(\hat{\mathbf{T}}) = \frac{1}{N-1} \sum_{i=1}^N (\hat{T}_i - \hat{M}(\hat{\mathbf{T}}))^2.$$

Здесь T – значение длительности мертвого времени, известное из имитационной модели рассматриваемого потока.

Зададим параметры потока, соблюдая условие рекуррентности, следующим образом: $\lambda_2 = 1$; $\alpha_2 = 0,42$; $\delta = 0,8$; $p = 0,75$; $T_m = 1300$ ед. времени.

Рассмотрим три варианта задания параметра λ_1 : $\lambda_1 = 3$; $\lambda_1 = 7$; $\lambda_1 = 11$, для $T = 1; 1,5; 2; 2,5; 3; 3,5$ ед. времени.

Ниже представлены результаты оценивания (табл. 1–3, рис. 2–7).

Таблица 1

Результаты работы алгоритма для $\lambda_1 = 3$

T	1	1,5	2	2,5	3	3,5
$\hat{M}(\hat{\mathbf{T}})$	0,998584	1,492618	1,989632	2,488640	2,988937	3,479390
$\hat{V}(\hat{\mathbf{T}})$	0,000006	0,000065	0,000131	0,000210	0,000306	0,000920

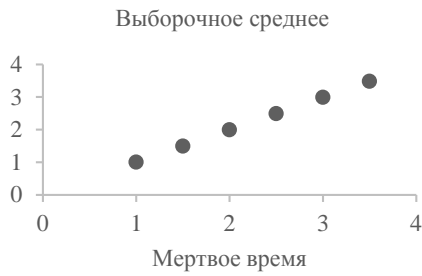


Рис. 2. Выборочное среднее $\hat{M}(\hat{T})$

Fig. 2. Sample average $\hat{M}(\hat{T})$

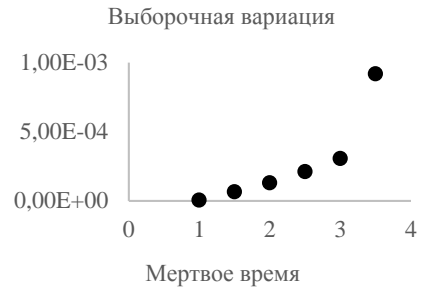


Рис. 3. Выборочная вариация $\hat{V}(\hat{T})$

Fig. 3. Sample variation $\hat{V}(\hat{T})$

Таблица 2

Результаты работы алгоритма для $\lambda_1 = 7$

T	1	1,5	2	2,5	3	3,5
$\hat{M}(\hat{T})$	1,000058	1,497625	1,997458	2,497454	2,996384	3,486561
$\hat{V}(\hat{T})$	0,000004	0,000014	0,000024	0,000071	0,000174	0,000501

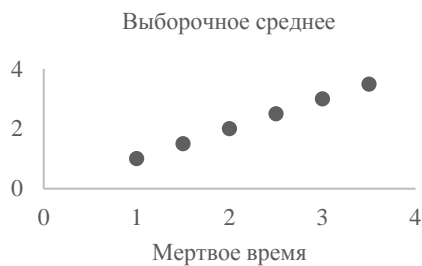


Рис. 4. Выборочное среднее $\hat{M}(\hat{T})$

Fig. 4. Sample average $\hat{M}(\hat{T})$

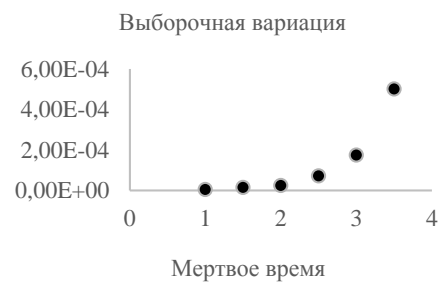


Рис. 5. Выборочная вариация $\hat{V}(\hat{T})$

Fig. 5. Sample variation $\hat{V}(\hat{T})$

Таблица 3

Результаты работы алгоритма для $\lambda_1 = 11$

T	1	1,5	2	2,5	3	3,5
$\hat{M}(\hat{T})$	1,000079	1,497675	1,997467	2,497575	2,996850	3,491620
$\hat{V}(\hat{T})$	0,000004	0,000014	0,000034	0,000074	0,000190	0,000510

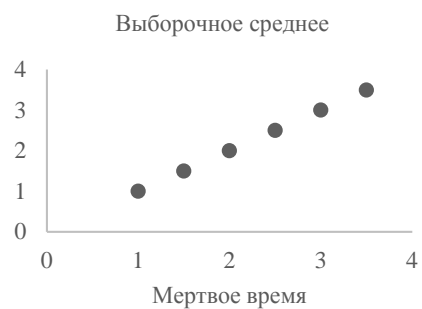


Рис. 6. Выборочное среднее $\hat{M}(\hat{T})$

Fig. 6. Sample average $\hat{M}(\hat{T})$

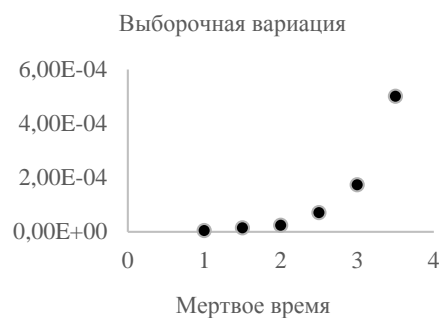


Рис. 7. Выборочная вариация $\hat{V}(\hat{T})$

Fig. 7. Sample variation $\hat{V}(\hat{T})$

Вывод: результаты статистических экспериментов, представленные в табл. 1–3, показывают, что выборочные средние значения \hat{T} оценок \hat{T} достаточно близки к истинным значениям T . При увеличении T увеличивается значение выборочной вариации оценки \hat{T} . Это связано с увеличением общего периода ненаблюдаемости потока событий, что влечет за собой большую потерю событий обобщенного полусинхронного потока при наличии продлевающегося мертвого времени.

Заключение

В настоящей работе рассмотрен обобщенный полусинхронный поток событий, функционирующий в условиях продлевающегося мертвого времени фиксированной длительности T . Для рекуррентного наблюдаемого потока событий построено преобразование Лапласа плотности вероятности общего периода ненаблюдаемости и преобразование Лапласа плотности вероятности длительности интервалов между соседними событиями в наблюдаемом потоке. Получено уравнение моментов для определения значения \hat{T} оценки \hat{T} длительности мертвого времени T . Проведены статистические эксперименты, показавшие приемлемое качество оценок.

Список источников

1. Горцев А.М., Назаров А.А., Терпугов А.Ф. Управление и адаптация в системах массового обслуживания. Томск : Изд-во Том. ун-та, 1978. 208 с.
2. Вишневецкий В.М., Дудин А.Н., Клименок В.И. Стохастические системы с коррелированными потоками. Теория и применение в телекоммуникационных сетях. М. : Техносфера, 2018. 564 с.
3. Башарин Г.П., Кокотушкин В.А., Наумов В.А. О методе эквивалентных замен расчета фрагментов сетей связи. Ч. 1 // Известия АН СССР. Техническая кибернетика. 1979. № 6. С. 92–99.
4. Башарин Г.П., Кокотушкин В.А., Наумов В.А. О методе эквивалентных замен расчета фрагментов сетей связи. Ч. 2 // Известия АН СССР. Техническая кибернетика. 1980. № 1. С. 55–61.
5. Neuts M.F. A versatile Markov point process // Journal of Applied Probability. 1979. V. 16. P. 764–779.
6. Нежелская Л.А. Оценка состояний дважды стохастических потоков событий. Томск : Изд-во Том. гос. ун-та, 2020. 210 с.
7. Апанасович В.В., Коляда А.А., Чернявский А.Ф. Статистический анализ случайных потоков в физическом эксперименте. Минск : Университетское, 1988. 256 с.
8. Горцев А.М., Калягин А.А., Нежелская Л.А. Оптимальная оценка состояний обобщенного полусинхронного потока событий // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2010. № 2 (11). С. 66–81.
9. Нежелская Л.А., Степаненко И.Д. Оценка длительности мертвого времени в рекуррентном обобщенном полусинхронном потоке событий с продлевающимся мертвым временем // Труды Томского государственного университета. Сер. физико-математическая. 2022. Т. 307. С. 37–46.
10. Горцев А.М., Калягин А.А., Нежелская Л.А. Совместная плотность вероятностей длительности интервалов обобщенного полусинхронного потока событий при непродлеваемом мертвом времени // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2014. № 2 (27). С. 19–27.
11. Шуленин В.П. Математическая статистика. Томск : Изд-во НТЛ, 2012. Ч. 1. 540 с.
12. Нежелская Л.А. Оценка состояний и параметров дважды стохастических потоков событий : дис. ... д-ра физ.-мат. наук. Томск, 2016. 341 с.
13. Нежелская Л.А., Степаненко И.Д. Статистические эксперименты на имитационной модели обобщенного полусинхронного потока событий с продлевающимся мертвым временем фиксированной длительности // Труды Томского государственного университета. Сер. физико-математическая. 2021. Т. 306. С. 82–89.

References

1. Gortsev, A.M., Nazarov, A.A. & Terpugov, A.F. (1978) *Upravlenie i adaptatsiya v sistemakh massovogo obsluzhivaniya* [Management and adaptation in queuing systems]. Tomsk: Tomsk State University.
2. Vishnevsky, V.M., Dudin, A.N. & Klimenok, V.I. (2018) *Stokhasticheskie sistemy s korrelirovannymi potokami. Teoriya i primeneniye v telekommunikatsionnykh setyakh* [Stochastic systems with correlated flows. Theory and application in telecommunication networks]. Moscow: Tekhnosfera.
3. Basharin, G.P., Kokotushkin, V.A. & Naumov, V.A. (1979) O metode ekvivalentnykh zamen rascheta fragmentov setey svyazi. [On the equivalent substitutions method for computing fragments of communication networks]. *Izvestiya AN SSSR. Tekhnicheskaya kibernetika*. 6. pp. 92–99.

4. Basharin, G.P., Kokotushkin, V.A. & Naumov, V.A. (1980) O metode ekvivalentnykh zamen rascheta fragmentov setey svyazi. [On the equivalent substitutions method for computing fragments of communication networks]. *Izvestiya AN SSSR. Tekhnicheskaya kibernetika*. 1. pp. 55–61.
5. Neuts, M.F. (1979) A versatile Markov point process. *Journal of Applied Probability*. 16. pp. 764–779. DOI: 10.2307/3213143
6. Nezhelskaya, L.A. (2020) *Otsenka sostoyaniy dvazhdy stokhasticheskikh potokov sobyitiy* [Estimation of states of doubly stochastic flows of events]. Tomsk: Tomsk State University.
7. Apanasovich, V.V., Kolyada, A.A. & Chernyavsky, A.F. (1988) *Statisticheskii analiz sluchaynykh potokov v fizicheskom eksperimente* [Statistical analysis of random flows in a physical experiment]. Minsk: Universitetskoe.
8. Gortsev, A.M., Kalyagin, A.A. & Nezhelskaya, L.A. (2010) Optimal estimation of the states of a generalized semisynchronous flow of events. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 2(11). pp. 66–81.
9. Nezhelskaya, L.A. & Stepanenko, I.D. (2022) Estimation of the duration of the fixed dead time of a generalized semi-synchronous stream of events functioning in conditions of prolonged dead time. *Trudy Tomskogo gosudarstvennogo universiteta. Ser. fiziko-matematicheskaya*. 307. pp. 37–47.
10. Gortsev, A.M., Kalyagin, A.A. & Nezhelskaya, L.A. (2014) Joint probability density of the duration of intervals of generalized semisynchronous flow of events with non-lasting dead time. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 2(27). pp. 19–27.
11. Shulenin, V.P. (2012) *Matematicheskaya statistika* [Mathematical Statistics]. Vol. 1. Tomsk: NTL.
12. Nezhelskaya, L.A. (2016) *Otsenka sostoyaniy i parametrov dvazhdy stokhasticheskikh potokov sobyitiy* [Evaluation of states and parameters of doubly stochastic flows of events]. Physics and Mathematics Dr. Diss. Tomsk.
13. Nezhelskaya, L.A. & Stepanenko, I.D. (2021) Statistical experiments on a simulation model of a generalized semi-synchronous flow of events with an extended dead time of fixed duration. *Trudy Tomskogo gosudarstvennogo universiteta. Ser. fiziko-matematicheskaya – Proceedings of Tomsk State University. Ser. Phys.-Math*. 306. pp. 82–89.

Информация об авторах:

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

Степаненко Илья Денисович – магистрант кафедры прикладной математики Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: 97step@mail.ru

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

Information about the authors:

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

Stepanenko Ilya D. (Master's Student, Department of Applied Mathematics, Institute of Applied Mathematics and Computer Science, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: 97step@mail.ru

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

Поступила в редакцию 13.10.2022; принята к публикации 01.03.2023

Received 13.10.2022; accepted for publication 01.03.2023

Научная статья

УДК 519.234:621.391

doi: 10.17223/19988605/62/8

Алгоритм обнаружения звукорядов в спектре шумоизлучения морских судов

Игорь Михайлович Рудько

Институт проблем управления им. В.А. Трапезникова РАН, Москва, Россия, igor-rudko@mail.ru

Аннотация. Рассматривается алгоритм обнаружения звукорядов в спектре шумоизлучения морских судов, который основан на модифицированном энергетическом приемнике с использованием на первом этапе фильтра на основе усеченных порядковых статистик для выделения дискретных составляющих спектра шумоизлучения, а на втором этапе – энергетического приемника и критерия максимального правдоподобия, что позволяет оценить наличие звукорядов и определить их основную частоту.

Ключевые слова: дискретные составляющие; звукоряды; порядковая статистика; системы обнаружения; статистическое моделирование.

Для цитирования: Рудько И.М. Алгоритм обнаружения звукорядов в спектре шумоизлучения морских судов // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2023. № 62. С. 76–84. doi: 10.17223/19988605/62/8

Original article

doi: 10.17223/19988605/62/8

Algorithm for detecting harmonic series in the noise emission spectrum of sea vessels

Igor M. Rudko

V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Science, Moscow, Russian Federation, igor-rudko@vail.ru

Abstract. An algorithm for detecting harmonic series (HS) in the noise emission spectrum of sea vessels is considered, which is based on a modified energy receiver using, at the first stage, a filter based on truncated order statistics (TOS-filter) to isolate discrete components (DC) of the noise emission spectrum, and at the second stage with using an energy receiver and the maximum likelihood criterion, which makes it possible to assess the presence of harmonic series and determine its fundamental frequency.

Keywords: discrete components; harmonic series; order statistics; detection systems; statistical modeling.

For citation: Rudko, I.M. (2023) Algorithm for detecting harmonic series in the noise emission spectrum of sea vessels. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 62. pp. 76–84. doi: 10.17223/19988605/62/8

Введение

Гидроакустические средства обнаружения морских судов работают в широкой полосе частот и производят обработку сигналов как во временной, так и в частотной области [1, 2]. Однако наиболее значимым является низкочастотный диапазон, так как в энергетическом спектре этого диапазона наблюдаются узкие локальные максимумы – дискретные составляющие (ДС), обусловленные работой вала, гребного винта и наиболее шумящих механизмов и машин, несущие информацию о типе и режиме работы движителя судна, количестве лопастей у винта и т.д., на основании которой возможно разделить морские суда на основные классы и определить параметры их движения.

Дискретные составляющие, обусловленные вращением винтов, проявляются на частотах, кратных основной гармонике. ДС, связанные с работой механизмов, обычно кратны половине частоты сети электропитания (25 Гц при частоте сети 50 Гц и 30 Гц при частоте сети 60 Гц) [3].

На низких частотах, как правило, ширина ДС существенно меньше, чем на высоких. Часто ДС, возбуждаемые от одного и того же источника, бывают синхронизированы между собой и образуют звукояды (ЗР), т.е. наборы ДС, частоты которых кратны между собой (гармоники).

Так, например, вальный ЗР может содержать ДС на частотах, кратных частоте вращения вала:

$$f_{B_k} = kv/60 = kf_{B_1} \quad (1)$$

где $k = 1, 2, 3, \dots, K$ – номера гармоник; v – скорость вращения вала (об./мин); f_{B_1} – частота вращения вала (Гц). Лопастной ЗР может содержать ДС на частотах, кратных произведению частоты вращения линии вала и числа лопастей винта: $f_{L_k} = zk v/60 = zk f_{L_1}$, где z – число лопастей винта.

Тональные компоненты ЗР являются высокостабильными и имеют очень узкие полосы частот [1, 2, 4]. В отличие от отдельных ДС шум в виде ЗР, несмотря на то что содержит узкополосные составляющие, в целом представляет собой широкополосный шумовой процесс. Узкополосный гидроакустический шум (т.е. ДС общего суммарного шума) для узкополосных систем шумопеленгования является полезным сигналом, обнаруживаемым на фоне помех.

Характеристикой, описывающей ДС шума гидроакустических целей, является их спектральная плотность мощности (СПМ). Эффективная ширина полосы (ЭП) СПМ ДС Δf зависит от значения частоты, на которой она образуется [5], прямо пропорциональна центральной частоте ДС f_1 и составляет 0,03–0,3% от значения f_1 . Относительная полоса СПМ каждой отдельной ДС (или ДС, входящих в состав ЗР, образованного одним общим источником) – постоянная величина. ЭП ДС является линейной функцией ее центральной частоты: $\Delta f_{\Sigma} = \beta f_1$. Соответственно, ЭП каждой k -й гармоники в звукоядре также является линейной функцией частоты: $\Delta f_{\Sigma k} = k\beta f_1$, а общая СПМ всего звукоядра из K ДС обладает масштабными (мультипликативными) свойствами в области частоты. Пример СПМ шума корабля [3] представлен на рис. 1. СПМ представляет суперпозицию ряда ДС и непрерывной части СПМ, имеющей спад частотной характеристики –6 дБ/октаву и принимаемой на фоне широкополосного шума моря, также имеющего спад частотной характеристики –6 дБ/октаву. При обнаружении ДС и шума моря, и непрерывная часть СПМ корабля – помехи, маскирующие ДС.

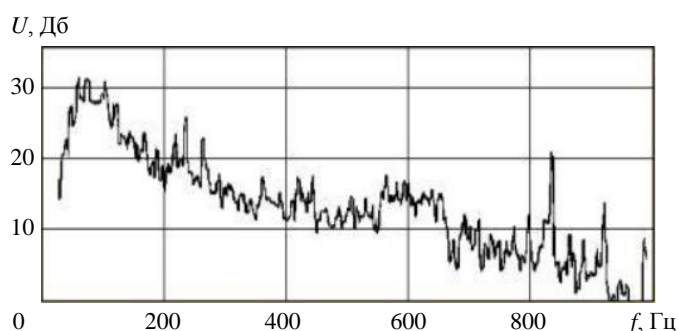


Рис. 1. Пример СПМ шума корабля
Fig. 1. An example of the power spectral density of ship noise

Алгоритм обнаружения ДС предполагает следующую последовательность операций [6]:

- широкополосная полосовая фильтрация (для формирования общего частотного диапазона) в полосе ΔF и «отбеливающий» фильтр [7], преобразующий СПМ помехи в белый шум в полосе ΔF ;
- многоканальная узкополосная полосовая фильтрация (формирование отдельных частотных каналов), т.е. дискретное преобразование Фурье (БПФ) реализации на отрезке длительности T_0 ;
- квадратичное детектирование, т.е. расчет энергетического спектра;
- интегрирование, т.е. накопление (усреднение) энергетических спектров по N реализациям;

– определение порога обнаружения α по правилу Неймана–Пирсона при задаваемой вероятности ложной тревоги $P_{лт}$ и сравнения с порогом (в каждом частотном канале).

Операция обнаружения – это задача проверки двух статистических гипотез: нулевой гипотезы H_0 , когда данные относятся только к шуму (непрерывной части СПМ шумоизлучения), и альтернативной гипотезы H_1 , когда данные относятся к совместному воздействию сигнала (ДС) и шума.

Данный алгоритм описывает многоканальный энергетический приемник [2]. В случае превышения порога в каком либо канале принимается решение об обнаружении ДС. Затем из обнаруженных ДС выбираются ДС, образующие ЗР. Так как каждая ДС обнаруживается независимо от других, при малых отношениях сигнал/помеха обнаружение ЗР становится проблематичным. Известны алгоритмы [8], основанные на вычислении интегрального вейвлет-спектра, которые непосредственно вычисляют ЗР. Вычисление интегрального вейвлет-спектра основано на непрерывном вейвлет-преобразовании сигнала, которое требует больших вычислительных затрат [9].

В настоящей работе рассматривается алгоритм, непосредственно обнаруживающий ЗР, который основан на модифицированном энергетическом приемнике с использованием фильтра на основе усеченных порядковых статистик (УПС-фильтра), и имеющий быстродействие существенно выше, чем вычисление интегрального вейвлет-спектра.

1. Предварительная фильтрация с использованием УПС-фильтра

Пусть на входе системы обнаружения принимается сигнал $S(t)$, который обрабатывается по алгоритму БПФ со следующими параметрами: Δt – интервал дискретизации по времени, $T_0 = L\Delta t$ – длина реализации, L – число временных отсчетов, $\Delta f = 1/T_0$ – разрешающая способность БПФ, $\Delta F = \Delta f L/2$ – полоса анализа БПФ (полоса пропускания входного фильтра системы обнаружения).

Известно [2], что для белого шума оценки СПМ в узкой полосе Δf имеют центральное χ^2 -распределение с $2N$ степенями свободы, где N – число сглаженных (накопленных) СПМ.

За время $T_N = N T_0$ формируется матрица \mathbf{M}_{ij} размерностью $(N \times L/2)$ в координатах (время – i , частота – j), каждый элемент которой x_{ij} имеет центральное χ^2 -распределение с двумя степенями свободы. В многоканальном энергетическом приемнике для каждого элемента разрешения по частоте Δf_j производится накопление (усреднение) СПМ по N реализациям:

$$X_j = \sum_{i=1}^N x_{ij}. \quad (2)$$

Закон распределения энергии принимаемого сигнала, заданный плотностью вероятности $f(x)$ случайной величины X , позволяет математически описать вероятностные характеристики обнаружения – вероятность правильного обнаружения и вероятность ложной тревоги:

$$P_{лт} = \int_h^{\infty} f_{ш}(x) dx, \quad P_{обн} = \int_h^{\infty} f_{с+ш}(x) dx, \quad (3)$$

где h – порог обнаружения; $f_{ш}(x)$ – плотность вероятности случайной величины X при гипотезе H_0 ; $f_{с+ш}(x)$ – плотность вероятности X при гипотезе H_1 .

Плотность распределения статистики X для гауссовского процесса $S(t)$ описывается плотностью χ^2 -распределения $f(x) = k_n(x/\sigma^2)/\sigma^2$, $x \geq 0$, где n – число степеней свободы.

Поскольку распределение для гипотезы H_0 отличается от распределения для гипотезы H_1 только величиной дисперсии, то для гипотезы H_0 : $\mu_0 = 2N\sigma_{ш}^2$, $\sigma_0^2 = 4N\sigma_{ш}^2$, а для гипотезы H_1 : $\mu_1 = 2N(\sigma_{ш}^2 + \sigma_c^2) = 2N\sigma_{ш}^2(1+\rho)$, $\sigma_1^2 = 4N(\sigma_{ш}^2 + \sigma_c^2) = 4N\sigma_{ш}^2(1+\rho)$, где $\rho = \sigma_c^2/\sigma_{ш}^2$ – отношение сигнал/помеха в полосе Δf , а $2N$ – число степеней свободы.

Определим вероятность обнаружения как $P_{обн} = \beta$, тогда $P_{необн} = 1 - \beta$ и в соответствии с (2), (3)

$$P_{необн} = 1 - \beta = \int_0^h \frac{1}{\sigma_1^2} k_{2N} \left(\frac{x}{\sigma_1^2} \right) dx = \int_0^{h/\sigma_1^2} k_{2N}(u) du = K_{2N}(h/\sigma_1^2), \quad (4)$$

где $K_n(x)$ – функция χ^2 -распределения. Порог h определяется из первого уравнения (3) путем задания вероятности ложной тревоги

$$P_{\text{лт}} = \alpha = 1 - \int_0^{h/\sigma_0^2} k_{2N}(u) du = K_{2N}(h/\sigma_0^2) = 1 - K_{2N}(h_0),$$

$$P_{\text{обн}} = \beta = 1 - K_{2N}(h/\sigma_1^2) = 1 - K_{2N}(h_1) = 1 - K_{2N}(h/(1+\rho)).$$

В рассматриваемом алгоритме интегратор (2) заменяется на УПС-фильтр [10, 11] (в каждом частотном канале Δf_j , $1 \leq j \leq L/2$).

Алгоритм работы УПС-фильтра [10] (для одного частотного канала):

1. Вектор-столбец матрицы \mathbf{M}_{ij} разбивается на m одинаковых интервалов, в каждом из которых согласно (2) вычисляется последовательность выборок $\mathbf{X} \sim \{X_1, \dots, X_i, \dots, X_m\}$. Каждый элемент этой выборки X_i имеет центральное χ^2 -распределение с $2n$ степенями свободы, где $n = N/m$. Накапливается c оцениваемых выборок \mathbf{X} .

2. По накопленным выборкам строится матрица \mathbf{X}_{ik} размерностью m строк на c столбцов (c – «глубина» матрицы памяти) – $\{X_1, \dots, X_i, \dots, X_m\}_k$, где $1 \leq k \leq c$.

3. В каждом столбце матрицы \mathbf{X}_{ik} строится порядковая статистика $\mathbf{X}_{(i)k}$, где $1 \leq i \leq m$, – упорядоченные величины статистики X_i , такие что $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(i)} \leq \dots \leq X_{(m)}$. Если случайные величины X_i статистически независимы и одинаково распределены, то случайные величины $X_{(i)}$ зависимы из-за неравенств между ними.

4. В каждой строке полученной матрицы $\mathbf{X}_{(i)k}$ определяются оценки математических ожиданий (вектор \hat{m}) $\hat{m}_i = \frac{1}{c} \sum_{k=1}^c X_{(i)k}$, где $1 \leq i \leq m$.

5. Порог отсекация k (первый порог) определяется из условия

$$k = \arg \min_i |h_0 - \hat{m}_i|, \quad (5)$$

где $1 \leq i \leq m$, а h_0 определяется по формуле

$$\alpha = \int_0^{h_0} f_{\text{ш}}(x) dx = \frac{1}{2^{n/2} \sigma_0^n \Gamma(n)} \int_0^{h_0} x^{n-1} e^{-x/2\sigma_0^2} dx, \quad (6)$$

где α – заданный квантиль, σ_0^2 – дисперсия шума (известный параметр).

