

ISSN 2313-7304

ПРОЦЕССЫ УПРАВЛЕНИЯ И УСТОЙЧИВОСТЬ

ТОМ 11 (27)

**ТРУДЫ LV МЕЖДУНАРОДНОЙ НАУЧНОЙ
КОНФЕРЕНЦИИ АСПИРАНТОВ И СТУДЕНТОВ**

Санкт-Петербург

1 – 5 апреля 2024 года

ISSN 2313-7304

***CONTROL PROCESSES
AND STABILITY***

Vol. 11 (27)

**PROCEEDINGS OF THE LV
INTERNATIONAL CONFERENCE**

**St. Petersburg
April 1 – 5, 2024**

PUBLISHING HOUSE
FEDOROVA G.V.
2024

ББК 22.1

П84

Редакционная коллегия:

д-р физ.-мат. наук, проф. Л. А. Петросян (отв. редактор) (СПбГУ),
д-р физ.-мат. наук, проф. Н. В. Смирнов (зам. отв. редактора) (СПбГУ)
д-р физ.-мат. наук, проф. О. И. Дривотин (СПбГУ)
д-р физ.-мат. наук, проф. А. П. Жабко (СПбГУ)
д-р физ.-мат. наук, проф. Е. Д. Котина (СПбГУ)
д-р физ.-мат. наук, проф. А. Ю. Крылатов (СПбГУ)

Процессы управления и устойчивость. Том 11(27). № 1 / науч.
П84 редакторы тома Н. В. Смирнов, А. С. Еремин, Т. Е. Смирнова.
СПб.: Издательский Дом Федоровой Г.В., 2024. 372 с.
ISSN 2313-7304

«Процессы управления и устойчивость» – ежегодное научное издание. Издается на факультете прикладной математики – процессов управления СПбГУ с 1998 года. Том журнала традиционно формируется по результатам работы ежегодной Международной научной конференции аспирантов и студентов «Процессы управления и устойчивость» и включает статьи студентов, аспирантов, сотрудников СПбГУ и других высших учебных заведений, в том числе зарубежных, по математической теории процессов управления, математическим методам в механике и физике, математическому моделированию в медико-биологических системах, информационным и компьютерным технологиям, теории управления социально-экономическими системами.

В Том 11(27) включены статьи участников 55-й Международной научной конференции аспирантов и студентов «Процессы управления и устойчивость» (1–5 апреля 2024 года), прошедшие рецензирование и рекомендованные к публикации. Сборник предназначен для студентов старших курсов физико-математических факультетов университетов, аспирантов и научных работников.

ББК 22.1

*Журнал издается при финансовой поддержке
факультета прикладной математики – процессов управления
Санкт-Петербургского государственного университета*

ISSN 2313-7304

© Издательский Дом Федоровой Г. В., 2024
© Смирнов Н. В., 2024

*300-летию Санкт-Петербургского
государственного университета,
55-летию факультета ПМ–ПУ
посвящается!*

Содержание

Пленарная сессия	7
Приглашенные лекторы	7
<i>Костромин С. А.</i> Системы управления большими физическими установками	7
<i>Крылатов А. Ю.</i> Оптимизационные модели и методы равновесного распределения потоков в транспортных сетях.....	9
1. Математическая теория процессов управления	35
<i>Андрянова Н. Р.</i> Сверхфинитная устойчивость мультиагентной системы с переключениями	35
<i>Бабанов И. В.</i> Оптимальное управление трехмассовой колебательной системой	42
<i>Гапонов Д. Н.</i> Устойчивость линейных систем с запаздыванием. Подход вычислительной алгебры	48
<i>Каминцева В. С.</i> Реализация метода синтеза оптимального управления для динамических объектов.....	56
<i>Маковеева Р. Е., Nesterov S. D.</i> Lyapunov – Krasovskii functionals for a linear system with distributed parameters	63
2. Математические методы в механике и физике	71
<i>Беляева А. А., Князев Н. А.</i> Влияние модификаций вычислительных моделей постоянных дипольных магнитов источника синхротронного излучения на оптические характеристики пучка заряженных частиц	71
<i>Вековцев В. В., Иордан А. А., Никифоров К. А.</i> Разработка цифровой модели автоэмиссионной системы.....	81

<i>Вяльцева В. В., Пронина Ю. Г.</i> Об оценке трехмерной плотности параллельных трещин по их следам на поверхности	88
<i>Грошева А. Д., Знаменщикова Н. С.</i> Оценка возможности разработки трёхмерных вычислительных моделей прецизионных магнитных систем, состоящих из соленоидов и ферромагнитных экранов, с использованием интегральных и дифференциальных подходов	93
<i>Казакова А.</i> Минимальная дисперсия для экспоненциального распределения	99
<i>Кикенов А. А.</i> Математическая модель конвертоплана	107
<i>Кондратьева А. Д.</i> Выпучивание под собственным весом нанокантилевера с упругим поворотом	112
<i>Набиев Т. К., Шиманчук Д. В.</i> Расчёт траекторий движения группы БПЛА в строю	121
<i>Belenok A. I.</i> Ritz method for building parametric representation of cost optimal trajectory	129
<i>Mullayanov T. O.</i> Investigating elastic response of nanostructured solid surface using boundary perturbation method	134
<i>Xing J.</i> MS-stability comparison of Runge–Kutta–Chebyshev methods application to SDEs	139

3. Математические модели медико-биологических систем 145

<i>Давыденко Б. Н.</i> Прогнозирование результата коррекции гиперметропии с использованием CHAID анализа	145
<i>Легкова И. А.</i> Обнаружение отклонений и нарушений сердечного ритма в данных ЭКГ пациентов	150
<i>Халтобина М. М., Елизарова М. В.</i> Анализ выживаемости клеток в рамках RMR-модели	155
<i>Хоменко И. Е.</i> Преобразование количественных данных в бинарные с помощью адаптивной двусторонней оценки для построения модели выживаемости пациентов	163
<i>Valieva N. L., Goncharova A. B.</i> Breast cancer classification in mammograms	168

4. Информационные и компьютерные технологии 173

<i>Антропова Е. Г.</i> Решение оптимизационных задач при помощи нейронных сетей	173
<i>Бабкина А. В.</i> Очистка медицинских ультразвуковых изображений от спекл-шума	179
<i>Горбунов В. И., Салимов Т. А.</i> Проблемы практического применения результатов Data Science и подходы к их решению . . .	187
<i>Гусева М. Ю.</i> Расчет аналитического решения трехмерной системы обыкновенных дифференциальных уравнений с релейным гистерезисом и возмущением	195
<i>Ковалев А. Ю., Калинин П. Е.</i> Альтернативный алгоритм построения шума Перлина	200
<i>Колесов В. В.</i> Алгоритмы для классификации жанров фильма	205
<i>Кужло Р. А.</i> Применение метода кластеризации по времени и географическим данным для уменьшения объема передаваемых данных	212
<i>Мосина М. А.</i> Исследование методов распределения нагрузки на основе муравьиного алгоритма	218
<i>Набиев Т. Р.</i> Применение методов компьютерного зрения для обнаружения и отслеживания дронов на видео	226
<i>Неронов Р. М.</i> Решение задачи разведочного анализа текстов через кластеризацию и аппроксимацию числа кластеров	231
<i>Никитин А. В.</i> Применение нейросетевых методов для сегментации болезней культурных растений по цветным снимкам в точном земледелии	236
<i>Павлов А. В.</i> Методы точной настройки: RAG vs LoRA	241
<i>Павлов М. В.</i> Нейросетевые методы тестирования игрового процесса на основе обучения с подкреплением в разработке компьютерных игр	247
<i>Пономарев Ю. Ю.</i> Вопросы точных оценок характеристик деревьев при моделировании метаданных в информационных системах общего назначения	253
<i>Потешкин Е. П.</i> Выбор параметров в методе Монте-Карло SSA	262

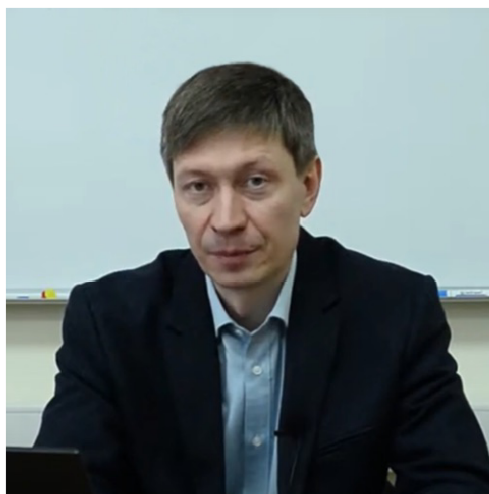
<i>Рудович К. В.</i> Библиотека для создания и имплементации генетических алгоритмов на языке C++	270
<i>Хромов Н. А., Голяндина Н. Э.</i> High-order MSSA для выделения сигнала	278
<i>Чайкин Г. А.</i> Разработка инструментария для трекинга экспериментов для задачи отслеживания знаний при непрерывном обучении	287
<i>Huang L.</i> Deep residual neural network based on feature extraction for ultrasound image classification of focal liver lesions	293
5. Управление социально-экономическими системами..	299
<i>Аникушин В. А., Редько А. Ю.</i> Применение модели ARIMA для прогнозирования численности населения Земли	299
<i>Бушужева А. С., Власова Т. В.</i> Применение эвристических алгоритмов для решения задачи маршрутизации с несколькими складами	307
<i>Калинин П. Е.</i> Модель кратного тестирования в эксперименте с бинарным исходом	312
<i>Карпычев А. В.</i> Исследование возможности моделирования спроса со стороны потребителя электроэнергии в электрической сети	317
<i>Король М. М.</i> Алгоритмы равновесного распределения потоков на улично-дорожной сети города	323
<i>Новгородцев В. А.</i> О задаче управления портфелем в случае модели Хестона	330
<i>Русакова К. В.</i> Сравнение стратегий с короткой памятью в игре «дилемма заключенного»	339
<i>Vukova M. S.</i> Payment schemes and their properties for cooperative stochastic games	345
<i>Lorits E. M.</i> Evolutionary model with environmental feedback and agent opinions	355
<i>Wang D., Guo Z.</i> Single-factor multi-stage production control model based on linear optimal control	361
Contents	366
Именной указатель	370

*Пленарная
сессия*



Костромин Сергей Александрович

*доктор физико-математических наук, профессор,
кафедра информационных и ядерных технологий,
факультет прикладной математики – процессов управления,
Санкт-Петербургский государственный университет*



Системы управления большими физическими установками

Аннотация. Ускорители заряженных частиц – инструмент исследований фундаментального характера. Системы управления современными крупными комплексами на базе ускорителей или коллайдеров, оперируют несколькими тысячами параметров одновременно. Таким комплексам необходимы программные средства для моделирования физических процессов, распределенные системы хранения и обработки больших объемов данных, а также системы управления с элементами искусственного интеллекта. Возможные пути решения этих задач иллюстрируются на примере комплекса NICA, создаваемого в Объединенном институте ядерных исследований.

Kostromin Sergey Aleksandrovich

*D.Sc., Professor,
Department of Information and Nuclear Technologies,
Faculty of Applied Mathematics and Control Processes,
Saint Petersburg State University*

Control processes in fundamental science

Abstract. Particles accelerator – facility of the fundamental researches. Control systems of modern particles accelerator or colliders control and operate several thousand of different parameters the same time. Each such facilities requires computer tools for simulation of physical processes, systems for distributed storage and processing of big amount of data, modern control systems with elements of artificial intelligence. Possible ways for solution of these tasks are illustrated on example of the NICA project, which is under realization in the Joint institute for nuclear research.

Крылатов Александр Юрьевич

*доктор физико-математических наук, профессор,
кафедра математической теории экономических решений,
факультет прикладной математики – процессов управления,
Санкт-Петербургский государственный университет*



Оптимизационные модели и методы равновесного распределения потоков в транспортных сетях

Аннотация. Оценку потоковой загрузки дуг транспортной сети можно осуществлять путём решения специальных задач условной нелинейной оптимизации. Удалось доказать, что для данного типа задач может быть получен проектирующий оператор в явном виде. В результате, поиск равновесного распределения потоков можно свести к запуску процедуры простой итерации с гарантированно сжимающим оператором, обеспечивающим квадратичную сходимость полученного итерационного процесса. Примечательно, что с математической точки зрения полученный оператор обобщает большинство ранее предложенных методов решения задачи, основанных на идее уравнивания времени движения по маршрутам. Разрабатываемые на базе полученного оператора алгоритмы находят шаблоны равновесного распределения потоков в транспортных сетях с высочайшим уровнем точности.

Krylatov Alexander Yurevich

*D.Sc., Professor,
Department of Mathematical Theory of Economic Decisions,
Faculty of Applied Mathematics and Control Processes,
Saint Petersburg State University*

**Optimization models and methods
for equilibrium flow assignment
in transportation networks**

Abstract. Congestion estimation of the transportation network arcs can be carried out by solving a special constrained nonlinear optimization problem. Recently, we were lucky to prove that for such kinds of problems, a projection operator can be obtained in explicit form. As a result, the equilibrium flow assignment search can be reduced to a fixed-point iteration with the contraction operator at quadratic convergence. Remarkable that from mathematical perspectives, the developed operator generalizes most path-equilibration operators already exploited by researchers. The obtained operator allows one to develop algorithms for equilibrium flow assignment search at the highest level of accuracy.

УДК 519.8

Крылатов А. Ю.

Оптимизационные модели и методы равновесного распределения потоков в транспортных сетях

Введение. Сегодня интеллектуальные системы позволяют поддерживать процессы принятия решений в различных сферах человеческой деятельности. В частности, эффективное управление современными крупными транспортными сетями представляется невозможным без соответствующих математических моделей распределения потоков. При этом ключевой задачей, решению которой способствует развитие моделей равновесного распределения потоков, является задача оптимизации топологии транспортной сети. В общем виде задача оптимизации топологии сети формулируется как задача двухуровневой оптимизации с потенциальной игрой участников движения на нижнем уровне. Решение подобной задачи требует создания специальных методов и подходов. В настоящей статье собраны ключевые математические постановки, возникающие вокруг проблемы оптимизации сетевых транспортных структур, а также обзор научной литературы, посвященной исследованиям соответствующих моделей и методов.

1. Оптимизация топологии сети. Задача оптимизации топологии сети (Network design problem, NDP) формулируется в виде игры Штакельберга на сети, представленной ориентированным графом $G = (V, E)$. Как правило, лидер (лицо, принимающее решения на верхнем уровне) обладает бюджетом B и вкладывает его в развитие топологии транспортной сети, в то время как участники движения нижнего уровня реагируют на изменения топологии, минимизируя индивидуальное время движения в этой сети [1]. В частности, лидер может вкладываться в увеличение значения пропускной способности c_e любой дуги e , $e \in E$, в заданных пределах $l_e > 0$ и $u_e > 0$, т. е. $l_e \leq c_e \leq u_e$ для любого $e \in E$, $c = (c_1, \dots, c_n)^T$. В силу того, что любое увеличение пропускной способности приводит к бюджетным затратам, вводятся функции стоимости влияния на пропускные

Крылатов Александр Юрьевич – профессор, Санкт-Петербургский государственный университет; e-mail: a.krylatov@spbu.ru, тел.: +7(812)428-71-59

Работа выполнена при поддержке Российского научного фонда, проект № 22-71-10063

способности в виде дифференцируемых, строго возрастающих неотрицательных на множестве действительных неотрицательных чисел функций $g_e(c_e)$ таких, что $g_e(l_e) = 0$ для всех $e \in E$. В таком случае область переменных решения верхнего уровня выглядит следующим образом:

$$C = \left\{ c \in \mathbb{R}^n \mid \sum_{e \in E} g_e(c_e) \leq B, \quad l_e \leq c_e \leq u_e, \quad \forall e \in E \right\}.$$

Как только лидер устанавливает пропускные способности дуг сети $c \in C$, участники движения начинают конкурировать за наименьшее время на перемещение между узлами отправления-прибытия. Действительно, каждой упорядоченной паре узлов w , $w \in V \times V$, можно поставить в соответствие величину спроса на перемещение $F^w \geq 0$ между узлами пары и набор возможных маршрутов R^w :

$$F^w = \sum_{r \in R^w} f_r^w \quad \forall w \in W, \quad (1)$$

где $W = \{w \in V \times V \mid F^w > 0\}$ и

$$f_r^w \geq 0 \quad \forall r \in R^w, w \in W, \quad (2)$$

являются переменными потока по маршрутам, в то время как величина потока $x_e \geq 0$ по дуге e представляет собой сумму потоков по всем маршрутам, проходящим через эту дугу:

$$x_e = \sum_{w \in W} \sum_{r \in R^w} f_r^w \delta_{e,r}^w \quad \forall e \in E, \quad (3)$$

где $\delta_{e,r}^w = 1$, если маршрут $r \in R^w$ включает дугу e , и $\delta_{e,r}^w = 0$ в противном случае; $x = (x_1, \dots, x_n)^T$. Таким образом, допустимые величины загрузки дуг определяются следующим образом:

$$X = \{x \in \mathbb{R}^n \mid x \text{ удовлетворяет (3) при (1), (2)}\}.$$

Более того, каждая дуга $e \in E$ имеет свою функцию задержки $t_e(x_e, c_e)$, строго возрастающую по x_e , а значит участники движения сталкиваются с конкурентной маршрутизацией (поскольку время в пути независимого пользователя сети зависит от стратегий всех остальных пользователей).

В случае дифференцируемых, строго возрастающих, неотрицательных функций задержки по переменным потока на множестве неотрицательных действительных чисел, значения потоковых нагрузок, возникающих в результате конкуренции неатомарных участников движения (которые стремятся минимизировать индивидуальные задержки при выборе своих маршруты), можно найти как единственное решение следующей задачи оптимизации [2]:

$$\hat{x} = \arg \min_{x \in X} \sum_{e \in E} \int_0^{x_e} t_e(u, c_e) du.$$

Более того, если $t_e(x_e, c_e)$, $e \in E$, является непрерывно-дифференцируемой, строго убывающей и положительной по c_e , то лидер при оптимизации сети может решать двухуровневую задачу [3]:

$$\min_{c \in C} \sum_{e \in E} t_e(x_e, c_e) x_e + \beta \sum_{e \in E} g_e(c_e), \quad (4)$$

при

$$x = \arg \min_{x \in X} \sum_{e \in E} \int_0^{x_e} t_e(u, c_e) du. \quad (5)$$

где $\beta \geq 0$ задает вес инвестиционных затрат в целевой функции верхнего уровня. Предложенная постановка задает ситуацию, при которой лидер стремится минимизировать общее время движения в сети за счет минимальных инвестиций в ее пропускную способность с учетом конкурентного поведения участников движения.

Формулировка (4), (5) является одной из самых первых формулировок задачи оптимизация топологии сети, поскольку она проистекает из важных практических причин [3]. Прежде всего, нижний уровень оптимизации адекватно моделирует некооперативное поведение участников движения при выборе маршрутов в реальных улично-дорожных сетях. Потенциальная игра (5) моделирует поведение участников движения, каждый из которых стремится минимизировать личное время в пути от пункта отправления до пункта прибытия [2, 4]. Во-вторых, менеджеры сети не могут напрямую влиять на поведение водителей, но могут влиять на топологию сети посредством установки режимов работы сигналов светофоров, управления парковками или пропускной способностью участков сети [1]. Несомненно, изменения в топологии сети корректируют выбор маршрутов и влияют на задержки в пути. Таким образом, лицо, принимающее решения, может управлять распределением потоков в сети,

воздействуя на пропускные способности ее дуг в рамках доступного бюджета (4). Отметим, что структура задачи (4), (5) близка хорошо известным задачам запрета (interdiction) [5]. Однако между этими двумя видами задач существует принципиальное различие. Действительно, лидер в задаче (4), (5) стремится минимизировать общее время движения в сети за счет увеличения пропускной способности, не выходя за доступный бюджет. Напротив, лидер в задачах запрета, как правило, максимизирует свою утилитарную функцию за счет введения запретов в рамках доступного бюджета. Таким образом, задача оптимизации топологии сети (4), (5) образует важный независимый класс задач, которые представляют собой наиболее вычислительно затратные задачи в области транспортного планирования [6, 7].

Благодаря практической значимости задачи оптимизации топологии сети (4), (5), многие исследователи пытались разработать эффективные вычислительные методы для ее решения. Подавляющее большинство разработанных методик, доступных сегодня в литературе, используют идею приближенного выражения решения нижнего уровня x как вектор-функцию от переменных решения верхнего уровня c . Другими словами, исследователи исходят из предположения о существовании некоторой непрерывной и дифференцируемой вектор-функции $\phi(c)$ такой, что $x = \phi(c)$, где $x \in X$ — решение (5) для $c \in C$. Несомненно, такое предположение упрощает задачу (4), (5), сводя ее к одноуровневой задаче оптимизации. Однако вид функции $\phi(c)$ неизвестен, как и ее свойства [8, 9]. Следовательно, используя различные аппроксимации $\phi(c)$, исследователи получают возможные решения, но не могут сказать, приводят ли они к глобальному оптимуму или нет. Действительно, поскольку $\phi(c)$ может быть невыпуклой, то и целевая функция (4) может быть невыпуклой. Тем не менее, самые первые алгоритмы продемонстрировали хорошую сходимость к локальному оптимуму, но только на небольших сетях [8, 10, 11].

Некоторые исследователи использовали приближение первого порядка для $\phi(c)$: $\phi(c) = \phi(c^0) + \nabla\phi(c^0)(c - c^0)$, где $\nabla\phi(c^0)$ получено в результате анализа чувствительности равновесных сетевых потоков [12, 13]. Такое приближение плодотворным образом сказалось на развитии различных проекционных методов решения задачи [14, 15]. Действительно, аппроксимация первого порядка позволяет эффективно идентифицировать точки локального оптимума с помощью применения градиентных методов [16]. Эти подходы демонстриру-

ют хорошую сходимость даже в сети среднего размера, такой как в Су-Фолс (Sioux Falls – город в США, улично-дорожная сеть которого используется для тестирования и апробации алгоритмов), но не могут достичь глобального оптимума. Как предполагают исследователи, множественные локальные оптимумы возникают из-за невыпуклости задачи (4), (5), и каждый новый запуск метода приводит к разным решениям [11,14]. Тем не менее, недавние результаты работы на небольшой тестовой сети показали, что локальные оптимумы, полученные методами проекции, довольно близки к глобальному [17]. К сожалению, тест проводился только на небольшой сети, поскольку разработанная методика глобальной оптимизации, хотя и представляется теоретически привлекательной, но имеет недостатки в части вычислительных затрат.

Таким образом, до сих пор существуют пробелы в изучении задачи оптимизации топологии сети (4), (5). Недавно авторам статьи удалось получить некоторое представление функции типа $\phi(c)$ в явном виде для случая сети с одной парой отправления-прибытия и непересекающимися маршрутами. Полученное в явном виде представление функции типа $\phi(c)$ позволило найти условия оптимальности для задачи (4), (5) в описанном выше случае сети [18]. Подобные результаты кажутся интригующими и могут быть развиты на более общие случаи сетей. При этом важно отметить, что полученные результаты во многом связаны с внимательным изучением задачи нижнего уровня (5), являющейся важной частью исследуемой двухуровневой задачи. В этом смысле, исследователям, заинтересованным в решении задачи оптимизации топологии сети, следует обратить особое внимание на задачу равновесного распределения потоков.

2. Равновесное распределение потоков в сети. Задача распределения потоков (Traffic Assignment Problem, TAP) – хорошо известная задача, методы решения которой широко используются исследователями и транспортными инженерами для оценки или прогнозирования заторов на городских дорогах. TAP основана на так называемом поведенческом принципе равновесия пользователей (UE), сформулированном Дж. Г. Вардропом следующим образом: *Время в пути на всех фактически используемых маршрутах одинаково и меньше того, которое затратило бы одно транспортное средство на любом неиспользуемом маршруте* [19]. Первая математическая формулировка TAP была дана Бекманом *et al* [20]. Они предположили, что функции времени в пути аддитивны по дугам (сепарабельные). Такой подход оказался плодотворным, поскольку первый

эффективный алгоритм решения задачи, независимо полученный ЛеБланом и Нгуеном, оказался основанным на варьировании переменных потоков на дугах [21, 22]. При этом, реализовав алгоритм Франк – Вулфа [23] для оценки потоков на дугах, ЛеБлан и Нгуен внесли значительный вклад в соответствующую область исследований. Они предложили искать оптимальное решение не во всем пространстве переменных, а лишь в подпространстве из «наиболее подходящих». В настоящее время хранение данных о переменных, касающихся ограниченного набора маршрутов и дуг (restricted set), стало наиболее распространенной (и даже центральной) практикой среди разработчиков алгоритмов в этой области.

Сегодня исследователи выделяют три основных типа алгоритмов решения ТАР в зависимости от переменных решения:

- дуговые алгоритмы (arc-based algorithms);
- маршрутные алгоритмы (path-based algorithms);
- кустовые алгоритмы (bush-based algorithms).

Комплексное эмпирическое исследование, сравнивающее алгоритмы распределения потоков из упомянутых выше групп, было проведено в работе [24]. Авторы пришли к выводу, что при выборе подходящего алгоритма, достигаемый уровень точности должен быть одним из наиболее важных факторов, которые следует учитывать. Самый высокий уровень точности продемонстрировали кустовые подходы, такие как алгоритм из [25, 26]. Методы, основанные на уравновешивании времени движения по маршрутам, и различные усовершенствованные алгоритмы, основанные на дугах, характеризовались средним уровнем точности за разумное время. Низкий уровень точности был продемонстрирован простыми алгоритмами на основе дуг, использующими базовую вычислительную технику Франк – Вулфа [2, 27]. Однако недавнее вычислительное исследование показало, что алгоритмы, основанные на методах уравновешивания времени движения по маршрутам, могут продемонстрировать высочайший уровень точности за разумное время [28].

Стоит отметить, что алгоритмы на основе дуг были самым популярным инструментом для решения ТАР в течение длительного периода из-за их низких требований к процессору и оперативной памяти. В настоящее время экономия вычислительных ресурсов и памяти по-прежнему имеет значение при разработке алгоритма, но она

уже не является принципиальной, поскольку современные вычислительные платформы чрезвычайно мощны и обладают огромным объемом оперативной памяти. Сегодня на первый план выходит вопрос точности, как основной с практической точки зрения [29]. Таким образом, популярные ранее дуговые алгоритмы теперь кажутся абсолютно устаревшими, поскольку не могут соответствовать современным стандартам точности. В свою очередь, алгоритмы на основе кустов и путей, несмотря на более высокие требования к процессору и оперативной памяти, заняли лидирующие позиции среди всех доступных подходов. Недавно были получены новые результаты в области разработки алгоритмов равновесного распределения потоков при помощи уравнивающего оператора с высочайшей точностью [30].

Первый подход к решению ТАР с использованием потоков по маршрутам в качестве переменных решения был предложен в работе [31]. В основу подхода легла идея переноса части потока между доступными маршрутами с помощью так называемого *оператора равновесия*. Направление спуска определялось путем смещения потока с маршрута с большим временем движения на маршрут с меньшим временем движения, что, как было доказано, снижает целевую функцию ТАР. Однако в явном виде были получены только два оператора, подходящие для работы с частными случаями ТАР. Тем не менее, данные теоретические основы существенно повлияли на развитие алгоритмов на основе уравнивания времени движения по используемым маршрутам. Еще один важный вклад в разработку алгоритмов решения ТАР сделан в работе [32]. Подход генерации столбцов был реализован, чтобы избежать априорного создания всех маршрутов, соединяющих каждую пару отправления-прибытия (OD), и, таким образом, работать только с ограниченным набором возможных маршрутов. Этот подход оказался решающим при работе с крупными сетями.

Альтернативный подход к решению задачи ТАР с точки зрения маршрутных потоков без использования процедуры балансировки пути предложен в [33]. Алгоритм симплицальной декомпозиции оказался подходящим для решения ТАР, дезагрегированного по парам OD, поскольку ограничения на маршрутные потоки, связанные с любой отдельной парой OD, можно было легко представить в виде стандартного симплекса. Техника уравнивания пути, основанная на проекции градиента, для работы со случаем сепарабельных функций времени движения для произвольных дорожных сетей

предложена в работах [34–37]. Дополнительный взгляд на процесс уравнивания времени движения на маршрутах как на некоторый процесс возмущения, вытекающий из концепции социального давления, был предложен в [38,39]. Концепция социального давления интуитивно основывалась на следующей идее: те водители, которые ездят по относительно дорогим маршрутам, более склонны менять маршруты, чем те, кто ездит по недорогим маршрутам. Вычислительное исследование подходов, основанных на реализации метода симплицальной декомпозиции для поиска распределения потоков по маршрутам и дугам, и подходов, основанных на уравнивании пути с помощью градиентной проекции, предложено в [40]. Новый взгляд на разработку методов оптимизации для решения ТАР дан в работах [28, 41, 42]. Эти исследования в конечном итоге показали, что маршрутные алгоритмы способны достигать современного уровня точности за разумное время. Усиление такого понимания возникло после получения в явном виде сжимающего оператора для равновесного распределения потоков по маршрутам [30].

Распределение потока по дугам x и распределение потока по маршрутам f , разрешающие (5), удовлетворяют следующим условиям:

$$t_r^w(f) = \sum_{e \in E} t_e(x_e) \delta_{e,r}^w \begin{cases} = t^w, & \text{если } f_r^w > 0, \\ \geq t^w, & \text{если } f_r^w = 0, \end{cases} \quad \forall r \in R^w, w \in W, \quad (6)$$

где t^w называется *равновесным временем движения* или временем движения по используемым маршрутам между парой $w \in W$, $\mathbf{t} = (t^1, \dots, t^m)^T$ [2, 43]. Опираясь на (6), можно предположить, что уравнивание времен движения на альтернативных маршрутах ведет к получению равновесного распределения потоков. Действительно, было доказано, что перемещение части потока с маршрута с большим временем движения на маршрут с меньшим временем движения ведет к уменьшению значения целевой функции в задаче (5) [44]. Таким образом, последовательное уравнивание времен движения по маршрутам между всеми парами отправления-прибытия представляет собой естественный метод решения задачи распределения потоков. Общее описание подобного алгоритма представлено в виде псевдокода (см. ниже, процедура уравнивания обозначена как \mathfrak{E}) [30]. Приведенный алгоритм сходится к равновесному распределению потоков по маршрутам, но единственность такого распределения не гарантирована. Действительно, хотя равновесное распределение потоков по дугам является единственным

при строго возрастающих функциях $t_e(x_e)$, $e \in E$, соответствующее распределение потоков по маршрутам, в общем случае, может быть получено не единственным образом.

```

1:  $f \leftarrow \mathbb{0}$  &  $\widehat{R} \leftarrow \emptyset$ 
2: for  $w = 1, \dots, |W|$  do
3:    $p \leftarrow$  кратчайший маршрут между парой  $w$ 
4:    $\widehat{R} \leftarrow \widehat{R} \cup p$ 
5: end for
6:  $f \leftarrow \mathfrak{E}(f, \widehat{R})$ 
7: while  $\exists p \in R^w \setminus \widehat{R}^w$ ,  $w \in W : t_p^w(f) < t_r^w(f) \forall r \in \widehat{R}^w$  do
8:   for  $w = 1, \dots, |W|$  do
9:      $p \leftarrow$  кратчайший маршрут между парой  $w$  при потоках  $f$ 
10:     $\widehat{R} \leftarrow \widehat{R} \cup p$ 
11:   end for
12:    $f \leftarrow \mathfrak{E}(f, \widehat{R})$ 
13:   for  $w = 1, \dots, |W|$  do
14:     if  $\exists r \in \widehat{R}^w : f_r^w < 0$  then
15:        $\widehat{R} \leftarrow \widehat{R} \setminus r$ 
16:     end if
17:   end for
18:    $f \leftarrow \mathfrak{E}(f, \widehat{R})$ 
19: end while
20: return  $f$  &  $\widehat{R}$ 

```

Условия единственности решения задачи равновесного распределения потоков в терминах дуговых потоков хорошо исследованы [2, 43, 45]. Между тем, условия единственности равновесного распределения потоков в терминах маршрутных потоков, согласно литературе, в основном исследовались вычислительно [46, 47] или на примере энтропийных моделей [48]. Определенные аналитические исследования представлены в работе [49]. Влияние зависимости/независимости маршрутов на единственность равновесного распределения потоков по маршрутам сети изучено в [30].

Рассмотрим понятие зависимости маршрутов на примере сети, изображенной на рис. 1. Сеть имеет пункт отправления 1 и пункт прибытия 7, матрица инцидентности между дугами и маршрутами дана в таблице 1. Видно, что четыре маршрута, соединяющие узлы

1 и 7, являются зависимыми, т. е. выразимы друг через друга, в то время как любые три из этих маршрутов являются независимыми (не выразимы друг через друга).

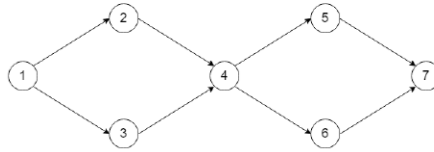


Рис. 1. Сеть с одной парой отправление-прибытие

Понимаемая таким образом зависимость маршрутов оказалась важной для разработки алгоритмов решения задачи распределения потоков на основе оператора уравнивания времен движения по используемым маршрутам.

Таблица 1. Матрица инцидентности

Маршрут	(1, 2)	(2, 4)	(4, 5)	(5, 7)	(1, 3)	(3, 4)	(4, 6)	(6, 7)
1	1	1	1	1	0	0	0	0
2	0	0	0	0	1	1	1	1
3	0	0	1	1	1	1	0	0
4	1	1	0	0	0	0	1	1

Было показано, что если все используемые маршруты независимы, то существует единственное равновесное распределение потоков по маршрутам сети между парой узлов отправления-прибытия.

В современных крупных городах множество маршрутов движения между парой узлов отправления-прибытия избыточно. При решении задачи распределения потоков и задачи оптимизации топологии сети важно работать с ограниченным множеством независимых маршрутов, востребованных среди участников движения. Согласно модели Вардропа, такое множество состоит из маршрутов, время движения по которым одинаково и меньше времени движения по всем прочим маршрутам, между выбранной парой узлов отправления-прибытия. Более того, выявление подобного множества маршрутов является значимым для решения задачи, обратной к задаче равновесного распределения потоков, которую принято называть задачей оценки спроса на перемещение.

3. Оценка спроса на перемещение. В современных крупных городах наблюдается возрастающая динамика автомобилизации, что

приводит к различным негативным последствиям, таким как пробки, аварии, снижение средней скорости, шум, нехватка парковочных мест, неудобства для пешеходов, загрязнение и ущерб окружающей среде. Эффективное управление дорожным движением оказывается единственным способом решения этих проблем, поскольку пропускная способность реальных дорожных сетей сегодня часто близка к пределу и уже не может быть увеличена. Однако реальные дорожные сети чрезвычайно огромны, и транспортным инженерам приходится использовать интеллектуальные системы для поддержки процесса принятия решений при планировании или управлении транспортными процессами. Надежность таких решений во многом зависит от точности информации о спросе на поездки [50]. С другой стороны, оценка спроса на поездки сама по себе является весьма сложной задачей. Она исследовалась в течение последних 40 лет, но все еще существуют практические и теоретические пробелы, которые необходимо заполнить [51].

Исследователи изучают задачу оценки спроса на перемещение с помощью различных подходов и методов. С математической точки зрения самый простой подход может относиться к гравитационным методам, которые хорошо описаны в контексте пространственного регионального анализа [52]. Другой большой класс подходов можно отнести к *априорному* экономическому теоретизированию [53]. Такие базовые концепции разрабатывались и различными способами реализовывались учеными из разных областей. Более того, разработка технических средств и устройств, таких как счетчики трафика или датчики сканирования номерных знаков, не решила полностью вышеупомянутую проблему, но инициировала исследования возникающих проблемных вопросов. Разрабатывались методы, которые были нацелены на преодоление неполноты данных, собираемых сканерами номерных знаков [54]. Фактически, неполные данные до сих пор являются обычным входным условием при решении задачи оценки спроса на перемещение. Один из самых распространенных подходов к решению задачи оценки спроса на перемещение основан на идее минимизации разницы между наблюдаемым трафиком и модельными оценками равновесного распределения потоков [55]. Другой подход решает задачу путем максимизации целевой функции, основанной на энтропии, при сохранении согласованности между априорными и модельными оценками распределения потоков [56, 57]. Ряд исследователей предложили сводить задачу к поиску наименьших квадратов [58, 59] и максимального правдоподобия [60].

Первые двухуровневые формулировки задачи оценки матрицы спроса с использованием принципа равновесного распределения потоков были предложены в [61–63]. Поскольку возникшие двухуровневые задачи оптимизации на самом деле были NP-сложными, исследователи начали искать способы сведения задачи к одноуровневой оптимизации [64–66]. Более того, возникли различные эвристические алгоритмы решения двухуровневой задачи поиска матрицы спроса на перемещение [67]. Впервые проблема оценки спроса на перемещение, как обратная к задаче распределения потоков, была сформулирована в [68]. Действительно, как правило, задача оценки спроса на поездки подразумевает, что значения спроса на поездки являются неизвестными переменными, которые можно найти с помощью значений трафика, наблюдаемого в дорожной сети. Следовательно, способ решения этой проблемы во многом зависит от технических устройств, доступных для наблюдения за дорожным движением и сбора данных о дорожном движении. Проблема сбора данных о трафике воспринимается исследователями как очень важная [69–71]. Например, счетчик трафика подсчитывает точное количество проезжающих транспортных средств. Другими словами, транспортный инженер может собирать точные данные о дорожном движении на участках дорог с помощью предустановленных счетчиков. Главными преимуществами счетчиков трафика являются их простота и дешевизна. При этом счетчики собирают информацию о количестве транспортных средств и не собирают никаких персональных данных об их владельцах. Многие исследователи основывают свои исследования проблемы оценки спроса на поездки на счетчиках трафика как источнике исходных данных [72]. Поднимаемая здесь вспомогательная задача – поиск оптимальной схемы расположения таких датчиков [73].

Датчик сканирования номерных знаков распознает номера проезжающих транспортных средств. С одной стороны, транспортный инженер может собирать точные данные о трафике на элементах сети с помощью датчиков сканирования номерных знаков. С другой стороны, он может восстановить реальные маршруты, по которым перемещались транспортные средства. Однако датчики сканирования номеров собирают как информацию о количестве транспортных средств, так и личные данные их владельцев, что является ключевым недостатком применения таких устройств. Тем не менее, крупные города уже оснащены большим количеством датчиков сканирования номерных знаков, которые в основном используются для

контроля правил дорожного движения. Поэтому неудивительно, что многие исследователи в своих работах основывают свои подходы на датчиках сканирования номерных знаков как источнике исходных данных [74]. Некоторые исследователи объединяют данные, собранные как счетчиками трафика, так и датчиками сканирования номерных знаков, при решении задачи оценки спроса на поездки или поиске оптимального расположения датчиков [75, 76]. Нет сомнений в том, что интеллектуальные инструменты, предназначенные для повседневного решения задачи оценки спроса на поездки, могут обрабатывать оба типа исходных данных. Современные сервисы, такие как Google.Maps или Яндекс.Карты, предоставляют информацию о пробках на дорогах онлайн. Фактически, благодаря таким сервисам можно в онлайн-режиме узнать среднюю скорость транспортных средств на любом участке дорожной сети. Насколько известно, впервые применимость онлайн-сервисов в качестве источников исходных данных при решении задачи оценки спроса на поездки была продемонстрирована в [77].

Оценка спроса на поездки и распределения потоков – две задачи, которые можно считать взаимно обратными. Действительно, переменные решения задачи оценки спроса на перемещение являются входными данными для задачи распределения потоков, а переменные решения задачи распределения потоков являются входными данными для задачи оценки спроса на перемещение (таблица 2). Первое рассмотрение проблемы оценки спроса на поездки как обратной задачи назначения трафика было сделано в [68]. Однако обратная задача была сформулирована в очень строгих условиях, например, считалось, что даже матрица маршрутов задана. Одной из основных целей наших исследований является устранение чрезмерного использования заранее предоставленной и *априорной* информации при оценке спроса на поездки. Действительно, при использовании априорной матрицы в целевой функции задачи оценки спроса задача оценки спроса на перемещение сводится к поиску среднеквадратичного отклонения.

Таблица 2. Входные данные и переменные решения

	Расположения пар узлов спроса	Величины спроса	Потоки на дугах
Задача распределения потоков	Дано	Дано	Переменные
Задача оценки спроса	Переменные	Переменные	Дано

Фактически, проблема оценки спроса на поездки сводится к проблеме итеративного обновления информации о спросе на поездки для заданного набора пар отправления-прибытия.

Будем считать, что предположение о существовании *априорной* матрицы спроса является слишком строгим при рассмотрении задачи оценки спроса на поездки. В связи с чем, сформулируем эту задачу, избегая заранее заданной или *априорной* информации, как обратную к задаче распределения потоков [78]. В результате, *задача оценки спроса на поездки* имеет следующую форму двухуровневой программы оптимизации:

$$\min_F \|\chi(F) - \bar{x}\| \quad (7)$$

при ограничениях

$$F^w \geq 0 \quad \forall w \in V \times V, \quad (8)$$

где отображение $\chi(F)$ задается через (5). Однако можно видеть, что задача (7), (8) имеет тривиальное решение, такое, что $F^w = \bar{x}_w$ для всех $w \in W$ при $W = E$. Кроме того, если рассмотреть транспортную сеть, представленную ориентированным графом с четырьмя узлами и четырьмя дугами, то можно заметить неединственность решения.

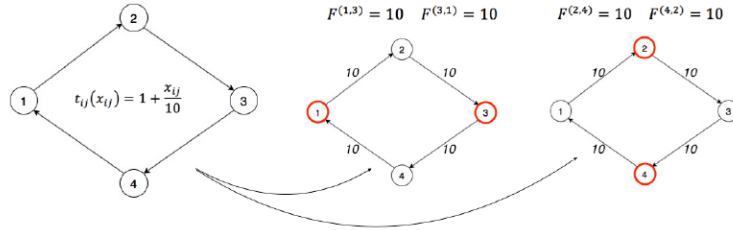


Рис. 2. Транспортная сеть с четырьмя узлами

Предположим, что на каждой дуге имеется поток, равный 10. Решение соответствующей обратной задачи о распределении трафика существует. Однако в этом случае есть несколько решений задачи оценки спроса на поездки. Например, $F^{(1,3)} = 10$ и $F^{(3,1)} = 10$ или $F^{(2,4)} = 10$ и $F^{(4,2)} = 10$ (рис. 2). Таким образом, можно видеть, что задача оценки спроса на поездки может иметь несколько решений. Собственно, ранее уже выяснилось, что [78]:

- вообще говоря, обратная задача распределения потоков не имеет однозначного решения;

- если расположение OD-пар известно, то обратная задача распределения равновесного трафика может иметь единственное решение.

Другими словами, *априорная* информация о местоположении пар узлов отправления-прибытия оказывается весьма важной для оценки спроса на поездки, и это все же не такое строгое требование, как *априорная* информация о матрице спроса. Следовательно, можно улучшить формулировку (7), (8) следующим образом:

$$\min_F \|\chi(F) - \bar{x}\|$$

при ограничениях

$$F^w > 0 \quad \forall w \in W,$$

$$F^w = 0 \quad \forall w \in V \times V \setminus W,$$

где W является заданной.

4. Заключение. Отметим, что описанные постановки задач и используемые модели имеют широкое поле применимости, выходящее далеко за пределы исключительно транспортных процессов в улично-дорожных сетях [79–81]. Среди задач, которые можно формулировать и решать с применением описанного в настоящей статье аппарата, следует выделить следующие:

- оптимизация мультимодальных транспортных сетей;
- оптимизация цепей поставок;
- поиск экономического равновесия Вальраса;
- поиск невальрасовского экономического равновесия.

Литература

1. Farahani R. Z., Miandoabchi E., Szeto W. Y., Rashidi H. A review of urban transportation network design problems // European Journal of Operational Research. 2013. Vol. 229. P. 281–302.
2. Patriksson M. The traffic assignment problem: models and methods. The series “Topics in Transportation” by VSP. Utrecht: Dover Publications, 1994. 240 p.

3. Migdalas A. Bilevel programming in traffic planning: Models, methods and challenge // *Journal of Global Optimization*. 1995. Vol. 7. No 4. P. 381–405.
4. Krylatov A., Zakharov V., Tuovinen T. Principles of Wardrop for traffic assignment in a road network // *Springer Tracts on Transportation and Traffic*. 2020. Vol. 15. P. 17–43.
5. Smith J. C., Song Y. A survey of network interdiction models and algorithms // *European Journal of Operational Research*. 2020. Vol. 283. No 3. P. 797–811.
6. Suh S., Kim T. Solving nonlinear bilevel programming models of the equilibrium network design problem: a comparative review // *Annals of Operations Research*. 1992. Vol. 34. No 1. P. 203–218.
7. Yang H., Bell M. G. H. Models and algorithms for road network design: a review and some new developments // *Transport Reviews*. 1998. Vol. 18. No 3. P. 257–278.
8. Abdulaal M., LeBlanc L. J. Continuous equilibrium network design models // *Transportation Research Part B*. 1979. Vol. 13. No 1. P. 19–32.
9. Marcotte P. Network optimization with continuous control parameters // *Transportation Science*. 1983. Vol. 17. No 2. P. 181–197.
10. Marcotte P., Marquis G. Efficient implementation of heuristics for the continuous network design problem // *Annals of Operations Research*. 1992. Vol. 34. No 1. P. 163–176.
11. Suwansirikul C., Friesz T. L., Tobin R. L. Equilibrium decomposed optimization: a heuristic for the continuous equilibrium network design problem // *Transportation Science*. 1987. Vol. 21. No 4. P. 227–292.
12. Tobin R. L., Friesz T. L. Sensitivity analysis for equilibrium network flow // *Transportation Science*. 1988. Vol. 22. No 4. P. 231–293.
13. Tobin R. L. Sensitivity analysis for variational inequalities // *Journal of Optimization Theory and Applications*. 1986. Vol. 48. No 1. P. 191–204.

14. Chiou S. Bilevel programming for the continuous transport network design problem // *Transportation Research Part B*. 2005. Vol. 39. No 4. P. 361–383.
15. Friesz T. L., Cho H.-J., Mehta N. J., Tobin R. L., Anandalingam G. A Simulated annealing approach to the network design problem with variational inequality constraints // *Transportation Science*. 1992. Vol. 26. No 1. P. 1–68.
16. Meng Q., Yang H., Bell M. G. H. An equivalent continuously differentiable model and a locally convergent algorithm for the continuous network design problem // *Transportation Research Part B*. 2001. Vol. 35. P. 83–105.
17. Li C., Yang H., Zhu D., Meng Q. A global optimization method for continuous network design problems // *Transportation Research Part B*. 2012. Vol. 46. No 9. P. 1144–1158.
18. Krylatov A. Global optimum search in the network design problem // *Computational Mathematics and Mathematical Physics*. 2024. (in press)
19. Wardrop J. G. Some theoretical aspects of road traffic research // *Proc. Institution of Civil Engineers*. 1952. Vol. 2. P. 325–378.
20. Beckmann M., McGuire C., Winsten C. *Studies in the economics of transportation*. New Haven: Yale University Press, 1956. 249 p.
21. LeBlanc L. *Mathematical programming algorithms for large scale network equilibrium and network design problems*. PhD Dissertation, Department of Industrial Engineering and Management Science, Northwestern University. 1973.
22. Nguyen S. *A mathematical programming approach to equilibrium methods of traffic assignment with fixed demands*. Publication 138, Department d'Informatique et de Recherche Operationelle, Universite de Montreal, Montreal. 1973.
23. Frank M., Wolfe P. An algorithm for quadratic programming // *Naval Research Logistics Quarterly*. 1956. Vol. 3. P. 95–110.

24. Perederieieva O., Ehrgott M., Raith A., Wang J. Y. T. A framework for and empirical study of algorithms for traffic assignment // *Computers & Operations Research*. 2015. Vol. 54. P. 90–107.
25. Dial R. A path-based user-equilibrium traffic assignment algorithm that obviates path storage and enumeration // *Transportation Research Part B*. 2006. Vol. 40. P. 917–936.
26. Bar-Gera H. Traffic assignment by paired alternative segments // *Transportation Research Part B*. 2010 Vol. 44. No 8-9. P. 1022–1046.
27. Lupi M. Convergence of the frank-wolfe algorithm in transportation networks // *Civil Engineering Systems*. 1985. Vol. 19. P. 7–15.
28. Galligari A., Sciandrone M. A computational study of path-based methods for optimal traffic assignment with both inelastic and elastic demand // *Computers and Operations Research*. 2019. Vol. 103. P. 158–166.
29. Boyce D., Ralevich-Dekic B., Bar-Gera H. Convergence of traffic assignment: how much is enough? // *Journal of Transportation Engineering*. 2004. Vol. 130. No 1. P. 49–55.
30. Krylatov A. Y. Sequential path-equilibration algorithm for highly accurate traffic flow assignment in an urban road network // *Optimization Methods and Software*. 2023. Vol. 38. No 5. P. 1081–1104.
31. Dafermos S., Sparrow F. The traffic assignment problem for a general network // *Journal of Research of the National Bureau of Standards*. 1969. Vol. 73B. No 2. P. 91–118.
32. Leventhal T., Nemhauser G., Trotter L. J. A column generation algorithm for optimal traffic assignment // *Transportation Science*. 1973. Vol. 7. No 2. P. 168–176.
33. Larsson T., Patriksson M. Simplicial decomposition with disaggregated representation for the traffic assignment problem // *Transportation Science*. 1992. Vol. 26. No 1. P. 4–17.
34. Jayakrishnan R., Tsai W., Prashker J., Subodh R. A faster path-based algorithm for traffic assignment // *Transportation Research Record*. 1994. Vol. 1443. P. 75–83.