6. Вычисляется оценка

$$W = \sum_{i=k}^m X_{(i)k}. \quad (7)$$

На выходе УПС-фильтра получаем последовательность отфильтрованных оценок W , задача обнаружения по которым решается по «классическому» алгоритму проверки двух гипотез.

УПС-фильтр работает по принципу скользящего окна, т.е. каждый новый вектор \mathbf{X}_k с индексом $c + 1$ вытесняет из матрицы $\mathbf{X}_{(i)k}$ вектор \mathbf{X}_k с индексом 1.

Сравним статистические свойства случайных величин Z и W , сформированных из случайной выборки X_i , $1 \leq i \leq m$, двумя различными способами:

– **Энергетический приемник:** $Z = \sum_{i=1}^m X_i$ – имеет χ^2 -распределение с $2N = 2nm$ степенями свободы и математическим ожиданием $\mu_Z = 2nm\sigma^2$ и дисперсией $\sigma_Z^2 = 4nm\sigma^2$, и в силу центральной предельной теоремы при достаточно больших значениях m ее функция плотности распределения нормализуется: $Z \sim N(\mu_Z, \sigma_Z^2)$.

– **УПС-фильтр:** $W = \sum_{i=k}^m X_{(i)}$, где $X_{(i)}$ при достаточно больших значениях m – $W \sim N(\mu_W, \sigma_W^2)$.

Известно [12], что при больших значениях m достаточно точным приближением к математическому ожиданию элемента порядковой статистики с номером μ_j является значение, удовлетворяющее уравнению $P(x) = j/(m+1)$, где $P(x)$ – функции распределения. Если ввести в рассмотрение функцию Q , такую что $Q[P(x)] = x$, то получим асимптотическую формулу $\mu_j \sim Q(j/(m+1))$, где $1 \leq j \leq m$, т.е. математические ожидания μ_j приближенно равны квантилям функции распределения $P(x)$, взятым с равномерным шагом (в рассматриваемом случае, как следует из (6), $P(x) = k_n(x)$).

Таким образом, для гипотезы H_0 $\mu_j(H_0) = \mu_j \sigma_0^2$, а для гипотезы H_1 $\mu_j(H_1) = \mu_j \sigma_0^2(1 + \rho)$.

Согласно п. 5 алгоритма работы УПС-фильтра порог k определяется по дисперсии шума $\sigma_0^2 - k = k_0$. При наличии ДС ($\rho > 0$) $\mu_j(H_0) < \mu_j(H_1)$ и определяемый по (5) порог $k_1 < k_0$.

Пример. $N = 64, m = 8, n=8, \rho = 0,5, \sigma_0^2 = 1, \alpha = 0,5$.

Для энергетического приемника:

– гипотеза H_0 : $\mu_{z_0} = 2nm\sigma_0^2 = 128, \sigma_{z_0}^2 = 4nm\sigma_0^2 = 256$;

– гипотеза H_1 : $\mu_{z_1} = 2nm\sigma_0^2(1 + \rho) = 192, \sigma_{z_1}^2 = 4nm\sigma_0^2(1 + \rho) = 384$;

Для УПС-фильтра:

$\mu_j(H_0) = \{9,5534, 11,4976, 13,0832, 14,5804, 16,1225, 17,8403, 19,9557, 23,0991\}$;

$\mu_j(H_1) = \{14,3301, 17,2465, 19,6248, 21,8707, 24,1837, 26,7604, 29,9336, 34,6486\}$.

По формуле (6) $h_0 = 15,3385$. По формуле (5) определяется порог отсеечения k : для $\mu_j(H_0)$ $k_0 = 5$.

Пока в полосе Δf_j принимается только помеха $k_0 = 5$ и $W = \sum_{i=5}^8 X_{(i)}$, а $\mu_{w_0} = \sum_{i=5}^8 \mu_i(H_0) = 96,1$.

При появлении ДС с $\rho = 0,5$ математические ожидания $\mu_j(H_1)$ увеличиваются и согласно формуле (5)

порог отсеечения уменьшается до $k_1 = 1$ и $W = \sum_{i=1}^8 X_{(i)}$, а $\mu_{w_1} = \sum_{i=1}^8 \mu_i(H_1) = 192$. Таким образом, после

УПС-фильтра уровень широкополосной помехи уменьшается с 128 до 96, т.е. на 1,25 дБ. На рис. 2 приведены результаты статистического моделирования (усреднение по 10 000 реализациям).

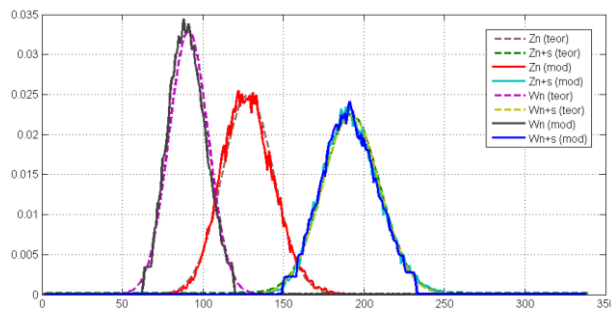


Рис. 2. Плотности вероятности процессов Z и W и их нормированные гистограммы для гипотез H_0 и H_1
 Fig. 2. Probability densities of Z and W processes and their normalized histograms for hypotheses H_0 and H_1

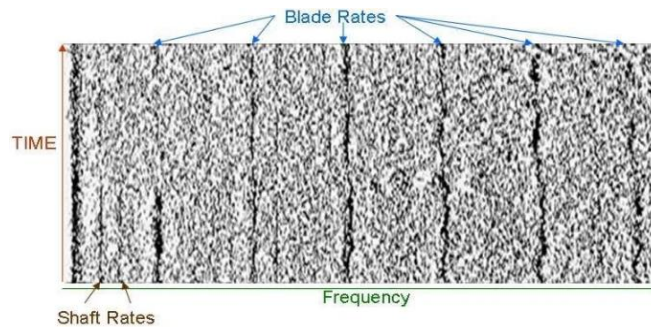


Рис. 3. LOFAR, пример частотно-временной СПМ
 Fig. 3. LOFAR, an example of the time-frequency spectral power density (SPD)

Для всех частотных каналов Δf_j , $1 \leq j \leq L/2$, получаем вектор (СПМ) $\overline{W^j} = (W^1, \dots, W^j, \dots, W^{L/2})$. Временная последовательность оценок СПМ может использоваться для визуального представления спектра низкочастотных звуков в частотно-временном анализе, как сделано, например, в системе гидроакустического наблюдения LOFAR [13]. Пример такой визуализации – на рис. 3. Использование в качестве временных оценок СПМ векторов $\overline{W^j}$, в которых существенно уменьшена широкополосная составляющая исходного СПМ, позволяет обнаруживать слабые ДС.

2. Обнаружение звукояда

Оценки СПМ вычисляются с шагом Δf по частоте. Очевидно, что основная частота ЗР не кратна Δf и может меняться в диапазоне $[f_{\min}, f_{\max}]$, поэтому в ЗР необходима корректировка частоты f_1 .

В заданном диапазоне $[f_{\min}, f_{\max}]$ зададим K возможных значений основной частоты f_1 с шагом Δf и сформируем множество $\{f_1^k\}$ основных частот: $f_1^k = f_{\min}^k + k\Delta f$, $0 \leq k \leq K-1$, $K = (f_{\max} - f_{\min})/\Delta f$.

Применив соотношение (1) для каждой основной частоты f_1^k вычислим последовательность кратных частот $f_i^k = lf_1^k$. Скорректируем значения частот f_i^k путем их округления до ближайших целых значений, соответствующих частотам оценок СПМ в БПФ, $f_i^k = \arg \min_j |j\Delta f - lf_1^k|$, где $1 \leq j \leq L/2$. В результате получим множество $\{\overline{F_k}\}$ векторов $\overline{F_k} = (f_1^k, \dots, f_{L/2}^k)$ скорректированных частот (K векторов). Размер вектора $\overline{F_k} - L_k = \lfloor \Delta f L/2 f_1^k \rfloor$ (целая часть). Для вектора наблюдений $\overline{W^j}$ по каждому значению основных частот из множества $\{f_1^k\}$ вычислим оценку СПМ всего ЗР, используя соответствующие векторы $\overline{F_k}$ в качестве адресов, $W_k^\Sigma = \sum_{i=1}^{L_k} W_k(f_i^k)$. Построив по всем значениям k оценку

условной плотности вероятности $p(W_k^\Sigma | f_1^k) = W_k^\Sigma / \sum_{k=1}^K W_k^\Sigma$, можно по критерию максимального правдоподобия оценить основную частоту ЗР – f_1^{\max} .

После получения оценки W_{\max}^Σ операция обнаружения сводится к задаче проверки двух гипотез:

- H_0 , когда данные относятся только к шуму, для которого известны оценки μ_{w0} и σ_{w0}^2 ;
- H_1 , когда данные относятся к совместному воздействию сигнала (ДС) и шума.

Таким образом, задача сводится к определению порога обнаружения α по правилу Неймана–Пирсона при заданной $P_{\text{лт}}$ и сравнению с порогом. Очевидно, что вероятность $P_{\text{обн}}$ по всему ЗР будет больше, чем $P_{\text{обн}}$ по отдельным ДС. Для статистики Z шума (H_0) $\mu_{z0} = 2nm\sigma_0^2 L_k$, $\sigma_{z0}^2 = 4nm\sigma_0^2 L_k$, а для статистики Z шума и сигнала (H_1) $\mu_{z1} = 2nm\sigma_0^2(1 + \rho')$, $\sigma_{z1}^2 = 4nm\sigma_0^4(1 + \rho')$, где ρ' определяется обобщенным отношением сигнал/помеха в широкой полосе ΔF [7]: $\rho' = \sigma_{1\Sigma}^2 / \sigma_{0\Sigma}^2 = \sum_{i=1}^{L_k} \sigma_{1i}^2 / \sum_{i=1}^{L_k} \sigma_{0i}^2$, а σ_{0i}^2 и σ_{1i}^2 – дисперсии соответственно сигнала и шума в узкой полосе Δf_i .

Для статистики W формулы аналогичны, но σ_{0i}^2 и σ_{1i}^2 определяются на выходе УПС-фильтра.

3. Результаты математического моделирования

В качестве модели шумоизлучения используется модель, описанная в [5, 14]:

$$S(t) = \sum_{k=1}^K A_k \cos \left[\omega_0 t + \int_0^t f(kt) dt + \varphi_0 \right].$$

Данное выражение описывает колебание, модулируемое по частоте, а функция $f(t)$ называется модулирующим сигналом, позволяющим изменять основную частоту ЗР по специальным случайным законам в пределах $5 \div 10\%$ от стационарного режима.

Плотности вероятности процессов Z и W и их нормированные гистограммы для гипотез H_0 и H_1 для одной ДС приведены в *примере* выше и на рис. 2. ЗР состоит из 9 гармоник с одинаковыми амплитудами, основная частота $f_1 = 10,17$ (у.е.), $\Delta F = 128$, $\Delta f = 1$, $N = 64$, $m = 8$, $n = 8$, $\sigma_0^2 = 1$, $\alpha = 0,5$ и для всех гармоник $\rho = 0,5$. На рис. 4 приведены СПМ шума, ЗР и смеси шума и ЗР до и после УПС-фильтра (усреднение по 50 реализациям СПМ).

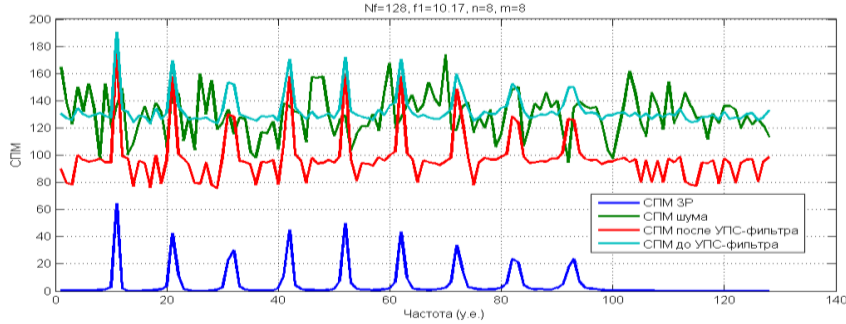


Рис. 4. СПМ шума, ЗР и смеси шума и ЗР до и после УПС-фильтра

Fig. 4. SPD noise, SPD harmonic series (HS) and mixture of noise and HS, before and after TOS-filter

На рис. 5 представлен частотно-временной СПМ до УПС-фильтра (*a*) и после УПС-фильтра (*b*). На рис. 6, *a* для рассматриваемого примера приведены результаты статистического моделирования (усреднение по 10 000 реализациям). На рис. 6, *b* показаны оценки вероятности обнаружения $P_{\text{обн}}$ в зависимости от отношения сигнал/помеха ρ (для ДС) и ρ' (для ЗР).

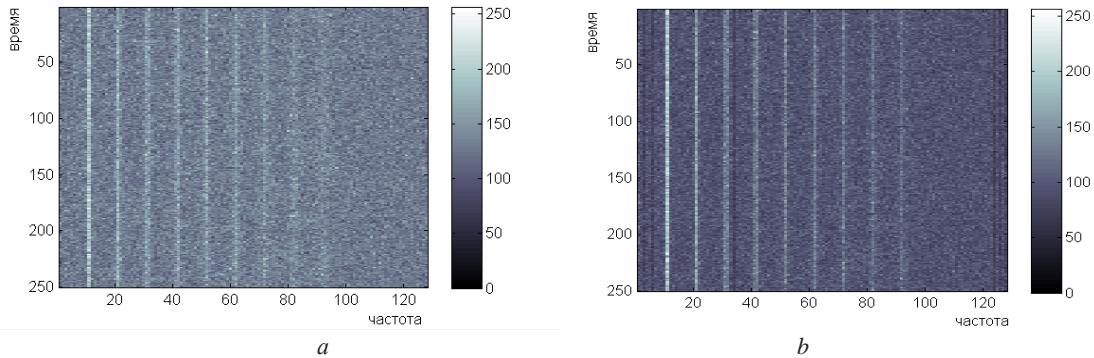


Рис. 5. Частотно-временной СПМ до УПС-фильтра (*a*) и после УПС-фильтра (*b*)

Fig. 5. Time-frequency SPD before the TOS-filter (*a*) and after the TOS-filter (*b*)

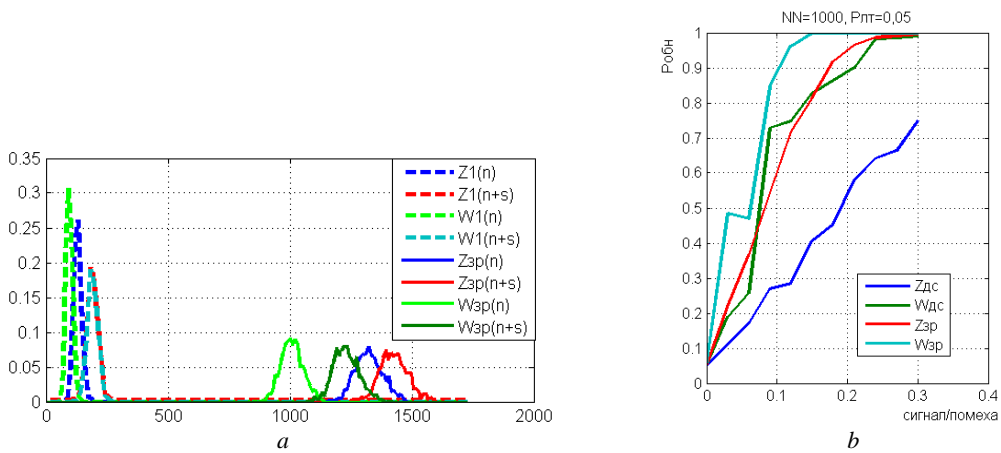


Рис. 6. Нормированные гистограммы процессов Z и W для гипотез $H_0(n)$ и $H_1(n+s)$ (*a*); оценки зависимости $P_{\text{обн}}$ от отношения сигнал/помеха (*b*)

Fig. 6. Normalized histograms of processes Z and W for hypotheses $H_0(n)$ and $H_1(n+s)$ (*a*); estimates of the P_{det} dependence on the signal-to-noise ratio (*b*)

Заключение

Исследован алгоритм обнаружения сигналов (ДС и ЗР) на фоне шума, основанный на свойствах усеченных порядковых статистик (УПС-фильтр), который позволяет обеспечить большую $P_{обн}$ при заданной $P_{лт}$ по сравнению с «классическим» алгоритмом проверки двух гипотез. Выигрыш достигается за счет введения дополнительного порога, отсекающего малые значения обрабатываемого сигнала, и использования для построения оценки этого порога информации, не использующейся в «классическом» алгоритме и содержащейся в предыдущих реализациях сигнала.

Список источников

1. Болгов В.М., Плахов Д.Д., Яковлев В.Е. Акустические шумы и помехи на судах. Л. : Судостроение, 1984. 192 с.
2. Бурдик В.С. Анализ гидроакустических систем. Л. : Судостроение, 1988. 392 с.
3. Горбань И.И. Обработка гидроакустических сигналов в сложных динамических условиях. Киев : Наукова думка, 2008. 276 с.
4. Шаталов Г.В. Моделирование узкополосного шума в виде отдельных дискретных составляющих и широкополосного шума в виде звукоядов // Морская радиоэлектроника. 2010. № 2 (32). С. 38–41.
5. Глазов Ю.Е., Ракитина И.С., Эпендиев М.Б., Глазкова В.Н., Мельникова Е.А. Методы снижения заметности узкополосных процессов в гидроакустическом поле морских подвижных объектов (краткий обзор) // Проблемы машиностроения и автоматизации. 2009. № 3. С. 69–76.
6. Деев В.В., Забродин Ю.М., Пахомов А.П., Тенетко В.А., Титов М.С. Анализ информации оператором-гидроакустиком. Л. : Судостроение, 1990. 192 с.
7. Зарайский В.А., Тюрин А.М. Теория гидролокации. Л. : ВМА, 1975. 604 с.
8. Рохманийко А.Ю., Якунин К.В., Сапрыкин В.А., Малый В.В., Есипов В.С. Устройство обнаружения шумовых гидроакустических сигналов в виде звукояда на основе вычисления интегрального вейвлет-спектра : патент на изобретение РФ № 2464588 от 20.10.2012; приоритет от 15.06.2011.
9. Яковлев А.Н. Введение в вейвлет-преобразования. Новосибирск : Изд-во НГТУ, 2003. 104 с.
10. Рудько И.М. Применение порядковых статистик в задачах обнаружения в частотной области // Управление большими системами. М. : ИПУ РАН, 2016. Вып. 62. С. 6–29.
11. Рудько И.М. Исследование влияния параметров фильтра на основе усеченных порядковых статистик (УПС-фильтра) на его выходные характеристики // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2022. № 60. С. 42–51.
12. David H.A., Nagaraja H.N. Order Statistics. John Wiley & Sons, 2003. 458 p. (Wiley Series in Probability and Statistics).
13. Lampert T.A., O'Keefe S.E.M. On the Detection of Tracks in Spectrogram Images // Pattern Recognition. 2013. V. 46 (5). P. 1396–1408.
14. Рудько И.М. Стохастическая модель звукояда // Управление большими системами. М. : ИПУ РАН, 2014. Вып. 49. С. 81–95.

References

1. Bolgov, V.M., Plakhov, D.D. & Yakovlev, V.E. (1984) *Akusticheskie shumy i pomekhi na sudakh* [Acoustic noise and interference on ships]. Leningrad: Sudostroenie.
2. Burdic, W.S. (1988) *Analiz gidroakusticheskikh sistem* [Acoustic system analysis]. Translated from English. Leningrad: Sudostroenie.
3. Gorban, I.I. (2008) *Obrabotka gidroakusticheskikh signalov v slozhnykh dinamicheskikh usloviyakh* [Processing of hydroacoustic signals in complex dynamic conditions]. Kyiv: Naukova Dumka.
4. Shatalov, G.V. (2010) Modeling of narrow-band noise in the form of separate discrete components and wide-band noise in the form of sound sequences. *Morskaya radioelektronika – Marine radioelectronics*. 2(32). pp. 38–41.
5. Glazov, Yu.E., Rakitina, I.S., Ependiev, M.B., Glazkova, V.N. & Melnikova, E.A. (2009) Metody snizheniya zametnosti uzkoplosnykh protsessov v gidroakusticheskom pole morskikh podvizhnykh ob"ektov (kratkiy obzor) [The techniques for hiding the narrow-band components of the primary underwater sound fields of the sea moving objects (a brief review)]. *Problemy mashinostroeniya i avtomatizatsii*. 3. pp. 69–76.
6. Deev, V.V., Zabrodin, Yu.M., Pakhomov, A.P., Tenetko, V.A. & Titov, M.S. (1990) *Analiz informatsii operatorom-gidroakustikom* [Analysis of information by hydroacoustic operator]. Leningrad: Sudostroenie.
7. Zarayskiy, V.A. & Tyurin, A.M. (1975) *Teoriya gidrolokatsii* [Theory of Hydrolocation]. Leningrad: Naval Academy.
8. Rokhmaniyko, A.Yu., Yakunin, K.V., Saprykin, V.A., Maly, V.V. & Esipov, V.S. (2012) *Ustroystvo obnaruzheniya shumovykh gidroakusticheskikh signalov v vide zvukoryada na osnove vychisleniya integral'nogo veyvlet-spektra* [A device for detecting noise hydroacoustic signals in the form of a scale based on the calculation of the integral wavelet spectrum]. Patent No. 2464588 for Invention of the Russian Federation dated October 20, 2012, with priority dated June 15.
9. Yakovlev, A.N. (2003) *Vvedenie v veyvlet-preobrazovaniya* [Introduction to wavelet transforms]. Novosibirsk: NSTU.

10. Rudko, I.M. (2016) Applying Order Statistics to Detection Problems in the frequency domain. *Upravlenie bol'shimi sistemami*. 62(3). pp. 6–29.
11. Rudko, I.M. (2022) Investigating the effect of filter parameters based on truncated order statistics (tos-filter) on its output characteristics. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 60. pp. 42–51. DOI: 10.17223/19988605/60/5
12. David, H.A. & Nagaraja, H.N. (2003) *Order Statistics*. John Wiley & Sons.
13. Lampert, T.A., O'Keefe, S.E.M. (2013) On the Detection of Tracks in Spectrogram Images. *Pattern Recognition*. 46(5). pp. 1396–1408.
14. Rudko, I.M. (2014) Stokhasticheskaya model' zvukoryada [Stochastic model of harmonic series]. *Upravlenie bol'shimi sistemami*. 49(14). pp. 81–95.

Информация об авторе:

Рудко Игорь Михайлович – кандидат технических наук, старший научный сотрудник Института проблем управления им. В.А.Трапезникова РАН (Москва, Россия). E-mail: igor-rudko@mail.ru

Автор заявляет об отсутствии конфликта интересов.

Information about the author:

Rudko Igor M. (PhD, Tech. Senior Researcher, V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russian Federation). E-mail: igor-rudko@mail.ru

The author declares no conflicts of interests.

Поступила в редакцию 30.07.2022; принята к публикации 01.03.2023

Received 30.07.2022; accepted for publication 01.03.2023

Original article

doi: 10.17223/19988605/62/9

Robust extrapolation for systems with unknown input and interval uncertainty in system and observations

Valery I. Smagin¹, Konstantin S. Kim²

^{1,2} Tomsk State University, Tomsk, Russian Federation

¹ vsm@mail.tsu.ru

² kks93@rambler.ru

Abstract. The problem of robust extrapolation for discrete linear system with unknown input and uncertain interval parameters in system and model of observations is considered. The probabilistic approach is used, which is based on replacing uncertain parameters of interval type by independent random variables with uniform distribution in recursive Kalman schemes. The LSM algorithms and nonparametric smoothing procedures are applied for estimating unknown input. The proposed algorithms can be used in control systems with incomplete information. Simulation results are presented and discussed.

Keywords: robust extrapolation; interval parameters; unknown input; nonparametric smoothing.

For citation: Smagin, V.I., Kim, K.S. (2023) Robust extrapolation for systems with unknown input and interval uncertainty in system and observations. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 62. pp. 85–91. doi: 10.17223/19988605/62/9

Научная статья

УДК 681.5

doi: 10.17223/19988605/62/9

Робастная экстраполяция для систем с неизвестным входом и интервальной неопределенностью в объекте и наблюдениях

Валерий Иванович Смагин¹, Константин Станиславович Ким²

^{1,2} Томский государственный университет, Томск, Россия

¹ vsm@mail.tsu.ru

² kks93@rambler.ru

Аннотация. Рассмотрена задача робастной экстраполяции для дискретного объекта с неизвестным входом и интервальными параметрами в модели объекта и наблюдениях. Используется вероятностный подход, в основе которого лежат замена неопределенных параметров интервального типа независимыми случайными величинами с равномерным распределением, алгоритмы оценивания неизвестного входа, рекуррентные схемы экстраполяции на один такт (экстраполятор Калмана), метод наименьших квадратов и сглаживающие непараметрические процедуры. Представлены результаты моделирования.

Ключевые слова: робастная экстраполяция; интервальные параметры; неизвестный вход; непараметрическое сглаживание.

Для цитирования: Смагин В.И., Ким К.С. Робастная экстраполяция для систем с неизвестным входом и интервальной неопределенностью в объекте и наблюдениях // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2023. № 62. С. 85–91. doi: 10.17223/19988605/62/9

Introduction

The problem of synthesis of filters, extrapolators and observers for dynamical systems with uncertain parameters, in particular, with interval parameters, was considered in [1–5]. These papers use methods of robust data processing. The robust Kalman filter is obtained for time-varying discrete-time linear systems by solving an optimization problem such that the upper bound on the variance of estimation error to be minimized [1]. In [2], the problem of state estimation and determining at any moment the smallest set containing all the possible values of the state vector, simultaneously compatible with the state equations and with a priori known bounds of the uncertain parameters, is considered. In [3], a problem was considered in the case when all the matrices of both the system and the observations model are subjected to norm-limited parametric uncertainties. The robust regularization is implemented using the penalty function method. Robust Kalman filter is proposed in [4], where the problem was solved by using a linear matrix inequality optimization problem. In [5], was proposed the robust Kalman filtering framework for systems with probabilistic uncertainty in system parameters. The uncertainty is propagated using conditional expectations and polynomial chaos expansion framework. Methods of data processing using estimates of unknown input are given in [6, 7]. In these papers Least Squares Method (LSM) was used to obtain estimates of unknown input. In [8–10] it was proposed to use a compensatory approach to calculate estimates of unknown input. In [11, 12] there were used additionally the algorithms of nonparametric smoothing to increase the accuracy of estimating unknown input.