35. Sun C., Jayakrishnan R., Tsai W. Computational study of a path-based algorithm and its variants for static traffic assignment // *Transportation Research Record*. 1996. Vol. 1537. No 1. P. 106–115.
36. Chen A., Jayakrishnan R. A path-based gradient projection algorithm: effects of equilibration with a restricted path set under two flow update policies // *Transportation research board annual meeting*. 1998. P. 1–19.
37. Florian M., Constantin I., Florian D. A new look at projected gradient method for equilibrium assignment // *Transportation Research Record*. 2009. Vol. 2090. P. 10–16.
38. Kupsizewska D., Van Vliet D. 101 uses for path-based assignment // *Transport Planning Methods: Proc., PTRC Planning and Transport Summer Ann. Meeting, Univ. Sussex*, 1999. P. 121–132.
39. Kumar A., Peeta S. An improved social pressure algorithm for the static deterministic user equilibrium traffic assignment problem // *In: Transportation research board of the national academics*. Washington DC. 2011.
40. Chen A., Lee D. H., Jayakrishnan R. Computational study of state-of-the-art path-based traffic assignment algorithms // *Mathematics and Computers in Simulation*. 2002. Vol. 59. P. 509–518.
41. Di Lorenzo D., Galligari A., Sciandrone M. A convergent and efficient decomposition method for the traffic assignment problem // *Computational Optimization and Applications*. 2015. Vol. 60. No 1. P. 151–170.
42. Galligari A., Sciandrone M. A convergent and fast path equilibration algorithm for the traffic assignment problem // *Optimization Methods and Software*. 2018. Vol. 33. No 2. P. 354–371.
43. Sheffi Y. *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*. NJ. Englewood Cliffs: Prentice-Hall Inc, 1985. 399 p.
44. Devarajan S. A note on network equilibrium and noncooperative games // *Transportation Research Part B*. 1981. Vol. 15. No 6. P. 421–426.

45. LeBlanc L., Morlok E., Pierskalla W. An efficient approach to solving the road network equilibrium traffic assignment problem // *Transportation Research*. 1975. Vol. 9. P. 309–318.
46. Bar-Gera H., Luzon A. Differences among route flow solutions for the user-equilibrium traffic assignment problem // *Journal of Transportation Engineering*. 2007. Vol. 133. No 4. P. 232–239.
47. Bar-Gera H., Luzon A. Non-unique route flow solutions for user-equilibrium assignments // *Traffic Engineering and Control*. 2007. Vol. 48. No 9. P. 408–412.
48. Kumar A., Peeta S. Entropy weighted average method for the determination of a single representative path flow solution for the static user equilibrium traffic assignment problem // *Transportation Research Part B*. 2015. Vol. 71. P. 213–229.
49. Chao S., Chenga L., Xub T. Range of user-equilibrium route flow with applications // *Procedia – Social and Behavioral Sciences*. 2014. Vol. 138. P. 86–96.
50. Heydecker B. G., Lam W. H. K., Zhang N. Use of travel demand satisfaction to assess road network reliability // *Transportmetrica*. 2007. Vol. 3. No 2. P. 139–171.
51. Frederix R., Viti F., Tampere C. M. J. Dynamic origin-destination estimation in congested networks: theoretical findings and implications in practice // *Transportmetrica A: Transport Science*. 2013. Vol. 9. No. 6. P. 494–513.
52. Isard W. *Methods of Regional Analysis: An Introduction to Regional Science*. New York: John Wiley and Sons, Inc., 1960. 823 p.
53. Fisher F. *A Priori Information and Time Series Analysis*. Amsterdam: North-Holland Publishing Co., 1962. 184 p.
54. Makowski G. G., Sinha K. C. A statistical procedure to analyze partial license plate numbers // *Transportation Research*. 1976. Vol. 10. P. 131–132.
55. Carey M., Hendrickson C., Siddharthan K. A method for direct estimation of origin/destination trip matrices // *Transportation Science*. 1981. Vol. 15. No 1. P. 32–49.

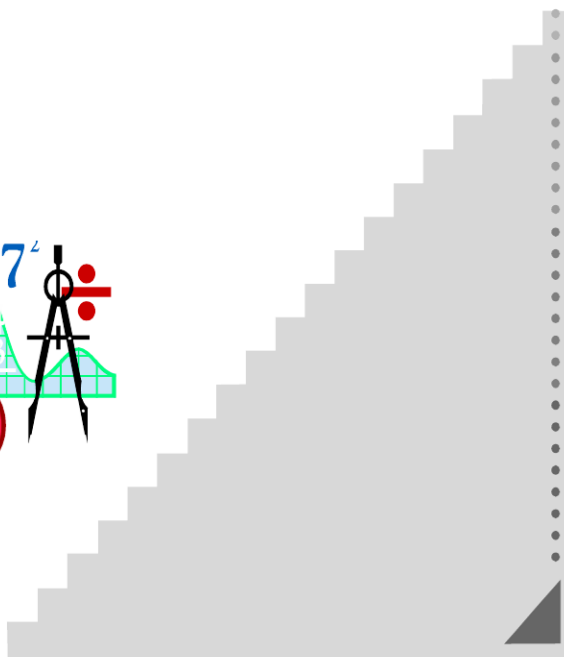
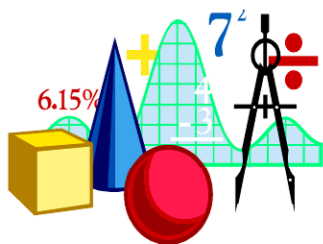
56. Van Zuylen H. J., Willumsen L. G. The most likely trip matrix estimated from traffic counts // *Transportation Research Part B*. 1980. Vol. 14. P. 281–293.
57. Bell M. G. H. The estimation of an origin-destination matrix from traffic counts // *Transportation Science*. 1983. Vol. 17. No 2. P. 198–217.
58. Cascetta E. Estimation of trip matrices from traffic counts and survey data: a generalized least squares estimator // *Transportation Research Part B*. 1984. Vol. 18. P. 289–299.
59. McNeil S., Hendrickson C. A regression formulation of the matrix estimation problem // *Transportation Science*. 1985. Vol. 19. No 3. P. 278–292.
60. Watling D. P. Maximum likelihood estimation of an origin-destination matrix from a partial registration plate survey // *Transportation Research Part B*. 1994. Vol. 28. No 4. P. 289–314.
61. Nguyen S. Estimating an OD matrix from network data: a network equilibrium approach. Publication 60, Centre de Recherche sur les Transports, Université de Montréal. 1977.
62. Fisk C. On combining maximum entropy trip matrix estimation with user optimal assignment // *Transportation Research Part B*. 1988. Vol. 22. No 1. P. 69–73.
63. Yang H., Sasaki T., Iida Y., Asakura Y. Estimation of origin-destination matrices from link traffic counts on congested networks // *Transportation Research Part B*. 1992. Vol. 26. No 6. P. 417–434.
64. Bell M. G., Shield C. M., Busch F., Kruse C. A stochastic user equilibrium path flow estimator // *Transportation Research Part C*. 1997. Vol. 5. No 34. P. 197–210.
65. Bar-Gera H. Primal method for determining the most likely route flows in large road network // *Transportation Science*. 2006. Vol. 40. No 3. P. 269–286.

66. Shen W., Wynter L. A new one-level convex optimization approach for estimating origin-destination demand // *Transportation Research Part B*. 2012. Vol. 46. P. 1535–1555.
67. Lundgren J. T., Peterson A. A heuristic for the bilevel origin-destination matrix estimation problem // *Transportation Research Part B*. 2008. Vol. 42. P. 339–354.
68. Bierlaire M. The total demand scale: a new measure of quality for static and dynamic origin-destination trip tables // *Transportation Research Part B*. 2002. Vol. 36. P. 837–850.
69. Quandt R. E., Baumol W. J. The demand for abstract transport modes: theory and measurement // *Journal of Regional Science*. 1966. Vol. 6. No 2. P. 13–26.
70. Yang Y., Fan Y. Data dependent input control for origin-destination demand estimation using observability analysis // *Transportation Research Part B*. 2015. Vol. 78. P. 385–403.
71. Yang Y., Fan Y., Wets R. J. B. Stochastic travel demand estimation: Improving network identifiability using multi-day observation sets // *Transportation Research Part B*. 2018. Vol. 107. P. 192–211.
72. Viti F., Rinaldi M., Corman F., Tampere C. Assessing partial observability in network sensor location problems // *Transportation Research Part B*. 2014. Vol. 70. P. 65–89.
73. Bianco L., Cerrone C., Cerulli R., Gentili M. Locating sensors to observe network arc flows: exact and heuristic approaches // *Computers and Operation Research*. 2014. Vol. 46. P. 12–22.
74. Li X., Ouyang Y. Reliable sensor deployment for network traffic surveillance // *Transportation Research Part B*. 2011. Vol. 45. P. 218–231.
75. Parry K., Hazelton M. L. Estimation of origin-destination matrices from link counts and sporadic routing data // *Transportation Research Part B*. 2012. Vol. 46. P. 175–188.

76. Castillo E., Nogal M., Rivas A., Sanchez-Cambronero S. Observability of traffic networks. Optimal location of counting and scanning devices // *Transportmetrica B: Transport Dynamics*. 2013. Vol. 1. No. 1. P. 68–102.
77. Krylatov A., Shirokolobova A., Zakharov V. OD-matrix estimation based on a dual formulation of traffic assignment problem // *Informatica*. 2016. Vol. 40. No 4. P. 393–398.
78. Krylatov A., Raevskaya A., Zakharov V. Travel demand estimation in urban road networks as inverse traffic assignment problem // *Transport and Telecommunication*. 2021. Vol. 22. No 3. P. 287–300.
79. Аббасов М. Э., Шарлай А. С. Метод поиска оптимальной по стоимости траектории дороги на поверхности местности // *Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления*. 2023. Т. 19. № 2. С. 139–147.
80. Suna P., Parilina E. M., Gaob H. W. Two-stage network games modeling the Belt and Road Initiative // *Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes*. 2022. Vol. 18. No 1. P. 87–98.
81. Васильев В. А., Кузнецова Д. С. Алгоритм роя частиц и анализ его эффективности для решения задач безусловной однокритериальной оптимизации // *Процессы управления и устойчивость*. 2022. Т. 9. № 1. С. 190–196.



*1. Математическая
теория процессов
управления*



УДК 517.977

Андрянова Н. Р.

Сверхфинитная устойчивость мультиагентной системы с переключениями

Рекомендовано к публикации профессором Александровым А. Ю.

1. Введение. Задача равноудаленного расположения мобильных агентов на фиксированном отрезке пользуется интересом [1–5]. Предполагается, что агенты обмениваются информацией друг с другом. Они получают информацию о расстояниях до какого-либо одного агента-информатора слева и одного – справа [1]. Зачастую существенной оказывается возможность переключаться между информаторами.

Имеется ряд работ, где сформировано децентрализованное управление, при котором гарантируется стабилизация к требуемому расположению [2–4]. Для достижения нужных положений за конечное время строятся так называемые сверхфинитные протоколы управления [1, 5].

Как и в [1], в предложенной работе считается, что движение агентов задается скалярными дифференциальными уравнениями первого порядка. В отличие от [1], в данной работе количество информаторов может быть увеличено вплоть до включения всех участников в число информаторов. Сверхфинитная устойчивость равноудаленных положений будет обеспечена при достаточно больших длинах промежутков между моментами переключений.

2. Постановка задачи. Пусть n пронумерованных агентов движутся по прямой и их динамика задается дифференциальными уравнениями

$$\dot{x}_i = u_i, \quad i = 1, \dots, n. \quad (1)$$

Каждому участнику системы известны расстояния до некоторых других агентов-информаторов. Граф взаимодействий, отражающий связи всех агентов с информаторами называется коммуникационной

Андрянова Наталья Романовна – аспирант, Санкт-Петербургский государственный университет; e-mail: st040174@student.spbu.ru, тел.: +7(981)940-80-79

Исследование выполнено за счет гранта Российского научного фонда № 24-21-00091, <https://rscf.ru/project/24-21-00091/>

топологией. Будем предполагать, что имеются S топологий и закон переключений между ними – кусочно-постоянная непрерывная справа функция $\sigma : [0, +\infty) \rightarrow \{1, \dots, S\}$.

Считаем заданными моменты переключений s_1, s_2, \dots – точки разрыва первого рода функции σ . Случай Зено [6], при котором допускается бесконечное число переключений, совершаемых за конечный промежуток времени, исключается из рассмотрения.

Обозначим через $N_{il}^{(\sigma)}$ и $N_{ir}^{(\sigma)}$ множества номеров левых и правых информаторов i -го агента. Множество $N_i^{(\sigma(t))} = N_{il}^{(\sigma(t))} \cup N_{ir}^{(\sigma(t))}$ представляет собой совокупность номеров соседей i -го агента в момент t . Определим множество $\bar{N}_i^{(s)} = \{|i - j|, j \in N_i^{(s)}\}$ и назовем его множеством относительных порядковых номеров информаторов. Выбирается отрезок $[a, b]$, на котором планируется разместить агентов. Его границы, когда это удобно, будут интерпретироваться как не меняющие своего положения агенты с номерами 0 и $n + 1$. Задачей этой работы является достижение размещения агентов в точках

$$\tilde{x}_i = a + \frac{i}{n+1}(b-a), \quad i = 1, \dots, n, \quad (2)$$

за конечное время. Сформируем вектор $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_n)^T$.

3. Описание протокола. Покажем, что в качестве протоколов u_i подойдут функции $u_i = \varphi_i(z_i^{(\sigma)})$,

$$z_i^{(\sigma)}(t) = \sum_{j \in N_i^{(\sigma)}} a_{ij}^{(\sigma)}(x_j(t) - x_i(t)),$$

где $\varphi_i(z) = d_1 z^p + d_2 z^q$, с положительными постоянными d_1 и d_2 . Показатели $p \in (0, 1)$ и $q > 1$ – рациональные числа с нечетными числителями и знаменателями. Коэффициенты $a_{ij}^{(\sigma)}$ определяются по формулам из работы [4], где построен протокол, обеспечивающий асимптотическую устойчивость положений равновесия:

$$a_{ij}^{(\sigma)} = \frac{\omega_i^{(\sigma)}}{(i-j)q_{il}^{(\sigma)}}, \quad j \in N_{il}^{(\sigma)}, \quad q_{il}^{(\sigma)} = \text{card}(N_{il}^{(\sigma)}),$$

$$a_{ij}^{(\sigma)} = \frac{\omega_i^{(\sigma)}}{(j-i)q_{ir}^{(\sigma)}}, \quad j \in N_{ir}^{(\sigma)}, \quad q_{ir}^{(\sigma)} = \text{card}(N_{ir}^{(\sigma)}),$$

$$\omega_i^{(\sigma)} \left(\sum_{j \in N_{il}^{(\sigma)}} \frac{1}{(i-j)q_{il}^{(\sigma)}} + \sum_{j \in N_{ir}^{(\sigma)}} \frac{1}{(j-i)q_{ir}^{(\sigma)}} \right) = 1,$$

в остальных случаях $a_{ij}^{(\sigma)} = 0$. Обозначение $\text{card}(A)$ отвечает за количество элементов множества A .

Рассмотрим систему (1), замкнутую оговоренным управлением:

$$\dot{x}_i = \varphi_i(z_i^{(\sigma)}), \quad i = 1, \dots, n. \quad (3)$$

Будем обозначать ее решение с начальным условием $x(t_0) = x_0$ через $x(t, x_0, t_0)$.

Определение [5]. Решение (2) системы (3) *сверхфинитно устойчиво*, если 1) оно глобально асимптотически устойчиво, 2) для любой x_0 найдется $T(x_0) \geq 0$, такое что $x(t, x_0, t_0) = \tilde{x}$ для $t \geq t_0 + T(x_0)$, 3) существует $T_{\max} > 0$ такое, что $T(x_0) \leq T_{\max}$ для любых x_0 .

4. Анализ устойчивости. Сформулируем основное утверждение данной работы.

Утверждение. Пусть агентам известны относительные порядковые номера информаторов, и пусть $N_{il}^{(s)} \neq \emptyset$ и $N_{ir}^{(s)} \neq \emptyset$ для всех $s = 1, \dots, S$, $i = 1, \dots, n$. Тогда существует $\tau > 0$, такое что если длины отрезков между переключениями $s_{j+1} - s_j > \tau$, $j = 1, 2, \dots$, то решение (2) системы (3) *сверхфинитно устойчиво*.

Доказательство. Проведем замену $y(t) = x(t) - \tilde{x}$ и перейдем к системе в отклонениях

$$\dot{y}(t) = \Phi \left((A^{(\sigma(t))} - I)y(t) \right),$$

где $A^{(\sigma(t))} = \{a_{ij}^{(\sigma(t))}\}_{i,j=1,\dots,n}$, $\Phi(y) = (\varphi_1(y_1), \dots, \varphi_n(y_n))^T$.

Для каждой подсистемы

$$\dot{y}(t) = \Phi \left(\tilde{A}^{(s)} y(t) \right), \quad s = 1, \dots, S, \quad (4)$$

матрицы $\tilde{A}^{(s)} = A^{(s)} - I$ метцлеровы. Кроме того, как было показано в работе [4], существует вектор с положительными компонентами $\theta_i = 1 - \frac{1}{2^i}$, такой что $A^{(s)}\theta < \theta$ для всех $s \in \{1, \dots, S\}$. Тогда $\tilde{A}^{(s)}$ гурвицевы, а значит существуют положительные векторы $\eta^{(s)}$, такие что $(\tilde{A}^{(s)})^T \eta^{(s)} < 0$, $s = 1, \dots, S$. С найденными векторами строятся

функции Ляпунова

$$V_s(y) = |\tilde{A}^{(s)}y|^T \eta^{(s)}, \quad (5)$$

где модуль вектора – вектор модулей компонент. Про такие функции Ляпунова известно [1], что они положительно определены, глобально липшецевы, и для них находится коэффициент $\mu \geq 1$ из неравенств

$$V_i(y) \leq \mu V_j(y), \quad i = 1, \dots, S, \quad j = 1, \dots, S.$$

Достаточно рассмотреть $\mu > 1$, так как при $\mu = 1$ сверхфинитная устойчивость гарантируется при произвольном выборе закона переключения в связи с существованием общей функции Ляпунова [1].

Воспользуемся производной Дини

$$D^+V_s(y) = \overline{\lim}_{h \rightarrow +0} \frac{V_s(y(t+h)) - V_s(y(t))}{h}.$$

Согласно теореме Иосидзавы о производной Дини [7] для функции (5) вдоль решений подсистемы (4)

$$D^+V_s(y) \leq \overline{\lim}_{h \rightarrow +0} \frac{V_s(y(t) + h\Phi(\tilde{A}^{(s)}y(t))) - V_s(y(t))}{h}.$$

Подробнее распишем выражение, стоящее под знаком предела:

$$\begin{aligned} & \frac{1}{h} \left(|\tilde{A}^{(s)}y + h\tilde{A}^{(s)}\Phi(\tilde{A}^{(s)}y)|^T \eta^{(s)} - |\tilde{A}^{(s)}y|^T \eta^{(s)} \right) = \\ & = \frac{1}{h} \sum_{i=1}^n \left[\left| \sum_{j=1}^n \tilde{a}_{ij}^{(s)} y_j + h \sum_{j=1}^n \tilde{a}_{ij}^{(s)} \varphi_j \left(\sum_{k=1}^n \tilde{a}_{jk}^{(s)} y_k \right) \right| - \left| \sum_{j=1}^n \tilde{a}_{ij}^{(s)} y_j \right| \right] \eta_i^{(s)}. \end{aligned}$$

Каждое слагаемое в квадратных скобках оценим сверху выражением

$$h \left| \sum_{j \neq i}^n \tilde{a}_{ij}^{(s)} \varphi_j(\mathfrak{S}_j^{(s)}) \right| + \left| \mathfrak{S}_i^{(s)} + h \tilde{a}_{ii}^{(s)} \varphi_i(\mathfrak{S}_i^{(s)}) \right| - |\mathfrak{S}_i^{(s)}|,$$

где $\mathfrak{S}_i^{(s)} = \sum_{j=1}^n \tilde{a}_{ij}^{(s)} y_j$. Пользуясь тем, что $\tilde{a}_{ii}^{(s)} = -1$, получим верхнюю оценку D^+V_s :

$$\overline{\lim}_{h \rightarrow +0} \frac{1}{h} \sum_{i=1}^n \left[h \sum_{j \neq i}^n \tilde{a}_{ij}^{(s)} |\varphi_j(\mathfrak{S}_j^{(s)})| + |\mathfrak{S}_i^{(s)} - h \varphi_i(\mathfrak{S}_i^{(s)})| - |\mathfrak{S}_i^{(s)}| \right] \eta_i^{(s)}.$$

Заметим, что функция φ_i секторного типа, т.е. $y\varphi_i(y) > 0$ при $y \neq 0$, для всех $i = 1, \dots, n$. Тогда за счет малости h верно равенство

$$|\mathfrak{S}_i^{(s)} - h\varphi_i(\mathfrak{S}_i^{(s)})| - |\mathfrak{S}_i^{(s)}| = -h|\varphi_i(\mathfrak{S}_i^{(s)})|.$$

Пользуясь свойством вектора $\eta^{(s)}$, и тем, что $|\varphi_i(x_i)| = \varphi_i(|x_i|)$ получим

$$\begin{aligned} D^+V_s(y) &\leq \overline{\lim}_{h \rightarrow +0} \left(|\Phi(\tilde{A}^{(s)}y)|^T (\tilde{A}^{(s)})^T \eta_s \right) \leq -\lambda \sum_{i=1}^n |\varphi_i(\tilde{A}^{(s)}y)| = \\ &= -\lambda \sum_{i=1}^n \left(d_1 \left| \sum_{j=1}^n \tilde{a}_{ij}^{(s)} y_j \right|^p + d_2 \left| \sum_{j=1}^n \tilde{a}_{ij}^{(s)} y_j \right|^q \right), \end{aligned}$$

где $\lambda > 0$.

В итоге, по свойствам однородных функций найдутся положительные α и β , такие что

$$D^+V_s(y) \leq -\alpha V_s^p(y) - \beta V_s^q(y).$$

Теорема 1 из [1] позволяет получить τ – оценку длительности переключений, гарантирующую сверхфинитную устойчивость, и, более того, оценку времени переходного процесса.

Утверждение доказано.

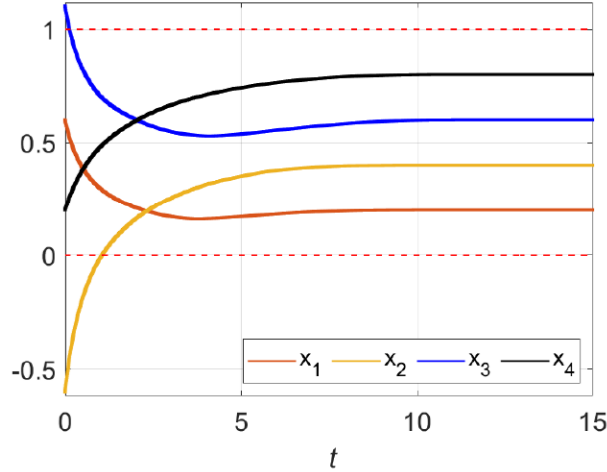


Рис. 1. Компоненты решения системы (3) в случае а)

5. Результаты моделирования. Будем распределять четыре агента на отрезке $[0, 1]$. В протоколе выбраны параметры $d_1 = 1/5$, $d_2 = 1$, $p = 3/5$, $q = 7/3$. Переключения происходят в моменты $1, 2, \dots$ между двумя коммуникационными топологиями, заданными матрицами смежности

$$M_1 = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{pmatrix}.$$

Рассмотрены две ситуации, когда начальные данные а) $x(0) = (0,6 -0,6 \ 1,1 \ 0,2)^T$ (см. рис. 1); б) $x(0) = (11 \ -3 \ -8 \ 5)^T$ (см. рис. 2).