This paper considers the problem of robust extrapolation in discrete systems with additive perturbations with unknown input and with interval parameters. It is based on a probabilistic approach to solving problems for model with interval parameters, which consists in the fact that the interval parameter is replaced by a uniformly distributed random variable [13].

The results of the work generalize results of the paper [14] to the case of the presence interval parameters not in the systems model only but in the observations model also.

1. The problem statement

Consider the linear discrete system with interval parameters, described by the difference equation

$$x(k+1) = \tilde{A}x(k) + f(k) + \tilde{B}q(k), \quad x(0) = x_0, \quad (1)$$

where $x(k) \in \mathbb{R}^n$ is the state vector at time instant k , $f(k)$ is the unknown input vector; x_0 is the random vector with known mathematical expectation and covariance matrix $N_0 = E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T]$; \tilde{A} is the state transition matrix with interval uncertainty (with the given lower and upper bounds of \underline{A} and \bar{A} , respectively), $q(k) \in \mathbb{R}^n$ is the random perturbations with the following characteristics: $E[q(k)] = 0$, $E[q(k)q^T(j)] = I\delta_{kj}$, \tilde{B} is the perturbations transition matrix with interval uncertainty (with the given lower and upper bounds of \underline{B} and \bar{B}), I is the identity matrix. Here δ_{kj} is the Kronecker symbol.

The observations model is determined by the formula

$$y(k) = \tilde{S}x(k) + \tilde{F}v(k), \quad (2)$$

where $y(k) \in \mathbb{R}^l$ is the observations vector, \tilde{S} is the observations transition matrix with interval uncertainty (with the given lower and upper bounds of \underline{S} and \bar{S}), $v(k) \in \mathbb{R}^l$ is the observations noise ($E[v(k)] = 0$, $E[v(k)v^T(j)] = I\delta_{kj}$), \tilde{F} is the observations transition matrix with interval uncertainty (with the given lower and upper bounds of \underline{F} and \bar{F}).

It is assumed that the sequences $q(k)$, $v(k)$ and x_0 are mutually independent, system (1), (2) is observable under parametric perturbations of the dynamics matrix \tilde{A} and observations matrix \tilde{S} . Using the infor-

mation available to the time $k \in [0; T]$, it is required to construct the forecast $\hat{x}(k+1)$ by minimizing the following criterion

$$J(0;T) = E[\sum_{k=0}^T (x(k) - \hat{x}(k))^T (x(k) - \hat{x}(k))]. \quad (3)$$

2. Robust Extrapolator

To solve the problem, we will use the recurrent extrapolator Kalman (EK), while using the probabilistic approach to find its transition matrix. The essence of the method lies in the fact that the interval parameters are replaced by independent random variables that are uniformly distributed over the uncertainty intervals.

Then, using the probabilistic approach, we will replace the uncertain interval matrices $\tilde{A}, \tilde{B}, \tilde{S}, \tilde{F}$ by the matrices whose elements depend on random variables

$$\begin{aligned} A(\theta) &= (A + \sum_{i=1}^m A_i \theta_i), \quad B(\theta) = (B + \sum_{i=m+1}^{m+m_1} B_i \theta_i), \\ S(\theta) &= (S + \sum_{i=m+m_1+1}^{m+m_1+m_2} S_i \theta_i), \quad F(\theta) = (F + \sum_{i=m+m_1+m_2+1}^{m+m_1+m_2+m_3} F_i \theta_i), \end{aligned} \quad (4)$$

where θ_i are independent uniformly distributed random variables according on the interval $[-1, +1]$ ($-1 \leq \theta_i \leq 1$ ($i = \overline{1, m+m_1+m_2+m_3}$)). We will assume that the random variables θ_i are independent of x_0 , $q(k)$ and $v(k)$. In (4), the matrices $A = \frac{1}{2}(A + \bar{A})$, $B = \frac{1}{2}(B + \bar{B})$, $S = \frac{1}{2}(S + \bar{S})$ and $F = \frac{1}{2}(F + \bar{F})$ are the medians of the interval matrices $\tilde{A}, \tilde{B}, \tilde{S}$ and \tilde{F} . The matrices A_i, B_i, S_i and F_i can be set so that one element corresponding to the uncertain element of the matrices $\tilde{A}, \tilde{B}, \tilde{S}$ and \tilde{F} remains nonzero. Its value can be determined by the width of the interval uncertainty of the elements of the matrices $\tilde{A}, \tilde{B}, \tilde{S}$ and \tilde{F} .

In this case, the model of the system (1) and observations (2) takes the form

$$x(k+1) = A(\theta)x(k) + f(k) + B(\theta)q(k), \quad x(0) = x_0, \quad (5)$$

$$y(k) = S(\theta)x(k) + F(\theta)v(k). \quad (6)$$

However, we restrict ourselves to characterizing the first two moments of $x(k)$ as defined below, since we apply the EK. To obtain the estimate, we use the recurrent algorithm

$$\hat{x}(k+1) = A\hat{x}(k) + f(k) + K(k)(y(k) - S\hat{x}(k)), \quad \hat{x}(0) = \bar{x}_0, \quad (7)$$

where the matrix of the extrapolator transition coefficients $K(k)$ is determined by the optimization of criterion (3), taking into account the type of distribution of the parameter θ and assuming that the vector $f(k)$ is known.

Using the property of the trace operation ($y^T A y = \text{tr} A y y^T$) and the rules for differentiating the trace function from the matrix multiplication [15]:

$$\frac{\partial \text{tr} A X B}{\partial X} = A^T B^T, \quad \frac{\partial \text{tr} A^T X B^T}{\partial X} = B A, \quad (8)$$

from the equation

$$\frac{\partial J(0;T)}{\partial K} = 0, \quad (9)$$

we obtain an analytical expression for the matrix $K(k)$

$$K(k) = A N(k) S^T (F F^T + S N(k) S^T + V(k))^{-1}, \quad (10)$$

where the matrix $V(k)$ is determined by the formula

$$V(k) = \frac{1}{3} \sum_{i=m+m_1+1}^{m+m_1+m_2} S_i N(k) S_i^T + \frac{1}{3} \sum_{i=m+m_1+1}^{m+m_1+m_2} S_i \hat{x}(k) \hat{x}(k)^T S_i^T + \frac{1}{3} \sum_{i=m+m_1+m_2+1}^{m+m_1+m_2+m_3} F_i F_i^T, \quad (11)$$

and $N(k)$ satisfies the difference matrix equation

$$N(k+1) = (A - K(k)S)N(k)(A - K(k)S)^T + \frac{1}{3} \sum_{i=1}^m A_i N(k) A_i^T + \frac{1}{3} \sum_{i=1}^m A_i \hat{x}(k) \hat{x}(k)^T A_i^T + BB^T + \frac{1}{3} \sum_{i=m+1}^{m+m_2} B_i B_i^T + K(k)(FF^T + SN(k)S^T + V(k))K(k)^T, \quad N(0) = N_0. \quad (12)$$

In model (5) due to the fact that the median of the interval matrix \tilde{A} is used as the dynamics matrix EK, the vector of the unknown input will change (this vector is denoted by $r(k)$)

$$r(k) = f(k) + \sum_{i=1}^m A_i \theta_i x(k), \quad -1 \leq \theta_i \leq 1 (i = \overline{1, m}), \quad (13)$$

where the second term is an additional unknown input arising from the uncertainty of the state transition matrix.

As an algorithm for estimating the unknown input $r(k)$, we will use the LSM method, in this case, the estimate can be constructed on the basis of minimizing the additional criterion [6, 7]

$$I_1 = \sum_{t=1}^k \left\{ \|y(t) - S(A\hat{x}(t-1) + r(t-1))\|_C^2 + \|r(t-1)\|_D^2 \right\}, \quad (14)$$

In (14) C, D are positive definite weight matrices. The LSM estimates of the unknown input, based on the minimization of criterion (14), will take the form

$$\hat{r}^{(LSM)}(k) = [S^T C S + D]^{-1} S^T C [y(k) - S A \hat{x}(k-1)]. \quad (15)$$

To increase the accuracy of estimating an unknown input, we will additionally use nonparametric algorithms [11, 12] for smoothing the innovation process $y(k) - S A \hat{x}(k-1)$

$$\hat{r}^{(NP)}(k) = [S^T C S + D]^{-1} S^T C \hat{\Omega}, \quad (16)$$

where the j component of the vector $\hat{\Omega}(k)$ can be calculated by

$$\hat{\Omega}_j(k) = \frac{\sum_{t=1}^k [y(t) - S(A\hat{x}(t-1))]_j G\left(\frac{k-t+1}{\mu_j}\right)}{\sum_{t=1}^k G\left(\frac{k-t+1}{\mu_j}\right)}. \quad (17)$$

In formulas (17) $G(\cdot)$ is a kernel function, μ_j is a bandwidth parameters.

Robust extrapolation estimates in discrete systems with interval parameters were determined from the recurrent equation

$$\hat{x}(k+1) = A\hat{x}(k) + \hat{r}(k) + K(k)(y(k) - S\hat{x}(k)), \quad \hat{x}(0) = \bar{x}_0, \quad (18)$$

where the matrix transition coefficients $K(k)$ was calculated by formulas (10)–(12), and the estimates of $\hat{r}(k)$ was determined by formulas (16), (17). Note that the medians of the matrices \tilde{A}, \tilde{S} are used to calculate the estimates of $\hat{r}(k)$.

3. Simulation Results

The simulation was performed for the following data ($m=2, m_1=1, m_2=2, m_3=2$):

$$A = \begin{pmatrix} 0 & 1 \\ 0,02 & 0,73 \end{pmatrix}, \quad S = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 0 & 0 \\ 0 & 0,4 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 \\ 0,04 & 0 \end{pmatrix}, \quad S_3 = \begin{pmatrix} 0,05 & 0 \\ 0 & 0 \end{pmatrix},$$

$$B = \begin{pmatrix} 0,1 & 0 \\ 0 & 0,15 \end{pmatrix}, \quad B_4 = \begin{pmatrix} 0,05 & 0 \\ 0 & 0 \end{pmatrix}, \quad B_5 = \begin{pmatrix} 0 & 0 \\ 0 & 0,1 \end{pmatrix},$$

$$F = \begin{pmatrix} 0,5 & 0 \\ 0 & 0,6 \end{pmatrix}, \quad F_6 = \begin{pmatrix} 0,4 & 0 \\ 0 & 0 \end{pmatrix}, \quad F_7 = \begin{pmatrix} 0 & 0 \\ 0 & 0,4 \end{pmatrix},$$

$$C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, D = \begin{pmatrix} 0,1 & 0 \\ 0 & 0,1 \end{pmatrix},$$

$$f(k) = \begin{pmatrix} 0,1 + 0,25 \sin(0,1k) \\ 0,1 + 0,2 \sin(0,15k) \end{pmatrix}.$$

The initial conditions are:

$$x(0) = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \hat{x}(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, N(0) = \begin{pmatrix} 0,1 & 0 \\ 0 & 0,1 \end{pmatrix}.$$

In (17), a Gaussian kernel function was used.

Fig. 1 shows the results of comparing the standard errors of the deviations of the estimates of the state vector

$$\sigma_i = \sqrt{\frac{\sum_{k=1}^N (x_i(k) - \hat{x}_i(k))^2}{N-1}}, (i = \overline{1,2})$$

for five algorithms (with different implementations of the components of a random vector θ):

- optimal EK for systems with median matrices A, B, S and F , when $f(k)$ is known (OEK);
- optimal EK for systems with median matrices A, B, S and F , when $f(k)$ is unknown, estimate $f(k)$ is not used (OEKUN);
- optimal EK for systems with median matrices A, B, S and F , LSM method (15) was used to calculate estimates of an unknown input, nonparametric smoothing was not applied for estimate unknown input (EK-LSM);
- optimal EK for systems with median matrices A, B, S and F , LSM method and nonparametric smoothing (formulas (16), (17)) was applied (EK-LSM-NP);
- proposed robust EK (10)–(12) for systems with interval parameters, LSM method and nonparametric smoothing were used (REK).

The table presents 10 realizations of the values of the components of the random vector θ distributed with uniform density. Simulation results are obtained for extrapolation algorithms ($N = 200$) and by averaging 100 realizations.

Realizations of the values of the components of the random vector θ distributed with uniform density

n/n	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	θ_7
1	-0,97	0,99	-0,39	-0,73	-0,63	-0,24	0,35
2	-0,83	0,84	0,07	0,39	0,15	-0,18	-0,02
3	-0,82	0,30	0,57	-0,06	0,25	-0,48	0,18
4	-0,18	0,88	-0,56	-0,24	0,52	0,08	0,72
5	0,12	0,07	0,01	-0,73	-0,71	-0,04	0,38
6	-0,87	-0,06	0,25	0,19	-0,03	-0,73	0,92
7	-0,48	0,57	-0,15	-0,87	-0,98	-0,84	0,64
8	-0,80	0,38	-0,69	-0,94	0,08	-0,08	-0,94
9	0,56	0,72	0,08	-0,80	0,59	-0,13	0,92
10	0,13	-0,20	0,45	-0,85	-0,07	-0,77	-0,25

Fig. 1 shows that the procedures with robust extrapolation (REK) have the advantages in the accuracy compared to the known algorithms using the LSM estimates and LSM estimates with nonparametric smoothing. The advantage (REK) in accuracy compared to the (EK-LSM-NP) algorithm is from 3% to 15%.

The worst results were obtained using estimates of an unknown input using the LSM method without smoothing (EK-LSM). The reason for this is the high level of intensity of measurements errors, which was used in the example, which led to a low quality of the estimation of $f(k)$ by the LSM method and, as a consequence, to a low accuracy of the estimation of the state vector. Smoothing the innovation process and proposed robust algorithm improve accuracy estimates of the state vector.

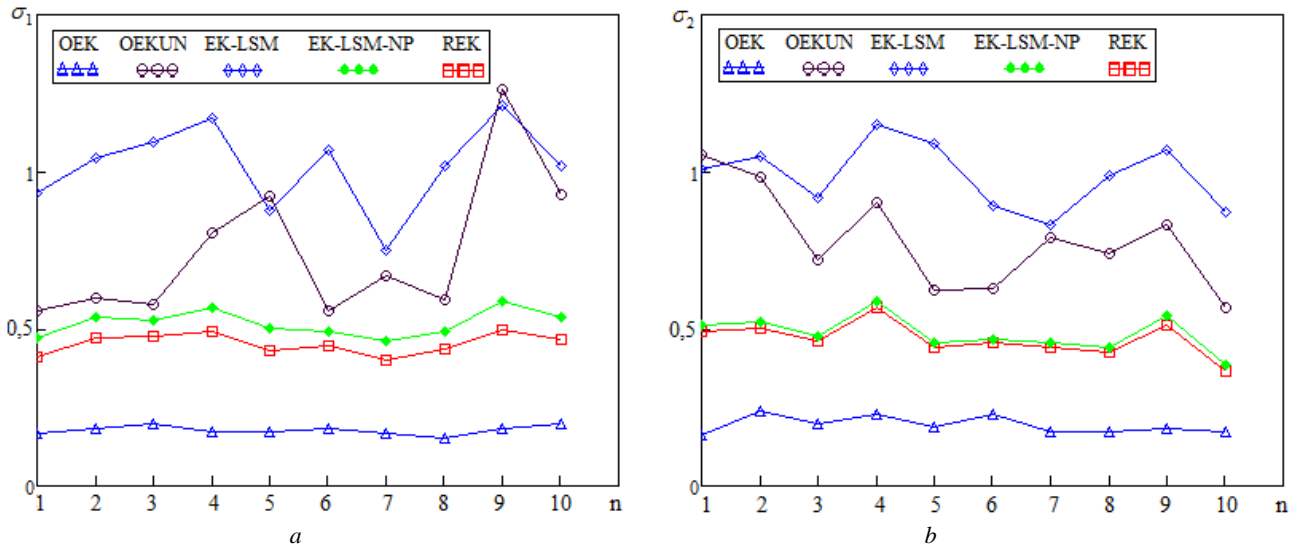


Fig. 1. Standard errors for extrapolation algorithms for ten simulation results
 a) first component, b) second component

The algorithm (OEK) determines the potential capabilities of the EK algorithm, but algorithm (OEK) requires the exact values of the input vector $f(k)$. In our problem, the input vector $f(k)$ is not available for observation.

Conclusions

Using the probabilistic approach, algorithm for the synthesis of the robust extrapolator for discrete systems with unknown input and with interval parameters in the model and observations is proposed.

The problem is solved using recurrent algorithms, the LSM method and nonparametric smoothing procedures. The proposed method implements a decrease in the influence of uncertainties in the model and observations using replacing an interval uncertainty by the probabilistic uncertainty and taking into account estimates of the unknown input with additional smoothing.

The numerical example shows that the joint use of smoothing algorithms and robust approach can improve the estimation accuracy of the state vector.

References

1. Abolhasani, M. & Rahmani, M. (2017) Robust Kalman filtering for discrete-time systems with stochastic uncertain time-varying parameters. *Electronics Letters*. 53(3). pp. 146–148.
2. Ichalal, D., Marx, B., Maquin, D. & Ragot, J. (2018) State estimation of system with bounded uncertain parameters: interval multi-model approach. *International Journal of Adaptive Control and Signal Processing*. 32(3). pp. 480–493.
3. Rocha, K.D.T. & Terra, M.H. (2021) Robust Kalman filter for systems subject to parametric uncertainties. *Systems and Control Letters*. 157. Art. 105034.
4. Abolhasani, M. & Rahmani, M. (2018) Robust Kalman filtering for discrete-time time-varying systems with stochastic and norm-bounded uncertainties. *Journal of Dynamic Systems, Measurement, and Control*. 140(3). Art. No. 030901.
5. Kim, S., Deshpande, V.M. & Bhattacharya, R. (2021) Robust Kalman filtering with probabilistic uncertainty in system parameters. *IEEE Control Systems Letters*. 5(1). pp. 295–300.
6. Janczak, D. & Grishin, Yu. (2006) State estimation of linear dynamic system with unknown input and uncertain observation using dynamic programming. *Control and Cybernetics*. 4. pp. 851–862.
7. Witczak, M. (2014) Fault diagnosis and fault-tolerant control strategies for non-linear systems. Chapter 2. Unknown input observers and filters. *Lecture Notes in Electrical Engineering*. Springer International Publishing, Switzerland. pp. 19–56.
8. Smagin, V.I. & Smagin, S.V. (2011) Filtering for linear not stationary discrete system with unknown disturbances. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 16(3). pp. 43–50.
9. Smagin, V.I. (2014) State estimation for linear discrete systems with unknown input using compensations. *Russian Physics Journal*. 57(5). pp. 682–690. DOI: 10.1007/s11182-014-0291-x

10. Smagin, V.I. & Koshkin, G.M. (2015) Kalman filtering and control algorithms for systems with unknown disturbances and parameters using nonparametric technique. *Proceedings 20th International Conference on Methods and Models in Automation and Robotics (MMAR 2015)*. Miedzyzdroje, Poland. pp. 247–251.
11. Koshkin, G.M. & Smagin, V.I. (2016) Kalman filtering and forecasting algorithms with use of nonparametric functional estimators. *Springer Proceedings in Mathematics & Statistics*. 2nd Conference of the International-Society-for-Nonparametric-Statistics (ISNPS). Vol. 175. pp. 75–84. DOI: 10.1007/978-3-319-41582-6_6
12. Smagin V., Koshkin G. & Udod V. (2015) State estimation for linear discrete-time systems with unknown input using nonparametric technique. *Proceedings of the International Conference on Artificial Intelligence and Control Automation (AICA 2015)*. Atlantis Press, Bangkok, Thailand. pp. 675–677. DOI: 10.2991/cisia-15.2015.184
13. Barmish, B.R. & Polyak, B.T. (1996) A new approach to open robustness problems based on probabilistic predication formulae. Vol. H. *Proceedings 13th World IFAC Congr.* San Francisco. USA. pp. 1–6.
14. Kim, K.S. & Smagin, V.I. (2022) Robust extrapolation in discrete systems with interval parameters using algorithms for estimating unknown input. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*. 59. pp. 47–54. DOI: 10.17223/19988605/59/5
15. Athans, M. (1968) The matrix minimum principle. *Informat. and Contr.* 11. pp. 592–606.

Information about the authors:

Smagin Valery I. (Doctor of Technical Science, Professor, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: vsm@mail.tsu.ru

Kim Konstantin S. (Candidate of Physics and Mathematics, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: kks93@rambler.ru

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

Информация об авторах:

Смагин Валерий Иванович – профессор, доктор технических наук, профессор кафедры прикладной математики Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: vsm@mail.tsu.ru

Ким Константин Станиславович – кандидат физико-математических наук, доцент Высшей IT школы Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: kks93@rambler.ru

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

Поступила в редакцию 16.10.2022; принята к публикации 01.03.2023

Received 16.10.2022; accepted for publication 01.03.2023

Научная статья

УДК 681.51:519.6

doi: 10.17223/19988605/62/10

Вычисление коэффициента корреляции случайных величин с использованием дополнительной информации о форме их взаимосвязи

Виктор Анатольевич Удод

Томский государственный университет, Томск, Россия, pr.udod@mail.ru

Аннотация. Задача вычисления коэффициента корреляции двух случайных величин рассматривается в предположении, что математическое ожидание их произведения неизвестно, но известны первые два момента распределения каждой из них. Решение предлагается для случая, когда набор возможных значений одной величины определенным образом зависит от фиксированного значения другой величины. Показано, что с использованием дополнительной информации о форме взаимосвязи исследуемых случайных величин можно существенно упростить расчет коэффициента корреляции.

Ключевые слова: случайные величины; коэффициент корреляции; условное математическое ожидание.

Для цитирования: Удод В.А. Вычисление коэффициента корреляции случайных величин с использованием дополнительной информации о форме их взаимосвязи // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2023. № 62. С. 92–100. doi: 10.17223/19988605/62/10

Original article

doi: 10.17223/19988605/62/10

Calculation of the correlation coefficient of random variables using additional information about the form of their relationship

Victor A. Udod

Tomsk State University, Tomsk, Russian Federation, pr.udod@mail.ru

Abstract. The problem of calculating the correlation coefficient of two random variables is considered under the assumption that the mathematical expectation of their product is unknown, but the first two moments of the distribution of each of them are known. The solution is proposed for the case when the set of possible values of one quantity depends in a certain way on a fixed value of another quantity. It is shown that with the use of additional information about the form of the relationship of the studied random variables, it is possible to significantly simplify the calculation of the correlation coefficient.

Keywords: random variables; correlation coefficient; conditional mathematical expectation.

For citation: Udod, V.A. (2023) Calculation of the correlation coefficient of random variables using additional information about the form of their relationship. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 62. pp. 92–100. doi: 10.17223/19988605/62/10

Введение

В настоящее время корреляционный анализ широко используется практически во всех сферах науки и техники: в физике [1], химии [2], биологии [3], медицине [4], экономике [5], строительстве [6], а также в образовании [7] и т.д. При этом значительная часть исследований при решении той или

иной задачи зачастую сводится к проведению парного корреляционного анализа, а именно к вычислению коэффициента корреляции Пирсона r_{xy} случайных величин X и Y :

$$r_{xy} = \frac{\text{cov}(X, Y)}{\sigma_x \sigma_y} = \frac{M(XY) - M(X)M(Y)}{\sqrt{D(X)D(Y)}}. \quad (1)$$

Здесь $\text{cov}(X, Y) = M(XY) - M(X)M(Y)$ – ковариация случайных величин X и Y ; $M(XY)$ – математическое ожидание произведения случайных величин X и Y ; $M(X)$ и $M(Y)$ – математические ожидания, $\sigma_x = \sqrt{D(X)}$ и $\sigma_y = \sqrt{D(Y)}$ – средние квадратические отклонения, а $D(X)$ и $D(Y)$ – дисперсии случайных величин X и Y соответственно.

При вычислении коэффициента корреляции возможна ситуация, когда известны лишь числовые характеристики $M(X)$, $M(Y)$, $D(X)$ и $D(Y)$ исследуемых случайных величин X и Y , а $M(XY)$ неизвестно. В этом случае применение формулы (1) становится затруднительным либо вовсе невозможным. Определенным выходом из такой ситуации может служить получение статистической оценки параметра r_{xy} , т.е. вычисление выборочного коэффициента корреляции r_{xy}^* . Подобный подход часто применяется при решении как общетеоретических, так и прикладных задач. Между тем если для вычисления выборочного коэффициента корреляции r_{xy}^* необходимо проведение большого объема физических (химических, биологических и т.п.) экспериментов или имитационного моделирования, то эффективность такого подхода заметно снижается.

Таким образом, актуальной является задача вычисления коэффициента корреляции r_{xy} на основе первых двух моментов распределения случайных величин X и Y и некоторой дополнительной (априорной) информации, характеризующей взаимосвязь между этими величинами. Решению данной задачи применительно к некоторым случайным величинам с определенной формой их взаимосвязи и посвящена настоящая работа.

1. Постановка задачи

Пусть X и Y есть случайные величины, удовлетворяющие определенным предположениям, описание которых приведено ниже.

Если X и Y – непрерывные случайные величины, то предполагается, что величина X может принимать только неотрицательные значения, а при заданном значении x величины X множество возможных значений величины Y представляет собой отрезок $[0, x]$.

Если же X и Y – дискретные случайные величины, то предполагается, что величина X может принимать только неотрицательные целочисленные значения, а при заданном значении x_i величины X величина Y может принимать значения $0, 1, 2, \dots, x_i$ (неотрицательные целочисленные значения, не превосходящие значения x_i).

Если X и Y – непрерывные случайные величины, то формализованное представление описанной взаимосвязи между величинами X и Y имеет вид:

$$A(X) \subset [0, \infty); A(Y|x) = [0, x], \quad (2)$$

если X и Y – дискретные случайные величины, то

$$A(X) \subset Z_0; A(Y|x_i) = \{0, 1, 2, \dots, x_i\}. \quad (3)$$

Здесь $A(X)$ – множество всех возможных значений случайной величины X ; $A(Y|x)$ и $A(Y|x_i)$ – множество всех возможных значений случайной величины Y при $X = x$ и $X = x_i$ соответственно; Z_0 – множество неотрицательных целых чисел.

Требуется вычислить коэффициент корреляции случайных величин X и Y , удовлетворяющих ограничению (2) или (3) в предположении, что математическое ожидание их произведения неизвестно, но известны первые два момента распределения каждой из них.

Следует заметить, что взаимосвязь между случайными величинами, описываемая соотношениями (2) и (3), является весьма распространенной. В подтверждение этому приведем два примера из физики:

1. Случайная величина X – энергия кванта тормозного рентгеновского излучения, а случайная величина Y – энергия, переданная этим квантом поглотителю в результате взаимодействия с ним (соответствует ограничению (2)).

2. Случайная величина X – число квантов рентгеновского излучения, упавших на детектор за определенный промежуток времени, а случайная величина Y – число квантов, зарегистрированных детектором за тот же самый промежуток времени (соответствует ограничению (3)).

2. Подход к решению задачи и полученные результаты

При ограничении (2) или (3) случайная величина Y представляет собой некую случайную долю случайной величины X . Вследствие этого для решения поставленной задачи представляется целесообразным применить подход, основанный на предположении следующей формальной взаимосвязи данных величин:

$$Y = \alpha X, \quad (4)$$

где α – независимая от X случайная величина, распределенная на промежутке $[0, 1]$.

Взяв математическое ожидание от обеих частей равенства (4) и учитывая независимость величин α и X , получим

$$M(Y) = M(\alpha X) = M(\alpha)M(X),$$

откуда

$$M(\alpha) = \frac{M(Y)}{M(X)}. \quad (5)$$

Из (4) и независимости случайных величин α и X вытекает равенство

$$M(XY) = M(X\alpha X) = M(\alpha X^2) = M(\alpha)M(X^2).$$

Отсюда, учитывая (5), получаем

$$M(XY) = \frac{M(Y)}{M(X)}M(X^2). \quad (6)$$

При подстановке (6) в (1) будем иметь

$$\begin{aligned} r_{xy} &= \frac{\frac{M(Y)}{M(X)}M(X^2) - M(X)M(Y)}{\sqrt{D(X)D(Y)}} = \frac{\frac{M(Y)}{M(X)}(M(X^2) - M^2(X))}{\sqrt{D(X)D(Y)}} = \\ &= \frac{\frac{M(Y)}{M(X)}D(X)}{\sqrt{D(X)D(Y)}} = \frac{M(Y)}{M(X)}\sqrt{\frac{D(X)}{D(Y)}}. \end{aligned} \quad (7)$$

Таким образом, на основе предположения (4) согласно (7) нами получена следующая формула для вычисления коэффициента корреляции случайных величин X и Y , удовлетворяющих ограничению (2) или (3):

$$r_{xy} = \frac{M(Y)}{M(X)}\sqrt{\frac{D(X)}{D(Y)}}.$$

Между тем предположение (4) для случайных величин X и Y , удовлетворяющих ограничению (2) или (3), в действительности может и не выполняться. Тогда формулу (6) следует воспринимать как приближенную, а значит, и полученную выше формулу для вычисления коэффициента корреляции этих величин также следует воспринимать как приближенную, т.е.

$$r_{xy} \approx \frac{M(Y)}{M(X)}\sqrt{\frac{D(X)}{D(Y)}}. \quad (8)$$

При этом вполне закономерно возникает задача оценки погрешности формулы (8). Однако ее решение весьма затруднительно вследствие неограниченного многообразия законов распределения. В то же время можно указать условие (более общее по сравнению с (4)), при соблюдении которого формула (8) является точной, т.е. коэффициенты корреляции, вычисляемые по формуле (8) и по «истинной» формуле (1), будут одинаковыми.

Формула (8) является логическим следствием из (1) и (6), а значит, учитывая справедливость формулы (1), она будет точной, если формула (6) будет верна (будет точной).

Нетрудно убедиться, что формула (6) будет верна (будет точной), если условное математическое ожидание (УМО) $M(Y|x)$ имеет вид:

$$M(Y|x) = \theta x \quad (9)$$

для непрерывных случайных величин X и Y и, соответственно, вид:

$$M(Y|x_i) = \theta x_i \quad (10)$$

для дискретных случайных величин X и Y , где θ – постоянный множитель.

Покажем это на примере непрерывных случайных величин X и Y (для дискретных случайных величин ход рассуждений будет аналогичным). С этой целью введем обозначения: $f(x, y)$ – плотность вероятности непрерывной двумерной случайной величины (X, Y) ; $\varphi(x)$ – плотность вероятности случайной величины X ; $\psi(y|x)$ – условная плотность вероятности случайной величины Y при $X = x$.

С учетом принятых обозначений формула (6) примет вид:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x, y)dxdy = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} yf(x, y)dxdy}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xf(x, y)dxdy} \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^2 f(x, y)dxdy. \quad (11)$$

Здесь и всюду в дальнейшем предполагается, что интегралы сходятся и при этом знаменатель в правой части (11) отличен от нуля.

Воспользуемся разложением

$$f(x, y) = \varphi(x)\psi(y|x). \quad (12)$$

При подстановке (12) в (11) будем иметь

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy\varphi(x)\psi(y|x)dxdy = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y\varphi(x)\psi(y|x)dxdy}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x\varphi(x)\psi(y|x)dxdy} \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^2\varphi(x)\psi(y|x)dxdy. \quad (13)$$

Переходя в (13) от двойных интегралов к повторным интегралам получим

$$\int_{-\infty}^{\infty} x\varphi(x) \left(\int_{-\infty}^{\infty} y\psi(y|x)dy \right) dx = \frac{\int_{-\infty}^{\infty} \varphi(x) \left(\int_{-\infty}^{\infty} y\psi(y|x)dy \right) dx}{\int_{-\infty}^{\infty} x\varphi(x) \left(\int_{-\infty}^{\infty} \psi(y|x)dy \right) dx} \cdot \int_{-\infty}^{\infty} x^2\varphi(x) \left(\int_{-\infty}^{\infty} \psi(y|x)dy \right) dx. \quad (14)$$

Условная плотность вероятности (так же как и безусловная) обладает свойством нормировки, т.е.

$$\int_{-\infty}^{\infty} \psi(y|x)dy = 1. \quad (15)$$

С учетом (15) равенство (14) примет вид:

$$\int_{-\infty}^{\infty} x\varphi(x) \left(\int_{-\infty}^{\infty} y\psi(y|x)dy \right) dx = \frac{\int_{-\infty}^{\infty} \varphi(x) \left(\int_{-\infty}^{\infty} y\psi(y|x)dy \right) dx}{\int_{-\infty}^{\infty} x\varphi(x)dx} \cdot \int_{-\infty}^{\infty} x^2\varphi(x)dx. \quad (16)$$

Каждый из внутренних интегралов в (16) представляет собой УМО $M(Y|x)$. Следовательно, (16) можно переписать в виде:

$$\int_{-\infty}^{\infty} x\varphi(x)M(Y|x)dx = \frac{\int_{-\infty}^{\infty} \varphi(x)M(Y|x)dx}{\int_{-\infty}^{\infty} x\varphi(x)dx} \cdot \int_{-\infty}^{\infty} x^2\varphi(x)dx. \quad (17)$$

Наконец, подставляя (9) в (17), получим тождество (учитывая, что $\theta = \text{const}$)

$$\int_{-\infty}^{\infty} \theta x^2\varphi(x)dx = \frac{\int_{-\infty}^{\infty} \theta x\varphi(x)dx}{\int_{-\infty}^{\infty} x\varphi(x)dx} \cdot \int_{-\infty}^{\infty} x^2\varphi(x)dx,$$

что и доказывает справедливость формулы (6) при условии (9).

3. Примеры условных законов распределения, когда полученная формула является точной

Приведем сначала два примера условных законов распределения непрерывной случайной величины Y на промежутке $[0, x]$, когда равенство (9) верно, а значит, формула (8) является точной. Предварительно заметим, что условие нормировки условной плотности вероятности $\psi(y|x)$ и выражение для УМО $M(Y|x)$ при этом соответственно будут иметь следующий вид:

$$\int_0^x \psi(y|x)dy = 1; \quad (18)$$

$$M(Y|x) = \int_0^x y\psi(y|x)dy. \quad (19)$$

В качестве первого примера рассмотрим распределение случайной величины Y на промежутке $[0, x]$, которое обладает симметрией относительно центра данного промежутка, т.е. когда условная плотность вероятности $\psi(y|x)$ характеризуется свойством

$$\psi(y|x) = \psi(x - y|x) \text{ для } \forall y \in [0, x] \text{ и } \psi(y|x) = 0, \text{ если } y \notin [0, x]. \quad (20)$$

Заметим, что распространенным частным случаем симметричного распределения (20) является равномерное распределение

$$\psi(y|x) = \begin{cases} \frac{1}{x}, & \text{если } y \in [0, x], \\ 0, & \text{если } y \notin [0, x]. \end{cases}$$

С учетом (20) и (18) будем иметь для УМО (19) (произведя в одном из интегралов, приведенных ниже, замену переменной y на переменную z по формуле $y = x - z$):

$$\begin{aligned} M(Y|x) &= \int_0^x y\psi(y|x)dy = \frac{1}{2} \left[\int_0^x y\psi(y|x)dy + \int_0^x y\psi(y|x)dy \right] = \\ &= \frac{1}{2} \left[\int_0^x y\psi(y|x)dy - \int_x^0 (x-z)\psi(x-z|x)dz \right] = \frac{1}{2} \left[\int_0^x y\psi(y|x)dy + \int_0^x (x-z)\psi(x-z|x)dz \right] = \\ &= \frac{1}{2} \left[\int_0^x y\psi(y|x)dy + \int_0^x (x-y)\psi(x-y|x)dy \right] = \frac{1}{2} \left[\int_0^x y\psi(y|x)dy + \int_0^x (x-y)\psi(y|x)dy \right] = \\ &= \frac{1}{2} \int_0^x (y+x-y)\psi(y|x)dy = \frac{1}{2} \int_0^x x\psi(y|x)dy = \frac{1}{2} x \int_0^x \psi(y|x)dy = \frac{1}{2} x. \end{aligned}$$

Таким образом, в данном случае равенство (9) верно, а коэффициент $\theta = 1/2$.

В качестве второго примера рассмотрим условное распределение случайной величины Y на промежутке $[0, x]$ по степенному закону, т.е.

$$\psi(y|x) = \begin{cases} \frac{a+1}{x^{a+1}} y^a, & \text{если } y \in [0, x], \\ 0, & \text{если } y \notin [0, x], \end{cases}$$

где параметр $a > 0$. При подстановке данной плотности вероятности в (19) получим

$$M(Y|x) = \frac{a+1}{a+2} x.$$

Таким образом, в данном случае равенство (9) также верно, и при этом коэффициент $\theta = a + 1/a + 2$.

Приведем теперь три примера условных законов распределения дискретной случайной величины Y , удовлетворяющей ограничению (3), когда (10) верно, а значит, формула (8) является точной. Предварительно заметим, что согласно ограничению (3) случайная величина X принимает только целочисленные неотрицательные значения, т.е. все x_i ($i = 1, 2, \dots, n$) – целые неотрицательные числа, а при $X = x_i$ случайная величина Y распределена на множестве целых чисел $\{0, 1, 2, \dots, x_i\}$, т.е. имеет следующие возможные значения:

$$y_j = j, \quad (j = 0, 1, 2, \dots, x_i). \quad (21)$$

Введем обозначение вероятности того, что случайная величина Y примет значение $y_j = j$ ($j = 0, 1, 2, \dots, x_i$) при условии, что случайная величина X приняла значение x_i ($i = 1, 2, \dots, n$):

$$S(j|i) = P\{Y = y_j | X = x_i\} \quad (22)$$

Заметим при этом, что свойство нормировки условной вероятности $S(j|i)$, учитывая (21), будет иметь вид:

$$\sum_{j=0}^{x_i} S(j|i) = 1. \quad (23)$$

Рассмотрим теперь в качестве первого примера распределение случайной величины Y на множестве $\{0, 1, 2, \dots, x_i\}$, которое обладает симметрией относительно медианы (серединного элемента) данного множества, т.е. когда условная вероятность $S(j|i)$ характеризуется свойством

$$S(j|i) = S(x_i - j|i), \quad \forall j = 0, 1, 2, \dots, x_i. \quad (24)$$

Отметим, что распространенным частным случаем указанного условного распределения является дискретное равномерное распределение случайной величины Y на множестве $\{0, 1, 2, \dots, x_i\}$:

$$S(j|i) = \frac{1}{x_i + 1}, \quad \forall j = 0, 1, 2, \dots, x_i.$$

С учетом (21), (23) и (24) будем иметь для УМО $M(Y|x_i)$ (произведя в одной из сумм, приведенных ниже, замену переменной j на переменную k по формуле $j = x_i - k$):

$$\begin{aligned} M(Y|x_i) &= \sum_{j=0}^{x_i} y_j S(j|i) = \sum_{j=0}^{x_i} j S(j|i) = \frac{1}{2} \left[\sum_{j=0}^{x_i} j S(j|i) + \sum_{j=0}^{x_i} j S(j|i) \right] = \\ &= \frac{1}{2} \left[\sum_{j=0}^{x_i} j S(j|i) + \sum_{k=x_i}^0 (x_i - k) S(x_i - k|i) \right] = \frac{1}{2} \left[\sum_{j=0}^{x_i} j S(j|i) + \sum_{k=0}^{x_i} (x_i - k) S(x_i - k|i) \right] = \\ &= \frac{1}{2} \left[\sum_{j=0}^{x_i} j S(j|i) + \sum_{j=0}^{x_i} (x_i - j) S(x_i - j|i) \right] = \frac{1}{2} \left[\sum_{j=0}^{x_i} j S(j|i) + \sum_{j=0}^{x_i} (x_i - j) S(j|i) \right] = \frac{1}{2} \sum_{j=0}^{x_i} x_i S(j|i) = \\ &= \frac{1}{2} x_i \sum_{j=0}^{x_i} S(j|i) = \frac{1}{2} x_i. \end{aligned}$$

Таким образом, в данном случае равенство (10) верно, а коэффициент $\theta = 1/2$.

В качестве второго рассмотрим следующий пример. Случайная величина Y при $X = x_i$ распределена по биномиальному закону с параметрами x_i (число испытаний Бернулли) и p (вероятность успеха (некоторого события) в отдельном испытании Бернулли). Тогда согласно [8]

$$M(Y|x_i) = x_i p,$$

а значит, в данном случае равенство (10) также верно, а коэффициент $\theta = p$.

Рассмотрим теперь третий пример. Пусть имеется некоторая генеральная совокупность объема N , в которой M ($0 < M < N$) объектов обладают заданным свойством. Предположим, что объем выборки, подлежащей извлечению из этой генеральной совокупности, есть случайная величина X , которая может принимать значения $1, 2, \dots, n$. Тогда случайная величина Y – число объектов в выборке, обладающих заданным свойством, при $X = x_i = i$ ($i = 1, 2, \dots, n$) будет иметь гипергеометрическое распределение с параметрами x_i, M, N . Следовательно, согласно [8]

$$M(Y|x_i) = x_i \frac{M}{N},$$

а значит, в данном случае равенство (10) также верно, а коэффициент $\theta = \frac{M}{N}$. Заметим, что условие

$A(Y|x_i) = \{0, 1, 2, \dots, x_i\}$ из ограничения (3) при этом будет выполнено, если положить $M \geq n$; $N - M \geq n$.

Применим теперь формулу (6) для приближенного вычисления коэффициента корреляции двух случайных слагаемых случайной величины.

4. Вычисление коэффициента корреляции двух случайных слагаемых случайной величины

Пусть случайная величина X разбивается случайным образом на сумму двух случайных величин Y и Z , т.е.

$$X = Y + Z. \quad (25)$$

Требуется вычислить коэффициент корреляции случайных величин Y и Z .

Из (25) следует

$$\text{cov}(Y, Z) = M(XY) - M(X)M(Y) - D(Y).$$

Отсюда, принимая во внимание приближенную (в общем случае, т.е. без предположения (4)) формулу (6), будем иметь

$$\text{cov}(Y, Z) \approx \frac{M(Y)}{M(X)} D(X) - D(Y). \quad (26)$$

Из (26) окончательно получаем искомую приближенную формулу для вычисления коэффициента корреляции двух случайных слагаемых случайной величины

$$r_{yz} \approx \frac{\frac{M(Y)}{M(X)} D(X) - D(Y)}{\sqrt{D(Y)D(Z)}}. \quad (27)$$

Заметим, что полученная формула (27) является точной при соблюдении условия (9) для непрерывных величин или условия (10) для дискретных величин.

5. Пример использования полученных результатов

Пусть случайная величина X распределена по закону Пуассона с параметром λ . Случайные величины Y и Z удовлетворяют уравнению (25) и при $X = n$ распределены по биномиальному закону соответственно с параметрами n, p и $n, 1 - p$. Тогда (в чем несложно убедиться) безусловные законы распределения случайных величин Y и Z будут представлять собой законы распределения Пуассона

с параметрами λp и $\lambda(1-p)$ соответственно. Очевидно, что пары случайных величин X и Y , а также X и Z удовлетворяют ограничению (3).

Отметим как один из возможных вариантов следующий физический смысл рассматриваемых случайных величин: случайная величина X – число квантов моноэнергетического излучения, упавших на поглотитель за фиксированный промежуток времени, случайные величины Y и Z – число квантов среди тех, которые упали на поглотитель и при этом соответственно испытали / не испытали взаимодействие с поглотителем.

Применяя формулы (8) и (27) к рассматриваемым величинам, будем иметь

$$r_{xy} \approx \sqrt{p}; r_{xz} \approx \sqrt{1-p}; r_{yz} \approx 0. \quad (28)$$

С целью проверки полученных результатов (28) было проведено компьютерное моделирование этих случайных величин с последующим расчетом выборочных коэффициентов корреляции r^* при разных сочетаниях параметров λ и p . При этом для каждой пары значений данных параметров разыгрывалось 10 000 реализаций исследуемых случайных величин. Результаты моделирования приведены в таблице. Там же для наглядности приведены теоретические коэффициенты корреляции r , вычисленные по формулам (28).

Выборочные и теоретические коэффициенты корреляции рассматриваемых случайных величин при разных сочетаниях параметров λ и p

λ	p	r_{xy}^*	r_{xy}	r_{xz}^*	r_{xz}	r_{yz}^*	r_{yz}
1 000	0,1	0,3230	0,3162	0,9479	0,9487	0,0048	0
1 000	0,5	0,7164	0,7071	0,6964	0,7071	-0,0018	0
1 000	0,9	0,9496	0,9487	0,3072	0,3162	-0,0065	0
100	0,1	0,3089	0,3162	0,9504	0,9487	-0,0023	0
100	0,5	0,7100	0,7071	0,7056	0,7071	0,0020	0
100	0,9	0,9487	0,9487	0,3416	0,3162	0,0270	0
10	0,1	0,3189	0,3162	0,9475	0,9487	-0,0009	0
10	0,5	0,7054	0,7071	0,7065	0,7071	-0,0032	0
10	0,9	0,9492	0,9487	0,3184	0,3162	0,0038	0

Как видно из таблицы, значения теоретических коэффициентов корреляции хорошо согласуются с соответствующими значениями выборочных коэффициентов корреляции, что делает целесообразным использование полученных в настоящей работе результатов при исследовании тесноты связи между случайными величинами, удовлетворяющими ограничению (2) или (3). Эти результаты могут быть использованы, в частности, при исследовании взаимосвязи выходных сигналов сэндвич-детекторов излучения, применяемых в рентгеновских досмотровых системах с целью обнаружения несанкционированных вложений в контролируемых объектах [9, 10].

Заключение

Получена формула (8) для вычисления коэффициента корреляции двух случайных величин при неизвестном математическом ожидании их произведения для случая, когда эти величины удовлетворяют ограничению (2) или (3) и известны первые два момента распределения каждой из них. Показано, что данная формула является точной при соблюдении условия (9) или (10). Приведены примеры законов распределения, для которых эти условия выполняются.

Получена формула (27) для вычисления коэффициента корреляции случайных слагаемых случайной величины, которая, как и формула (8), является точной при соблюдении тех же условий (9) или (10).

На примерах показано, как с использованием дополнительной информации о форме взаимосвязи исследуемых случайных величин можно существенно упростить расчет коэффициента корреляции.

Список источников

1. Базулин Е.Г. Два подхода к решению задач ультразвуковой дефектометрии: анализ высококачественного изображения отражателей и корреляционный анализ измеренных эхосигналов // Дефектоскопия. 2016. № 2. С. 11–32.
2. Жданов Ю.А., Минкин В.И. Корреляционный анализ в органической химии. Ростов н/Д : Изд-во Ростов. унта, 1966. 471 с.
3. Дунаев А.А. Адаптивный корреляционный анализ биоэлектрических сигналов // Российский медико-биологический вестник имени академика И.П. Павлова. 2003. № 1-2. С. 157–161.
4. Окатов Д.А., Минеева Т.А. Парный корреляционный анализ статистической обработки данных в медицине // Тенденции развития науки и образования. 2021. № 72-1. С. 87–91.
5. Постников В.П., Трубинова К.А. Корреляционный анализ влияния венчурного инвестирования на инновационное развитие экономики // Финансы и кредит. 2020. Т. 26, № 8 (800). С. 1767–1784.
6. Кургузов К.В., Фоменко И.К. Пространственно-корреляционный анализ инженерно-геологических данных на примере строительства логистического комплекса // Вестник МГСУ. 2019. Т. 14, № 8. С. 976–990.
7. Богомолова Е.В. Корреляционный анализ в современном педагогическом эксперименте // Информатика и прикладная математика. 2020. № 26. С. 20–25.
8. Кремер Н.Ш. Теория вероятностей и математическая статистика : учебник для вузов. М. : ЮНИТИ-ДАНА, 2001. 543 с.
9. Удод В.А., Воробейчиков С.Э., Назаренко С.Ю. Математические модели радиационных прозрачностей объекта контроля при использовании сэндвич-детекторов рентгеновского излучения // Дефектоскопия. 2020. № 2. С. 31–41.
10. Khan S.U., Khan I.U., Ullah I., Saif N., Ullah I. A review of airport dual energy X-ray baggage inspection techniques: Image enhancement and noise reduction // Journal of X-Ray Science and Technology. 2020. V. 28 (3). P. 481–505.

References

1. Bazulin, E.G. (2016) Two approaches to solving the problems of ultrasonic defectometry: an analysis of a high -quality image of reflectors and a correlation analysis of measured echosignals. *Defektoskopiya – Russian Journal of Non-Destructive Testing*. 2. pp. 11–32.
2. Zhdanov, Yu.A. & Minkin, V.I. (1966) *Korreljatsionnyy analiz v organicheskoj khimii* [Correlation analysis in organic chemistry]. Rostov on Don: Rostov State University.
3. Dunaev, A.A. (2003) *Adaptivnyy korreljatsionnyy analiz bioelektricheskikh signalov* [Adaptive correlation analysis of bioelectric signals]. *Rossiyskiy mediko-biologicheskij vestnik imeni akademika I.P. Pavlova – I.P. Pavlov Russian Medical Biological Herald*. 1-2. pp. 157–161.
4. Okatov, D.A. & Mineeva, T.A. (2021) *Parnyy korreljatsionnyy analiz statisticheskoy obrabotki dannykh v meditsine* [Pair correlation analysis of statistical data processing in medicine]. *Tendentsii razvitiya nauki i obrazovaniya*. 72-1. pp. 87–91. DOI: 10.18411/lj-04-2021-19
5. Postnikov, V.P. & Trubinova, K.A. (2020) A correlation analysis of the impact of venture investment on innovative economic development. *Finansy i kredit – Finance and Credit*. 26(8). pp. 1767–1784. DOI: 10.24891/fc.26.8.176
6. Kurguzov, K.V. & Fomenko, I.K. (2019) Spatial and correlation analysis of engineering-geological survey data for logistics center construction. *Vestnik MGSU – VESTNIK MGSU (Monthly Journal on Construction and Architecture)*. 14(8). pp. 976–990. DOI: 10.22227/1997-0935.2019.8.976-990
7. Bogomolova, E.V. (2020) *Korreljatsionnyy analiz v sovremennom pedagogicheskom eksperimente* [Correlation analysis in a modern pedagogical experiment]. *Informatika i prikladnaya matematika*. 26. pp. 20–25.
8. Kremer, N.Sh. (2001) *Teoriya veroyatnostey i matematicheskaya statistika* [Theory of Probability and Mathematical Statistics]. Moscow: Unity-Dana.
9. Udod, V.A., Vorobeychikov, S.E. & Nazarenko, S.Yu. (2020) *Matematicheskie modeli radiatsionnykh prozrachnostey ob'ekta kontrolya pri ispol'zovanii sendvich-detektorov rentgenovskogo izlucheniya* [Mathematical models of radiation transparency of the control object when using sandwich detectors of X-ray radiation]. *Defektoskopiya – Russian Journal of Non-Destructive Testing*. 2. pp. 31–41.
10. Khan, S.U., Khan, I.U., Ullah, I., Saif, N. & Ullah, I. (2020) A review of airport dual energy X-ray baggage inspection techniques: Image enhancement and noise reduction. *Journal of X-Ray Science and Technology*. 28(3). pp. 481–505. DOI: 10.3233/xst-200663

Информация об авторе:

Удод Виктор Анатольевич – профессор, доктор технических наук, профессор кафедры информационных технологий и бизнес-аналитики Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: pr.udod@mail.ru

Автор заявляет об отсутствии конфликта интересов.

Information about the author:

Udod Victor A. (Doctor of Technical Sciences, Professor, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: pr.udod@mail.ru

The author declares no conflicts of interests.

Поступила в редакцию 30.08.2022; принята к публикации 01.03.2023

Received 30.08.2022; accepted for publication 01.03.2023

Научная статья

УДК 519.246.5

doi: 10.17223/19988605/62/11

Метод устранения мешающего параметра в статистике пуассоновского потока точек

Гурами Шалвович Цициашвили¹, Владимир Николаевич Бочарников²,
Сергей Михайлович Краснопеев³

¹ Институт прикладной математики ДВО РАН, Владивосток, Россия

^{2,3} Тихоокеанский институт географии ДВО РАН, Владивосток, Россия

¹ guram@iam.dvo.ru

² vbocharnikov@mail.ru

³ sergeikr@tigdvo.ru

Аннотация. Сопоставление общего числа зарегистрированных следов с имеющимися в литературе оценками числа тигров показало наличие существенных различий на территории Приморского края. Для устранения вероятных ошибок исходные данные представляются неоднородным точечным пуассоновским потоком. Рассматривается задача обработки данных по расчету плотности населения амурского тигра в районах Приморского края. Первичным основанием для такого расчета послужила авторская выборка сведений регистрации следов хищника в зимний период. Рабочая методика учета следа моделируется процедурой раскраски точек потока в зависимости от вероятности обнаружения следа. Эта вероятность становится мешающим параметром, устранение которого производится заменой числа следов на долю следов в различных районах. Доказывается теорема о сходимости доли следов к предельному значению при устремлении к бесконечности параметра пуассоновского распределения, характеризующего общее число следов. Полученные результаты иллюстрируются картой, характеризующей расчетные доли следов амурского тигра в Приморском крае на учетный период 2005 г.