Графики были получены с помощью метода Рунге–Кутты четвертого порядка с шагом 0,0001. После момента $t = 11$ в обоих случаях норма разности численно полученного решения и желаемого размещения не превосходит 10^{-5} .

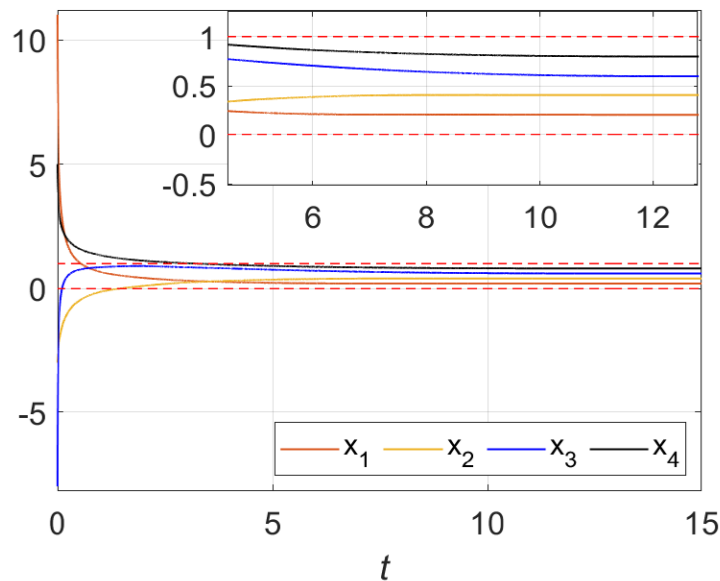


Рис. 2. Компоненты решения системы (3) в случае б)

6. Заключение. Рассмотрена задача равномерного размещения агентов, движение которых задано интеграторами первого порядка. При этом предполагается наличие переключений между коммуникационными топологиями. Показано, что при достаточно больших длинах промежутков между моментами переключений мультиагентная система будет иметь сверхфинитно устойчивое распределение. Результат подкреплён численным примером.

Литература

1. Aleksandrov A. Y., Andriyanova N. R. Fixed-time stability of switched systems with application to a problem of formation control // *Nonlinear Analysis: Hybrid Systems*. 2021. Vol. 40. Art. no 101008.
2. Aleksandrov A. Yu., Arakelov A. I. A problem of the equidistant deployment for discrete-time multiagent systems // *Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes*. 2022. Vol. 18. No 1. P. 171–178.
3. Воробьёва А. А. Управление мультиагентными системами с запаздыванием // *Процессы управления и устойчивость*. 2019. Т. 6. № 1. С. 31–35.
4. Aleksandrov A. Y., Andriyanova N. R. Distributed algorithms for mobile agent deployment on a line segment under switching topology and communication delays. *IEEE Control Systems Letters*. 2022. Vol. 6. P. 3218–3223.
5. Парсегов С. Э., Поляков А. Е., Щербаков П. С. Достижение равноудаленного расположения агентов на отрезке за заданное время // *Доклады Академии наук*. 2013. Т. 448. № 5. С. 524–524.
6. Liberzon D. *Switching in Systems and Control*, Boston, MA: Birkhauser, 2003. 233 p.
7. Руш Н., Абетс П., Лалуа М. Прямой метод Ляпунова в теории устойчивости / пер. с англ. В. Н. Рубановского и др. / под ред. В. В. Румянцева. М.: Мир, 1980. 300 с.

УДК 517.977

Бабанов И. В.

Оптимальное управление трехмассовой колебательной системой

Рекомендовано к публикации профессором Смирновым Н. В.

1. Введение. В данной работе исследован метод Р. Габасова [1] по сведению задачи оптимального управления канонического вида к задаче линейного программирования (ЛП). Для этого была разработана программа на языке С++ с использованием библиотеки линейной алгебры [2]. Проанализирован пример колебательной системы, для которого было построено оптимальное управление с целью её уравновешивания. Численные эксперименты показали состоятельность данного подхода.

2. Постановка задачи. Рассмотрим каноническую задачу оптимального управления, следуя работе [1]. Пусть объект управления описывается системой дифференциальных уравнений

$$\dot{x}(t) = A(t)x(t) + b(t)u(t), \quad x(t_0) = x_0, \quad (1)$$

$$t \in [t_0, t_1] = T \subset \mathbb{R}, \quad t_0 < t_1 < +\infty$$

где $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}$, $A(t)$ – $(n \times n)$ -матрица, $b(t)$ – n -мерный вектор, $t \in [t_0, t_1] = T$. Разобьем отрезок $[t_0, t_1]$ на N частей, $h = \frac{t_1 - t_0}{N}$ – шаг дискретизации.

Определение 1. Функция $u(t)$, $t \in [t_0, t_1]$ называется дискретным управлением с периодом квантования h , если

$$\begin{cases} u(t) = u(t_0 + kh) = u_{kh}, & t \in [t_0 + kh, t_0 + (k+1)h], \\ k = 0, \dots, N-1. \end{cases} \quad (2)$$

и кроме того $|u(t)| \leq M$, $M \in \mathbb{R}$. Для оценки решений уравнения (1) на множестве допустимых управлений (2) введем целевую функцию

$$c^T x(t_1) \rightarrow \max, \quad (3)$$

Бабанов Илья Владимирович – студент, Санкт-Петербургский государственный университет; e-mail: st097273@student.spbu.ru, тел.: +7(950)500-21-07

при условии

$$Hx(t_1) = g. \quad (4)$$

В (4) $g \in \mathbb{R}^m$, $\text{rang } H = m \leq n$.

Задача (1)–(4) является канонической задачей оптимального управления. Размерность n вектора состояний $x(t)$ называется порядком задачи.

3. Сведение к задаче ЛП. Воспользуемся формулой Коши для решения (1)

$$x(t, t_0, x_0) = Y(t) \left[Y^{-1}(t_0)x_0 + \int_{t_0}^t Y^{-1}(\tau)b(\tau)u(\tau)d\tau \right], \quad (5)$$

где $Y(t)$ – фундаментальная матрица соответствующей однородной системы.

Подставим (5) в целевую функцию (3):

$$c^T x(t_1) = c^T Y(t_1) \left[Y^{-1}(t_0)x_0 + \int_{t_0}^{t_1} Y^{-1}(\tau)b(\tau)u(\tau)d\tau \right].$$

При этом, условие (3) будет эквивалентно следующему

$$\int_{t_0}^{t_1} c^T Y(t_1) Y^{-1}(\tau) b(\tau) u(\tau) d\tau \rightarrow \max. \quad (6)$$

Целевую функцию (6) можно переписать в виде

$$\int_{t_0}^{t_1} z_c^T(\tau) b(\tau) u(\tau) d\tau \rightarrow \max,$$

где $z_c^T(t) = c^T Y(t_1) Y^{-1}(t)$. При этом $Z^T(t) = Y^{-1}(t)$ – фундаментальная матрица сопряженной системы $\dot{z} = -A^T(t)z$.

Далее рассмотрим терминальное ограничение (4) с учетом (5)

$$HY(t_1) \left[Y^{-1}(t_0)x_0 + \int_{t_0}^t Y^{-1}(\tau)b(\tau)u(\tau)d\tau \right] = g. \quad (7)$$

Введем вектор g_0 , чтобы переписать соотношение (7) в виде

$$\int_{t_0}^{t_1} HY(t_1)Y^{-1}(\tau)b(\tau)u(\tau)d\tau = g_0, \quad (8)$$

где $g_0 = g - Hx_0(t_1)$ — m -мерный вектор, $x_0(t)$ — решение соответствующей однородной системы (1), т. е. $x_0(t) = Y(t)Y^{-1}(t_0)x_0$.

Обозначим h_1, \dots, h_m — строки матрицы H , для i -ой строки рассмотрим решение $z_i(t)$ сопряженной системы $\dot{z} = -A^T(t)z$ с начальными условиями

$$z_i(t_1) = h_i^T.$$

В силу того, что

$$z_i(t) = Z(t)(Z^T)^{-1}(t_1)h_i^T, \\ z_i^T(t) = h_i Z^{-1}(t_1)Z^T(t) = h_i Y(t_1)Y^{-1}(t),$$

$z_i^T(t)$ есть i -ая строка матрицы $HY(t_1)Y^{-1}(t)$. Обозначим матрицу размерности $(m \times n)$, строками которой являются $z_i^T(t)$, $i = 1, \dots, m$ за $G(t)$. Тогда (8) запишется как

$$\int_{t_0}^{t_1} G(\tau)b(\tau)u(\tau)d\tau = g_0,$$

причем

$$G(t_1) = H.$$

Следовательно, условия (3), (4) соответственно перепишутся в виде

$$\sum_{t \in T_u} u(t) \int_t^{t+h} z_c^T(\tau)b(\tau)d\tau \rightarrow \max, \\ \sum_{t \in T_u} u(t) \int_t^{t+h} G(\tau)b(\tau)d\tau = g_0,$$

где $T_u = \{t_0, t_0 + h, \dots, t_0 + kh, \dots, t_1 - h\}$.

Введем обозначения

$$C_h(t) = \int_t^{t+h} z_c^T(\tau) b(\tau) d\tau,$$

$$d_h(t) = \int_t^{t+h} G(\tau) b(\tau) d\tau,$$

$$U = (u_{0h} \quad u_{1h} \quad \dots \quad u_{(N-1)h})^T,$$

$$C = (C_h(t_0) \quad C_h(t_0 + h) \quad \dots \quad C_h(t_1 - h))^T,$$

$$D = (d_h(t_0) \quad d_h(t_0 + h) \quad \dots \quad d_h(t_1 - h))^T.$$

Тогда задачу (1)–(4) можно записать в виде

$$\begin{cases} C^T U \rightarrow \max, \\ DU = g_0, \\ |u_i| \leq M, \quad i = 1, \dots, N, \quad M \in \mathbb{R}. \end{cases} \quad (9)$$

В результате динамическая задача оптимального управления (1)–(4) сведена к интервальной задаче ЛП (9).

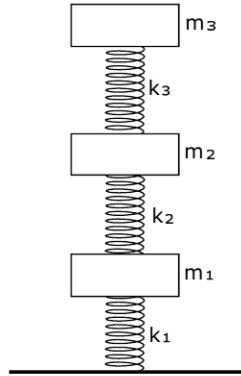


Рис. 1.

4. Численный эксперимент. Рассмотрим трехмассовую систему (см. рис. 1), на которую будем действовать управлением $u = u(t)$. Она задается системой линейных дифференциальных уравнений

$$\begin{cases} \ddot{x}_1 = -\frac{k_1+k_2}{m_1}x_1 + \frac{k_2}{m_1}x_2, \\ \ddot{x}_2 = \frac{k_2}{m_2}x_1 - \frac{k_3+k_2}{m_2}x_2 + \frac{k_3}{m_2}x_3, \\ \ddot{x}_3 = \frac{k_3}{m_3}x_2 - \frac{k_3}{m_3}x_3 + u. \end{cases}$$

Для примера возьмем $k_1 = k_2 = k_3 = m_1 = m_2 = m_3 = 1$. Поместим эту задачу в рамки нахождения кусочно-постоянного управления $u(t)$ с периодом квантования $h = 1$,

$0 \leq u(t) \leq 1$. Получаем систему (1), в которой

$$A(t) = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -2 & 1 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \end{pmatrix}, \quad b(t) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

$$x(t_0) = x_0 = (1 \ 1 \ 1 \ 0 \ 0 \ 0)^T, \quad t \in [t_0 = 0, t_1 = 10].$$

Зададим функционал, относительно которого будем оценивать оптимальность управления

$$\int_{t_0}^{t_1} u(t) dt \rightarrow \min.$$

Тогда вектор весов целевой функции (5) есть $c = (1 \ 1 \ \dots \ 1 \ 1)^T$, а условие (4) будет задаваться единичной матрицей $H = E_6$, и вектором $g = (0 \ 0 \ 0 \ 0 \ 0 \ 0)^T$.

Для построения оптимального управления написана программа на языке C++. Приведем графики поведения, к примеру, отклонения от положения равновесия первого груза ($x_1(t)$, см. рис. 2, 3).

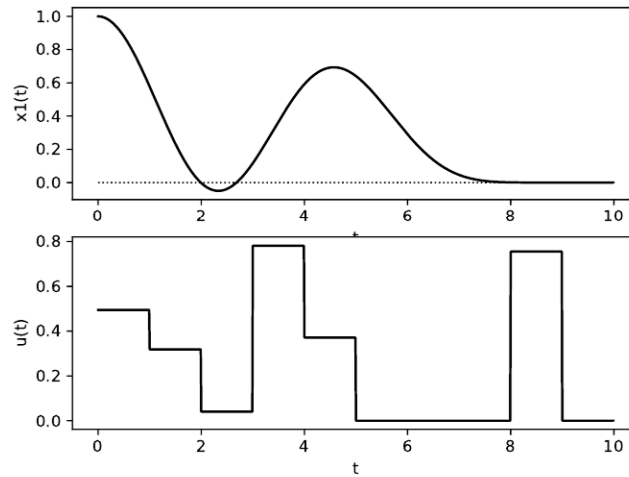


Рис. 2. $h = 1$, $t_1 = 10$.

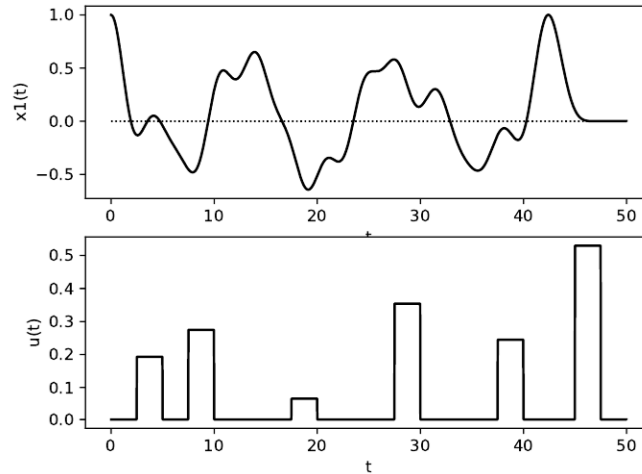


Рис. 3. $h = 2,5$, $t_1 = 50$.

5. Заключение. В рамках данной работы был изучен метод решения канонической задачи оптимального управления, представленной в работе [1]. В качестве реального примера была взята задача управления трехмассовой колебательной системой. Проведен численный эксперимент, по результатам которого можно сделать вывод, что рассмотренный подход обладает высоким потенциалом и может быть использован для различных приложений.

Литература

1. Балашевич Н. В, Габасов Р., Кириллова Ф. М. Численные методы программной и позиционной оптимизации для линейных систем // Журн. вычисл. математики и мат. физики. 2000. Т. 40. № 6. С. 838–859.
2. ALGLIB – C++/C#/Java numerical analysis library [Электронный ресурс]: URL:<https://www.alglib.net/> (дата обращения: 20.03.24).

УДК 517.929.4

Гапонов Д. Н.

Устойчивость линейных систем с запаздыванием. Подход вычислительной алгебры

Рекомендовано к публикации ассистентом Чашниковым М. В.

1. Введение. Исследование устойчивости системы является одним из ключевых вопросов в теории управления. Для случая линейной системы сформулировано множество утверждений о характере устойчивости, критериях в работах [1–5], достаточных условиях в работах [6–13]. Алгебраический подход очень удачно применяется в данной задаче. Действительно, объектами изучения являются полиномы и матрицы – основные понятия в алгебре. Однако добавление даже одного запаздывания сильно ломает структуру задачи.

В работах [14, 15] сформулированы обобщения методов Ляпунова и показан функциональный подход для исследования устойчивости в случае системы с запаздываниями, также в работе [16] показан критерий устойчивости с конечным числом шагов. Данные результаты имеют огромную теоретическую значимость, являются почвой для формулирования новых теорем и решают большой класс задач. Основным недостатком этих подходов является конструктивная проверка положительной определенности функционала и экспоненциальная вычислительная сложность.

В данной статье рассмотрен менее общий случай, а именно линейная дифференциальная система с соизмеримыми запаздываниями. Показан надежный алгоритм с полиномиальной вычислительной сложностью.

2. Постановка задачи. Рассмотрим линейную стационарную систему с соизмеримыми запаздываниями

$$\dot{x}(t) = \sum_{k=0}^m A_k x(t - kh), \quad t \geq 0,$$

где $h > 0$ – запаздывание, $A_k \in \mathbb{R}^{n \times n}$ – вещественные матрицы.

Гапонов Дмитрий Николаевич – магистрант, Санкт-Петербургский государственный университет; e-mail: st076870@student.spbu.ru, тел.: +7(911)686-20-18

Этой системе соответствует квазиполином вида

$$P(s) = \det \left(sI - A_0 - \sum_{k=1}^m A_k \exp(-skh) \right).$$

Определение 1. Квазиполином $P(s)$ называют устойчивым в смысле Гурвица, если все его корни расположены строго в левой полуплоскости комплексной плоскости.

Цель настоящей работы – сформулировать алгоритм для нахождения всех h , при которых квазиполином $P(s)$ имеет свойство устойчивости в смысле Гурвица.

3. Предварительные сведения. Введем базовые утверждения.

Лемма 1 [17]. Величина $s = i\hat{\omega}$ является решением уравнения

$$\det \left(i\hat{\omega}I - A_0 - \sum_{k=1}^m A_k \exp(-i\hat{\omega}kh) \right) = 0, \quad \hat{\omega} > 0,$$

тогда и только тогда, когда

$$\det \left(i\hat{\omega}I - A_0 - \sum_{k=1}^m A_k \left(\frac{1 - i\hat{\omega}\hat{T}}{1 + i\hat{\omega}\hat{T}} \right)^{2k} \right) = 0$$

при $h = \frac{4}{\hat{\omega}} \left(\arctg(\hat{\omega}\hat{T}) + j\frac{\pi}{2} \right)$, $j \in \mathbb{Z}^+$, $\hat{T} > 0$.

Лемма 2. Задача хранения полинома от двух переменных в памяти компьютера равносильна задаче хранения взвешенного неориентированного графа.

Действительно, пусть имеется полином от двух переменных $f(x, y) = \sum_{i=0}^n \sum_{j=0}^m a_{ij} x^i y^j$. Тогда слагаемое вида $a_{ij} x^i y^j$ можно воспринимать как две вершины x^i и y^j , связанные неориентированным ребром с весом a_{ij} . Получаем однозначное соответствие.

Известно [18], что хранить граф можно тремя способами:

- 1) матрица смежности – за $O(|V|^2)$ по памяти;
- 2) список смежности – за $O(|V| + |E|)$ по памяти;
- 3) список ребер – за $O(|E|)$ по памяти.

Способ выбирается в зависимости от плотности графа. Будем считать, что перед каждым элементом $x^i y^j$ стоит коэффициент, тогда стоит хранить в виде матрицы смежности за $O(nm)$ по памяти.

Определение 2. Пусть имеется два полинома от двух переменных $f(x, y), g(x, y) \in \mathbb{R}[x, y]$, $\deg_y(f) = n$, $\deg_y(g) = m$, $n \geq m$. Разложим их по степеням переменной y :

$$f(x, y) = \sum_{i=0}^n f_i(x) y^i, \quad g(x, y) = \sum_{j=0}^m g_j(x) y^j.$$

Составим матрицу $B(x) = (b_{ij}(x))_{i,j=0}^n$ из коэффициентов

$$b_{ij}(x) = \sum_{k=0}^{\min(i, n-1-j)} (f_{j+k+1}(x) g_{i-k}(x) - f_{i-k}(x) g_{j+k+1}(x)).$$

Тогда $\text{Res}_y(f(x, y), g(x, y)) = (-1)^{n(n-1)/2} f_n^{n-m}(x) \det B(x)$ будем называть результатом полиномов $f(x, y)$ и $g(x, y)$ по переменной y .

Замечание 1. Наивное произведение/деление двух полиномов $f(x)$ и $g(x)$, $\deg(f) = n$, $\deg(g) = m$, $n \geq m$, может быть выполнено за время порядка $O(nm)$. Используя [19], можно улучшить асимптотику до $O(n \log(n) \log(\log(n)))$.

Лемма 3. Наивное вычисление результата (по определению 2) двух полиномов $f(x, y)$ и $g(x, y)$, $\deg_y(f) = n$, $\deg_y(g) = m$, $n \geq m$, имеет асимптотику порядка $O(n^7)$ по времени и $O(n^4)$ по памяти.

Доказательство. Пусть произведение/деление двух полиномов $f(x)$ и $g(x)$, $\deg(f) = n$, $\deg(g) = m$, $n \geq m$, выполняется наивно за время порядка $O(nm)$, тогда построение матрицы $B(x)$ выполняется за время порядка $T_B(n) = \sum_{i=0}^n \sum_{j=0}^m \sum_{k=0}^r O(n^2) = O(n^5)$,

где $r = \min(i, n-1-j)$. Получаем $(n \times n)$ -матрицу, элементами которой являются полиномы одной переменной степени порядка $O(n)$, соответственно она занимает порядка $O(n^3)$ памяти. Оценим вычисление $\det B(x)$. Используя алгоритм Барейса–Монтанте [20], имеем три вложенных цикла, в которых выполняем произведение/деление полиномов. Изначально имеем полиномы степени $2n$ в худшем случае, но на каждой итерации степень будет расти, учтем это. Тогда $T_{\det}(n) = \sum_{k=0}^n (n-k)^2 2(k(n-1)+1)^2 +$

$+2((k+1)(n-1)+1)((k-1)(n-1)+1)) = O(n^7)$. В итоге получим полином степени $O(n^2)$, значит вычисление занимает порядка $O(n^4)$ по памяти.

Замечание 2. В статье [21] рассмотрены другие алгоритмы взятия определителя матрицы, элементами которой являются полиномы и проведен более детальный асимптотический анализ.

Замечание 3. В статье [22] представлен алгоритм вычисления результата от полиномов двух переменных с более хорошей асимптотикой, не хуже чем $O(n^3 \log^2(n) \log(\log(n)))$.

Теорема 1 [23]. *Зафиксируем $x = \alpha$, полиномы от переменной y : $f(\alpha, y)$ и $g(\alpha, y)$ имеют общий корень тогда и только тогда, когда $\text{Res}_y(f(\alpha, y), g(\alpha, y)) = 0$.*

4. Пересечение мнимой оси. Рассмотрим квазиполином

$$P(s) = \det \left(sI - A_0 - \sum_{k=1}^m A_k \exp(-skh) \right).$$

Не умаляя общности, пусть при $h = 0$ он устойчив в смысле Гурвица. Если мы начнем увеличивать h , то квазиполином может стать неустойчивым, т. е. один из его корней пересечет мнимую ось и окажется в правой полуплоскости комплексной плоскости. Тогда, если можно определить при каком \hat{h} корень квазиполинома стал чисто мнимым, то при $\forall h \in [0, \hat{h})$ квазиполином $P(s)$ устойчив в смысле Гурвица.

Замечание 4. Используя $\frac{ds}{dh} = -\frac{\partial P(s)}{\partial h} / \frac{\partial P(s)}{\partial s}$, можно находить интервалы $(\hat{h}_{j-1}, \hat{h}_j)$, так как зная $\text{sign} \left(\text{Re} \left(\frac{ds}{dh} \Big|_{\substack{s=i\hat{\omega}_j \\ h=\hat{h}_j}} \right) \right)$ можно определить, в какую полуплоскость перешел корень.

5. Переход к алгебре. Имеем уравнение

$$\det \left(sI - A_0 - \sum_{k=1}^m A_k \exp(-skh) \right).$$

Используя лемму 1, получим

$$\det \left(sI - A_0 - \sum_{k=1}^m A_k \left(\frac{1-sT}{1+sT} \right)^{2k} \right) = 0.$$

Преобразуем

$$\begin{aligned} \det \left((sI - A_0)(1 + sT)^{2m} - \sum_{k=1}^m A_k (1 - sT)^{2k} (1 + sT)^{2m-2k} \right) = \\ = \sum_{i=0}^{n(1+2m)} \sum_{j=0}^{2nm} p_{ij} s^i T^j = f(s, T). \end{aligned}$$

Получаем полином от двух переменных. Рассмотрим его на мнимой оси, т. е. $s = iw$, $w > 0$,

$$f(iw, T) = F_r(w^2, T) + iwF_i(w^2, T).$$

Замечание 5. В статье [24] на этом этапе начинают применять численный метод для решения аналогичного уравнения.