Ключевые слова: пуассоновский поток; раскраска точек потока; число следов; предельная теорема для доли числа следов.

Для цитирования: Цициашвили Г.Ш., Бочарников В.Н., Краснопеев С.М. Метод устранения мешающего параметра в статистике пуассоновского потока точек // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2023. № 62. С. 101–106. doi: 10.17223/19988605/62/11

Original article

doi: 10.17223/19988605/62/11

The method of eliminating the interfering parameter in the statistics of the Poisson flow of points

Gurami Sh. Tsitsiashvili¹, Vladimir N. Bocharnikov², Sergei M. Krasnopeev³

¹ Institute for Applied Mathematics, FEB RAS, Vladivostok, Russian Federation

^{2,3} Institute of Pacific Ocean Geography FEB RAS, Vladivostok, Russian Federation

¹ guram@iam.dvo.ru

² vbocharnikov@mail.ru

³ sergeikr@tigdvo.ru

Abstract. Comparison of the total number of registered tracks with the estimates of the number of tigers available in the literature showed the presence of significant differences in the territory of Primorsky Krai. To eliminate possible errors, the source data is represented by an inhomogeneous point Poisson flow. The problem of processing data on the calculation of the population density of the Amur tiger in the regions of Primorsky Krai is considered. The primary basis

for this calculation was the author's sample of information on the registration of predator tracks in winter. The working method of trace accounting is modeled by the procedure of coloring flow points depending on the probability of trace detection. This probability becomes an interfering parameter, which is eliminated by replacing the number of traces with the proportion of traces in different areas. The theorem on the convergence of the fraction of traces to the limiting value is proved when the parameter of the Poisson distribution characterizing the total number of traces tends to infinity. The results obtained are illustrated by a map characterizing the estimated proportions of Amur tiger tracks in the Primorsky Territory for the accounting period of 2005.

Keywords: Poisson flow; coloring of flow points; number of traces; limit theorem for the fraction of the number of traces.

For citation: Tsitsiashvili, G.Sh., Bocharnikov, V.N., Krasnopeev, S.M. (2023) The method of eliminating the interfering parameter in the statistics of the Poisson flow of points. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 62. pp. 101–106. doi: 10.17223/19988605/62/11

Введение

В задачах из области наук о Земле нередко используется понятие неоднородного пуассоновского потока точек. Такой поток точек применялся в геологических исследованиях [1–3]. Подобная вероятностная модель также возникает при изучении пространственного распределения подвижных животных [4]. Близкие по приложениям результаты, связанные с применением пуассоновского потока точек в экологических исследованиях, приведены в монографии [5]. В настоящей работе неоднородный пуассоновский поток точек применен к решению важной региональной задачи – исследованию плотности населения амурского тигра, рассчитываемой на основе зарегистрированных на учетных маршрутах следов хищника на территории Приморского края [6–9]. Полагаем, что отсутствие необходимой математической коррекции может приводить к большим ошибкам при сопоставлении результатов в различные годы. В этой связи естественным будет обратить специальное внимание на расчеты вероятности обнаружения следов, которые зависят от многочисленных природных и экономических характеристик.

Полевой учет амурского тигра в зимний период 2004–2005 гг. осуществлялся Управлением Росприроднадзора по Приморскому и Хабаровскому краям, Приморкрайохотуправлением, Хабаровсккрайохотуправлением, Тихоокеанским институтом географии и Биолого-почвенным институтом ДВО РАН с финансовой поддержкой и участием Общества сохранения диких животных (WCS) и WWF России. По данным специалистов, суммарное число обнаруженных следов амурского тигра в 2005 г. составило 6 869, в то же время сообщалось, что общее число тигров на территории Приморского края в 2005 г. составляло 357–425 особей [10]. Различия между числом зарегистрированных «в поле» следов и общим числом обитающих тигров свидетельствуют о необходимости осуществления специальных математических процедур при обработке данных. Такие ошибки становятся при статистической обработке мешающими параметрами [11–13].

Содержательная важность и математическая актуальность предложенного анализа состоит в отсутствии необходимого приложения, полезного в использовании при изменяющихся процедурах расчетов следов животных. Важнейшим элементом является установление плотности населения амурского тигра, что не требует расчета общей численности животных, но позволяет избежать субъективных мнений при анализе ситуации. С тем чтобы освободить результаты обработки данных от этих мешающих параметров, в работе предлагается вместо числа следов оценивать доли следов в различных районах Приморского края. Укажем, что представляемое здесь исследование было осуществлено на основе организованного Тихоокеанским институтом географии ДВО РАН перерасчета данных, полученных большим коллективом исследователей в 2005 г. [14]. Следует также отметить, что в последнее время техника устранения мешающих параметров из статистических выборок существенно востребована в задачах квантовой информатики [15], что свидетельствует об общей перспективе данного математического подхода.

1. Основные результаты

Для решения поставленной задачи удобно воспользоваться следующей теоремой [5. С. 73] о раскраске точек пуассоновского потока.

Теорема 1. Пусть Π – пуассоновский процесс на пространстве S с мерой интенсивности λ . Раскрасим точки процесса Π случайным образом в k цветов так, что вероятности покрасить точку в i -й цвет равны p_i и цвета различных точек независимы (друг от друга и от положения точек). Тогда Π_i – независимые пуассоновские процессы с мерами интенсивности $\lambda_i = p_i \lambda$.

Используя эту теорему, можно произвести анализ числа точек пуассоновского потока в ограниченных и измеримых по Лебегу подобластях $G_k, k=1, \dots, m$, двумерного евклидова пространства. Предположим, что задан пуассоновский поток точек π с интенсивностью $\lambda = \lambda(x, y)$ и выполнены соотношения

$$\lambda_k = \int_{G_k} \lambda(x, y) dx dy < \infty, \quad k=1, \dots, m, \quad \lambda = \sum_{k=1}^m \lambda_k.$$

Обозначим $\Lambda_k = \frac{\lambda_k}{\lambda}$ и всюду далее полагаем $\Lambda_k = const, k=1, \dots, m$, допуская предельное соотношение $\lambda \rightarrow \infty$. Пусть каждая точка потока π независимо от других точек и от своих координат с вероятностью p входит в поток Π . Тогда поток Π является пуассоновским с интенсивностью $p\lambda(x, y)$. Следовательно, число n_k точек потока Π в подобласти G_k имеет пуассоновское распределение с параметром $p\lambda_k$, а сумма $n = \sum_{k=1}^m n_k$ имеет пуассоновское распределение с параметром $p\lambda$. Поэтому выполняются соотношения

$$Mn_k = Dn_k = p\lambda_k, \quad k=1, \dots, m, \quad Mn = Dn = p\lambda. \quad (1)$$

Отсюда следует, что относительная ошибка оценки параметра $p\lambda_k$ случайной величиной n_k имеет вид:

$$\sqrt{D \frac{n_k}{p\lambda_k}} = \frac{1}{\sqrt{p\lambda_k}}. \quad \text{Аналогично получаем, что относительная ошибка оценки параметра } p\lambda \text{ случайной}$$

$$\text{величиной } n \text{ имеет вид: } \sqrt{D \frac{n}{p\lambda}} = \frac{1}{\sqrt{p\lambda}}.$$

Из неравенства Чебышева следует, что при любом $\varepsilon, 0 < \varepsilon < 1$, и при $\lambda \rightarrow \infty$

$$P\left(1 - \varepsilon \leq \frac{n_k}{p\lambda_k} \leq 1 + \varepsilon\right) \geq 1 - \frac{1}{p\lambda_k \varepsilon^2} \rightarrow 1, \quad k=1, \dots, m, \quad P\left(1 - \varepsilon \leq \frac{n}{p\lambda} \leq 1 + \varepsilon\right) \geq 1 - \frac{1}{p\lambda \varepsilon^2} \rightarrow 1. \quad (2)$$

Из формул (1), (2) следует, что случайная величина n_k является несмещенной и состоятельной оценкой параметра $p\lambda_k$, а случайная величина n является несмещенной и состоятельной оценкой параметра $p\lambda$ при $\lambda \rightarrow \infty$.

Обозначим $N_k = \frac{n_k}{n}$, используя формулы (1), (2), нетрудно при любом $\varepsilon, 0 < \varepsilon < 1/2$, получить неравенство

$$P\left(1 - 2\varepsilon \leq \frac{N_k}{\Lambda_k} \leq 1 + 4\varepsilon\right) \geq P\left(\frac{1 - \varepsilon}{1 + \varepsilon} \leq \frac{N_k}{\Lambda_k} \leq \frac{1 + \varepsilon}{1 - \varepsilon}\right) \geq 1 - \frac{1}{p\lambda_k \varepsilon^2} - \frac{1}{p\lambda \varepsilon^2}. \quad (3)$$

Используя неравенство $\Lambda_k \leq 1$, из формулы (3) получаем соотношение

$$P(-2\varepsilon \leq N_k - \Lambda_k \leq 4\varepsilon) \geq 1 - \frac{1}{p\lambda_k \varepsilon^2} - \frac{1}{p\lambda \varepsilon^2} \rightarrow 1, \quad \lambda \rightarrow \infty. \quad (4)$$

Таким образом, доказано следующее утверждение.

Теорема 2. Справедлива сходимость по вероятности случайной величины N_k к параметру Λ_k при $\lambda \rightarrow \infty$.

Тем самым отношение N_k является состоятельной оценкой параметра Λ_k , свободного от вероятности p , являющейся в этой задаче мешающим параметром. Неравенство (4) является достаточно грубым и вдобавок включает неизвестный мешающий параметр p . Однако при увеличении λ (или при большом n) точность оценки (4) повышается.

2. Доли числа следов за 2005 г. в районах Приморского края

В настоящем разделе рассчитаны доли числа следов в 24 районах Приморского края за 2005 г. Вычислена разность между долями следов в различных районах края; полученное пространственное распределение показано на карте (рис. 1).

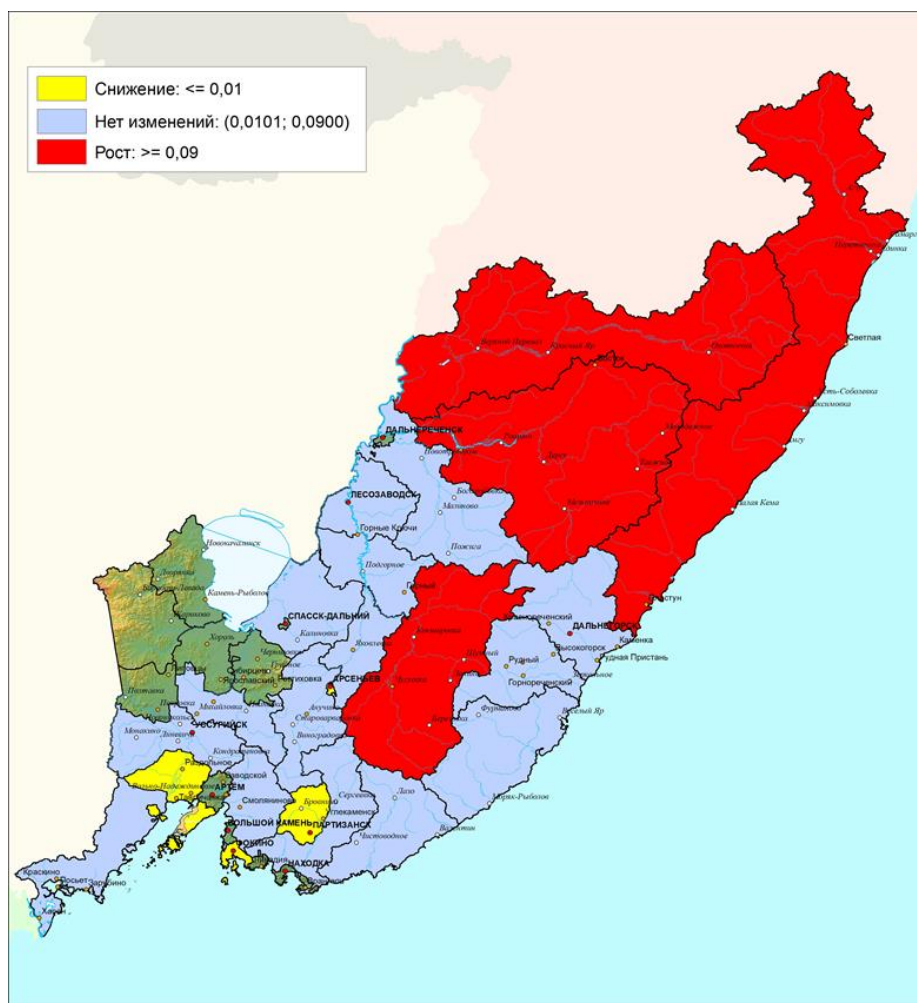


Рис. 1. Карта районов Приморского края с окраской, определяемой долями следов тигра в 2005 г.
Fig. 1. Map of Primorsky Krai districts with color determined by proportions of tiger tracks in 2005

Из рис. 1 видно, что в северных районах Приморского края доля числа следов максимальна. На юге края в основном располагаются районы со средней долей численности следов. При этом на юге края находится совсем небольшое число районов с малой долей числа следов. Районы (общим числом 4), в которых доли числа следов больше 0,09, окрашены красным цветом. Районы (общим числом 15), в которых доли числа следов меньше 0,09, но больше 0,01, окрашены голубым цветом. Районы (общим числом 5), у которых доли числа следов меньше 0,01, окрашены желтым цветом.

Полагаем, что наглядность представления географического распределения и тенденций ее изменения для амурского тигра чрезвычайно важна в практическом отношении, следовательно, существует необходимость системного привлечения методов математической обработки данных.

Заключение

Считаем, что рассмотренный в сообщении опыт может быть принят для рассмотрения при совершенствовании государственной методики учета следов животных. Исследование динамики пространственного распределения амурского тигра по территории Приморского края ныне существенно затруднено без специальной математической обработки первичных подсчетов следов. Авторы рекомендуют полевым специалистам более широко использовать для оценочных работ специализированные математические методы. В этой связи можно отметить, что хорошо зарекомендовала себя процедура устранения мешающих параметров путем перехода от абсолютного числа следов к доле следов в различных районах.

Список источников

1. Амбарцумян Р.В., Мекке Й., Штойян Д. Введение в стохастическую геометрию. М. : Наука, 1989. 400 с.
2. Stoyan D. On some qualitative properties of the Boolean model of stochastic geometry // *Z. angew. Math. Mech.* 1979. V. 59. P. 447–454.
3. Stoyan D. Stereological formulae for size distribution through marked point processes // *Prob. and Mat. Statist.* 1982. V. 2. P. 161–166.
4. Суханов В.В. Модель пространственного распределения подвижных животных в разреженной популяции // *Математическое моделирование природных систем : сб. ст. / ред. Е. В. Золотов, Б. И. Семкин. Владивосток : Дальнаука, 1981. С. 10–22.*
5. Кингман Дж. Пуассоновские процессы. М. : МЦНМО, 2007. 136 с.
6. Пикунов Д.Г. Организация учета численности диких животных в Приморье // *Вопросы производственного охотоведения Сибири и Дальнего Востока. Иркутск : ИСХИ, 1970. С. 165–173.*
7. Пикунов Д.Г., Микелл Д.Г., Серёдкин И.В., Николаев И.Г., Дунишенко Ю.М. Зимние следовые учеты амурского тигра на Дальнем Востоке России (методика и история проведения учетов). Владивосток : Дальнаука, 2014. 132 с.
8. Bocharnikov V.N., Fomenko P.V., Krasnopeev S.M. Assessment of dynamics of Amur Tiger habitat quality influenced by natural and anthropogenic factors // *Integrated Tools For Natural Resources Inventories In The 21st Century.* 1998. P. 51–55.
9. Абрамов К.Г. К методике учета тигра // *Вопросы организации и методы учета ресурсов фауны наземных позвоночных животных. М. : Изд-во АН СССР, 1961. С. 53–54.*
10. Мурзин А.А. Построение модели динамики популяции Амурского тигра и прогнозные расчеты его численности на период с 2015 по 2022 г. // *Международный журнал прикладных и фундаментальных исследований.* 2018. № 11. С. 333–340.
11. Кокс Д., Хинкли Д. Теоретическая статистика. М. : Мир, 1978. 560 с.
12. Young G.A., Smith R.L. *Essentials of Statistical Inference.* Cambridge University Press, 2005. 236 p. (Cambridge Series in Statistical and Probabilistic Mathematics).
13. Холево А.С. Статистические структуры квантовой механики и скрытые параметры. М. : Знание, 1985. 32 с.
14. Пикунов Д.Г., Серёдкин И.В., Солкин В.А. Амурский тигр (история изучения, динамика ареала, численности, экология и стратегия охраны). Владивосток : Дальнаука, 2010. 104 с.
15. Холево А.С. Квантовая информатика: прошлое, настоящее, будущее // *В мире науки.* 2008. № 7. С. 69–75.

References

1. Ambartsumyan, R.V., Mekke, Y. & Stoyan, D. (1989) *Vvedenie v stokhasticheskuyu geometriyu* [Introduction to stochastic geometry]. Moscow: Nauka.
2. Stoyan, D. (1979) On some qualitative properties of the Boolean model of stochastic geometry. *ZAMM – Journal of Applied Mathematics and Mechanics.* 59. pp. 447–454.
3. Stoyan, D. (1982) Stereological formulae for size distribution through marked point processes. *Probability and Mathematical Statistics.* 2. pp. 161–166.
4. Sukhanov, V.V. (1981) Model' prostranstvennogo raspredeleniya podvizhnykh zhiivotnykh v razrezhennoy populyatsii [Model of spatial distribution of mobile animals in a sparse population]. In: Zolotov, E.V. & Semkin, B.I. (eds) *Matematicheskoe modelirovanie prirodnykh sistem* [Mathematical Modeling of Natural Systems]. Vladivostok: Dalnauka. pp. 10–22.
5. Kingman, J. (2007) *Puassonovskie protsessy* [Poisson Processes]. Translated from English. Moscow: ICNMO. 136 p.
6. Pikunov, D.G. (1970) Organizatsiya ucheta chislenosti dikikh zhiivotnykh v Primor'e [Organization of accounting for the number of wild animals in Primorye]. In: *Voprosy proizvodstvennogo okhotovedeniya Sibiri i Dal'nego Vostoka* [Questions of Industrial Hunting in Siberia and the Far East]. Irkutsk: ISKhI. pp. 165–173.
7. Pikunov, D.G., Mikell, D.G., Sereдкин, I.V., Nikolaev, I.G. & Dunishenko, Yu.M. (2014) *Zimnie sledovye uchety amurskogo tigra na Dal'nem Vostoke Rossii (metodika i istoriya provedeniya uchetov)* [Winter trace records of the Amur tiger in the Russian Far East (methodology and history of accounting)]. Vladivostok: Dalnauka.

8. Bocharnikov, V.N., Fomenko, P.V. & Krasnopeev, S.M. (1998) Assessment of dynamics of Amur Tiger habitat quality influenced by natural and anthropogenic factors. In: Hansen, M.H. & Burk, T.E. (eds) *Integrated Tools For Natural Resources Inventories In The 21st Century*. General Technical Report (GTR). pp. 51–55.
9. Abramov, K.G. (1961) К методике учета тигра [To the tiger accounting methodology]. In: *Voprosy organizatsii i metody ucheta resursov fauny nazemnykh pozvonochnykh zivotnykh* [Problems of organization and methods of accounting for the fauna of terrestrial vertebrates]. Moscow: USSR Academy of Sciences. pp. 53–54.
10. Murzin, A.A. (2018) Postroenie modeli dinamiki populyatsii Amurskogo tigra i prognoznye raschety ego chislennosti na period s 2015 po 2022 g. [Constructing a model of the dynamics of the Amur tiger population and predictive calculations of its number for the period from 2015 to 2022]. *Mezhdunarodnyy zhurnal prikladnykh i fundamental'nykh issledovaniy – International Journal of Applied and Fundamental Research*. 11. pp. 333–340.
11. Cox, D. & Hinckley, D. (1978) *Teoreticheskaya statistika* [Theoretical Statistics]. Moscow: Mir.
12. Young, G.A. & Smith, R.L. (2005) *Essentials of Statistical Inference*. Cambridge University Press.
13. Holevo, A.S. (1985) *Statisticheskie struktury kvantovoy mekhaniki i skrytye parametry* [Statistical structures of quantum mechanics and hidden parameters]. Translated from English. Moscow: Znanie.
14. Pikunov, D.G., Seredkin I.V. & Solkin V.A. (2010) *Amurskiy tigr (istoriya izucheniya, dinamika areala, chislennosti, ekologiya i strategiya okhrany)* [Amur tiger (history of study, range dynamics, abundance, ecology and conservation strategy)]. Vladivostok: Dalnauka.
15. Holevo, A.S. (2008) *Kvantovaya informatika: proshloe, nastoyashchee, budushchee* [Quantum informatics: past, present, future]. *V mire nauki*. 7. pp. 69–75.

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

Бочарников Владимир Николаевич – доктор биологических наук, ведущий научный сотрудник Тихоокеанского института географии ДВО РАН (Владивосток, Россия). E-mail: vbocharnikov@mail.ru

Краснопеев Сергей Михайлович – кандидат физико-математических наук, научный сотрудник Тихоокеанского института географии ДВО РАН (Владивосток, Россия). E-mail: sergeikr@tigdvo.ru

Вклад авторов: все авторы сделали эквивалентный вклад в подготовку публикации. Цициашвили Г.Ш. разработал математическую часть статьи. Бочарников В.Н. сформулировал содержательную постановку задачи и предоставил фактический материал. Краснопеев С.М. представил результаты вычислений в виде карты. Авторы заявляют об отсутствии конфликта интересов.

Information about the authors:

Tsitsiashvili Gurami Sh. (Doctor of physical and mathematical sciences, main researcher of Institute for Applied Mathematics, FEB RAS, Vladivostok, Russian Federation) E-mail: guram@iam.dvo.ru

Bocharnikov Vladimir N. (Doctor of biological sciences, leading researcher of Institute of Pacific Ocean Geography FEB RAS, Vladivostok, Russian Federation). E-mail: vbocharnikov@mail.ru

Krasnopeev Sergei M. (Candidate of physical and mathematical sciences, researcher of Institute of Pacific Ocean Geography FEB RAS, Vladivostok, Russian Federation). E-mail: sergeikr@tigdvo.ru

Contribution of the authors: the authors contributed equally to this article. Tsitsiashvili G.Sh. developed the mathematical part of the article. Bocharnikov V.N. formulated a meaningful statement of the problem and provided factual material. Krasnopeev S.M. presented the results of calculations in the form of a map. The authors declare no conflicts of interests.

Поступила в редакцию 28.10.2022; принята к публикации 01.03.2023

Received 28.10.2022; accepted for publication 01.03.2023

ИНФОРМАТИКА И ПРОГРАММИРОВАНИЕ

INFORMATICS AND PROGRAMMING

Научная статья

УДК 519.713

doi: 10.17223/19988605/62/12

**Оптимизация клиентской компоненты по критерию отсутствия
лишних диалогов с сервером****Максим Леонидович Громов¹, Светлана Анатольевна Прокопенко²,
Александр Павлович Сотников³, Наталия Владимировна Шабалдина⁴***^{1, 2, 3, 4} Томский государственный университет, Томск, Россия**¹ maxim.leo.gromov@gmail.com**² s.prokopenko@sibmail.com**³ sotnikhtc@gmail.com**⁴ nataliamailbox@mail.ru*

Аннотация. Рассматривается задача оптимизации клиентской компоненты взаимодействующих по клиент-серверной архитектуре приложений. Критерием оптимизации выступает уменьшение количества взаимодействий, т.е. уточнение клиентской спецификации таким образом, чтобы по возможности не осуществлялось лишних запросов от клиента к серверу. Предлагается алгоритм модификации спецификации клиента по дереву достижимости.

Ключевые слова: оптимизация; клиент; сервер; конечный автомат; композиция.

Для цитирования: Громов М.Л., Прокопенко С.А., Сотников А.П., Шабалдина Н.В. Оптимизация клиентской компоненты по критерию отсутствия лишних диалогов с сервером // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2023. № 62. С. 107–114. doi: 10.17223/19988605/62/12

Original article

doi: 10.17223/19988605/62/12

**Optimization of the client component by the criterion of the absence of redundant
dialogues with the server****Maxim L. Gromov¹, Svetlana A. Prokopenko², Aleksandr P. Sotnikov³, Natalia V. Shabaldina⁴***^{1, 2, 3, 4} Tomsk State University, Tomsk, Russian Federation**¹ maxim.leo.gromov@gmail.com**² s.prokopenko@sibmail.com**³ sotnikhtc@gmail.com**⁴ nataliamailbox@mail.ru*

Abstract. In this paper the problem of client component optimization within client-server interaction is considered. The optimization criterion is the reduction of client-server interactions number. Client's specification is clarified in way to minimize extra requests by client to server. The algorithm for modifying the client's specification according to the reachability tree is proposed.

Keywords: optimization; client; server; finite state machine; composition.

For citation: Gromov, M.L., Prokopenko, S.A., Sotnikov, A.P., Shabaldina, N.V. (2023) Optimization of the client component by the criterion of the absence of redundant dialogues with the server. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 62. pp. 107–114. doi: 10.17223/19988605/62/12

Введение

Клиент–сервер – классическая архитектура организации информационных систем. В этой архитектуре система представляется двумя программами: клиентом и сервером. Сервер владеет ресурсами (файлами, базами данных и т.п.) и обслуживает запросы на доступ к ним. Клиент запускается, например, в браузере пользователя информационной системы и через свой графический интерфейс (кнопки, текстовые поля и т.п.) общается с пользователем. Между собой клиент и сервер общаются через телекоммуникационную сеть (например, Интернет). Всякое сообщение (запрос) от клиента серверу или наоборот проходит через сеть. В общем случае путь сообщения пролегает через некоторое количество промежуточных узлов (маршрутизаторов, коммутаторов и т.д.), и каждый узел затрачивает энергию на обработку сообщения. Кроме того, время, затрачиваемое на прохождение запроса к серверу, формирование ответа и прохождение ответа от сервера к клиенту, имеет значительные величины, заметные человеку. Поэтому манипуляции с интерфейсом в браузере, например нажатия кнопки, которые требуют взаимодействия клиента с сервером, выглядят с точки зрения человека как некоторое «под тормаживание» системы, что ухудшает впечатление пользователя (user experience) от системы. Таким образом, уменьшение количества взаимодействий между клиентом и сервером является задачей, актуальной как с точки зрения экологии (уменьшении совокупного потребления ресурсов системой), так и с точки зрения пользовательского опыта.