Таким образом, уравнение $f(iw, T) = 0$ сводится к системе

$$\begin{cases} F_r(w^2, T) = 0, \\ F_i(w^2, T) = 0. \end{cases}$$

Мы можем «избавиться» от w^2 , используя теорию исключений [23]. Применим результат

$$\text{Res}_{w^2} (F_r(w^2, T), F_i(w^2, T)) = g(T).$$

Осталось только найти вещественные решения уравнение $g(T) = 0$, в левой части которого стоит полином степени $2n^2m$. Затем подставить полученные корни в систему, найти вещественные корни системы и получить пары решений (\hat{w}_k, \hat{T}_k) . Тогда искомое запаздывание, при котором нарушается устойчивость, $\hat{h} = \min_k \left\{ \frac{4}{\hat{\omega}_k} \arctg(\hat{\omega}_k \hat{T}_k) \right\}$.

Замечание 6. Безусловно, уравнение $g(T) = 0$ не решается алгебраическими методами, но эта задача проще исходной, так как существует множество оценок вещественных корней, способов локализации корней и численных итеративных методов поиска вещественных корней полинома.

Замечание 7. В статье [25] предложен иной алгебраический способ нахождения искомого \hat{h} .

6. Описание алгоритма. Для определения \hat{h} необходимо

- 1) сделать замену из леммы 1 и сформировать $f(s, T)$;
- 2) перейти к системе из двух полиномов от двух переменных;
- 3) применить результат к системе по переменной w^2 и перейти к уравнению $g(T) = 0$;
- 4) найти корни $g(T) = 0$, подставить их в систему и получить пары решений;
- 5) используя соотношение из леммы 1, получить искомое \hat{h} .

Самая трудоемкая часть алгоритма – это вычисление результата. Если вычислять результат наивно, то получим асимптотику порядка $O(n^7 m^7)$ по времени, но используя [22], можно добиться асимптотики не хуже $O(n^3 m^3 \log^2(nm) \log(\log(nm)))$.

Замечание 8. Данный подход можно обобщить для случая систем с несколькими несоизмеримыми запаздываниями, тогда каждое $\exp(-sk\tau_j)$ необходимо заменить на соответствующее $\left(\frac{1-sT_j}{1+sT_j}\right)^{2k}$. В итоге получим $f(iw, T_0, T_1, \dots, T_m)$. К сожалению, исследовать такой полином вычислительно трудно, поэтому имеет смысл применять данный алгоритм, когда количество τ_j невелико.

7. Заключение. В работе предложен алгоритм определения запаздываний, при которых линейная стационарная система с соизмеримыми запаздываниями имеет свойство устойчивости в смысле Гурвица. Проведен асимптотический анализ данного алгоритма.

Литература

1. Hurwitz A. Ueber die Bedingungen, unter welchen eine Gleichung nur Wurzeln mit negativen reellen Theilen besitzt // *Mathematische Annalen*. 1895. Vol. 46. P. 273–284.
2. Routh E. J. *A Treatise on the Stability of a Given State of Motion: Particularly Steady Motion*. London: Macmillan, 1877.
3. Strelitz S. On the Routh–Hurwitz problem // *American Mathematical Monthly*. 1977. Vol. 84. P. 542–544.
4. Харитонов В. Л. Об асимптотической устойчивости положения равновесия семейства систем линейных дифференциальных уравнений // *Дифференциальные уравнения*. 1978. Т. 10. № 11. С. 2086–2088.

5. Bartlett A. C., Hollot C. V., Huang L. Root locations of an entire polytope of polynomials: It suffices to check the edges // *Mathematics of Control, Signals and Systems*. 1988. Vol. 1. P. 61–71.
6. Xie X. A new criterion of linear system stability // *Special Issue on Basic Theory of Transaction of Northeast Institute of Technology of China*. 1963. Vol. 1. P. 26–30.
7. Nie Y. A new criterion of Hurwitz polynomial // *Mechanics*. 1976. Vol. 2. P. 110–116.
8. Липатов А. В., Соколов Н. И. О некоторых достаточных условиях устойчивости и неустойчивости линейных непрерывных стационарных систем // *Автоматика и телемеханика*. 1978. Вып. 9. С. 30–37.
9. Масленников В. В. Гипотеза о существовании простого аналитического достаточного условия устойчивости // *Автоматика и телемеханика*. 1984. Вып. 2. С. 160–161.
10. Клепцын А. Ф. Об одном достаточном условии устойчивости многочлена // *Автоматика и телемеханика*. 1984. Вып. 10. С. 175–177.
11. Nie Y., Xie X. New criteria for polynomial stability // *IMA Journal of Mathematical Control and Information*. 1987. Vol. 4. No 1. P. 1–12.
12. Hou X., Wang X. A generalization of Xie–Nie stability criterion: *Computer Algebra and Geometric Algebra with Applications*. Berlin: Springer, 2004. P. 65–71.
13. Жабко И. А. Достаточные условия устойчивости полиномов с вещественными коэффициентами // *Актуальные проблемы математики и информатики. Труды математического факультета*. 2010. № 1. С. 52–62.
14. Kharitonov V. L., Zhabko A. P. Lyapunov–Krasovskij approach to the robust stability analysis of time-delay systems // *Automatica*. 2003. Vol. 30. No 1. P. 15–20.
15. Kharitonov V. L. *Time-delay systems: Lyapunov functionals and matrices*. Basel: Birkhäuser, 2013. 311 p.

16. Mondie S., Egorov A., Gomez M. A. Lyapunov stability tests for linear time-delay systems // *Annual Reviews in Control*. 2022. Vol. 54. P. 68–80.
17. Thowsen A. The Routh-Hurwitz method for stability determination of linear differential-difference systems // *International Journal of Control*. 1981. Vol. 33. No 5. P. 991–995.
18. Cormen T., Leiserson C., Rivest R., Stein C. *Introduction to Algorithms*, Fourth Edition. Cambridge, Massachusetts: MIT Press, 2022. 1312 p.
19. Cantor D. G., Kaltofen E. On fast multiplication of polynomials over arbitrary algebras // *Acta Informatica*. 1991. Vol. 28. P. 693–701.
20. Bareiss E. H. Sylvester’s identity and multistep integer preserving Gaussian elimination // *Mathematics of Computation*. 1968. Vol. 22. No 103. P. 565–578.
21. Horowitz E., Sahni S. On computing the exact determinant of matrices with polynomial entries // *Journal of the ACM*. 1975. Vol. 22. No 1. P. 38–50.
22. Villard G. On computing the resultant of generic bivariate polynomials // *ISSAC’18: Proceedings of the 2018 ACM International Symposium on Symbolic and Algebraic Computation*. 2018. No 43. P. 391–398.
23. Утешев А. Ю., Калинина Е. А. *Теория исключения: Учеб. пособие*. СПб: НИИ химии СПбГУ, 2002. 72 с.
24. Ebenbauer C., Allgower F. Stability analysis for time-delay systems using Rekasius’s substitution and sum of squares // *IEEE Conference on Decision and Control*. 2007. No 45. P. 5376–5381.
25. Харитонов В. Л. К определению максимально допустимого запаздывания в задаче стабилизации // *Дифференциальные уравнения*. 1982. Т. 18. № 4. С. 723–724.

УДК 517.977

Каминцева В. С.

Реализация метода синтеза оптимального управления для динамических объектов

Рекомендовано к публикации профессором Смирновым Н. В.

1. Введение. Проблемы оптимального управления интенсивно исследуются в мировой научной литературе с середины двадцатого века. За это время результаты данной теории нашли свое применение в различных областях науки, техники и экономики. И это влияние продолжает возрастать, становясь все более значимым.

Одним из способов решения проблемы построения оптимального управления является подход Р. Габасова [1], в основе которого лежит алгоритм сведения динамической задачи оптимального управления к интервальной задаче линейного программирования (ИЗЛП). Используя данный подход, несколько изменяя и дополняя его в зависимости от специфики задачи и требуемых результатов, решается множество прикладных проблем [2–6].

В данной статье рассматривается подход Габасова нахождения оптимальных управлений для линейной неоднородной системы произвольной размерности. Результаты иллюстрируются задачей оптимального управления электродвигателем.

2. Общая постановка задачи. Каноническая задача оптимального управления. Следуя работе [1], рассмотрим метод построения оптимальных управлений в линейных системах. В классе дискретных управлений рассмотрим линейную задачу терминального управления

$$\begin{aligned} J(u) = c^T x(t^*) \longrightarrow \max, \quad \dot{x} = A(t)x + b(t)u + q(t), \quad x(t_*) = x_0, \\ Hx(t^*) = g, \quad |u(t)| \leq 1, \quad t \in T, \\ x \in \mathbb{R}^n, \quad u \in \mathbb{R}, \quad g \in \mathbb{R}^n, \quad \text{rank} H = m \leq n, \end{aligned} \quad (1)$$

где $A(t), b(t), q(t), t \in T$, – кусочно-непрерывные $(n \times n)$ -матричная и n -векторные функции, $T = [t_*, t^*]$, $t_* < t^* < \infty$.

Каминцева Виктория Сергеевна – студент, Санкт-Петербургский государственный университет; e-mail: st097171@student.spbu.ru, тел.: +7(904)399-49-72

Положим $h = (t_* - t^*)/N$, N – натуральное число. Функцию $u(t), t \in T$, назовем *дискретным управлением* (с периодом дискретизации h), если $u(t) = u(t_* + kh) = u_{kh}$, $t \in [t_* + kh, t_* + (k+1)h]$, $k = \overline{0, N-1}$.

Данная задача является *канонической задачей* оптимального управления. Размерность n вектора состояний $x(t)$ называется порядком задачи.

3. Переход к интервальной задаче линейного программирования. Модифицируем задачу (1), чтобы иметь возможность использовать методы линейного программирования [1].

Используя общую формулу Коши

$$x(t, t_*, x_0) = Y(t)[Y^{-1}(t_*)x_0 + \int_{t_*}^t Y^{-1}(\tau)(b(\tau)u(\tau) + q(\tau)) d\tau] \quad (2)$$

преобразуем функционал, подставив значение программного движения в конечный момент времени $x(t^*) = x(t^*, t_*, x_0)$. С учетом (2), в результате преобразований новый функционал примет вид

$$\int_{t_*}^{t^*} c^T Y(t^*)Y^{-1}(\tau)b(\tau)u(\tau) d\tau \longrightarrow \max. \quad (3)$$

Также подставим значение $x(t^*)$ в терминальные ограничения. Тогда они примут вид

$$\int_{t_*}^{t^*} G(\tau)b(\tau)u(\tau) d\tau = g_0, \quad (4)$$

где $G(\tau) = HY(t^*)Y^{-1}(\tau)$ – матрица размерности $(m \times n)$, а $g_0 = g - HY(t^*)[Y^{-1}(t_*)x_0 - \int_{t_*}^{t^*} Y^{-1}(\tau)q(\tau) d\tau]$ – m -мерный постоянный вектор.

Поскольку в данной задаче рассматривается управление из класса кусочно-постоянных функций, то интеграл из функционала (3) и интеграл из равенства (4) разобьются на сумму интегралов вида

$$\sum_{t \in T_u} u(t) \int_t^{t+h} c^T Y(t^*)Y^{-1}(\tau)b(\tau) d\tau \longrightarrow \max, \quad (5)$$

$$\sum_{t \in T_u} u(t) \int_t^{t+h} G(\tau)b(\tau) d\tau = g_0 \quad (6)$$

соответственно, где $T_u = \{t_*, t_* + h, \dots, t^* - h\}$.

Обращая внимание на формулы (5) и (6), можно заметить, что все интегралы по конечным элементарным промежуткам зависят лишь от известных функций, и их можно вычислить заранее. Обозначим

$$c_h(t) = \int_t^{t+h} c^T Y(t^*) Y^{-1}(\tau) b(\tau) d\tau, \quad t \in T_u, \quad (7)$$

$$d_h(t) = \int_t^{t+h} G(\tau)b(\tau) d\tau, \quad t \in T_u. \quad (8)$$

Введем обозначения: $U = (u_0, u_h, \dots, u_{kh}, \dots, u_{(N-1)h})^T$ – N -мерный вектор, состоящий из управлений на каждом промежутке дискретизации; $C = (c_h(t_*), \dots, c_h(t_* + kh), \dots, c_h(t^* - h))^T$ – N -мерный вектор, состоящий из величин (7), вычисленных в соответствующие моменты времени; $D = (d_h(t_*), \dots, d_h(t_* + kh), \dots, d_h(t^* - h))$ – $(m \times N)$ -мерная матрица, состоящая из векторов (8), вычисленных в соответствующие моменты времени.

В результате задача (1) сведена к задаче линейного программирования

$$\begin{cases} C^T U \rightarrow \max, \\ DU = g_0, \\ |u_i| \leq 1, \quad i = \overline{1, N}. \end{cases} \quad (9)$$

4. Управление электродвигателем. В качестве примера рассмотрим задачу оптимального управления электродвигателем [7]. Уравнение двигателя постоянного тока имеет вид

$$I\ddot{\varphi} = i_{\text{я}} k_{\Phi} \Phi - M_c,$$

где I – момент инерции вращающейся части двигателя, φ – угол поворота вала двигателя, $i_{\text{я}}$ – ток в якорной цепи, k_{Φ} – конструктивная постоянная, Φ – магнитный поток, M_c – момент сопротивления.

Используя обозначения

$$b = k_{\Phi}\Phi/I, \quad x_1 = \varphi, \quad x_2 = \dot{\varphi}, \quad u = i_{я}, \quad q_c = M_c/I,$$

его можно записать в векторной форме

$$\dot{x} = Ax + Bu + q, \quad (10)$$

где

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ b \end{pmatrix}, \quad q = \begin{pmatrix} 0 \\ -q_c \end{pmatrix}.$$

Рассмотрим задачу поворота вала двигателя на заданный угол за время T при минимальном расходе энергии с последующей его остановкой. В качестве критерия оптимальности рассмотрим функционал

$$\int_{t_*}^{t^*} u(t) dt \longrightarrow \min, \quad (11)$$

где t_* и t^* фиксированы, $t^* - t_* = T$. Краевые условия в этом случае примут вид

$$x_1(t_*) = x_1^0, \quad x_2(t_*) = x_2^0, \quad x_1(t^*) = x_1^*, \quad x_2(t^*) = 0. \quad (12)$$

Задачу (10)–(12) можно свести к виду (1), при этом многообразии $Hx(t^*) = g$ будет иметь вид

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1(t^*) \\ x_2(t^*) \end{pmatrix} = \begin{pmatrix} x_1^* \\ 0 \end{pmatrix}. \quad (13)$$

Итоговая задача линейного программирования (9) примет вид

$$\begin{cases} C^T U \longrightarrow \min, \\ DU = g_0, \end{cases} \quad (14)$$

где $C^T = (h \dots h)$ – N -мерный вектор, а D – $(m \times N)$ -мерная матрица, вычисленная способом, описанным в предыдущем пункте.

5. Численный эксперимент. Для реализации и тестирования алгоритма были выбраны следующие параметры:

$$\Phi = 0,07 \text{ Вб}, \quad k_{\Phi} = 28,66, \quad I = 2 \text{ кг} \cdot \text{м}^2, \quad M_c = 2 \text{ Н} \cdot \text{м},$$

начальное и конечное положения: $x(t_*) = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$, $x(t^*) = \begin{pmatrix} 6 \\ 0 \end{pmatrix}$. Задача рассмотрена на промежутке времени $[t_*, t^*] = [0, 10]$. Этот интервал разбит на $N = 6$ равных временных отрезков и на нем реализовано кусочно-постоянное управление.

Результат применения алгоритма для данной задачи можно наглядно оценить с помощью графиков (см. рис. 1, 2).

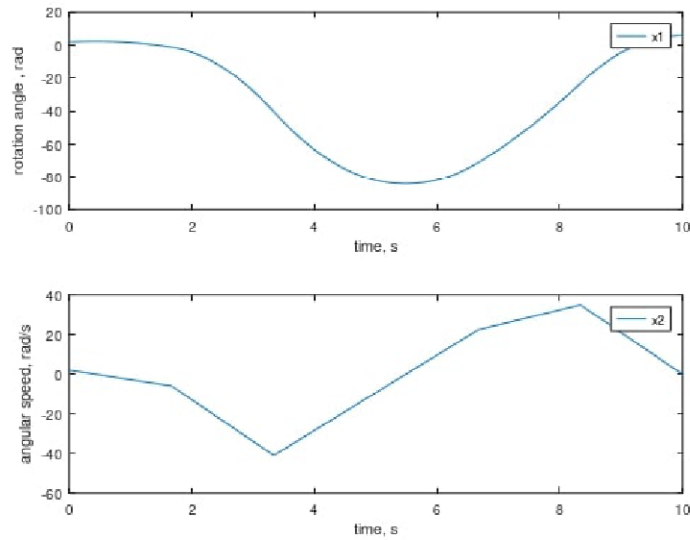


Рис. 1. Графики изменения положения вала и его скорости

На верхнем графике рис. 1 показано изменение положения вала из начального в конечное состояние при реализации оптимального управления. На нижнем графике отражено изменение угловой скорости. Можно убедиться, что конечные значения соответствуют требуемым значениям из постановки задачи.

На рис. 2 представлен график найденного оптимального управления при заданных условиях. Как видно из графиков, найденное управление за отведенное время переводит вал из начального положения в конечное с последующей остановкой.

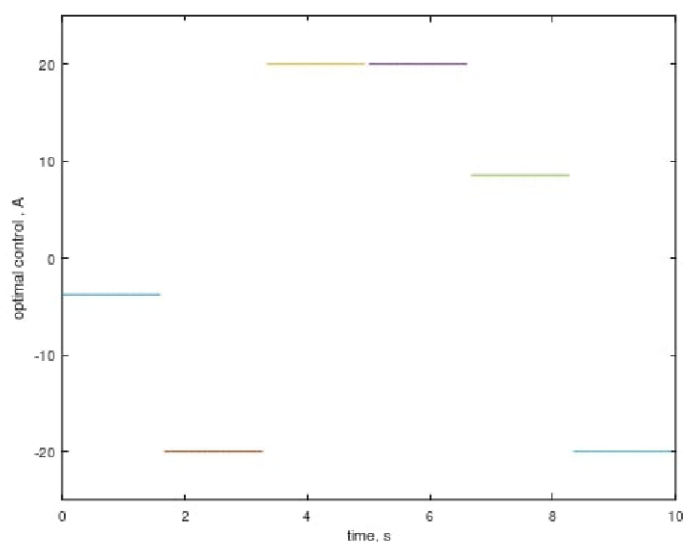


Рис. 2. График оптимального управления

6. Заключение. Рассмотренный в данной статье подход может применяться для широкого класса задач, при этом изменение параметров внутри самой задачи (например, число точек переключения управления, начальных данных и т. д.) не потребует каких-либо изменений в алгоритме и программной реализации, что является его явным преимуществом.

Алгоритм очень хорошо показал себя на линейных системах с линейным функционалом, однако при наличии нелинейности данный метод в его канонической форме применить нельзя, потребуются существенная модификация [8]. Несмотря на это, подход Габасова имеет очень высокий потенциал для развития теории в разных направлениях: нахождение управления с обратной связью, управление в дискретных системах и т. д.

В дальнейшем планируется исследовать проблему нахождения управления с обратной связью, решение которой также строится на основе данного подхода.

Литература

1. Балашевич Н. В., Габасов Р., Кириллова Ф. М. Численные методы программной и позиционной оптимизации линейных систем управления // Журнал вычислительной математики и математической физики. 2000. Т. 40. № 6. С. 836–859.
2. Смирнов Н. В., Пересада В. П., Гирдюк Д. В., Постнов К. В., Попков А. С. Модель межотраслевого баланса – один из базовых элементов информационного обеспечения работы ситуационных центров регионов // Информатизация и связь. 2019. № 3. С. 20–25.
3. Межотраслевой баланс: анализ динамики и управление макроэкономическими тенденциями : учебное пособие для вузов / Н. В. Смирнов, В. П. Пересада, Т. Е. Смирнова [и др.] ; Под редакцией Н. В. Смирнова. Санкт-Петербург : Лань, 2021. 180 с.
4. Баранов О. В., Попков А. С., Смирнов Н. В. Оптимальная стабилизация квадрокоптера в режиме реального времени // Устойчивость и процессы управления: Материалы III международной конференции. СПб.: Издательский дом Федоровой Г. В, 2015. С. 115–116.
5. Смирнов Н. В., Соловьева И. В. Применение метода позиционной оптимизации для многопрограммной стабилизации билинейных систем // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2009. № 3. С. 251–259.
6. Смирнов Н. В., Смирнова Т. Е. Синтез многопрограммных управлений в билинейных системах // Прикладная математика и механика. 2000. Т. 64. № 6. С. 929–932.
7. Теория автоматического управления. Ч. 2. Теория нелинейных и специальных систем автоматического управления. Изд. 2-е / под ред. А. А. Воронова. М.: Высшая школа, 1986. 504 с.
8. Балашевич Н. В., Габасов Р., Калинин А. И., Кириллова Ф. М. Оптимальное управление нелинейными системами // Журнал вычислительной математики и математической физики. 2002. Т. 42. № 7. С. 969–995.

UDC 517.95

Makoveeva P. E., Nesterov S. D.

Lyapunov – Krasovskii functionals for a linear system with distributed parameters

Recommended for publication by associate professor Egorov A. V.

1. Introduction. An important problem in the study of retarded type delay differential equations is the analysis of stability of solutions. One of the methods for such analysis is based on the Lyapunov – Krasovskii functionals and has been widely applied in various fields.

In this paper, the method is applied for the delay dynamical system with distributed parameters. Via the Lyapunov matrix, the functional with a prescribed derivative is constructed, that allows analyze stability and robust stability. We show that for the considered system the Lyapunov matrix is represented as a trigonometric series based on the formula in article [2].

2. Problem statement. Some mathematical models describing the population dynamics [3] can be represented by a dynamical system of the following form:

$$\begin{cases} u_t(x, t) = au_{xx}(x, t) + bu(x, t - h), & x \in (0, l), \quad t \geq 0, \\ u(0, t) = u(l, t) = 0, & t \geq 0, \\ u(x, \theta) = \varphi(x, \theta), & \theta \in [-h, 0], \quad x \in [0, l]. \end{cases} \quad (1)$$

It consists of a partial differential equation and homogeneous boundary and corresponding initial conditions. Here φ is the initial function, for simplicity we assume that it is absolutely continuous, $a > 0$, $b \neq 0$ are the model parameters, $h > 0$ is the delay. In what follows, we assume the length of the segment $l = 1$, since the problem can be reduced to this case by scaling the coordinate x .

Makoveeva Polina Evgenevna – graduate student, Saint Petersburg State University; e-mail: st076937@student.spbu.ru, phone: +7(912)241-52-32

Nesterov Sergey Dmitrievich – graduate student, Saint Petersburg State University; e-mail: st071609@student.spbu.ru, phone: +7(910)822-73-48

The work was performed with financial support of Russian Science Foundation, grant No. 23-71-10099

3. Solution. To find the solution of system (1), we apply the Fourier method of separation of variables. Let $u(x, t) = f(x)g(t)$. Substituting this product into the nominal system, we obtain two equations. The first one is an ordinary differential equation (ODE) of the second order, the second one is a differential-difference equation. Thus, we obtain the system

$$\begin{cases} af''(x) = -\mu f(x), \\ g'(t) - bg(t-h) = -\mu g(t), \end{cases}$$

where μ is a constant. First ODE with the boundary conditions $f(0) = f(l) = 0$ has the solutions $f_i(x) = \sin(\pi ix)$, $i \in \mathbb{N}$.

3.1. Solution of the differential-difference equation. The solution of original system (1) takes the following form:

$$u(x, t) = \sum_{i=1}^{\infty} g_i(t) \sin(\pi ix), \quad (2)$$

where g_i are the coefficients of the Fourier series for the solution u . In order to satisfy the initial conditions, we need to require the following equalities: $\varphi_i(\theta) = g_i(\theta)$, $i = 1, 2, \dots$, $\theta \in [-h, 0]$. Here $\varphi_i(\theta)$ are coefficients of the Fourier series for the function φ , which can be constructed by the formula

$$\varphi_i(\theta) = 2 \int_0^1 \sin(\pi ix) \varphi(x, \theta) dx, \quad i \in \mathbb{N}, \quad \theta \in [-h, 0].$$

Thus, we need to solve the following system:

$$\begin{cases} \dot{g}_i(t) = -\mu_i g_i(t) + b g_i(t-h), & t \geq 0, \\ g_i(\theta) = \varphi_i(\theta), & \theta \in [-h, 0]. \end{cases} \quad (3)$$

Definition 1. $K_i(t)$, $t \in [-h, \infty)$, is said to be the fundamental solution if it satisfies equation

$$\frac{dK_i(t)}{dt} = -\mu_i K_i(t) + b K_i(t-h), \quad t \geq 0,$$

and $K_i(t) = 0$ for all $t \in [-h, 0]$, $K_i(0) = 1$.

Lemma 1. The solution of system (3) can be represented by the Cauchy formula

$$g_i(t, \varphi) = K_i(t) \varphi_i(0) + b \int_{-h}^0 K_i(t-h-\theta) \varphi_i(\theta) d\theta, \quad t \geq 0. \quad (4)$$

We substitute the obtained Cauchy formula (4) into equation (2). Then the solution of system (1) takes the form

$$u(x, t, \varphi) = \sum_{i=1}^{\infty} \sin(\pi i x) \left[K_i(t) \varphi_i(0) + b \int_{-h}^0 K_i(t-h-\theta) \varphi_i(\theta) d\theta \right].$$

The function $u(\cdot, t, \varphi)$ is from the Lebesgue space $L_2((0, 1))$.

4. Construction of the functional v_0 . In this paragraph, an explicit formula for the functional is derived, as well as the Lyapunov matrix, which is the key element of the functional.

Definition 2. Derivative along the solutions of system (1) is determined by the following formula:

$$\left. \frac{d}{dt} v_0(t_0, \varphi) \right|_{(1)} = \limsup_{\alpha \rightarrow +0} \frac{v_0(u_\alpha(\varphi)) - v_0(\varphi)}{\alpha},$$

where $u_t(\varphi) : \theta \in [-h, 0] \rightarrow u(\cdot, \theta, \varphi)$.

Theorem 1. Let system (1) be exponentially stable. If a non-positive quadratic derivative along the solutions of the system (1) is

$$\left. \frac{d}{dt} v_0(\varphi) \right|_{(1)} = - \|\varphi(\cdot, 0)\|_{\mathcal{L}_2}^2, \quad (5)$$

then the functional has the form

$$\begin{aligned} v_0(\varphi) &= \int_0^1 \varphi(y_2, 0) \int_0^1 U(0, y_1, y_2) \varphi(y_1, 0) dy_1 dy_2 \\ &+ 2b \int_0^1 \varphi(y_1, 0) \int_0^1 \int_{-h}^0 U(-h-\theta, y_1, y_2) \varphi(y_2, \theta) d\theta dy_2 dy_1 \\ &+ b^2 \int_0^1 \int_{-h}^0 \varphi(y_1, \theta_1) \int_0^1 \int_{-h}^0 U(\theta_1-\theta_2, y_1, y_2) \varphi(y_2, \theta_2) d\theta_2 dy_2 d\theta_1 dy_1, \end{aligned} \quad (6)$$

where the Lyapunov matrix has the following form:

$$U(\tau, y_1, y_2) = 2 \sum_{i=1}^{\infty} \sin(\pi i y_1) \sin(\pi i y_2) \int_0^{\infty} K_i(t) K_i(t+\tau) dt, \quad (7)$$

$$\tau \in \mathbb{R}, \quad y_1 \in [0, 1], \quad y_2 \in [0, 1].$$

Proof. Consider functional v_0 which satisfies the condition of the theorem.

We integrate equality (5) on the segment $[0, T]$. As T go to infinity, we obtain

$$v_0(\varphi) = \int_0^\infty \int_0^1 \left(\sum_{i=1}^\infty \sin(\pi i x) \left[K_i(t) \varphi_i(0) + b \int_{-h}^0 K_i(t-h-\theta) \varphi_i(\theta) d\theta \right] \right)^2 dx dt.$$

With the change of order of integration and some simple transformation we arrive at the sum of three terms: $v_0(\varphi) = R_1 + R_2 + R_3$. Next, we consider separately each terms that was obtained by squaring the integrand. We assume convergence of all series, it is necessary to change order of summation and integration.

To construct a the functional with a prescribed derivative, we will apply an idea similar to the one presented in [4] for time delay equation without partial derivatives.

For R_1 double summation reduces to unary summation due to the orthogonality of the system of sinuses. The Lyapunov matrix consists of the series and the improper integral.

Final form of the first term:

$$\begin{aligned} R_1 &= 2 \int_0^1 \varphi(y_2, 0) \int_0^1 \sum_{i=1}^\infty \left[\sin(\pi i y_1) \sin(\pi i y_2) \int_0^\infty K_i^2(t) dt \right] \varphi(y_1, 0) dy_1 dy_2 \\ &= \int_0^1 \varphi(y_2, 0) \int_0^1 U(0, y_1, y_2) \varphi(y_1, 0) dy_1 dy_2. \end{aligned}$$

Now consider R_2 – the second term of the functional v_0 . We apply transformations similar to those performed in R_1 and obtain

$$R_2 = 2b \int_0^1 \varphi(y_1, 0) \int_0^1 \int_{-h}^0 U(-h-\theta, y_1, y_2) \varphi(y_2, \theta) d\theta dy_2 dy_1.$$

Repeat the same steps for the third term

$$R_3 = b^2 \int_0^1 \int_{-h}^0 \varphi(y_1, \theta_1) \int_0^1 \int_{-h}^0 U(\theta_1 - \theta_2, y_1, y_2) \varphi(y_2, \theta_2) d\theta_2 dy_2 d\theta_1 dy_1.$$

Thus, we have proved that a functional with a non-positive quadratic derivative can be represented as a sum $v_0(\varphi) = R_1 + R_2 + R_3$.

5. Lyapunov matrix. As can be seen from the presented theorem, the Lyapunov matrix U plays a key role in the construction of the functional v_0 with a prescribed derivative. It is very difficult to calculate it directly by formula (7). This is due to the fact that the Lyapunov matrix contains an improper integral and a series. The following theorem simplifies the problem considerably. For simplicity, here we consider the case $h = 1$, to which the original problem can be reduced by time scaling.

Theorem 2. *The Lyapunov matrix (7) is represented as the following series absolutely and uniformly convergent on the set $\{\tau \in [-1, 1], y_1, y_2 \in [0, 1]\}$:*

$$U(\tau, y_1, y_2) = \sum_{i=1}^{\infty} \sin(\pi i y_1) \sin(\pi i y_2) u_i(|\tau|).$$

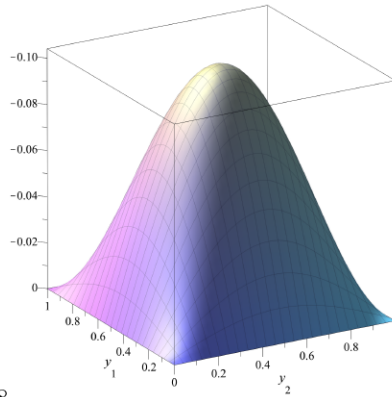
Here, each of the functions $u_i(\tau)$, $\tau \in [0, 1]$, $i \in \mathbb{N}$, is defined by the corresponding formula from Theorem 2 in [2], where the number $-\mu_i$ is taken as the parameter a . In particular, for sufficiently large i such that $\mu_i > |b|$ and $|b| \cosh \lambda_i > \mu_i$, where $\lambda_i = \sqrt{\mu_i^2 - b^2}$, $\mu_i = -a(\pi i)^2$,

$$u_i(\tau) = \frac{b \sinh \lambda_i (1 - \tau) + \mu_i \sinh \lambda_i \tau - \lambda_i \cosh \lambda_i \tau}{\lambda_i (b \cosh \lambda_i - \mu_i)}, \quad \tau \in [0, 1].$$

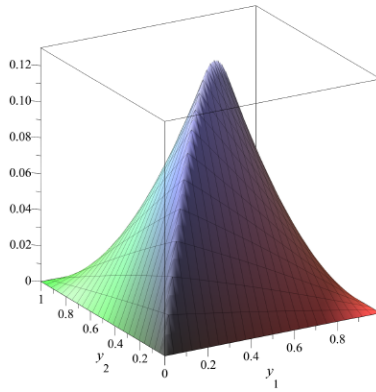
This theorem shows that we can take the partial sum of series (7) as a uniform approximation of the Lyapunov matrix. By increasing the number of terms of the partial sum, we can obtain an approximation with an arbitrary accuracy.

6. Example. Using this theorem we can graph the Lyapunov matrix with fixed parameter τ . This example considers the following system (1) parameters: $a = 1$, $b = 2$, $h = 1$.

In one case, when $\tau > 0$ we obtain a smooth surface. In the other case, when the parameter $\tau = 0$ – a non-smooth surface with a kink on the line $y_1 = y_2$. In these cases, 10 terms were enough to get a good approximation.



$\tau = 0.8$



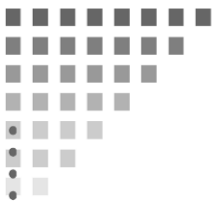
$\tau = 0$

Figure 1. Lapunov matrix with fixed parameter τ

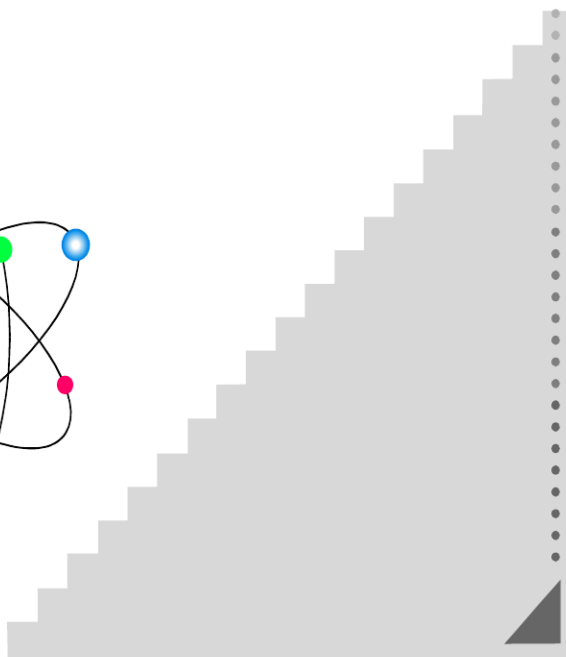
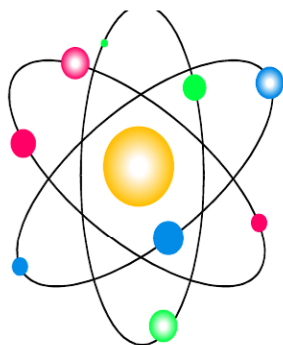

7. Conclusion. In the case of exponential stability of the time delay system with distributed parameters, we propose a method for constructing the Lapunov – Krasovskii functional with a given non-positive derivative. The explicit form of the Lyapunov matrix defining this functional is deduced and the formula for calculation is presented.

References

1. Kharitonov V. L., Zhabko A. P. Lyapunov–Krasovskii approach to the robust stability analysis of time-delay systems // *Automatica*. 2003. Vol. 39. No 1. P. 15–20.
2. Egorov A. V., Mondie S. A stability criterion for the single delay equation in terms of the Lyapunov matrix // *Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes*. 2013. Vol. 10. No 1. P. 106–115.
3. Shuxia P. Asymptotic behavior of travelling fronts of the delayed Fisher equation // *Nonlinear Analysis: Real World Applications*. 2009. No 10. P. 1173–1182.
4. Kharitonov V. L. *Time-delay systems: Lyapunov functionals and matrices*. Boston: Birkhauser, 2013. 311 p.



*2. Математические
методы в механике и
физике*



УДК 621.384.639

Беляева А. А., Князев Н. А.

**Влияние модификаций вычислительных моделей
постоянных дипольных магнитов
источника синхротронного излучения
на оптические характеристики
пучка заряженных частиц**

1. Введение. С развитием науки и технологий синхротронное излучение стало активно использоваться в различных областях: биологии, химии, медицине, физике и т. д. Учитывая сложность разработки оборудования, необходимого для создания излучения данного вида, важную роль обретают математические расчеты, обеспечивающие тонкую настройку синхротронной установки.

Объектом исследования является источник синхротронного излучения 4-го поколения ускорительно-накопительный комплекс (УНК), разрабатываемый в Национальном исследовательском центре «Курчатовский институт» [1–5]. Подобный проект, Европейский источник синхротронного излучения (ESRF) [6, 7], уже был построен и успешно запущен в 2020 году в Гренобле, Франция. В обоих ускорителях используются высококоэрцитивные постоянные магниты, что обеспечивает высокую индукцию поля и отсутствие необходимости в питании, в отличие от электромагнитов. Это повышает требования к точности оптических расчетов на данном этапе и, следовательно, обуславливает важность рассматриваемой задачи.

2. Постановка задачи. УНК – это комбинация синхротронного накопительного кольца и лазера на свободных электронах. Такая конфигурация позволяет сочетать достоинства обеих систем и при этом минимизировать затраты на их создание и эксплуатацию за счет использования единого инжекционного комплекса и общей инфраструктуры. Интересующие нас дипольные магниты входят в

Беляева Александра Антоновна – студент, Санкт-Петербургский государственный университет; e-mail: st088235@student.spbu.ru, тел.: +7(927)433-05-05

Князев Никита Андреевич – научный сотрудник, АО НИИЭФА им. Д. В. Ефремова; e-mail: knyazev@sintez.niiefa.spb.su, тел.: +7(908)027-92-59

Работа выполнена в рамках договора № 24576/201 между АО «НИИЭФА им. Д. В. Ефремова» и СПбГУ

состав накопительного кольца, поэтому более подробно остановимся на нем.

Накопительное кольцо комплекса состоит из повторяющихся участков (периодов), включающих в себя как магнитные элементы (диполи, квадруполи, сектуполи и т. д.), так и другие (промежутки, мониторы, корректоры, ВЧ-резонаторы). Иллюстрация структуры периода представлена на рис. 1 [8].

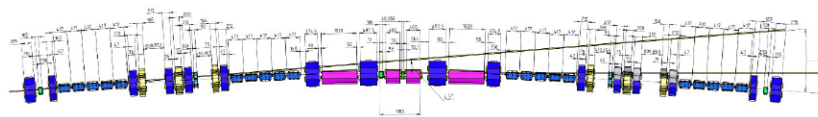


Рис. 1. Иллюстрация одного периода накопительного кольца с учетом поворота опорной траектории

Цель работы заключается в разработке моделей магнитной системы источника синхротронного излучения, в частности, в изучении влияния модификаций оптических моделей дипольных магнитов на характеристики пучка заряженных частиц для проектирования и создания данных магнитов.

3. Алгоритмическая реализация. В данном разделе выполняются реализация и оптимизация вычислительной модели оптической системы источника синхротронного излучения в комплексе программ MAD-X [9]. Рассматриваются способы модификации моделей дипольных магнитов и их влияние на оптические характеристики пучка и точность вычислительной модели. Также разработан алгоритм, позволяющий учитывать распределение магнитного поля при выполнении расчетов в MAD-X.

3.1. Оптические расчеты по техническим требованиям.

На данном этапе все магнитные элементы (диполи, квадруполи и т. д.) описывались как приближенные к идеальным, с постоянными полями без учета краевых эффектов. Полученные функции амплитуды (бета-функции) представлены на рис. 2. Вычисления проводились в соответствии с представленными исходными данными и техническими требованиями (ТТ) на одном периоде кольца.

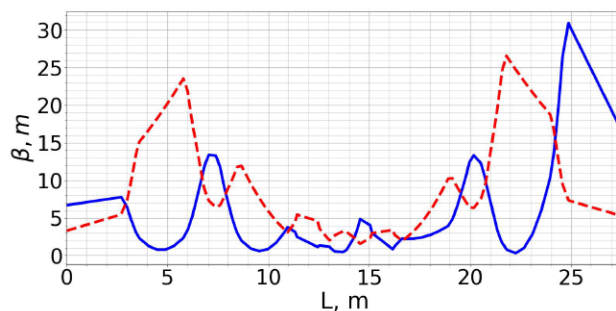


Рис. 2. Оптические функции, рассчитанные по ТТ. Горизонтальная ось соответствует длине траектории в метрах, вертикальная – значениям бета-функций в метрах. Сплошной линией изображена горизонтальная бета-функция, пунктиром – вертикальная бета-функция

Как можно заметить, полученные функции не являются симметричными. Более того, рассчитанные значения кратности бетатронных колебаний (Q_x, Q_y) отличаются от требуемых, из чего следует, что модель синхротрона с данными квадрупольными коэффициентами не обеспечивает фокусировку пучка и частицы теряются.

3.2. Оптимизация квадрупольных коэффициентов. Необходимо улучшить оптическую модель УНК путем оптимизации квадрупольных коэффициентов. В качестве критерия оптимизации установим соответствие расчетных кратностей бетатронных колебаний заданным. Мерой качества будет эмиттанс. Оптимизация проводилась в MAD-X с помощью модуля МАСН.

В общем случае фазовый объем определяется в шестимерном фазовом пространстве, но, в случае разделения переменных в уравнениях движения, могут независимо сохраняться проекции фазового объема на независимые фазовые плоскости. Площадь проекции фазового объема на каждую поперечную фазовую плоскость (горизонтальную и вертикальную – в случае декартовых координат) является эмиттансом пучка, и фазовый объем может быть им охарактеризован.

В таблице 1 представлены квадрупольные линзы и относительная погрешность изменения их квадрупольных коэффициентов (1). Видно, что больше всего изменились коэффициенты квадруполь под номерами 4, 5 и 3 (22,98%, 2,65%, 2,62% соответственно).

Таблица 1. Относительная погрешность изменений квадрупольного коэффициента линз

Квадруполь	1	2	3	4	5	6	7	8
$\varepsilon, \%$	0,88	0,46	2,62	22,98	2,65	1,36	1,42	0,66

$$\varepsilon = \left| \frac{K_1^{\text{тех}} - K_1^{\text{опт}}}{K_1^{\text{опт}}} \right|, \quad (1)$$

где $K_1^{\text{тех}}$ – квадрупольный коэффициент, рассчитанный по ТТ; $K_1^{\text{опт}}$ – подобранный квадрупольный коэффициент.

Сравним результаты оптических функций при оптимизации квадрупольных коэффициентов пар линз (4, 3) и (4, 5), параметры остальных магнитов возьмем из ТТ. Как видно из таблицы 2, наименьший эмиттанс был получен при оптимизации пары (4, 3). Более того, для данных коэффициентов квадрупольных расчетные кратности бетатронных колебаний совпали с требуемыми. Таким образом, наибольшую фокусировку выполняют квадрупольные (4, 3). Во всех дальнейших расчетах при моделировании установки использовались данные квадрупольные коэффициенты.

Таблица 2. Сравнение результатов оптимизации пар квадрупольных линз

Квадруполь	Эмиттанс, пм·рад	Q_x	Q_y
(4, 3)	56,89	2,38025	0,8335
(4, 5)	57,06	2,392238	0,851999

На рис. 3 изображены оптические функции для УНК с оптимизированными квадрупольными коэффициентами для линз (4, 3).

Из рис. 3 видно, что после оптимизации параметров квадрупольных линз (4, 3) бета-функции стали симметричными. Начальные значения совпадают с конечными, что означает сохранение периодичности в накопительном кольце. Рассчитанные значения кратности бетатронных колебаний соответствуют требуемым по ТТ.

3.3. Проведение оптических расчетов с использованием карты магнитного поля. Во время разработки короткого дипольного магнита в НИИЭФА им. Д. В. Ефремова [8] была рассчитана карта поля с учетом краевых эффектов данного элемента. Проанализируем, как изменится точность математической модели и оптические параметры пучка при добавлении соответствующей информации в расчетную модель MAD-X.

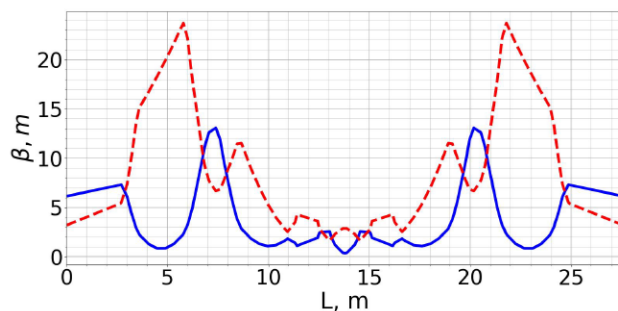


Рис. 3. Оптические функции, рассчитанные с оптимизацией квадрупольных коэффициентов для линз (4, 3). Сплошной линией изображена горизонтальная бета-функция, пунктирной – вертикальная бета-функция

В комплексе программ MAD-X не предусмотрена возможность использования карт электромагнитных полей, поэтому был разработан следующий способ моделирования. Область элемента дробится на N частей, каждая из которых имеет постоянное магнитное поле. При этом суммарный интеграл поля (и, соответственно, угол поворота траектории) сохраняется таким же, как и в случае постоянного магнита. Иллюстрация данного подхода представлена на рис. 4.

Результаты сравнения разного числа разбиений короткого магнита представлены в таблице 3. Можно сделать вывод, что для улучшения точности описания математической модели достаточно разделить магнит на небольшое количество частей (в данном случае на шесть).

Таблица 3. Оценка влияния числа разбиений короткого поворотного магнита на оптические характеристики пучка

Кол-во фрагментов	Эмиттанс, пм·рад	Q_x	Q_y
Без изменений	56,89	2,38025	0,8335
6	54,17	2,37984	0,8334485
16	54,39	2,37984	0,8334485
32	54,44	2,37984	0,8334485
76	54,44	2,37984	0,8334485

На рис. 5 представлены оптические функции УНК для модели короткого дипольного магнита с разбиением на 6 частей. Квадрупольные коэффициенты использовались такие же, как и в предыдущем

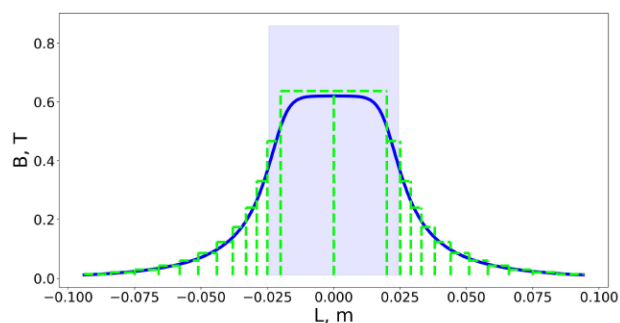


Рис. 4. Моделирование короткого диполя с учетом расчетной карты магнитного поля. Горизонтальная ось совпадает с длиной траектории относительно центра магнита в метрах, вертикальная – с величиной индукции поля в теслах. Сплошная линия соответствует рассчитанному полю, пунктирная – аппроксимированному полю. Закрашенная область соответствует модели с эффективным значением поля согласно ТТ

разделе. Симметричность оптических функций была достигнута за счет оптимизации начальных значений бета-функций.

3.4. Исследование влияния модификаций моделей длинных дипольных магнитов на оптические характеристики пучка. Период накопительного кольца установки содержит четыре дипольных магнита длиной около двух метров, разделенных на модули, каждый из которых имеет свое постоянное поле (в действительности с краевыми эффектами). Изготовление таких магнитов сложный и дорогостоящий процесс. Поэтому необходимо исследовать, как модификация вычислительных моделей данных модулей влияет на точность описания и оптические характеристики пучка.

Рассмотрим два варианта модификаций:

- усреднение – задание для каждого модуля одинакового поля, равного среднеарифметическому значению (рис. 6);
- корректировка – увеличение или уменьшение магнитной индукции каждого модуля на $x\%$ (рис. 7)

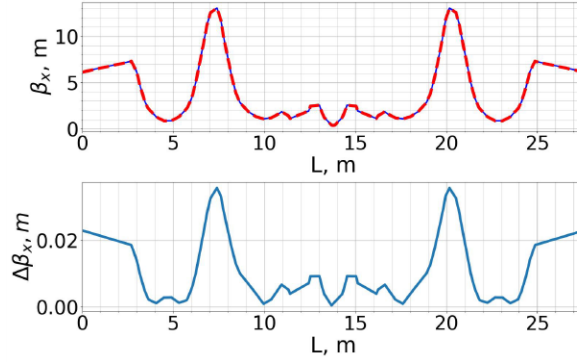


Рис. 5. Бета-функции на периоде для прецизионного поля короткого магнита (сплошная линия) и в случае аппроксимации кривой распределения поля с помощью шести отрезков (пунктирная линия). Нижний график демонстрирует отклонение бета-функции для рассматриваемых случаев

$$B_i^{\text{кор}} = B_i - \frac{x}{100\%}(B_i - \bar{B}), \quad (2)$$

где B_i – магнитная индукция i -го модуля, \bar{B} – среднее значение индукции поля, $B_i^{\text{кор}}$ – скорректированное значение индукции поля i -го модуля.

Результаты проведенных вычислительных экспериментов представлены в таблицах 4, 5.