При совместной работе клиентской и серверной реализаций протоколов возникают ситуации, когда клиент осуществляет лишние (избыточные) запросы к серверу, в то время как и без этих запросов ответ клиента конечному пользователю предопределен. Такая предопределенность может быть обусловлена как требованиями RFC, так и, в некоторых случаях, накопленной к этому этапу взаимодействия информацией, например о настройках сервера или о реализованных на сервере опциях.

Задача модификации («ограничения») спецификации клиентского приложения, правда, с несколько иной целью (по иному критерию), рассматривается в работе [1]. Предлагается ограничить спецификацию клиента, чтобы на серверное приложение не могли поступить входные последовательности, на которых поведение сервера не определено. Отметим, что в этой работе дана только идея; шаги определения недопустимой последовательности и ограничения спецификации клиента описаны общими формулировками.

Распространенным подходом к оптимизации одной из взаимодействующих компонент является решение автоматных уравнений [2, 3] и неравенств, после чего среди всего множества решений выбирается такое решение, которое соответствует заданному критерию. В нашем случае данный подход тоже можно было бы исследовать. Тогда необходимо решить уравнение

$$X \diamond Server \cong Spec,$$

где *Server* – спецификация сервера, *Spec* – спецификация взаимодействия клиента и сервера, которая может быть получена в результате композиции исходных спецификаций клиента и сервера, \diamond – операция композиции, \cong – отношение эквивалентности. Среди всех решений данного уравнения нужно выбрать такое решение *X* (такого клиента), которое порождает минимум диалогов с сервером. Возможно, такая постановка задачи окажется тесно связанной с поиском так называемых *l*-ограниченных решений [2].

В данной работе предлагается оптимизировать спецификацию клиента с точки зрения отсутствия лишних запросов к серверу путем построения модифицированного дерева достижимости. Такое дерево достижимости отличается от классического, введенного в работе [4] и используемого с некоторыми модификациями, например в работе [5]. В предлагаемом в данной работе варианте дерева достижимости вершинам ставится в соответствие помимо пар состояний клиента и сервера еще и таблица пе-

реходов-выходов сервера, уточненная согласно уже пройденному пути. В нашей постановке задачи спецификация сервера является недетерминированным автоматом, и при построении дерева некоторые ребра соответствуют фиксации переходов в таблице переходов-выходов сервера, поскольку мы полагаем, что реализация является детерминированной. Также в предлагаемом в работе модифицированном дереве достижимости стабильной объявляется не всякая вершина, при переходе в которую был выдан внешний выходной символ, а только такая, которая уже встречалась выше с той же таблицей переходов-выходов сервера.

1. Основные понятия и определения

1.1. Конечный автомат

Конечным автоматом называется пятерка $S = (S, I, O, T_S, s_0)$, где S – непустое конечное множество состояний с выделенным начальным состоянием s_0 , I – непустое множество входных символов, называемое входным алфавитом, O – непустое множество выходных символов, называемое выходным алфавитом, $T_S \subseteq I \times S \times S \times O$ – отношение переходов [2].

Автомат S называется полностью определенным, если для каждой пары $(i, s) \in I \times S$ существует по крайней мере одна пара $(o, s') \in O \times S$ такая, что $(i, s, s', o) \in T_S$. В противном случае автомат S называется частичным.

Автомат S называется наблюдаемым, если для любой тройки $(i, s, o) \in I \times S \times O$ существует не более одного состояния $s' \in S$ такого, что $(i, s, s', o) \in T_S$, в противном случае автомат называется ненаблюдаемым (т.е. в наблюдаемом автомате, зная текущее состояние, входное воздействие и наблюдая выходную реакцию, можно однозначно определить следующее состояние). Поскольку известно, что для любого недетерминированного ненаблюдаемого автомата существует эквивалентная наблюдаемая форма [6], для удобства далее будем рассматривать только наблюдаемые автоматы.

Автомат S называется детерминированным, если для любой пары $(i, s) \in I \times S$ существует не более одной пары $(o, s') \in O \times S$ такой, что $(i, s, s', o) \in T_S$, в противном случае автомат называется недетерминированным.

В настоящей работе мы рассматриваем полностью определенные, возможно, недетерминированные, но всегда наблюдаемые автоматы.

Отношение переходов обычным образом распространяется на входные и выходные последовательности. Пара α/β , где $\alpha = i_1 \dots i_k$, $\beta = o_1 \dots o_k$, называется вход-выходной последовательностью автомата в состоянии s , если существует последовательность состояний $s_0, \dots, s_k \in S$ такая, что $s_0 = s$ и $(i_j, s_{j-1}, s_j, o_j) \in T_S$, $j = 1, \dots, k$. Множество вход-выходных последовательностей в состоянии s называется языком автомата в состоянии s , обозначение: $L_S(s)$. Язык автомата в начальном состоянии называется просто языком автомата и обозначается L_S . Автомат $A = (S, I, O, T_A, s_0)$ есть редукция автомата $B = (Q, I, O, T_B, s_0)$ (обозначение: $A \leq B$), если $L_A \subseteq L_B$. Автоматы A и B эквивалентны (обозначение: $A \cong B$), если $A \leq B$ и $B \leq A$.

1.2. Параллельная композиция конечных автоматов

Рассмотрим параллельную композицию автоматов $client = (S, X \cup V, Y \cup U, T_{client}, s_0)$ и $server = (Q, U, V, T_{server}, q_0)$ (рис. 1), в которой:

- автомат $client$ имеет входной алфавит $X \cup V$, выходной алфавит $Y \cup U$;
- автомат $server$ имеет входной алфавит U , выходной алфавит V ;
- алфавиты X_1, X_2, Y_1, Y_2, V и U попарно не пересекаются.

Параллельная композиция автоматов $client$ и $server$ имеет:

- входной алфавит X ;
- выходной алфавит Y ;

– под действием входного символа $x \in X$ композиция «вырабатывает» внешний выходной символ $y \in Y$ или внутренний выходной символ из алфавита U , который является входным для другой компоненты;

– следующий входной символ может быть подан на композицию только после того, как компоненты закончили внутренний диалог.

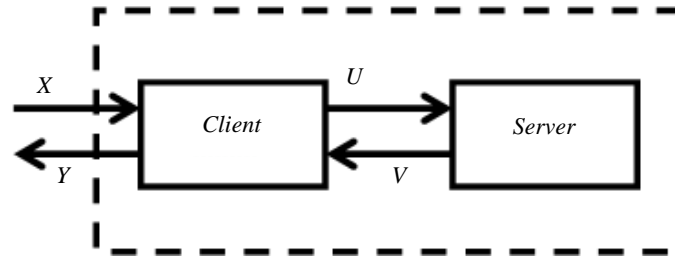


Рис. 1. Схема взаимодействия клиента и сервера
Fig. 1. Scheme of interaction between client and server

Иными словами, параллельная композиция представляет собой пару компонент, взаимодействующих в режиме диалога, где лишь внешняя компонента (*client*) может получать входные воздействия от среды (пользователя или тестировщика), в то время как внутренняя компонента (*server*) принимает внутренние выходные символы от первой компоненты и возвращает выходные символы, которые являются входными символами для внешней компоненты. Для того чтобы анализировать совместную работу таких компонент, их поведение часто моделируется одним конечным автоматом-композицией.

Диалогом считается непустое взаимодействие компонент, которое происходит между подачей внешнего входного символа и выдачей внешней выходной реакции.

Отметим, что при совместной работе компоненты могут попадать в бесконечный диалог, при котором никогда не достигается новое стабильное состояние. Такая ситуация называется осцилляцией, или заикливанием. Также при совместной работе компонент на компоненту может быть подано входное воздействие, при котором поведение компоненты не определено. Такие ситуации в данной работе не рассматриваются.

2. Построение модифицированного дерева достижимости

Введем понятие модифицированного дерева достижимости, отличающегося от классического дерева достижимости (*reachability graph*) следующим образом.

В классическом дереве достижимости [4] вершинам сопоставляются так называемые глобальные состояния – пары состояний взаимодействующих автоматов. Ребра – переходы между этими состояниями, помечены парой «входное воздействие / выходная реакция» в алфавитах активной на данный момент компоненты. В параллельной композиции компоненты активны по очереди, и ребро будет помечено парой x/y , если на внешнюю компоненту (на *client*) был подан внешний входной символ $x \in X$ и клиент ответил внешним выходным символом $y \in Y$ (диалога между клиентом и сервером в этом случае не произошло). Если же в ответ на внешний входной символ $x \in X$ клиент начинает диалог с сервером, то ребро будет помечено парой x/u , где $u \in U$. Символ $u \in U$ является одновременно выходным символом для клиента и входным символом для сервера, значит, следующее ребро будет помечено парой u/v – активной станет компонента *server*. Сервер является внутренней компонентой в данной композиции, поэтому ответ во внешнюю среду выдан быть не может. Ответ приходит на клиентскую компоненту. Далее снова активным становится клиент, поскольку $v \in V$ – входное воздействие на компоненту *client*. Теперь клиентская компонента может либо снова ответить выходным символом из алфавита U , и тогда ребро будет помечено парой v/u , что продолжит диалог между клиентом и сервером, либо клиент ответит внешним выходным символом $y \in Y$, ребро будет помечено парой v/y , автомат композиции перейдет в так называемое *стабильное состояние*. Напомним, что следующий

входной символ может быть подан на композицию только после того, как компоненты закончили внутренний диалог – это и есть стабильное состояние. Таким образом, построение дерева достижимости осуществляется путем моделирования взаимодействия автоматов-компонент. Корнем дерева является пара начальных состояний автоматов-компонент. Моделирование ведется по всем входным символам из внешнего алфавита X . Вершина объявляется листом, если она соответствует стабильному состоянию и уже встречалась. Построение дерева достижимости заканчивается, когда «раскрыты» все нелистовые вершины, которым соответствуют стабильные состояния.

В модифицированном дереве каждая вершина дополняется таблицей переходов-выходов сервера (внутренней компоненты). Напомним, что серверная компонента в нашем случае является недетерминированным автоматом. Пусть рассматривается нестабильное глобальное состояние st , и активной компонентой является сервер. Пусть в предыдущем состоянии на вход сервера был подан символ u . Тогда, раскрывая состояние st , мы перебираем все пары u/v , такие что из состояния t сервера есть переход по входу u с выходом v в таблице переходов-выходов, приписанной рассматриваемой вершине. Из текущей вершины (которой соответствует st) в дерево добавляем ребро с пометкой u/v в новую вершину st' (в таблице переходов-выходов текущей вершины есть переход из t в t' по u с выдачей v – в новую вершину в качестве пометки добавляем таблицу переходов-выходов из текущей вершины, но удалив в ней все остальные переходы из t по u , кроме выбранного нами для данного ребра и новой вершины перехода). Вершина объявляется листом, если ее пометка соответствует стабильному состоянию и уже встречалась в дереве именно как стабильное состояние с такой пометкой (обратим внимание: таблицы переходов-выходов у этих вершин должны совпадать). Построение дерева достижимости также заканчивается, когда «раскрыты» все нелистовые вершины, которым соответствуют стабильные состояния.

Если достигнуто стабильное состояние и не изменилась таблица переходов-выходов сервера (соответствующая данной паре состояний), то это стабильное состояние объявляется листом. Заканчиваем построение дерева достижимости, если все нераскрытые стабильные состояния объявлены листьями.

В качестве примера рассмотрим автоматы клиента и сервера, представленные на рис. 2. Модифицированное дерево достижимости представлено на рис. 3.

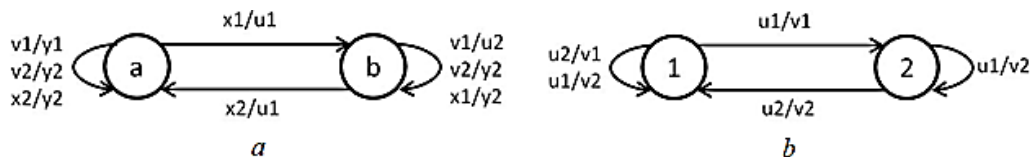


Рис. 2. Автомат клиента (a) и автомат сервера (b)
 Fig. 2. Client's Finite State Machine (a) and Server's Finite State Machine (b)

Можно видеть, что, как было описано выше, вершинам сопоставлены не только пары состояний автоматов *client* и *server*, но и обозначения T_0, T_1, T_2 , соответствующие трем различным таблицам переходов-выходов сервера (таблица).

Вершину дерева, соответствующую стабильному состоянию автомата композиции, назовем *стабильной вершиной*. Путь в дереве между двумя стабильными вершинами длины более одного, не проходящий через другие стабильные вершины, назовем *диалоговым путем*. Диалоговый путь, такой что в нем состояние сервера и таблица переходов-выходов сервера в начале пути совпадают с состоянием сервера и таблицей переходов-выходов сервера в конце пути, назовем *избыточным*.

Напомним, что идея модификации клиента заключается в том, чтобы избежать лишних диалогов между клиентом и сервером; такие диалоги на дереве соответствуют избыточным диалоговым путям. Тут нужно заметить, что значительная часть диалогов должна быть сохранена, т.е. в попытке оптимизировать мы не должны потерять «полезное» взаимодействие. В частности, нельзя избавляться от диалога, если в процессе диалога сервер изменил свое состояние.

Дуги в виде пунктирных линий, иллюстрируют работу алгоритма 1, приведенного ниже.

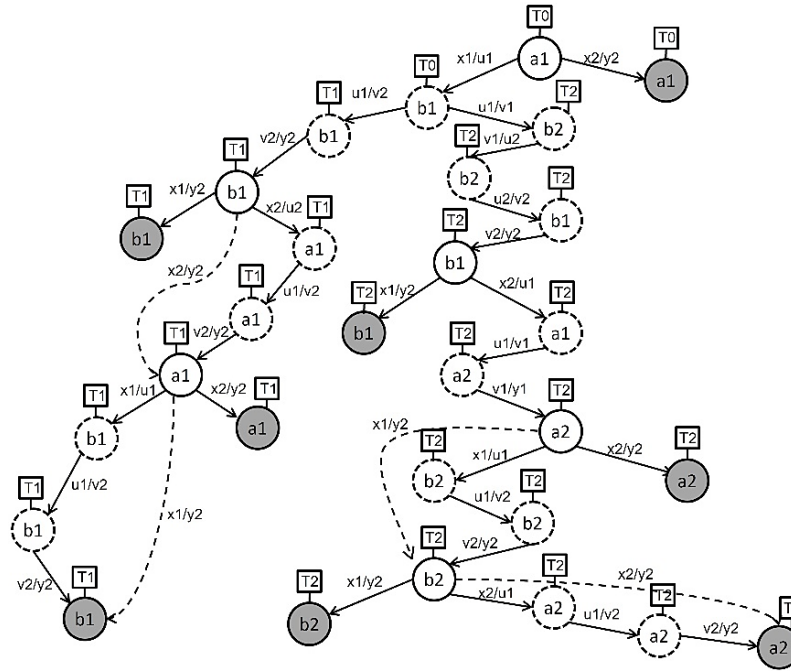


Рис. 3. Модифицированное дерево достижимости, построенное по автоматам *client* и *server*
 Fig. 3. Modified reachability tree based on *client* and *server* automata

Переходы-выходы сервера

<i>T0</i>	1	2
<i>u1</i>	1/ <i>v2</i>	2/ <i>v2</i>
<i>u2</i>	1/ <i>v1</i>	1/ <i>v2</i>

<i>T1</i>	1	2
<i>u1</i>	2/ <i>v1</i>	2/ <i>v2</i>
<i>u2</i>	1/ <i>v1</i>	1/ <i>v2</i>

<i>T2</i>	1	2
<i>u1</i>	1/ <i>v2</i>	2/ <i>v2</i>
<i>u2</i>	1/ <i>v1</i>	1/ <i>v2</i>

Алгоритм 1 построения модифицированной спецификации клиента по дереву достижимости.

Вход: Модифицированное дерево достижимости.

Выход: Автомат клиента, оптимизированный с точки зрения отсутствия лишних взаимодействий с сервером.

Шаг 1. Все избыточные пути в дереве «схлопываем» (оставляем входной символ первого ребра пути и выходной символ последнего ребра пути).

Шаг 2. Ставим в соответствие всем стабильным вершинам с одинаковыми пометками одно состояние клиента. Получим подмножество *Z* множества состояний нового клиента. Это подмножество будет дополнено промежуточными вершинами в процессе формирования графа переходов-выходов клиента.

Шаг 3. Начинаем перебирать все состояния из множества *Z*. Двигаемся по дереву.

Для каждого состояния $z \in Z$ определяем поведение под действием каждого внешнего входного символа *x*:

а) если вершина, соответствующая состоянию *z*, является началом ребра, помеченного парой *x/y*, $y \in Y$ и приводящего в вершину, соответствующую состоянию z' , то добавляем в автомат переход (x, z, z', y) ;

б) если вершина, соответствующая состоянию *z*, является началом ребра, помеченного парой *x/u*, $u \in U$, то определяем переход из состояния *z* в новое состояние $s \notin Z$, помеченный парой *x/u*. Далее в состоянии *s* определяем переход под действием входного символа $v \in V$ следующим образом:

– если вершина, соответствующая состоянию *s*, является началом ребра, помеченного парой *v/y*, $y \in Y$, то в автомат добавляется переход из состояния *s* в одно из состояний множества *Z* (согласно переходу в дереве), помеченный парой *v/y*;

– в противном случае повторяем шаг 3б.

В результате применения алгоритма автомат может получиться неприведенным, частичным, недетерминированным; число состояний может быть больше, чем у исходного автомата клиента.

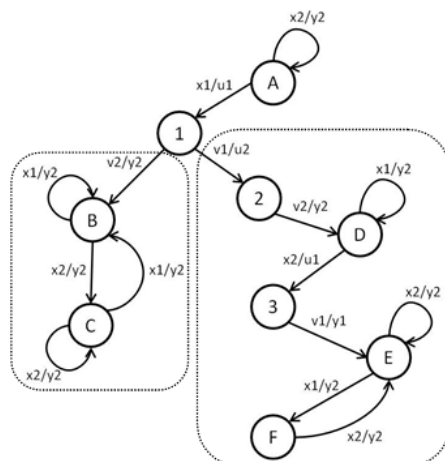


Рис. 4. Автомат клиента, построенный по алгоритму 1
Fig. 4. Client's Finite State Machine derived according to the Algorithm 1

На рис. 4 представлен автомат клиента, построенный по алгоритму 1. Пунктиром выделены различные варианты поведения клиента в зависимости от того, какую реакцию клиент получает от сервера в предшествующем диалоге. Отметим, что полученный в результате применения алгоритма 1 автомат клиента является частичным и неприведенным.

Заключение

В работе исследуется задача оптимизации спецификации клиентской компоненты взаимодействующих по клиент-серверной архитектуре приложений. Предложено в качестве критерия использовать отсутствие избыточных запросов к серверу. Оптимизация клиентской компоненты осуществляется путем построения модифицированного дерева достижимости и построения на основе этого дерева новой спецификации клиента. Отметим, что при решении задачи оптимизации число состояний клиента может увеличиться, кроме того, спецификация клиента может стать недетерминированным, частичным, неприведенным автоматом.

Результаты данной работы могут быть использованы при проектировании архитектуры информационных систем.

Список источников

1. Shirokova E. Checking Robustness of Web Services Based on the Parallel Composition of Partial Timed Finite State Machine // 2018 IEEE East-West Design & Test Symposium (EWDTS). 2018. P. 1–6.
2. Евтушенко Н.В., Рекун М.В., Тихомирова С.В. Недетерминированные автоматы: анализ и синтез : учеб. пособие. Томск : Том. гос. ун-т, 2009. Ч. 2: Решение автоматных уравнений. 111 с.
3. Castagnetti G., Piccolo M., Villa T., Yevtushenko N., Mishchenko A., Brayton R.K. Solving Parallel Equations with BALM-II : Technical Report No. UCB/EECS2012-181 // Electrical Engineering and Computer Sciences University of California at Berkeley. 2012. URL: <http://www.eecs.berkeley.edu/Pubs/TechRpts/2012/EECS-2012-181.pdf> (Дата обращения: 20.08.2022).
4. West C.H. An automated technique of communication protocols validation // IEEE Trans. Comm. 1978. V. 26. P. 1271–1275.
5. El-Fakih K., Trenkaev V., Spitsyna N., Yevtushenko N. FSM Based Interoperability Testing Methods for Multi Stimuli Model // Testing of Communicating Systems : 16th International Conference, TestCom 2004 : Proc. Oxford, 2004. P. 60–75. (Lecture Notes in Computer Science; v. 2978).
6. Starke P.H. Abstract Automata. New York : American Elsevier Publishing Company, 1972. 419 p.

References

1. Shirokova, E. (2018) Checking Robustness of Web Services Based on the Parallel Composition of Partial Timed Finite State Machines. *2018 IEEE East-West Design & Test Symposium (EWDTS)*. pp. 1–6. DOI: 10.1109/EWDTS.2018.8524850

2. Evtushenko, N.V., Rekun, M.V. & Tikhomirova, S.V. (2009) *Nedeterminirovannye avtomaty: analiz i sintez* [Nondeterministic finite state machines]. Vol. 2. Tomsk: Tomsk State University.
3. Castagnetti, G., Piccolo M., Villa, T., Yevtushenko, N., Mishchenko, A. & Brayton, R.K. (2012) Solving Parallel Equations with BALM-II. *Technical Report No. UCB/EECS2012-181*. Berkeley: Electrical Engineering and Computer Sciences University of California. [Online] Available from: <http://www.eecs.berkeley.edu/Pubs/TechRpts/2012/EECS-2012-181.pdf>. (Accessed: 20th August 2022).
4. West, C.H. (1978) An automated technique of communication protocols validation. *IEEE Trans. Comm.* 26. pp. 1271–1275.
5. El-Fakih, K., Trenkaev, V., Spitsyna, N. & Yevtushenko, N. (2004) FSM Based Interoperability Testing Methods. *Proceedings of the IFIP 16th International Conference on Testing of Communicating Systems*. 2978. pp. 60–75.
6. Starke, P.H. (1972) *Abstract Automata*. New York: American Elsevier Publishing Company.

Информация об авторах:

Громов Максим Леонидович – доцент, кандидат физико-математических наук, доцент кафедры информационных технологий в исследовании дискретных структур Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: maxim.leo.gromov@gmail.com

Прокопенко Светлана Анатольевна – доцент, кандидат технических наук, доцент кафедры информационных технологий в исследовании дискретных структур Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: s.prokopenko@sibmail.com

Сотников Александр Павлович – аспирант кафедры информационных технологий в исследовании дискретных структур Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: sotnikhtc@gmail.com

Шабалдина Наталия Владимировна – доцент, кандидат технических наук, доцент кафедры информационных технологий в исследовании дискретных структур Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: nataliamailbox@mail.ru

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

Information about the authors:

Gromov Maxim Leonidovich (Candidate of Physical and Mathematical Sciences, Associate Professor, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: maxim.leo.gromov@gmail.com

Prokopenko Svetlana Anatolievna (Candidate of Technical Sciences, Associate Professor, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: s.prokopenko@sibmail.com

Sotnikov Aleksandr Pavlovich (Post-graduate Student, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: sotnikhtc@gmail.com

Shabaldina Natalia Vladimirovna (Candidate of Technical Sciences, Associate Professor, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: nataliamailbox@mail.ru

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

Received 26.09.2022; accepted for publication 01.03.2023

Поступила в редакцию 26.09.2022; принята к публикации 01.03.2023

Научная статья

УДК 004.724.2 + 004.272.43

doi: 10.17223/19988605/62/13

Простая масштабируемая коммутируемая управляющая сеть

Михаил Фёдорович Каравай¹, Виктор Сергеевич Подлазов²,
Владимир Владимирович Соколов³

^{1, 2, 3} *Институт проблем управления им. В.А. Трапезникова РАН, Москва, Россия*

¹ *mkaravay@yandex.ru*

² *podlazov@gmail.com*

³ *sok25101946@gmail.com*

Аннотация. Рассматривается проектирование локальных управляющих сетей на базе квазиполных орграфов. Считается, что сеть содержит активное вычислительное ядро произвольных размеров и множество пассивных абонентов, последние не взаимодействуют друг с другом. Абоненты активного ядра имеют бесконфликтный параллельный доступ друг к другу и ко всем пассивным абонентам. В отличие от предложенных ранее сетей на базе квазиполных графов новая база позволяет существенно уменьшить сложность сетей и улучшить их масштабируемость. При этом сохраняются важнейшие функционалы сетей: бесконфликтность параллельных передач, отсутствие дедлоков, самомаршрутизируемость, схемная и канальная отказоустойчивость.

Ключевые слова: локальные управляющие сети; квазиполные графы и орграфы; параллельные системы; бесконфликтность передач.

Для цитирования: Каравай М.Ф., Подлазов В.С., Соколов В.В. Простая масштабируемая коммутируемая управляющая сеть // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2023. № 62. С. 115–123. doi: 10.17223/19988605/62/13

Original article

doi: 10.17223/19988605/62/13

Simple scalable switched control network

Mihail F. Karavay¹, Victor S. Podlazov², Vladimir V. Sokolov³

^{1, 2, 3} *V.A. Trapeznikov Institute of Control Sciences of Russian Academy of Science, Moscow, Russian Federation*

¹ *mkaravay@yandex.ru*

² *podlazov@gmail.com*

³ *sok25101946@gmail.com*

Abstract. The design of local control networks based on quasicomplete digraphs is considered. It is believed that the network contains an active computing core of arbitrary size and a set of passive subscribers, the latter do not interact with each other. Subscribers of the active core have conflict-free parallel access to each other and to all passive subscribers. Unlike previously proposed networks based on quasi-complete graphs, the new base can significantly reduce the complexity of networks and improve their scalability. At the same time, the most important functionals of networks are preserved: conflict-free parallel transmissions, no deadlocks, self-routing, circuit and channel fault tolerance.

Keywords: local control networks; quasicomplete graphs and digraphs; parallel systems; conflict-free transfers.

For citation: Karavay, M.F., Podlazov, V.S., Sokolov, V.V. (2023) Simple scalable switched control network. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 62. pp. 115–123. doi: 10.17223/19988605/62/13

Введение

В работе [1] управляющие сети (УС) рассматриваются как коммутируемые УС (КУС) в отличие от шинных и многошинных УС (ШУС). Основное преимущество КУС перед ШУС – более высокие помехоустойчивость и масштабируемость при существенно более низких частотах работы устройств и одинаковой производительности. В УС различаются два вида устройств – активные и пассивные. Активные устройства (АУ) управляют работой пассивных устройств (ПУ), как правило, в последовательном режиме. АУ имеют полные связи между собой и с ПУ, которые не имеют связей между собой и имеют связи только с АУ. В работе [1] впервые предложено строить КУС как сети с топологией квазиполных графов, которые состоят из $N = m(m - 1)/\sigma + 1$ m -портовых АУ, связанных дуплексными каналами через N коммутаторов $m \times m$. Соединения между коммутаторами и АУ задаются таблицами инцидентности симметричных блок-схем (block-designs) $B(N, m, \sigma)$, изучаемых в комбинаторике [2].