Таблица 4. Влияние усреднения поля для моделей длинных дипольных магнитов

	Эмиттанс, пм·рад	Q_x	Q_y
Без изменений	56,89	2,38025	0,8335
Диполь 1	115,32	2,380225	0,8335
Диполь 2	59,43	2,380248	0,8335
Диполи 1 и 2	130,85	2,380223	0,8335

Таблица 5. Влияние корректировки поля для моделей длинных дипольных магнитов

	Эмиттанс, пм·рад	Q_x	Q_y
Без изменений	56,89	2,38025	0,8335
Диполь 1	59,04	2,380246	0,8335
Диполь 2	56,2	2,380249	0,8335
Диполи 1 и 2	58,49	2,380245	0,8335

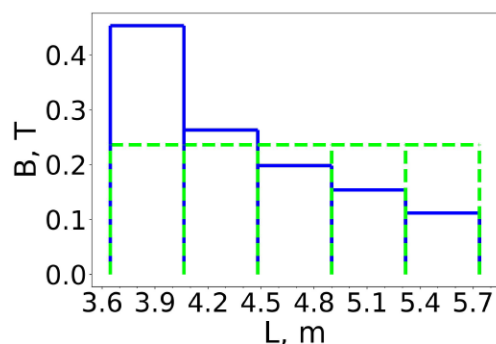


Рис. 6. Иллюстрация усреднения индукции поля длинного магнита. Сплошная линия соответствует полю по ТТ, пунктирная – усредненному

Из таблиц 4, 5 можно сделать следующие выводы:

1. Замена поля в модулях длинных дипольных магнитов на равное среднеарифметическому значению увеличивает эмиттанс пучка, в худшем случае более чем в два раза.
2. Увеличение или уменьшение в пределах 10% индукции поля в модулях дипольных магнитов согласно (2) с сохранением интеграла поля и угла поворота влияет на оптические характеристики пучка в пределах 5%.

Таким образом, на этапе проектирования длинных магнитов нет необходимости требовать обеспечения сложно достижимых прецизионных характеристик их поля согласно ТТ, при этом имеющие место изменения оптических характеристик пучка можно компенсировать, например, за счет оптимизации квадрупольных коэффициентов линз.

4. Заключение. Данная работа использует результаты проведенных ранее другими авторами исследований, касающихся в том числе спроектированной оптической схемы накопительного кольца источника синхротронного излучения НИЦ «Курчатовский институт». В ходе работы изучено влияние модификаций моделей дипольных магнитов, разрабатываемых на основе высококоэрцитивных постоянных магнитов, на оптические характеристики пучка заряженных частиц.

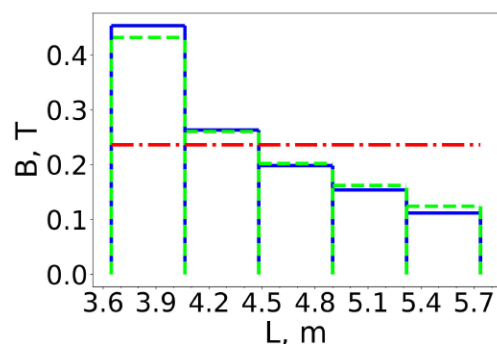


Рис. 7. Иллюстрация корректировки поля модулей длинного магнита на 10% относительно разности величины индукции элемента и среднеарифметического значения. Сплошная линия соответствует полю по ТТ, пунктирная – скорректированному полю, штрихпунктирная – среднему значению индукции

Численные эксперименты показали, что при создании оптических моделей дипольных магнитов в MAD-X возможно варьирование в допустимом диапазоне, по крайней мере в пределах 10%, уровня поля с сохранением интеграла поля. При этом имеющие место изменения оптических характеристик пучка можно компенсировать, например, за счет оптимизации квадрупольных коэффициентов линз.

Для короткого дипольного магнита разработан алгоритм, позволяющий учитывать распределение краевых эффектов поля в оптических расчетах с использованием MAD-X. Показано, что для улучшения точности описания модели при использовании указанного метода достаточно разделить элемент лишь на небольшое количество частей.

Результаты данных исследований используются при разработке и проектировании дипольных магнитов источника синхротронного излучения.

5. Благодарности. Авторы признательны за поддержку, дискуссии и плодотворные обсуждения Д. А. Овсянникову, В. А. Козыченко, Е. Д. Котиной (СПбГУ), Н. Ю. Казаринову, И. А. Иваненко (ОИЯИ ЛЯР СФУ), Н. И. Казаченко, В. П. Кухтину, Е. А. Ламзину, С. Е. Сычевскому (АО «НИИЭФА им. Д. В. Ефремова»).

Литература

1. Ashanin I. A., Polozov S. M., Pronikov A. I. et al. Beam dynamics simulation in a linear electron accelerator – injector for the 4th generation specialized synchrotron radiation source USSR // 27th Russian Particle Accelerator Conference. 2021. P. 149–152.
2. Ashanin I. A., Bashmakov Y. A., Budkin V. A. et al. Conceptual design of a dedicated fourth-generation specialized synchrotron radiation source (SSRS-4) at the Kurchatov Institute // Phys. Atom. Nuclei. 2018. No 11. P. 1646–1651.
3. Dyubkov V., Kulevoy T., Tsyplakov E. Lattice options with reverse bending magnets for USSR HMBA storage ring // 27th Russian Particle Accelerator Conference. 2021. P. 280–282.
4. Ашанин И. А., Ключевская Ю. Д., Махоро А. А. и др. Динамика пучка в линейном ускорителе-инжекторе специализированного источника синхротронного излучения 4-го поколения ИССИ-4 // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2019. Т. 15. С. 126–139.
5. Fomin Ye., Korchuganov V. Kurchatov synchrotron radiation source – from the 2nd to the 4th generation // Russian Particle Accelerator Conference. 2021. P. 84–87.
6. Raimondi P. Hybrid multiband achromat: From Super B to EBS // In Proc. IPAC'17. Copenhagen: Den mark. 2017. P. 3670–3675.
7. ESRF Upgrade Program Phase II [Электронный ресурс]: URL: https://www.esrf.fr/Apache_files/Upgrade/ESRF-orange-book.pdf (дата обращения:03.04.2024).
8. Arslanova D. et al. Feasibility study of permanent magnet dipoles for sila facility // Cybernetics and physics. 2023. No 12. P. 252–256.
9. The MAD-X Program (user's guide) [Электронный ресурс]: URL: <https://mad.web.cern.ch/mad/> (дата обращения:03.04.2024).

УДК 004.94:519.63:537.533

Вековцев В. В., Иордан А. А., Никифоров К. А.

Разработка цифровой модели автоэмиссионной системы

Введение. В данной работе рассматриваются особенности численного моделирования в рамках создания цифровой модели автоэмиссионной системы с катодно-модуляторным узлом, представленным ранее в двумерной постановке задачи [1]. С учетом осевой симметрии системы алгоритм траекторного анализа в [1] строился на основе значений потенциала и напряженности электрического поля в узлах на границе вычислительной области. Данный алгоритм может быть применен для экспресс-оценки токовых характеристик катодно-модуляторного узла с целью оптимизации модулей соответствующей электронно-оптической системы (ЭОС) [2], при этом используются либо численно-аналитические [3], либо проекционно-сеточные [1] методы вычисления электрического поля.

Однако рассмотрение подобной задачи в трехмерной постановке позволяет учесть более тонкие эффекты с более точным вычислением поля, хотя и является более ресурсоемким подходом. Важной особенностью использования вычислительных моделей для прогнозирования свойств и характеристик автоэмиссионных ЭОС является учет многомасштабности разнообразных процессов, протекающих при автоэлектронной эмиссии [4]. Одной из основных проблем многомасштабного моделирования является необходимость стыковки множества разнородных моделей, описывающих поведение и свойства сложных систем на различных уровнях детализации. Количество разномасштабных уровней детализации при моделировании зависит от сложности структуры и формы эмиттера.

В частности, многомасштабный подход необходим при решении задачи расчета электрического поля над эмиссионной поверхностью. Это обусловлено непосредственным влиянием на результат

Вековцев Вячеслав Владиславович – студент, Санкт-Петербургский государственный университет; e-mail: st088221@student.spbu.ru, тел.: +7(911)097-79-96

Иордан Андрей Алексеевич – студент, Санкт-Петербургский государственный университет; e-mail: st098445@student.spbu.ru, тел.: +7(996)777-26-36

Никифоров Константин Аркадьевич – доцент, Санкт-Петербургский государственный университет; e-mail: k.nikiforov@spbu.ru, тел.: +7(812)363-60-00

Работа выполнена при финансовой поддержке РФФ, грант № 23-29-00937

как микропараметров (распределения плотности атомной упаковки, значений работы выхода, локального радиуса кривизны поверхности и т. д.), так и макропараметров (геометрии системы электродов, определяющих распределение макроскопического электрического поля, формирование пространственного заряда). Поэтому на унифицированном языке моделирования [5] UML-диаграмма компонентов автоэмиссионной ЭОС (рис. 1) представляет комплексное рассмотрение процессов со столь разными масштабами времен и размеров, которое возможно только при помощи многомасштабного компьютерного моделирования.

Генерация автоэлектронного тока в процессе эмиссии системы происходит на наноуровне: под воздействием поля потенциальный порог на поверхности катода превращается в потенциальный барьер и появляется отличная от нуля вероятность туннелирования электронов, как следует из квантовой теории [4],

$$j(E) = \frac{AE}{\varphi t^2(y)} \exp\left(-B \frac{\varphi^{\frac{3}{2}}}{E} \nu(y)\right), \quad (1)$$

где j – плотность тока эмиссии, A и B – постоянные, связанные с фундаментальными физическими величинами, $A = \frac{e^3}{8\pi h}$, $B = \frac{8\pi\sqrt{2m}}{3eh}$, h – постоянная Планка, e – заряд электрона, m – масса покоя свободного электрона, $t(y)$, $\nu(y)$ – эллиптические функции Нордгейма аргумента $y = \sqrt{\frac{e^3 E}{4\pi\epsilon_0\varphi^2}}$, для которых используется аппроксимация $t^2(y) \approx 1,11619$, $\nu(y) \approx 0,95 - y^2$, ϵ_0 – электрическая постоянная, E – величина напряженности внешнего электрического поля, φ – работа выхода материала, являющаяся мерой энергии связи электронов с твердым телом.

Особенности численного моделирования. В UML-диаграмме компонентов ЭОС на рис. 1 отдельно выделяется интерфейс с начально-краевыми условиями электронного пучка с плотностью тока j после этапа генерации, наиболее специфического в случае автоэмиссии в соответствии с формулой (1). Выходные параметры этого этапа включают информацию о структуре поверхности, наличии адсорбированных слоев, значениях работы выхода, спектре энергий и начальных скоростей частиц, инжектируемых в вычислительную область.

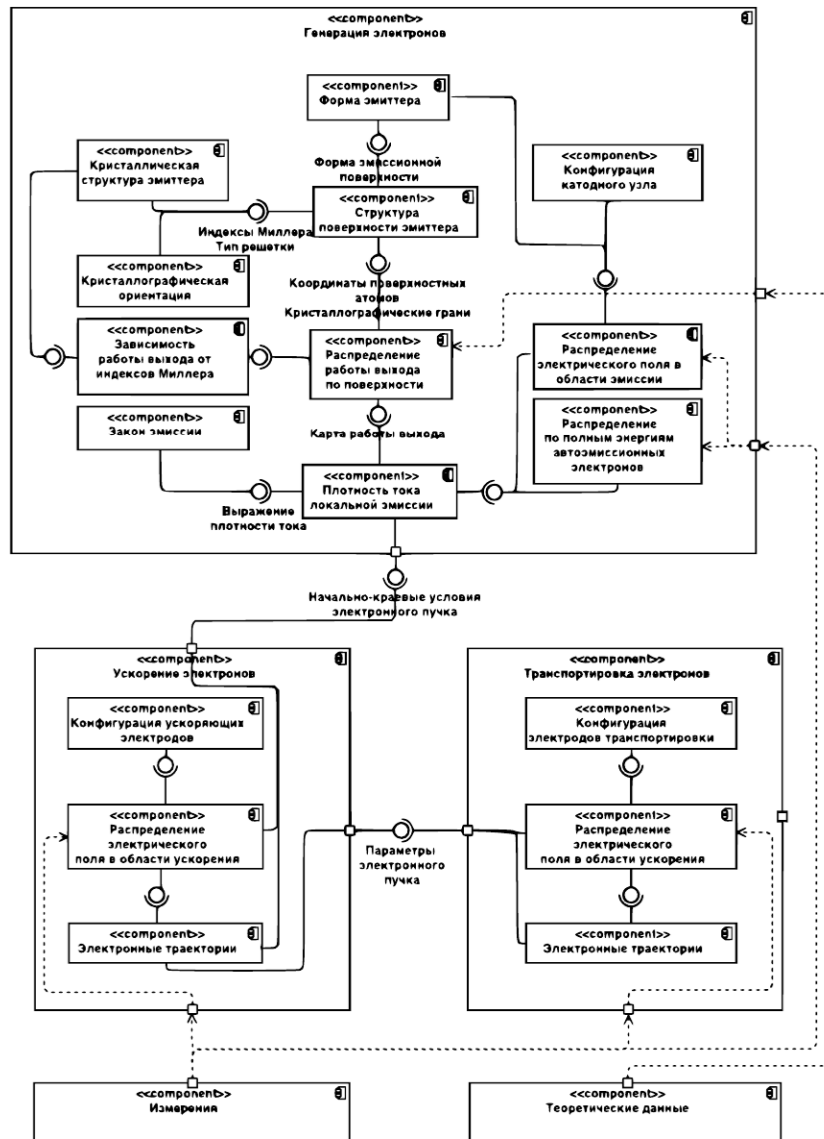


Рис. 1. UML-диаграмма компонентов автоэмиссионной электронно-оптической системы

Величина работы выхода φ в (1) изменяется по эмиссионной поверхности в связи с образованием на ней кристаллографических граней и неравномерным распределением плотности атомной упаковки, наличием адсорбатов. Поэтому осевая симметрия системы нарушается, хотя кристаллическая решетка может обладать каким-то другим типом симметрии. Следовательно, для учета этого в модели необходимо переходить от 2D- к 3D-постановке задачи. Например, кубический тип симметрии структуры катода позволяет рассматривать четверть объема вычислительной области с использованием соответствующих граничных условий периодичности.

UML-диаграмма компонентов в соответствии с вышеуказанными факторами определяет архитектуру программной реализации вычислительных модулей, устанавливает зависимости между ними при последующей интеграции в единую структуру цифровой модели автоэмиссионной ЭОС. Кроме того, представление и дискретизация вычислительной области цифровой модели должны отвечать следующим особенностям автоэмиссионных систем.

В макромасштабе система электродов, как правило, соответствует вычислительной области сложной формы, которая включает границу эмиттера с большой кривизной поверхности и малыми размерами, что приводит к значительному разбросу характерных размеров в одной геометрической конфигурации. При этом экспоненциальная зависимость плотности эмиссионного тока от напряженности поля (1) требует повышенной точности при учете граничных условий.

Дискретизация вычислительной области. Важным этапом является построение сетки, так как все дальнейшие вычисления выполняются на ее основе. Численные методы, используемые в COMSOL дискретизируют пространство моделирования именно с помощью заранее заданной вычислительной сетки. Пространственная дискретизация является необходимой процедурой для разбиения расчетной области на конечное число подобластей. В результате исходное непрерывное распределение искомой величины представляется в виде конечных распределений в дискретных объемах. Сетки, получаемые в результате пространственной дискретизации, разделяются на структурированные и неструктурированные [6]. От выбранного типа сетки зависит и численный метод дискретизации пространства. Так, метод контрольных объемов изначально предназначен для решения уравнений переноса и применяется для решения

задач на основе структурированных сеток. Преимуществом таких сеток является использование простых структур данных для хранения информации о координатах сеточных узлов и соответствующих значениях искомых переменных в них. Также данные сетки просты в построении и для решения уравнений на структурированных сетках существуют эффективные решатели, не требующие больших вычислительных ресурсов и работающие достаточно быстро. Недостатком же такого способа пространственной дискретизации является ограниченность в сложности геометрических форм. В свою очередь неструктурированные сетки отличаются большей гибкостью и адаптивностью, позволяя легко описывать сложные геометрические объекты и, соответственно, решать более широкий круг задач. При этом теряются преимущества, даваемые структурированными сетками на простой геометрии, а именно, требуются более сложные структуры данных, что приводит к более сложным алгоритмам решения уравнений и большему потреблению вычислительных ресурсов [7].

В рассматриваемом случае за построение необходимой трехмерной сетки отвечает метод конечных элементов. Дискретизация непрерывной задачи согласно этому численному методу производится на неструктурированной сетке с неравномерным пространственным шагом. Как отмечено в предыдущем разделе, автоэмиссионные системы требуют высокой точности вычисления в области, где происходит непосредственно эмиссия. Это связано с тем, что в данной области вычисляется плотность тока, экспоненциально зависящая от напряженности поля (1), т. е. погрешности вычисления напряженности экспоненциально распространяются на вычисленную плотность тока. В этой области требуется обеспечение максимально возможного точного вычисления напряженности. Еще одной особенностью является то, что для снижения погрешности интегрирования траекторий частиц частям вычислительной области, в которых движутся модельные частицы, должны соответствовать области с достаточно мелким разбиением сетки, в смысле обеспечения устойчивости методу интегрирования уравнений движения. Вычислительная сетка, удовлетворяющая описанным выше особенностям, построена следующим образом. Области наименьшего шага на рис. 2 соответствуют поверхности модулятора (рис. 2а) и катода (рис. 2в). На областях, куда в ходе симуляции работы системы не попадают эмитированные частицы, размеры ячеек сетки крупнее (рис. 2б), так как это снижает объем вычислений, при этом не влияя на точность решения.

Таким образом, элементами построенной сетки являются тетраэдры, общее количество которых превышает 650000 единиц (рис. 2). В результате осуществляется аппроксимация геометрии модели и приближенное решение задачи происходит в дискретных точках пространства, определяемых данной сеткой. При наличии сеточной сходимости по мере уточнения сетки численное решение имеет тенденцию более точно приближаться к решению граничной и начально-краевой задачи.

Результатом дискретизации является достаточно подробная сетка, предоставляющая решение с допустимой погрешностью (определенной точностью), но не перегружающая вычислительную систему, что обеспечивает эффективность вычислительного процесса по времени. В рассматриваемой модели выделены четыре отдельных под-области, на которых сетка имеет большую детализацию (рис. 2) в соответствии со значениями параметра, задающего максимальный размер сеточного элемента.

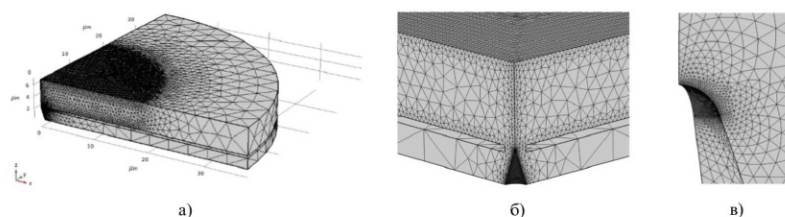


Рис. 2. а) дискретизация вычислительной области неструктурированной неравномерной сеткой, б) вид сетки на оси системы и поверхностях симметрии, в) дискретизация криволинейной поверхности эмиттера

Заключение. Разработка цифровой модели автоэмиссионной ЭОС связана с решением самосогласованной задачи, включающей расчет внешних и собственных электрических полей в областях со сложной конфигурацией границ. На унифицированном языке моделирования описаны особенности представления системы в виде UML-диаграммы компонентов, определяющей декомпозицию системы на ряд модулей с выделением типов связей между ними. Диффе-

ренциальные уравнения в частных производных дискретизируются по пространственным переменным методом конечных элементов, поэтому в статье рассмотрены важные вопросы дискретизации и построения сетки, от которых сильно зависит эффективность численных методов и алгоритмов по решению граничной, а также начально-краевой задачи.

Исследования проведены с использованием оборудования ресурсных центров Научного парка СПбГУ «Нанотехнологии» и «Вычислительный центр».

Литература

1. Соколов И. А., Никифоров К. А. Моделирование автоэмиссионного катодно-модуляторного узла // Процессы управления и устойчивость. 2023. Т. 10. № 1. С. 156–160.
2. Бугаев А. С., Виноградова Е. М., Егоров Н. В., Шешин Е. П. Автоэлектронные катоды и пушки. М.: Интеллект, 2017. 288 с.
3. Egorov N. V., Vinogradova E. M., Kurbatova G. I., Klemeshev V. A. Field emission system with two emitters mathematical modeling // Vestnik Sankt-Peterburgskogo Universiteta. Prikladnaya Matematika. Informatika. Protsessy Upravleniya. 2023. Vol. 19. No 4. P. 443–448.
4. Егоров Н. В., Шешин Е. П. Автоэлектронная эмиссия. Принципы и приборы. М.: Интеллект, 2011. 704 с.
5. Booch G., Rumbaugh J., Jacobson I. O. The Unified Modeling Language User Guide, second edition, 2005. 496 p.
6. Liseikin V. D. Grid generation methods. Netherlands: Springer, 2010. 390 p.
7. Василевский Ю. В., Данилов А. А., Липников К. Н. и др. Автоматизированные технологии построения неструктурированных расчетных сеток. М.: Физматлит, 2016. 216 с.

УДК 539.3

Вяльцева В. В., Пронина Ю. Г.

Об оценке трехмерной плотности параллельных трещин по их следам на поверхности

1. Введение. Наличие микротрещин в материалах приводит к изменению их эффективных упругих свойств [1, 2]. Для характеристики концентрации плоских круглых трещин, случайным образом расположенных внутри представительного объема V , Бристону [3] ввел скалярный параметр плотности трещин

$$\rho_{3D} = \frac{1}{V} \sum_k r_k^3, \quad (1)$$

где r_k – радиусы трещин. В случае наличия предпочтительной ориентации (ориентаций) в системе трещин скалярный параметр необходимо заменить на тензорный [2]: $\rho_{3D} = (1/V) \sum_k (r^3 \mathbf{n}\mathbf{n})_k$, где r и \mathbf{n} – радиус и нормаль k -й трещины. В двумерном случае скалярная плотность прямолинейных трещин с длинами $2l_k$, расположенных в плоскости площадью S , определяется как $\rho_{2D} = (1/S) \sum_k l_k^2$, а двумерный тензор плотности трещин как $\rho_{2D} = (1/S) \sum_k (l^2 \mathbf{n}\mathbf{n})_k$.

Определение трехмерной плотности трещин методами неразрушающего контроля весьма затруднительно. Поэтому актуален вопрос об ее оценке по следам трещин на поверхности образцов. Для геометрически изотропного распределения дефектов, когда ρ_{3D} и ρ_{2D} являются скалярами, этот вопрос решен в работе [4]. Целью данной работы является установление подобных зависимостей в случае системы параллельных трещин, ориентированных произвольно по отношению к выбранной системе координат, когда оба названных параметра являются тензорами второго ранга.

Вяльцева Виктория Витальевна – магистрант, Санкт-Петербургский государственный университет; e-mail: valcevav@gmail.com, тел.: +7(952)892-77-38

Пронина Юлия Григорьевна – профессор, Санкт-Петербургский государственный университет; e-mail: y.pronina@spbu.ru, тел.: +7(812)363-60-00 (9567)

Работа выполнена при финансовой поддержке РФФ, грант № 21-19-00100

2. Постановка задачи. Рассматривается материал, ослабленный системой параллельных плоских круглых (penny-shaped) трещин радиуса r , случайным образом равномерно распределенных по объему материала. Произведены сечения данного материала тремя взаимно перпендикулярными плоскостями, ориентированными вдоль координатных плоскостей выбранной системы координат $Ox_1x_2x_3$. Площади сечений примем равными S .

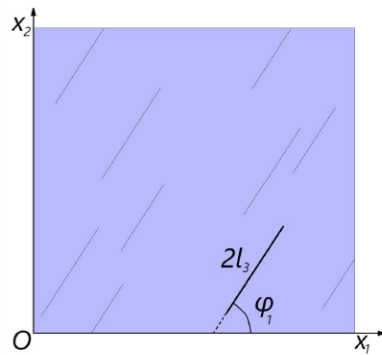


Рис. 1. Следы трещин в плоскости Ox_1x_2

Считаем известными длины следов трещин $2l_i$ на указанных плоскостях с нормальными Ox_i и углы φ_i , образуемые следами трещин в плоскости Ox_ix_{i+1} , с осью Ox_i (рис. 1).

Необходимо определить связь между трехмерной плотностью трещин в объеме материала и двумерными плотностями их следов на сечениях, а также провести серию численных экспериментов для оценки разброса расчетных оценок трехмерной плотности относительно ее реального значения в зависимости от числа наблюдаемых дефектов.