Пример таблицы инцидентности для $B(13, 4, 1)$ приведен в табл. 1, в которой 1-й столбец задает номера коммутаторов сети, выделенные курсивом, а строки в остальных столбцах задают номера АУ, подсоединенных к каждому коммутатору.

Сети с топологией квазиполных графов и орграфов использовались ранее авторами для оптимизации характеристик ряда системных сетей многопроцессорных вычислительных систем [3–7].

Пусть теперь в КУС имеется только три АУ с номерами 1–3, а остальные 10 устройств являются ПУ, между которыми соединения не нужны. Для исключения этих соединений в табл. 1 вычеркиваются строки, отмеченные заливкой, которые не содержат АУ с номерами 1–3, и производится перенумерация оставшихся коммутаторов. В результате образуется табл. 2, которая задает соединения в такой КУС.

Таблица 1

Таблица соединений для сети с топологией квазиполного графа при $m = 4$

<i>1</i>	1	13	11	5
<i>2</i>	2	1	12	6
<i>3</i>	3	2	13	7
<i>4</i>	4	3	1	8
<i>5</i>	5	4	2	9
<i>6</i>	6	5	3	10
<i>7</i>	7	6	4	11
<i>8</i>	8	7	5	12
<i>9</i>	9	8	6	13
<i>10</i>	10	9	7	1
<i>11</i>	11	10	8	2
<i>12</i>	12	11	9	3
<i>13</i>	13	12	10	4

Таблица 2

Таблица соединений для КУС с 3 АУ, 9 ПУ и 9 коммутаторами

<i>1</i>	1	13	11	5
<i>2</i>	2	1	12	6
<i>3</i>	3	2	13	7
<i>4</i>	4	3	1	8
<i>5</i>	5	4	2	9
<i>6</i>	6	5	3	10
<i>7</i>	10	9	7	1
<i>8</i>	11	10	8	2
<i>9</i>	12	11	9	3

Будем характеризовать КУС, создаваемые описанным образом, следующими параметрами: α – число АУ, $\pi = N - \alpha$ – число ПУ и κ – число коммутаторов. Обозначим $KUC(\alpha, \pi, \kappa)$ и определим

коммутационную сложность сети S количеством точек коммутации и канальную сложность L , выраженную числом дуплексных каналов. В такой КУС все АУ имеют бесконфликтные соединения на произвольных перестановках пакетов между ними, каждый АУ имеет соединение с любым ПУ и все АУ имеют бесконфликтные соединения с непересекающимися множествами ПУ.

Дополнительно предполагается, что множество портов каждого АУ создается их подсоединением через встречные демультиплексоры $1 \times m$ и мультиплексоры $1 \times m$ сложности m каждый. Тогда коммутационная сложность КУС(α, π, κ) задается как $S = \kappa m^2 + 2mN = \pi^\gamma$, где $\gamma = \log_\pi S$, а канальная сложность – как $L = mN = \pi^\lambda$, где $\lambda = \log_\pi L$, и по ним рассчитываются удельные сложности $s = S/\pi = \pi^{\gamma-1}$ и $l = L/\pi = \pi^{\lambda-1}$. В табл. 3 приводятся характеристики КУС($4, N - 4, \kappa$) при разных m .

Таблица 3

Характеристики КУС с $\alpha = 4$ и $\sigma = 1$

m	π	κ	S	s	L	l
4	9	10	$264 = \pi^{2,54}$	$\pi^{1,54}$	$40 = \pi^{1,68}$	$\pi^{0,68}$
5	17	14	$560 = \pi^{2,23}$	$\pi^{1,23}$	$70 = \pi^{1,5}$	$\pi^{0,5}$
7	Не существует					
6	27	19	$1\ 056 = \pi^{2,11}$	$\pi^{1,11}$	$114 = \pi^{1,44}$	$\pi^{0,44}$
8	53	27	$2\ 640 = \pi^{1,98}$	$\pi^{0,98}$	$216 = \pi^{1,35}$	$\pi^{0,35}$
9	69	31	$3\ 825 = \pi^{1,95}$	$\pi^{0,95}$	$279 = \pi^{1,33}$	$\pi^{0,33}$
10	87	39	$5\ 720 = \pi^{1,94}$	$\pi^{0,94}$	$390 = \pi^{1,33}$	$\pi^{0,33}$

К недостаткам КУС с топологией квазиполных графов можно отнести слабую масштабируемость, осуществляемую только за счет увеличения степени m , и довольно высокую сложность за счет большого числа коммутаторов κ . Здесь решается задача создания КУС большей масштабируемости и меньшей сложности на основе сетей с топологией квазиполных орграфов.

1. Квазиполные орграфы

Известен двумерный m -ичный гиперкуб [8. Р. 323] с $N = m^2$ узлами, в котором из каждого узла выходит набор ребер с длинами $(1, 2, \dots, m - 1, m, 2m, \dots, (m - 1)m)$. Длинной дуги мы называем разницу номеров по $\text{mod} N$ инцидентных ей узлов. В таком мультикольце все узлы имеют одинаковые степени $(m - 1)$ входных и выходных ребер (рис. 1).

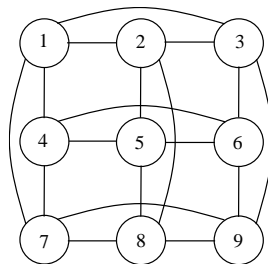


Рис. 1. Двумерный 3-ичный гиперкуб
Fig. 1. Two-dimensional 3-ary hypercube

В реальности каждый узел содержит абонента и коммутатор $m \times m$ полустепени m , связанных $2m$ дугами соответствующих длин (рис. 2). В таком представлении 2-мерный m -ичный гиперкуб является орграфом с $2N$ узлами двух сортов – абонентов и коммутаторов соответственно. В таком качестве он может быть представлен как двудольный квазиполный орграф с $N = m^2$ узлами в каждой доле (рис. 3). Соединения в нем задаются разными таблицами инцидентности для дуг от абонентов и дуг к абонентам (табл. 4).

В сетях с топологией квазиполного орграфа, в отличие от квазиполного графа, связи между абонентами во встречных направлениях проходят по симплексным каналам по разным маршрутам.

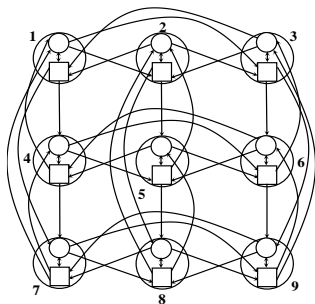


Рис. 2. 2-мерный 3-ичный гиперкуб как орграф
 Fig. 2. 2-dimensional 3-ary hypercube as a digraph

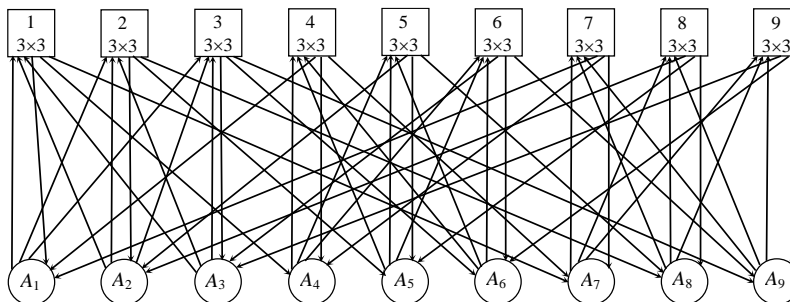


Рис. 3. Квазиполный орграф при $m = 3$
 Fig. 3. Quasicomplete digraph for $m = 3$

Таблица 4

Таблицы инцидентности для квазиполного орграфа при $m = 3$

Коммутаторы	Дуги от абонентов			Дуги к абонентам		
	1	2	3	1	4	7
1	1	2	3	1	4	7
2	1	2	3	2	5	8
3	1	2	3	3	6	9
4	4	5	6	1	4	7
5	4	5	6	2	5	8
6	4	5	6	3	6	9
7	7	8	9	1	4	7
8	7	8	9	2	5	8
9	7	8	9	3	6	9

2. Простая коммутлируемая управляющая сеть

На основе квазиполного орграфа можно построить коммутлируемую управляющую сеть $КУС(m, N - m, N)$ с двумя множествами абонентов – активных устройств A_i ($i \leq m$) и пассивных устройств P_i ($m < i \leq N$), которая представлена на рис. 4 для $m = 3$. В этой сети любое АУ может бесконфликтно передать пакет любому ПУ при произвольной перестановке пакетов между ними.

Как и ранее, предполагается, что множество входных и выходных портов каждого абонента создается посредством их подсоединения через мультиплексоры $1 \times m$ и демультимплексоры $1 \times m$ соответственно (см. рис. 4). В схеме $КУС(3, 6, 9)$ абоненты имеют двойную нумерацию – по таблице соединений (см. табл. 4) и как АУ и ПУ.

Соединения в $КУС(3, 6, 9)$ задаются табл. 4. Выделим в табл. 4 соединения от ПУ к коммутаторам с номерами больше 3 и удалим их (табл. 5). Тогда в редуцированной таблице соединений образуются группы из трех строк с входами от одинаковых наборов ПУ и с выходами к одному АУ, имеющему разные номера для разных строк в группе. Эти группы выделены разной заливкой. Каждую такую группу строк можно заменить на одну строку с одним коммутатором и со входами от ПУ

в группе и с выходами ко всем АУ. При этом образуется табл. 6, которая задает таблицу соединений уже для КУС(3, 6, 5) с минимальным числом коммутаторов, равным 5. На рис. 5 представлена схема КУС(3, 6, 5), которая имеет вид минимального квазиполного орграфа.

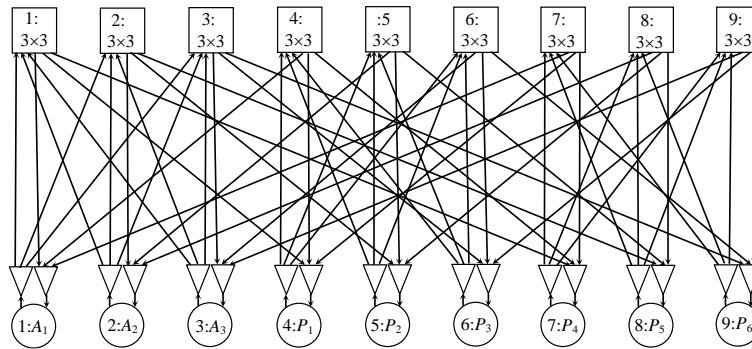


Рис. 4. КУС(3, 6, 9) с 1-портовыми абонентами
Fig. 4. Control network (3, 6, 9) with 1-port subscribers

Таблица 5

Редукция таблицы соединений для КУС(3, 6, 9)

Коммутаторы	Дуги от абонентов			Дуги к абонентам		
	1	2	3	1	4	7
1	1	2	3	1	4	7
2	1	2	3	2	5	8
3	1	2	3	3	6	9
4	4	5	6	1		
5	4	5	6	2		
6	4	5	6	3		
7	7	8	9	1		
8	7	8	9	2		
9	7	8	9	3		

Таблица 6

Таблицы соединений для редуцированной КУС(3, 6, 5)

Коммутаторы	Дуги от абонентов			Дуги к абонентам		
	1	2	3	1	4	7
1	1	2	3	1	4	7
2	1	2	3	2	5	8
3	1	2	3	3	6	9
4	4	5	6	1	2	3
5	7	8	9	1	2	3

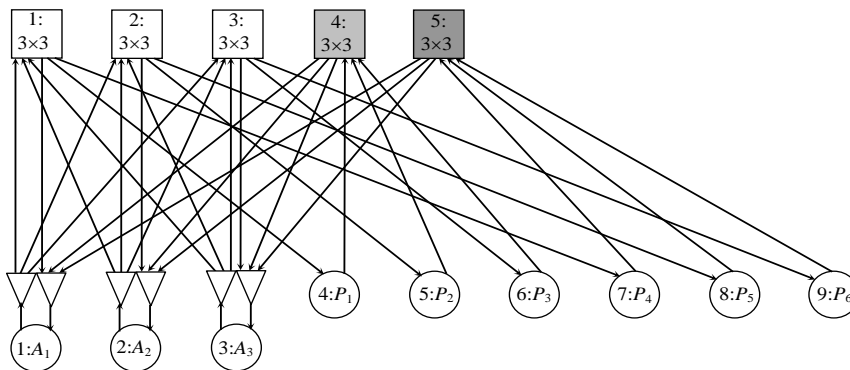


Рис. 5. Редуцированная КУС(3, 6, 5)
Fig. 5. Reduced control network (3, 6, 5)

Аналогичное редуцирование сети можно провести и для КУС(6, 30, 36). В результате описанного выше редуцирования образуется КУС(6, 30, 11) с минимальным числом коммутаторов.

В КУС($m, N - m, 2m - 1$) все каналы являются полудуплексными. Поэтому канальная сложность сети в дуплексных каналах задается как $L = m^2 + N - m = 2m^2 - m$. Коммутационная сложность этой сети задается как $S = m(m^2 + 2m) + (m - 1)m^2 = 2m^3 + m^2$. В частности, при $m = 4$ для КУС(4, 12, 7) имеем $L = 28 = \pi^{1,34}$, $l = \pi^{0,34}$ и $S = 144 = \pi^{2,0}$, $s = \pi^{1,0}$. Это существенно меньше, чем в КУС(4, 9, 10) с топологией квазиполного графа (см. табл. 3) при меньшем числе ПУ в последней.

Для простой КУС(m, π, κ) заголовки пакетов от АУ должны содержать номера выходных портов мультиплексоров и восходящих коммутаторов, а заголовки от ПУ – выходных портов нисходящих коммутаторов.

3. Масштабирование простой коммутируемой управляющей сети

Простая сеть КУС(m, π, κ) допускает увеличение числа ПУ в m^{r+1} раз, т.е. преобразование в сеть КУС(m, π_r, κ) с $\pi_r = \pi m^{r+1}$, посредством использования r ярусов демультимплексоров и мультиплексоров. Масштабирование посредством одного яруса состоит во вставке демультимплексора $1 \times m$ в каналы от ПУ и мультиплексора $m \times 1$ в каналы к ПУ, а их выходы и входы, соответственно, подсоединяются к m разным ПУ. На рис. 6 приведен пример частичного 1-ярусного масштабирования. При максимальном 1-ярусном масштабировании сеть КУС(3, 6, 5) преобразуется в КУС(3, 18, 5).

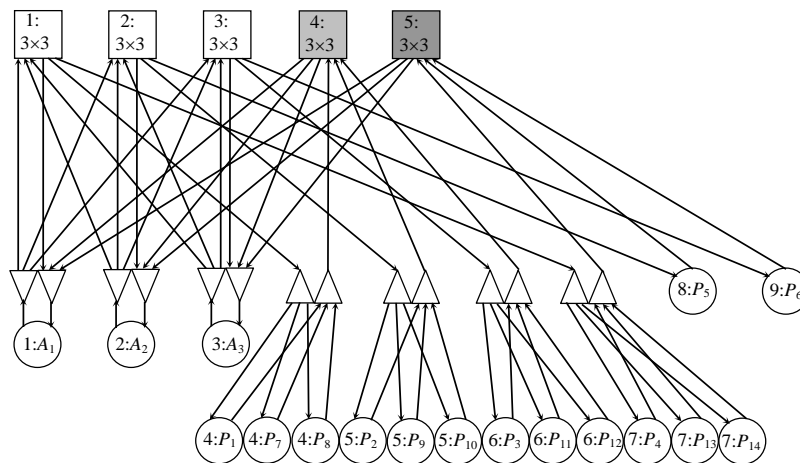


Рис. 6. Частичное 1-ярусное масштабирование КУС(3, 6, 5) в КУС(3, 14, 5)
Fig. 6. Partial 1-tier scaling control network (3, 6, 5) to control network (3, 14, 5)

Важно подчеркнуть, что в масштабированной сети КУС(m, π_r, κ) каждое АУ имеет двустороннюю связь с каждым из ПУ.

При 1-ярусном масштабировании канальная сложность КУС(m, π_1, κ) задается как $L_1 = L + \pi m = m^2 + N - m + (N - m)m = Nm + N - m = m^3 + m^2 - m$, а коммутационная сложность – как $S_1 = S + 2m\pi = 2m^3 + m^2 + 2(N - m)m = 4m^3 - m^2$. В частности, при $m = 4$ для КУС(4, 48, 8) имеем $L_1 = 76 = \pi^{1,12}$, $l_1 = \pi^{0,12}$ и $S_1 = 240 = \pi^{1,42}$, $s_1 = \pi^{0,42}$, что существенно меньше, чем в КУС(4, 53, 27) с топологией квазиполного графа (см. табл. 3) при близком числе ПУ.

Для масштабированной КУС(m, π_r, κ) заголовки пакетов от АУ должны дополнительно содержать номера выходных портов мультиплексоров в каждом добавленном ярусе.

4. Обобщение простой коммутируемой управляющей сети

В исходной постановке задачи число АУ α считается фиксированным и не зависящим от значения $m > \alpha$ ($\alpha = 4$ в [1]; см. табл. 3). Структуру простой КУС($m, N - m, 2m - 1$) можно подстроить под такую постановку, обеспечивая при этом увеличение числа ПУ на $m - \alpha$ единиц.

Таблица соединений для КУС(4, 32, 11)

Коммутаторы	Дуги от абонентов						Дуги к абонентам					
	1	2	3	4			1	7	13	19	25	31
1	1	2	3	4			1	7	13	19	25	31
2	1	2	3	4			2	8	14	20	26	32
3	1	2	3	4			3	9	15	21	27	33
4	1	2	3	4			4	10	16	22	28	34
5	1	2	3	4	31		31	11	17	23	29	35
6	1	2	3	4	32		32	12	18	24	30	36
7	7	8	9	10	11	12	1	2	3	4	5	6
8	13	14	15	16	17	18	1	2	3	4	5	6
9	19	20	21	22	23	24	1	2	3	4	5	6
10	25	26	27	28	29	30	1	2	3	4	5	6
11	31	32	33	34	35	36	1	2	3	4	5	6

Строится простая КУС($m, N - m, 2m$) и из нее удаляется $m - \alpha$ АУ со всеми входными / выходными каналами. В табл. 7 это показано для $m = 6$ и $\alpha = 4$. При этом освобождается $m - \alpha$ входных портов коммутаторов и по одному выходному порту у $m - \alpha$ коммутаторов. К ним подсоединяется $m - \alpha$ дополнительных ПУ, номера которых выделены заливкой. При этом часть входных портов m коммутаторов остается свободной. Таким образом, получается таблица соединения для обобщенной КУС($\alpha, N - \alpha, 2m - 1$).

5. Простая коммутируемая управляющая сеть с канальной отказоустойчивостью

Таблицу соединений для простой управляющей сети с $\alpha = m$ и $\pi = N - m$ можно формировать по восходящим и нисходящим каналам между АУ и ПУ через отдельные коммутаторы. В результате для $m = 3$ образуется простая КУС*($m, N - m, 2m$), которая содержит избыточные каналы, используемые для соединения восходящих и нисходящих коммутаторов (рис. 7).

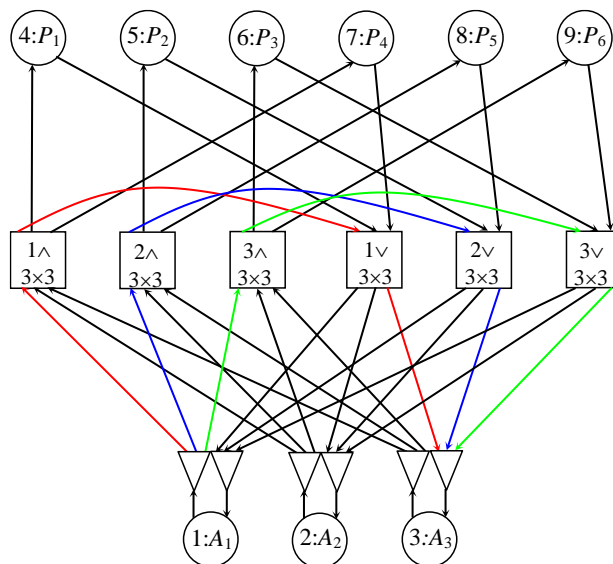


Рис. 7. Простая КУС(3, 6, 6), с 2-канальной отказоустойчивостью между АУ
 Fig. 7. Simple control network (3, 6, 6), with 2-channel fault tolerance between active units

КУС*($m, N - \alpha, 2m$) обладает $(m - 1)$ -канальной отказоустойчивостью связей между АУ. В этой сети коммутаторы делятся на восходящие (\wedge) на путях от АУ к ПУ и нисходящие (\vee) на путях от ПУ к АУ. При этом в коммутаторах остаются свободными по одному выходному и входному порту, которые соединяются указанными выше избыточными каналами, что и обеспечивает канальную отказоустойчивость.

Заключение

В работе предложена методика построения коммутационных управляющих сетей с топологией квазиполных орграфов. Даны формулы для расчета числа активных и пассивных устройств в них и расчета значений канальной и коммутационных сложностей. Показано, что они имеют существенно меньшую коммутационную сложность и существенно лучшую масштабируемость, чем ранее предложенные сети с топологией квазиполных графов.

Список источников

1. Каравай М.Ф., Михайлов А.М. Design of local heterogeneous system control networks of a new generation with the preservation of the optimality of the main topological functionals of the network // *Journal of Physics: Conference Series*. 2021. V. 2091. Art. 012038.
2. Каравай М.Ф., П.П. Пархоменко П.П., Подлазов В.С. Комбинаторные методы построения двудольных однородных минимальных квазиполных графов (симметричных блок-схем) // *Автоматика и телемеханика*. 2009. № 2. С. 153–170.
3. Каравай М.Ф., Подлазов В.С. Расширенные блок-схемы для идеальных системных сетей // *Проблемы управления*. 2012. № 4. С. 45–51.
4. Каравай М.Ф., Подлазов В.С. Топологические резервы суперкомпьютерного интерконекта // *Управление большими системами*. 2013. Вып. 41. С. 395–423.
5. Барabanова Е.А., Выговтов К.А., Подлазов В.С. Неблокируемые отказоустойчивые двухкаскадные дуальные фотонные коммутаторы // *Проблемы управления*. 2021. № 4. С. 82–91.
6. Подлазов В.С. Неблокируемые отказоустойчивые дуальные фотонные коммутаторы широкой масштабируемости // *Проблемы управления*. 2021. № 5. С. 70–87.
7. Barabanova E.A., Vitovtov K.A., Vishnevskii V.M., Podlazov V.S. High-capacity strictly non-blocking optical switches based on new dual principle // *Journal of Physics: Conference Series*. 2021. V. 2091. Art. 012040.
8. Bhuyan L.N., Agrawal D.P. Generalized Hypercube and Hyperbus Structures for a Computer Network // *IEEE Trans. on Computers*. 1984. V. C-33, № 4. P. 323–333.

References

1. Karavay, M.F. & Mikhailov, A.M. (2021) Design of local heterogeneous system control networks of a new generation with the preservation of the optimality of the main topological functionals of the network. *Journal of Physics: Conference Series*. 2091. Art. 012038. DOI: 10.1088/1742-6596/2091/1/012038
2. Karavay, M.F., Parkhomenko, P.P. & Podlazov, V.S. (2009) Combinatorial methods for constructing bipartite homogeneous minimal quasi-complete graphs (symmetric block diagrams). *Avtomatika i telemekhanika – Automation and Remote Control*. 2. pp. 312–327.
3. Karavay, M.F. & Podlazov, V.S. (2012) Extended block diagrams for ideal system networks. *Avtomatika i telemekhanika – Automation and Remote Control*. 74(4). pp. 710–724.
4. Karavay M.F. & Podlazov V.S. (2013) Topologicheskie rezervy superkomp'yuternogo interkonekta [Topological reserves of a supercomputer interconnect]. *Upravlenie bol'shimi sistemami*. 41. pp. 395–423.
5. Barabanova, E.A., Vitovtov, K.A. & Podlazov, V.S. (2021) Non-blocking fault-tolerant two-stage dual photonic switches. *Problemy upravleniya – Control Sciences*. 4. pp. 67–76.
6. Podlazov, V.S. (2021) Non-blocking fault-tolerant dual photonic switches of wide scalability. *Problemy upravleniya – Control Sciences*. 5. pp. 61–76.
7. Barabanova, E.A., Vitovtov, K.A., Vishnevskii, V.M. & Podlazov, V.S. (2021) High-capacity strictly non-blocking optical switches based on new dual principle. *Journal of Physics: Conference Series*. 2091. Art. 012040. DOI: 10.1088/1742-6596/2091/1/012040
8. Bhuyan, L.N. & Agrawal, D.P. (1984) Generalized Hypercube and Hyperbus Structures for a Computer Network. *IEEE Trans. on Computers*. C-33(4). pp. 323–333.

Информация об авторах:

Каравай Михаил Федорович – доктор технических наук, заведующий лабораторией технической диагностики Института проблем управления им. В.А. Трапезникова РАН (Москва, Россия). E-mail: mkaravay@yandex.ru

Подлазов Виктор Сергеевич – доктор технических наук, главный научный сотрудник Института проблем управления им. В.А. Трапезникова РАН (Москва, Россия). E-mail: podlazov@gmail.com

Соколов Владимир Владимирович – старший научный сотрудник Института проблем управления им. В.А. Трапезникова РАН (Москва, Россия). E-mail: sok25101946@gmail.com

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

Information about the authors:

Karavay Mikhail F. (Doctor of Technical Sciences, Head Laboratory Technical Diagnostics of the V.A. Trapeznikov Institute of Control Science of the RAS, Moscow, Russian Federation). E-mail: mkaravay@yandex.ru

Podlazov Viktor S. (Doctor of Technical Sciences, Chief Researcher of the V.A. Trapeznikov Institute of Control Science of the RAS, Moscow, Russian Federation). E-mail: podlazov@gmail.com

Sokolov Vladimir V. (Senior Researcher of the V.A. Trapeznikov Institute of Control Science of the RAS, Moscow, Russian Federation). E-mail: sok25101946@gmail.com

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

Received 10.09.2022; accepted for publication 01.03.2023

Поступила в редакцию 10.09.2022; принята к публикации 01.03.2023

ДИСКРЕТНЫЕ ФУНКЦИИ И АВТОМАТЫ

DISCRETE FUNCTION AND AUTOMATONS

Научная статья

УДК 519.7

doi: 10.17223/19988605/62/14

**Построение аппроксимирующих схем для синхронных автоматов
в рамках технологии троирования****Сергей Александрович Останин¹, Анжела Юрьевна Матросова²,
Валентина Валерьевна Андреева³***^{1, 2, 3} Томский государственный университет, Томск, Россия**¹ sergeiostanin@yandex.ru**² mau11@yandex.ru**³ avv.21@mail.ru*

Аннотация. Троирование (Triple-Modular Redundancy (TMR) technique) является одним из широко применяемых на практике подходов к обеспечению надежности функционирования логических схем. В рамках этой технологии используется три идентичных схемы, одноименные выходы которых поступают на схему голосования. В такой системе корректное функционирование обеспечивается при неисправности одной из трех схем. Однако появившиеся в последние годы возможности одновременного внесения в каждую копию и соответствующие линии вредоносных подсхем (Trojan Circuits) делают метод троирования уязвимым к таким действиям. Одним из выходов в этой ситуации является использование вместо идентичных трех синхронных последовательностных схем двух аппроксимирующих схем и одной рабочей схемы, выполняющей предписанное разработчиком функционирование синхронного автомата. Предлагаемый подход гарантирует отсутствие незащищенной области, характерной для известных методов применения аппроксимирующих схем в технологии троирования.

Ключевые слова: синхронные последовательностные схемы; безызбыточные системы ДНФ (БСДНФ); аппроксимирующие схемы; троирование; константные неисправности литер БСДНФ.

Для цитирования: 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

Original article

doi: 10.17223/19988605/62/14

**Deriving approximate circuits for TMR technique
applied to synchronous sequential circuits****Sergey A. Ostanin¹, Angela Yu. Matrosova², Valentina V. Andreeva³***^{1, 2, 3} Tomsk State University, Tomsk, Russian Federation**¹ sergeiostanin@yandex.ru**² mau11@yandex.ru**³ avv.21@mail.ru*

Abstract. Triple-Modular Redundancy (TMR) technique is one of the conventional approaches to provide reliable functioning of logical circuits. In the frame of this technique three identical circuits are applied, their like outputs are

connected with a voting circuit. Correct functioning of such system is guaranteed, if one of three circuits is fault. When using outsourcing it is possible to inject a Trojan Circuit in the same line of each identical circuits of TMR, and TMR technique becomes vulnerable. One way of confronting to vulnerability is connected with applying instead of identical sequential circuits two approximate circuits and one correct circuit that implements the proper functioning. The suggested approach guarantees absence of unprotected area that as a rule appears when using approximate circuits in TMR techniques.

Keywords: synchronous sequential circuits; irredundant system of SoPs; approximate circuits; TMR technique; literal constant faults of Irredundant system of SoPs.

For citation: 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'naja tehnika i informatika – Tomsk State University Journal of Control and Computer Science*. 62. pp. 124–131. doi: 10.17223/19988605/62/14

Введение

Схемы троирования являются одним из широко распространенных подходов к обеспечению надежности функционирования логических схем. Предполагается, что одна из трех схем может быть неисправной. В этом случае в условиях подключения к одноименным выходам трех схем значения значения выхода определяется значением большинства одноименных выходов. Схемы голосования предполагаются исправными. Однако 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. Вредоносная подсхема (Trojan Circuit) может изменить в нужный момент значение сигнала на этих линиях на противоположное значение с целью искажения работы схемы или с целью извлечения конфиденциальной информации из устройства, содержащего рассматриваемую схему. Это значит, что технология троирования оказывается уязвимой в условиях возможности включения вредоносных подсхем.

Одним из известных подходов к защите комбинационной схемы в рамках технологии троирования является использование аппроксимирующих схем, которые строятся из рабочей схемы, реализующей нужное разработчику поведение [1, 2]. Аппроксимирующие схемы применяются также при синтезе самопроверяемых логических схем [3]. В статье [3] аппроксимирующие схемы используются с целью сокращения аппаратных затрат в условиях, когда функционирование самопроверяемой схемы может незначительно отклоняться от функционирования рабочей схемы, что менее важно, чем сокращение аппаратных затрат. При использовании аппроксимирующих схем в схемах троирования отклонение от корректного функционирования возможно в присутствии неисправности в одной из трех схем в так называемой незащищенной области [1, 2]. При подобном использовании аппроксимирующих схем стремятся сократить незащищенную область. Если комбинационная схема является комбинационной составляющей последовательностной схемы, то ее поведение можно описать системой частичных булевых функций. Эти функции извлекаются либо из таблицы переходов-выходов синхронного автомата, либо из State Transition Graph (STG) описания путем кодирования символов алфавитов синхронного автомата. Использование системы частичных булевых функций открывает возможность для получения различных ее реализаций в качестве задания на синтез аппроксимирующих схем. Такой подход исключает появление незащищенной области в рабочей области функционирования синхронного автомата. Итак, еще на этапе создания рабочей схемы предлагается учесть необходимость ее защиты от внедрения вредоносных подсхем при использовании троирования, причем исключается незащищенная область, характерная при синтезе комбинационных схем, когда задание на синтез представляется системой полностью определенных булевых функций.

В рассматриваемом в данной работе подходе построение аппроксимирующих схем предлагается начать с использования избыточной системы ДНФ (БСДНФ), построенной по заданной системе частичных функций. БСДНФ является реализацией этой системы частичных функций. БСДНФ далее используется в качестве задания на синтез рабочей схемы синхронного автомата. Эта же БСДНФ яв-

ляется источником получения БСДНФ для каждой из аппроксимирующих схем, причем получаемые системы являются реализацией заданной системы частичных функций. Такой подход позволяет получать аппроксимирующие схемы, не порождающие незащищенной области. Кроме того, для получения аппроксимирующих схем не требуется введения в рабочую схему дополнительных элементов [1, 2] для обеспечения на выходах этих схем монотонного проявления искажений, вносимых в рабочую схему. Предлагаемый нами подход ориентирован на использование автоматизированного синтеза логических схем, применяемого в современных САПР. В частности, предполагается, что задание на синтез рабочей схемы в рамках САПР представляется БСДНФ. Ее искажения обеспечиваются введением в систему одиночных константных неисправностей литер [4, 5]. Речь идет о неисправностях константа 1, которые сводятся к снижению на единицу ранга конъюнкций отдельных ДНФ из БСДНФ рабочей схемы, то есть к b -неисправностям [4, 5] литер СБДНФ.

Выбор конъюнкций в БСДНФ рабочей схемы для внесения в них неисправностей с целью построения аппроксимирующих систем функций ориентирован на достижение значительного отличия аппроксимирующих систем от БСДНФ рабочей схемы, т.е. на обеспечение возможно больших затруднений при введении вредоносных подсхем в систему троирования. В то же время необходим баланс между вносимыми изменениями и сложностью получаемых аппроксимирующих схем.

1. Построение аппроксимирующих систем булевых функций

Пусть описание поведения рабочей схемы синхронного автомата представлено в виде БСДНФ. БСДНФ состоит из простых импликант системы булевых функций, т.е. ни одна из функций не может быть исключена из характеристик этих импликант. Предполагается, что БСДНФ построена по системе частичных булевых функций, полученной либо из таблицы переходов-выходов синхронного автомата, либо из State transition Graph (STG) описания поведения синхронного автомата. Система частичных функций представляется парами множеств M_{1i} , M_{0i} , задающих единичные и нулевые наборы частичной функции f_i системы. Каждая конъюнкция БСДНФ снабжается характеристикой, в ней перечислены функции системы, для которых эта конъюнкция является допустимой.

Конъюнкция допустима, если пересечение ее интервала с областью нулевых наборов значений переменных частичной функции системы, упомянутой в характеристике конъюнкции, пусто, а с областью единичных наборов значений переменных частичной функции не пусто.

Поскольку конъюнкция БСДНФ является простой импликантой системы булевых функций, то из нее нельзя исключить ни одной (любой) литеры в условиях сохранения характеристики этой конъюнкции.

Кроме того, так как список конъюнкций с соответствующими характеристиками представляет безызбыточную систему ДНФ, то из характеристики конъюнкции этой системы нельзя исключить ни одной функции.

Будем иметь в виду, что исключение переменной x_i из конъюнкции K системы приводит к пересечению с областью нулевых значений не каждой частичной функции из характеристики h этой конъюнкции, а только некоторых из них, причем обязательно хотя бы для одной из функций характеристики h . Функции характеристики h , для которых исключение переменной x_i не приводит к искажению соответствующей частичной функции, будем называть расширяемыми по рассматриваемой переменной для исследуемой конъюнкции.

Таким образом, для каждой переменной x_i конъюнкция K из БСДНФ множество функций ее характеристики h разбивается на два подмножества h_p и h_{np} , расширяемого по переменной и не расширяемого по переменной x_i .

Исключение переменной из конъюнкции K для функций из множества h_p сохраняет реализацию частичных функций множества h_p . Исключение переменной приводит лишь к расширению множества единичных наборов реализации частичных функций из характеристики h_p . При исключении переменной x_i получаем конъюнкцию системы $(K/x_i, h_p)$. Здесь конъюнкция K/x_i получается из конъюнкция K исключением переменной x_i .

Отметим, что при разбиении характеристики конъюнкции K на два подмножества получаемые из K конъюнкции имеют характеристики меньшей мощности, чем характеристика h , тогда каждая из них может быть расширена до простой импликанты заданной системы частичных функций. Итак, из простой импликанты (K, h) получаем две простые импликанты: K^*, h_{np} ; $K^*/x_i, h_p$.

В дальнейшем полученные простые импликанты будем называть продуктами операции расщепления простой импликанты (K, h) из БСДНФ.

Будем иметь в виду, что конъюнкция (K, h) расщепляема по переменной x_i , если переменная x_i порождает для этой конъюнкции непустую характеристику h_p .

Договоримся называть операцией расщепления конъюнкции (K, h) по x_i из БСДНФ замену ее конъюнкциями (K^*, h_{np}) ; $(K^*/x_i, h_p)$.

Утверждение 1. Операция расщепления конъюнкции K по переменной x_i приводит к расширению области единичных значений реализации системы частичных функций без искажения системы частичных функций, реализованной БСДНФ.

Действительно, замена простой импликанты (K, h) БСДНФ на допустимую конъюнкцию K^*, h_{np} сохраняет область единичных значений частичных функций, перечисленных в характеристике h_{np} при условии, что конъюнкция K^* совпадает с конъюнкцией K либо расширяет область единичных значений этих частичных функций при несовпадении конъюнкций. Замена простой импликанты (K, h) БСДНФ на допустимую конъюнкцию $(K^*/x_i, h_p)$ всегда приводит к расширению области единичных значений частичных функций, перечисленных в характеристике h_p . Это значит, что расщепление простой импликанты (K, h) по переменной x_i приводит к расширению области единичных значений реализации системы частичных функций без искажения системы частичных функций, представленной БСДНФ.

Пусть $F(x_1, \dots, x_n) = f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n)$ – система из m булевых функций от n переменных. Будем говорить, что система $G(x_1, \dots, x_n) = g_1(x_1, \dots, x_n), \dots, g_m(x_1, \dots, x_n)$ поглощается системой $F(x_1, \dots, x_n) = f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n)$, и обозначать это как $G(\alpha_1, \dots, \alpha_n) < F(\alpha_1, \dots, \alpha_n)$, если для любого набора $\alpha = \alpha_1, \dots, \alpha_n$

$$G(\alpha_1, \dots, \alpha_n) \leq F(\alpha_1, \dots, \alpha_n),$$

то есть для одноименных функций системы выполняется условие

$$g_i(\alpha_1, \dots, \alpha_n) \leq f_i(\alpha_1, \dots, \alpha_n).$$

Обозначим через $G(x_1, \dots, x_n)$ систему БСДНФ, реализуемую рабочей схемой, а через $F(x_1, \dots, x_n)$ – систему, полученную расщеплением некоторых ее простых импликант по тем или иным переменным.

Утверждение 2. Выполнение операций расщепления нескольких конъюнкций по переменной (одна и та же конъюнкция может расщепляться по нескольким переменным) приводит к получению системы $F(x_1, \dots, x_n)$, поглощающей систему $G(x_1, \dots, x_n)$: $G(x_1, \dots, x_n) < F(x_1, \dots, x_n)$.

Поскольку операция расщепления простой импликанты из $G(x_1, \dots, x_n)$ приводит к расширению области единичных значений некоторых функций этой системы, то многократное применение расщепления простых импликант системы $G(x_1, \dots, x_n)$ приводит к тому же результату.

Следствие 1. Поглощающая система $F(x_1, \dots, x_n)$ является реализацией той же системы частичных функций, что и система $G(x_1, \dots, x_n)$.

Следствие 2. Различные поглощающие системы $F(x_1, \dots, x_n)$, полученные расщеплением конъюнкций системы $G(x_1, \dots, x_n)$ по тем или иным переменным могут использоваться в качестве аппроксимирующих систем для системы $G(x_1, \dots, x_n)$, обеспечивая отсутствие незащищенной области синхронного автомата (синхронного последовательностного устройства).

Чем сильнее поглощающие системы $F(x_1, \dots, x_n)$, используемые для представления аппроксимирующих систем, отличаются друг от друга и от системы $G(x_1, \dots, x_n)$, тем труднее воспользоваться технологией введения вредоносной подсхемы в каждую из трех реализующих эти системы схем с целью обеспечения одинаковых искажений.

Отметим, что операции расщепления конъюнкций БСДНФ приводят к увеличению числа простых импликант, причем возможно одновременное уменьшение суммы рангов продуктов расщепления

по сравнению с рангом расщепляемой конъюнкции БСДНФ, а также уменьшение мощности характеристик h в простых импликантах системы, полученных в результате расщепления. Насколько вероятны такие ситуации, трудно сказать, и трудно без экспериментов на моделях реальных схем предсказать, как повлияет расщепление конъюнкций СБДНФ на сложность аппроксимирующей схемы, реализующей полученную безызбыточную систему ДНФ. Скорее всего, схема будет усложняться, что является естественной платой за защиту от внедрения вредоносных подсхем. Необходим компромисс между усложнением аппроксимирующих схем и сложностью введения в них вредоносной подсхемы при использовании технологии троирования.

Заметим, что, чем меньше ранг конъюнкции K простой импликанты (K, h) , тем больше искажаются функции, заданные БСДНФ, и в то же время меньше число искажаемых при расщеплении функций. Напротив, чем больше ранг конъюнкции K , тем меньше искажаются функции, заданные БСДНФ, и в то же время больше число искажаемых при расщеплении конъюнкций.

Необходимо построить две аппроксимирующие системы F_1, F_2 , поглощающие систему $G(x_1, \dots, x_n)$, и являющиеся реализацией исходной системы частичных функций.

Предлагается в каждой аппроксимирующей системе использовать результаты расщеплений конъюнкций как меньших, так и больших рангов БСДНФ для обеспечения разнообразия искажений этих систем при одинаковом числе расщеплений, используемых при получении аппроксимирующих систем.

К полученным системам и исходной безызбыточной системе применяем какую-либо из автоматизированных систем синтеза и используем три схемы в системе троирования.

2. Алгоритм получения систем F_1, F_2

Опишем основные шаги алгоритма. Пусть L – допустимое число расщеплений конъюнкций БСДНФ при построении каждой из аппроксимирующих систем. Предварительно упорядочим L простых импликант БСДНФ $G(x_1, \dots, x_n)$ по неубыванию рангов составляющих их конъюнкций и столько же конъюнкций по невозрастанию рангов. Получаем список простых импликант системы, порождающий обе аппроксимирующие системы. Список разделен на две части, в нем сначала перечисляются импликанты с конъюнкциями меньших рангов, а затем – с конъюнкциями больших рангов.

1. Извлекаем из списка простую импликанту БСДНФ.

2. Расщепляем ее по каждой переменной характеристики. Выбираем пары продуктов расщепления с максимальным значением характеристики h_p . Одна и та же простая импликанта БСДНФ в общем случае может порождать несколько пар продуктов расщепления по различным переменным с максимальным значением h_p .

3. Заменяем порождающую расщепление простую импликанту БСДНФ парой продуктов расщепления.

4. Выполняем пункты 1–3 для L расщеплений, чередуя выбор из первой и второй половин списка для каждой из аппроксимирующих систем. В результате получаем системы F_1^*, F_2^* .

Для каждой из полученных систем гарантировано, что все ее конъюнкции – простые импликанты, но не гарантирована безызбыточность по характеристикам конъюнкций. Поэтому необходима проверка на безызбыточность по характеристикам. После устранения лишних функций из характеристик получаем безызбыточные аппроксимирующие системы F_1, F_2 .

Полученные безызбыточные системы ДНФ для аппроксимирующих и рабочей систем функций поступают в качестве задания на синтез в систему САПР и используются для получения аппроксимирующих и рабочих схем в системе троирования.

Проиллюстрируем получение аппроксимирующей системы на простом примере. С этой целью выполним одно расщепление конъюнкции.

Задана система из двух булевых функций списком простых импликант системы в виде перечисления соответствующих конъюнкциям интервалов системы и характеристик. Интервалы пред-

ставлены троичными векторами, а характеристики – перечислением функций системы, для которых интервалы допустимы:

$$\begin{array}{cccccc}
 x_1 & x_2 & x_3 & x_4 & x_5 & \\
 1 & 0 & - & - & - & ; 1 \\
 - & 1 & 1 & 1 & - & ; 1, 2 \\
 - & 0 & - & - & 1 & ; 1 \\
 - & 1 & - & 1 & 0 & ; 1, 2 \\
 - & 0 & - & 0 & 0 & ; 2
 \end{array}$$

Извлекаем из системы ДНФ для двух функций:

$$f_1 = x_1 \bar{x}_2 \vee x_2 x_3 x_4 \vee \bar{x}_2 x_5 \vee x_2 x_4 \bar{x}_5,$$

$$f_2 = \bar{x}_2 \bar{x}_4 \bar{x}_5 \vee x_2 x_4 \bar{x}_5 \vee x_2 x_3 x_4.$$

Эти ДНФ представлены матрицами в коде Грея на рис. 1.

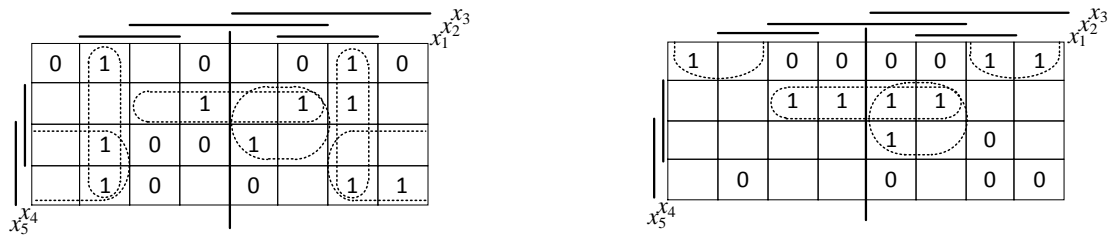


Рис. 1. Матричное представление ДНФ системы
Fig. 1. Matrix presentation of system SoPs

В системе расщепление возможно для двух конъюнкций: $K_1 = x_2 x_3 x_4$ и $K_2 = x_2 x_4 \bar{x}_5$, поскольку их характеристики содержат две функции системы. Для расщепления выбираем конъюнкцию системы $K_1 = x_2 x_3 x_4; 1, 2$. Для этой конъюнкции имеем:

$$x_2, h_p = \{f_1\}, h_{np} = \{f_2\},$$

$$x_3, h_p = \{f_2\}, h_{np} = \{f_1\},$$

$$x_4, h_p = \{\emptyset\}, h_{np} = \{f_1, f_2\}.$$

Рассматриваемая конъюнкция может расщепляться как по переменной x_2 , так и по переменной x_3 . Безразлично, какую переменную выбрать. Выбираем переменную x_2 . В этом случае конъюнкция системы $K_1 = x_2 x_3 x_4; 1, 2$ заменяется на конъюнкцию системы $x_3 x_4; 1$ которая не может быть расширена и, следовательно, является простой импликантой системы. Вторая конъюнкция системы $x_2 x_3 x_4; 2$ может быть расширена по переменной x_3 . В результате имеем простую импликанту $x_2 x_4; 2$. Новая система показана на рис. 2.

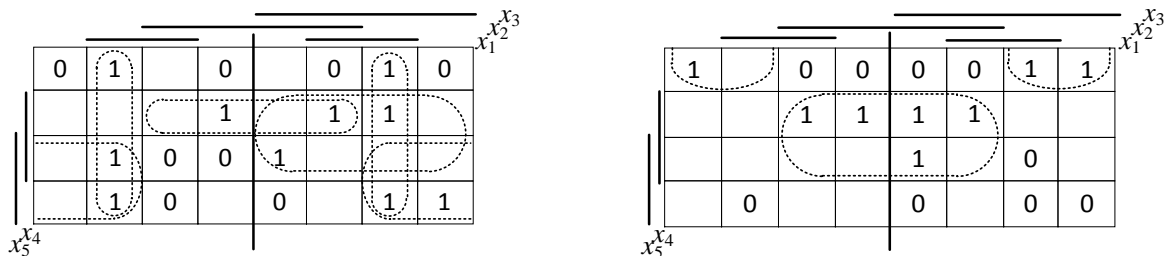


Рис. 2. Матричное представление аппроксимирующей системы
Fig. 2. Matrix presentation of approximate system SoPs

Выполнение операции расщепления требует затем проверки системы на безыбыточность по характеристикам. Так, в полученной системе для конъюнкции $x_2 x_4 \bar{x}_5; 1, 2$ функция f_2 в характеристике является избыточной, и ее следует удалить из характеристики этой конъюнкции. В результате имеем систему

x_1	x_2	x_3	x_4	x_5	
1	0	–	–	–	; 1
–	–	1	1	–	; 1
–	0	–	–	1	; 1
–	1	–	1	0	; 1
–	1	–	1	–	; 2
–	0	–	0	0	; 2

В ней число конъюнкций увеличилось на единицу. Сумма рангов также увеличилась на единицу. В системе отсутствуют общие конъюнкции. Ее можно рассматривать как аппроксимирующую систему. Для нее имеем следующую систему ДНФ:

$$f_1^{\text{анп}} = x_1 \bar{x}_2 \vee x_3 x_4 \vee \bar{x}_2 x_5 \vee x_2 x_4 \bar{x}_5 ,$$

$$f_2^{\text{анп}} = x_2 x_4 \vee \bar{x}_2 \bar{x}_4 \bar{x}_5 ,$$

которая представлена матрицами в коде Грея (см. рис. 2).

Заключение

Предложен подход к построению аппроксимирующих схем на основе предварительного получения аппроксимирующих систем булевых функций для синхронного автомата. Аппроксимирующие системы строятся из безызбыточной системы ДНФ (БСДНФ), являющейся реализацией системы частичных функций синхронного автомата. Полученные системы далее используются в качестве задания на синтез соответствующих схем, применяемых в схеме троирования наряду со схемой, построенной по БСДНФ. Аппроксимирующие системы ориентированы на защиту схемы троирования от внедрения вредоносных подсхем (Trojan Circuits). Этот подход позволяет исключить при троировании незащищенную область, которая обязательно присутствует при использовании известных ранее подходов к синтезу аппроксимирующих комбинационных схем.

Список источников

1. Sanchez Clemente A., Entrena L., Garsea-Valderas M., Lopez-Ongil C. Logic masking for SET mitigation using approximate logic circuits // Proc. of 18th IOLTS. 2012. P. 176–181.
2. Sanchez Clemente A.J. Transient error mitigation by means of approximate logic circuits : Thesis Doctoral / Universidad Carlos III de Madrid. Madrid, 2017.
3. Chaudhury M.R., Mohandram K.. Approximate logic circuits for low overhead non-intrusive concurrent error detection // Proceedings of DATE. 2008. P. 902–908.
4. Kohavi I., Kohavi Z. Detection of multiple faults in combinational logic networks // IEEE Transactions on Computers. 1972. V. C-21, № 6. P. 556–558.
5. Матросова А.Ю. Построение полного теста для схем, синтезированных методом факторизации // Автоматика и вычислительная техника. 1978. № 5. С. 42–45.

References

1. Sanchez Clemente, A., Entrena, L., Garsea Valderas, M. & Lopez-Ongil, C. (2012) Logic Masking for SET Mitigation Using Approximate Logic Circuits. *Proc. of 18th IOLTS*. pp. 176–181.
2. Sanchez Clemente, A.J. (2017) *Transient Error Mitigation by Means of Approximate Logic Circuits. Doctoral Thesis*. Universidad Carlos III de Madrid.
3. Chaudhury, M.R., Mohandram, K. (2008) Approximate Logic Circuits for Low Overhead Non-Intrusive Concurrent Error Detection. *Proc. of DATE*. pp. 902–908.
4. Kohavi, I. & Kohavi, Z. (1972) Detection of Multiple Faults in Combinational Logic Networks. *IEEE Transactions on Computers*. C-21(6). pp. 556–558.
5. Matrosova, A.Yu. (1978) Construction of a complete test for circuits synthesized by the factorization method. *Avtomatika i vychislitel'naya tekhnika – Automatic Control and Computer Sciences*. 5. pp. 42–45.

Информация об авторах:

Останин Сергей Александрович – доцент, кандидат технических наук, заведующий кафедрой компьютерной безопасности Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: sergeiostanin@yandex.ru

Матросова Анжела Юрьевна – профессор, доктор технических наук, профессор кафедры компьютерной безопасности Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: mau11@yandex.ru

Андреева Валентина Валерьевна – доцент, кандидат технических наук, доцент кафедры компьютерной безопасности Института прикладной математики и компьютерных наук Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: avv.21@mail.ru

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

Information about the authors:

Ostanin Sergey O. (Candidate of Technical Sciences, Associate Professor, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: sergeiostanin@yandex.ru

Matrosova Angela Yu. (Doctor of Technical Sciences, Professor, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: mau11@yandex.ru

Andreeva Valentina V. (Candidate of Technical Sciences, Associate Professor, National Research Tomsk State University, Tomsk, Russian Federation). E-mail: avv.21@mail.ru

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

Received 29.09.2022; accepted for publication 01.03.2023

Поступила в редакцию 29.09.2022; принята к публикации 01.03.2023

Научный журнал

**ВЕСТНИК
ТОМСКОГО
ГОСУДАРСТВЕННОГО
УНИВЕРСИТЕТА**

**УПРАВЛЕНИЕ,
ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА
И ИНФОРМАТИКА**

**TOMSK STATE UNIVERSITY
JOURNAL OF CONTROL AND COMPUTER SCIENCE**

2023. № 62

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 31.03.2023 г. Формат 60x84¹/₈.
Гарнитура Times. Усл. печ. л. 15,4.
Тираж 250 экз. Заказ № 5384. Цена свободная.

Дата выхода в свет 11.04.2023 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