3. Аналитическое решение. Сначала выведем интегральную зависимость между трехмерной плотностью трещин ρ_{3D} и двумерной плотностью их следов на сечениях $\rho_{2D,i}$. Положение каждой трещины определяется пятью величинами: расстояниями H_1, H_2, H_3 от центра трещины до плоскостей $Ox_2x_3, Ox_1x_3, Ox_1x_2$ соответственно, и двумя углами, например, α_3 (угол между плоскостью трещины и осью x_3) и θ (угол поворота трещины вокруг оси x_3 , отсчитываемый от оси x_2).

Квадрат полудлины следа трещины на поверхности с нормалью Ox_3 можно выразить как [4]

$$l_3^2 = r^2 - \frac{H_3^2}{\cos^2 \alpha_3}, \quad (2)$$

где

$$\alpha_3 = \arcsin \left(\frac{\tan^2 \varphi_2}{\cos^2 \varphi_1} + 1 \right)^{-\frac{1}{2}},$$

если трещина пересекает грань куба в плоскости Ox_1x_2 .

Интегрируя равенство (2) по H_3 от $-r$ до r (т. е. суммируя результат по всем трещинам в объеме материала), можно найти связь между ρ_{3D} и абсолютной величиной плотности следов трещин на поверхности Ox_1x_2 ($\rho_{2D,3} = (1/S) \sum_k (l_3^2)_k$) в виде

$$\rho_{3D} = \frac{3}{4 \cos \alpha_3} \rho_{2D,3} \mathbf{n}\mathbf{n}, \quad (3)$$

где нормаль \mathbf{n} однозначно выражается через углы φ_1, φ_2 .

Аналогичные формулы можно получить и для связи между ρ_{3D} и плотностями $\rho_{2D,1}$ и $\rho_{2D,2}$ следов трещин на двух других гранях. Среднее арифметическое значений, полученных по трем сечениям, дает (в среднем) более точную оценку искомой величины.

4. Оценка разброса экспериментальных данных. В пункте 3 выведена связь между интегральными характеристиками концентрации трещин. Поскольку на практике всегда имеются образцы лишь с конечным количеством трещин, то реальные значения ρ_{3D}^r будут отклоняться от математического среднего. Проведем серию численных экспериментов для оценки возможных отклонений плотности трещин, рассчитанной по формуле (3), от реального значения плотности, вычисленной по формуле (1), для конечного числа N случайно расположенных трещин.

Алгоритм расчета для каждого численного эксперимента состоит из следующих шагов:

1. Задание параметров r, α_3, θ и количества трещин N .
2. Вычисление реальной плотности трещин ρ_{3D}^r для заданных r, N .
3. Задание расстояний H_1, H_2, H_3 случайным образом для каждой трещины из набора.
4. Вычисление l_1, l_2 и l_3 для каждой трещины.
5. Вычисление $\rho_{2D,1}, \rho_{2D,2}$ и $\rho_{2D,3}$.
6. Вычисление ρ_{3D} по трем граням и их среднего значения.
7. Вычисление относительной разности ρ_{3D} и ρ_{3D}^r .

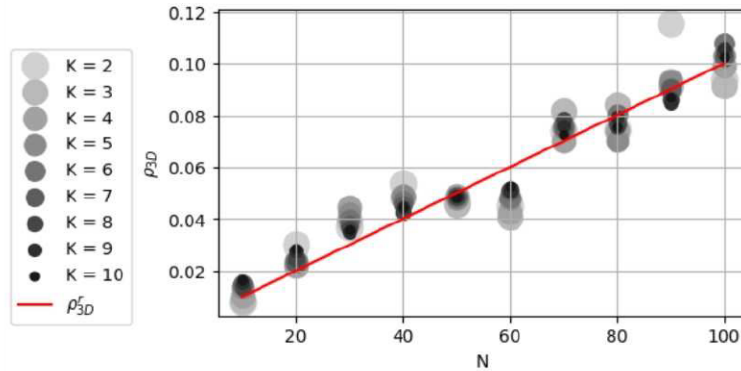


Рис. 2. Сопоставление реальной и расчетной плотности трещин для разного числа испытаний K при различных количествах трещин N ($\alpha = \pi/4, \theta = \pi/5, r/\sqrt{S} = 0,1$)

Для каждого набора параметров r, α_3, θ и N проводилась серия K экспериментов и итоговая оценка плотности трещин ρ_{3D} вычислялась как среднее арифметическое по всем полученным значениям. Заметим, что случай $K = 2$, можно рассматривать как осреднение по шести граням одного кубического образца, $K = 10$ – пяти образцов. На рис. 2 представлена зависимость расчетной плотности трещин от числа трещин N для различных K .

Сплошной линией показана реальная плотность ρ_{3D}^r , серыми кружками – средние арифметические плотностей, вычисленные (для каждого N) по следам на $3K$ сечениях. Как видно, с увеличением K в большинстве случаев значение расчетной плотности приближается к реальному. На рис. 3 представлена относительная разность между реальной плотностью ρ_{3D}^r и средними арифметическими расчетных оценок плотностей для $N = \overline{10, 100}$. Точечная линия соответствует $K = 2$, пунктирная линия – для $K = 6$, сплошная – $K = 10$. Видно, что, начиная с $N = 60$, плотность, рассчитанная для $K = 10$, не отклоняется от реальной более чем на 25%. Результаты не зависят от ориентации системы трещин.

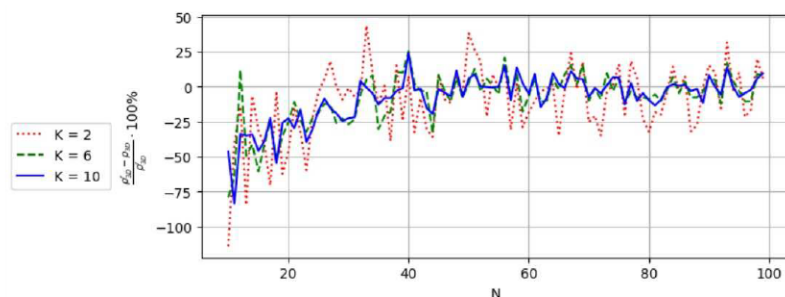


Рис. 3. Относительная разность между реальной и расчетной плотностью трещин для разного числа испытаний K при различных количествах трещин N ($\alpha = \pi/4, \theta = \pi/5, r/\sqrt{S} = 0,1$)

5. Заключение. В работе выведены формулы для нахождения трехмерной плотности системы параллельных плоских круглых трещин по их видимым следам на гранях образца. Проведена серия численных экспериментов для оценки разброса расчетной плотности трещин относительно ее реального значения при ограниченном количестве трещин.

Литература

1. Abakarov A., Pronina Y. On the effective elastic properties of a material with mutually perpendicular systems of parallel cracks // Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes. 2022. Vol. 18. No 1. P. 111–119.
2. Kachanov M. Elastic solids with many cracks and related problems // Advances in Applied Mechanics. 1993. No 30(C). P. 259–445.
3. Bristow J.R. Microcracks and the static and dynamic elastic constants of annealed heavily cold-worked metals // British J. Appl. Phy. 1960. No 11. P. 81–85.
4. Pronina Y., Kachanov M. Estimating concentrations of cracks and platelets from their traces in 2D cross-sections // Mechanics of Materials. 2023. Vol. 180. Art. no 104618.

УДК 621.318.43

Грошева А. Д., Знаменщикова Н. С.

Оценка возможности разработки трёхмерных вычислительных моделей прецизионных магнитных систем, состоящих из соленоидов и ферромагнитных экранов, с использованием интегральных и дифференциальных подходов

1. Введение. Моделирование магнитостатических полей является важным этапом при разработке и создании современных сложных магнитных систем. Такие системы находят применение в магниторезонансных томографах, катушках Гельмгольца, устройствах удержания плазмы (токамаках) и в других научно-исследовательских и промышленных установках.

Для численного моделирования возникающих магнитных полей и ponderomotorных сил применяются различные математические методы, а также используются разные программы, реализующие эти методы. Например, программный пакет КОМРОТ [1, 2] разработан для моделирования трёхмерных магнитостатических полей прецизионных магнитных систем со сложной геометрией в терминах модифицированного скалярного потенциала [3]. Программа использует представление конечных элементов [4] в дифференциальной формулировке и способна моделировать широкий спектр магнитных систем с ферромагнитными компонентами и активными катушками сложной конфигурации. Ещё один пример – программный пакет KLONDIKE [1, 2]. Он использует закон Био – Савара – Лапласа в интегральной формулировке для расчёта магнитного поля по заданному распределению токов. Ещё одной реализацией конечно-элементного метода является программа COMSOL. KLONDIKE и КОМРОТ являются отечественными продуктами и долгое время применяются в актуальных технических проектах, таких как ITER

Грошева Александра Дмитриевна – инженер-исследователь, АО «НИИЭФА им. Д. В. Ефремова»; аспирант, Санкт-Петербургский государственный университет; e-mail: a.d.grosheva@yandex.ru, тел.: +7(812)462-76-23

Знаменщикова Наталья Сергеевна – математик, АО «НИИЭФА им. Д. В. Ефремова»; e-mail: znamenatali@gmail.com, тел.: +7(995)889-03-26

(Франция), ГЛОБУС-М (Россия) и др. COMSOL в свою очередь является проприетарным коммерческим продуктом.

Целью данной работы является исследование возможности создания трёхмерных вычислительных моделей магнитных систем с применением интегральных и дифференциальных подходов, которые реализованы в программных пакетах KLONDIKE и КОМПОТ.

В качестве расчётной модели выбран соленоид с параметрами: внешний радиус $r_{\max} = 208$ мм, внутренний радиус $r_{\min} = 215,2$ мм, длина $\Delta z = 523$ мм, количество витков $N = 414$, сила тока в витке $I_1 = 1$ кА. Расчёт магнитного поля соленоида производился следующими способами:

- средствами программного пакета KLONDIKE;
- средствами программного пакета КОМПОТ с заданием геометрии в цилиндрической системе координат с узкими границами расчётной области;
- средствами программного пакета КОМПОТ с заданием геометрии в цилиндрической системе координат с широкими границами расчётной области;
- средствами программного пакета КОМПОТ с заданием геометрии в декартовой системе координат;
- средствами программного пакета COMSOL.

2. Модифицированный скалярный потенциал. Согласно уравнениям Максвелла, напряжённость магнитного поля \mathbf{H} удовлетворяет равенствам

$$\nabla \times \mathbf{H} = \mathbf{j}, \quad \nabla \cdot (\mu_0 \mu \mathbf{H}) = 0, \quad (1)$$

где \mathbf{j} – плотность тока, μ – магнитная проницаемость среды, $\mu_0 = 4\pi \cdot 10^{-7}$ Гн/м. Можно ввести векторный электрический потенциал \mathbf{P} , для которого

$$\nabla \times \mathbf{P} = \mathbf{j}. \quad (2)$$

Тогда поле \mathbf{H} может быть представлено как $\mathbf{H} = \mathbf{P} + \nabla\phi$, где ϕ – модифицированный скалярный потенциал, удовлетворяющий уравнению

$$\nabla \cdot (\mu_0 \mu (\mathbf{P} + \nabla\phi)) = 0. \quad (3)$$

Для нахождения скалярного потенциала и магнитного поля необходимо задать векторный электрический потенциал таким образом, чтобы

$$\oint_l P_l dl = I, \quad (4)$$

где I – сила тока, протекающая через контур l . Других ограничений на векторный электрический потенциал нет. Пусть вне соленоида $\mathbf{P} = 0$, а во внутренней области соленоида $\mathbf{P} = (0, 0, P_{\max})$. Тогда

$$\oint_l P_l dl = I = NI_1 = P_{\max} \Delta z, \quad P_{\max} = \frac{NI_1}{\Delta z}. \quad (5)$$

Нужный вид P_z можно обеспечить функцией вида

$$P_z(r) = P_{\max} \times \begin{cases} 1, & 0 \leq r \leq r_0, \\ \frac{r_{\max} - r}{r_{\max} - r_0}, & r_0 < r \leq r_{\max}, \\ 0, & r > r_{\max}. \end{cases} \quad (6)$$

3. Расчёт магнитного поля соленоида средствами программных пакетов КОМПОТ и KLONDIKE. Рассмотрим несколько способов задания геометрии соленоида в КОМПОТ. Поскольку он обладает осевой симметрией, то для расчётов естественной системой координат является цилиндрическая. Кроме того, в силу данной симметрии можно рассматривать не весь соленоид, а только его часть. С учётом вычислительной мощности оборудования и перспективы использования данной модели в последующих задачах соленоид разделён пополам осевым сечением, а для проведения расчётов выбрана одна из половин соленоида. На границе, соответствующей осевому сечению, на магнитное поле наложены граничные условия II рода, а на удалённых от расчётной модели границах для лучшей сходимости выбраны граничные условия I рода. При этом использованы два варианта размера расчётной области. В первом случае её границы отстоят от модели на один характерный размер, во втором случае – на десять характерных размеров. В качестве характерного размера выбрана длина соленоида.

Вообще говоря, остальные элементы системы могут не обладать осевой симметрией. В этом случае расчёт магнитного поля целесообразно проводить в декартовой системе координат. Вектор \mathbf{P} при

этом будет задаваться так же, как было описано выше, а граничные условия задаются аналогично цилиндрической системе: на сечении соленоида – условия II рода, на удалённых границах – условия I рода.

Программный пакет KLONDIKE является реализацией интегрального подхода и использует закон Био–Савара–Лапласа для вычисления магнитного поля. В отличие от КОМПОТ и COMSOL, в KLONDIKE при создании вычислительной модели нет необходимости задавать расчётную область и вещество, окружающее расчётную модель (например, воздух). Достаточно задать саму модель и точки, в которых необходимо вычислить вектор магнитной индукции \mathbf{B} .

Ниже приводятся расчёты магнитного поля \mathbf{B} , выполненные с помощью программных пакетов КОМПОТ (тремя способами: с заданием геометрии магнитной системы в цилиндрической системе координат в малой и большой расчётных областях, а также с заданием геометрии в декартовой системе), KLONDIKE и COMSOL. Максимальные значения модуля поля вдоль оси соленоида и соответствующие им координаты представлены в таблице 1.

Таблица 1. Максимальное значение модуля магнитного поля B вдоль оси соленоида и z -координаты, в которых оно достигается

	КОМПОТ малая область	КОМПОТ цилиндр. система координат	КОМПОТ декартова система координат	KLONDIKE	COMSOL
B , кГс	7,9806	7,6513	7,6515	7,6531	7,6520
z , см	25,825	25,825	25,825	26,600	25,638

Абсолютное и относительное отклонения значений поля, рассчитанных в КОМПОТ и KLONDIKE от расчёта COMSOL приведены на рис. 1.

Из результатов численного моделирования следует, что на расчёт в КОМПОТ большое влияние оказывает выбор размера расчётной области: большая расчётная область даёт лучшее согласие с KLONDIKE и COMSOL. Это объясняется тем, что в малой расчётной области на дальней от расчётной модели границе магнитное поле оказывается не близким к нулю, а потому граничные условия оказываются нарушены, что приводит к отклонению значений поля в целом.

Разность значений максимумов поля для разных размеров расчётной области в КОМПОТ составляет около 0,33 кГс, что составляет примерно 4%.

При этом отклонение максимума поля при выборе большой расчётной области от поля в KLONDIKE составляет 0,0005 кГс, т. е. менее 0,01%. Также в ходе проведённого исследования установлено, что

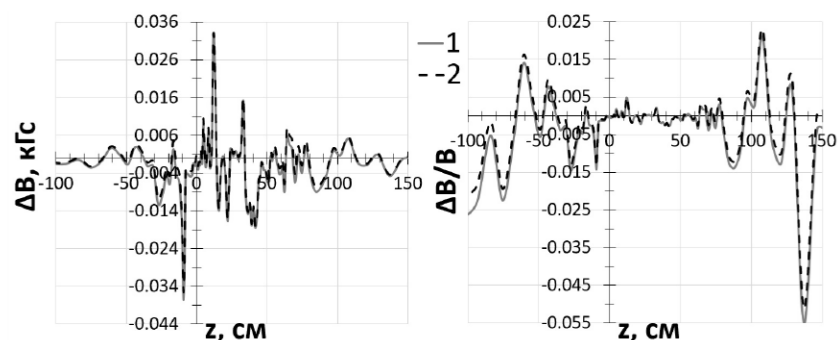


Рис. 1. Абсолютные (слева) и относительные (справа) отклонения модуля магнитного поля, рассчитанного в КОМПOT – 1 и KLONDIKE – 2 от COMSOL

при достаточно большой расчётной области выбор системы координат в КОМПOT не оказывает существенного влияния на результаты моделирования. Максимумы поля для разных систем координат отличаются на 0,0002 кГс (около 0,003%).

Различия между результатами, полученными с помощью пакета КОМПOT с декартовой системой координат и COMSOL, составило 0,0005 кГс (менее 0,01%). В свою очередь, разность между значениями полей, полученными с помощью KLONDIKE и COMSOL, составило 0,0011 кГс (около 0,014%).

Полученные результаты позволяют утверждать, что применяемые методы допустимы при создании трёхмерных моделей магнитных систем.

4. Заключение. Данная работа направлена на исследование возможности использования интегральных и дифференциальных подходов при разработке трёхмерных вычислительных моделей магнитных систем. Расчёты проведены на примере простой модели соленоида с помощью отечественных программных пакетов KLONDIKE

и КОМПОТ, а также с помощью программы COMSOL.

В ходе моделирования установлено, что на результаты расчётов в КОМПОТ в большей степени влияет размер расчётной области. При увеличении линейного размера области в 10 раз значение поля изменяется на величину порядка 4%.

Также продемонстрировано, что для КОМПОТ при выборе достаточно большой расчётной области выбор системы координат не играет существенной роли. Согласно результатам тестирования, отличие между максимумами магнитных полей составило порядка 0,003%. Результаты, полученные с помощью КОМПОТ и KLONDIKE, отличаются от расчета COMSOL на величину порядка 0,01%, что свидетельствует о применимости использованных методов в разработке трёхмерных вычислительных моделей прецизионных магнитных систем.

5. Благодарности. Авторы выражают искреннюю благодарность Д. А. Овсянникову (СПбГУ), А. Н. Мурашкину (ПИЯФ), О. А. Ковальчуку, В. П. Кухтину, Е. А. Ламзину, Д. Д. Мельникову, С. Е. Сычевскому (АО «НИИЭФА»).

Литература

1. Диоп Ш. Я., Капаркова М. В., Кухтин В. П. и др. Разработка вычислительных моделей для анализа магнитного поля в помещениях медицинских учреждений и госпиталей с целью нормализации его уровня. I. Искажение геомагнитного поля // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2022. Том 18. № 3. С. 365–378.
2. Amoskov V. M., Belov A. V., Belyakov V. A, et al. Computation technology based on КОМПОТ and KLONDIKE codes for magnetostatic simulations in tokamaks // Plasma Devices and Operations. 2008. Vol. 16. No 2. P. 89–103.
3. Френкель Я. И. Собрание избранных трудов, Т. 1: Электродинамика. М.: Изд-во Акад. наук СССР, 1956. 370 с.
4. Strang G., Fix G. An Analysis of the FE Method. New Jersey: Prentice-Hall, Englewood Cliffs, 1973. 306 p.

УДК 519.233.22

Казакова А.

Минимальная дисперсия для экспоненциального распределения

Рекомендовано к публикации профессором Буре В. М.

1. Введение. Целью работы является исследование точности оценивания параметра экспоненциального однопараметрического распределения при использовании метода максимального правдоподобия. Исследуются две формы записи распределения. Например, в [1, 2] описаны две формы записи:

$$\text{Тип 1. } F(t, \lambda) = 1 - e^{-\lambda t}, \quad f(t, \lambda) = \lambda e^{-\lambda t}, \quad t \geq 0, \lambda > 0,$$

$$\text{Тип 2. } F(t, \eta) = 1 - e^{-\frac{t}{\eta}}, \quad f(t, \eta) = \frac{1}{\eta} e^{-\frac{t}{\eta}}, \quad t \geq 0, \eta > 0.$$

Оба распределения связаны с двухпараметрическим распределением Вейбулла с параметром $\beta = 1$ (коэффициент формы). Распределение имеет функцию плотности и функцию распределения:

$$f(t, \eta, \beta) = \beta \frac{t^{\beta-1}}{\eta^\beta} \exp \left[- \left(\frac{t}{\eta} \right)^\beta \right], \quad F(t, \eta, \beta) = 1 - \exp \left[- \left(\frac{t}{\eta} \right)^\beta \right], \\ t \geq 0, \quad \eta > 0, \quad \beta > 0.$$

Одним из важных моментов рассмотрения приведенных распределений является применение схемы мгновенных отказов с постоянной интенсивностью отказов.

2. Оценивание параметра. Применим метод максимального правдоподобия [3] для получения оценок параметров для полных данных t_1, \dots, t_n .

Тип 1. Определим функцию правдоподобия

$$L_\lambda = \prod_{i=1}^n f(t_i, \lambda) = \lambda^n e^{-\lambda \sum_{i=1}^n t_i},$$

Казакова Анастасия – студент, Санкт-Петербургский государственный университет; e-mail: st092045@student.spbu.ru, тел.: +7(921)799-07-12

а также

$$(\ln L)_\lambda = \sum_{i=1}^n \ln f(t_i) = n \ln \lambda - \lambda n \bar{t}, \quad \bar{t} = \frac{1}{n} \sum_{i=1}^n t_i,$$

$$(\ln L)'_\lambda = \frac{\partial \ln L(t, \lambda)}{\partial \lambda} = \frac{n}{\lambda} - n \bar{t} = -n \left\{ \bar{t} - \frac{1}{\lambda} \right\}, \quad \bar{t} = \frac{1}{n} \sum_{i=1}^n t_i.$$

Решение уравнение правдоподобия S_λ дает оценку $\hat{\lambda}$. Контролируем переменную знака второй производной $(\ln L)''_\lambda$:

$$S_\lambda = (\ln L)'_\lambda = \frac{\partial \ln L(t, \lambda)}{\partial \lambda} = -n \left\{ \bar{t} - \frac{1}{\lambda} \right\} = 0, \quad (\ln L)''_\lambda < 0.$$

Тун 2. Аналогично получим решение для параметра η второго типа экспоненциального распределения:

$$L_\eta = L(t, \eta) = \prod_{i=1}^n f(t_i, \eta) = \frac{1}{\eta} e^{-\frac{1}{\eta} \sum_{i=1}^n t_i},$$

$$(\ln L)_\eta = \sum_{i=1}^n \ln f(t_i) = n \ln \frac{1}{\eta} - \frac{1}{\eta} \sum_{i=1}^n t_i = -n \ln \eta - \frac{1}{\eta} \sum_{i=1}^n t_i,$$

$$(\ln L)'_\eta = \frac{n}{\eta^2} \{ \bar{t} - \eta \},$$

$$S_\eta = (\ln L)'_\eta = \frac{n}{\eta^2} \{ \bar{t} - \eta \} = 0.$$

Тогда $\hat{\eta} = \bar{t}$. Также получаем $(\ln L)''_\eta < 0$.

Можно видеть, что оценками параметров являются статистики T от выборочного среднего \bar{t} наблюдений t_1, \dots, t_n .

3. Точность оценивания скалярного параметра θ . Для определения дисперсии оценок параметров λ, η воспользуемся подходом, изложенным в [3, 4]. Сначала решим задачу в общем виде для скалярного параметра θ , имеющего в качестве оценки $\hat{\theta}$ некоторую функцию $\tau(\theta)$, где $\theta = \lambda$ или $\theta = \eta$. Обозначим производные функции правдоподобия для параметра θ :

$$L' = \frac{\partial L(t, \theta)}{\partial \theta}, \quad L'' = \frac{\partial^2 L(t, \theta)}{\partial \theta^2},$$

$$(\ln L)' = \frac{\partial \ln L(t, \theta)}{\partial \theta}, \quad (\ln L)'' = \frac{\partial^2 \ln L(t, \theta)}{\partial \theta^2}.$$

Для обоих распределений выполняются условия регулярности для функции правдоподобия:

1. Вероятность совместного наблюдения n событий

$$\int_R L dt = 1, \quad L = L(t, \theta) = \prod_{i=1}^n f(t, \theta), \quad L > 0, \quad \theta \in \Theta \subset R.$$

2. Дифференцируемость функции правдоподобия и ее логарифма по θ :

- можно изменять порядок дифференцирования и интегрирования, когда пределы интегрирования не зависят от θ ;
- есть связь математического ожидания двух производных и самих производных.

Приведем связи в подинтегральных функциях:

$$\int_R L dt = 1, \quad \implies \quad \int_R L' dt = 0;$$

или для $\ln L$

$$\int_R \left(\frac{1}{L} L' \right) L dt = M \{ (\ln L)' \} = 0.$$

Продифференцируем последнее выражение и определим математические ожидания:

$$\int_R \left\{ \left(\frac{1}{L} L' \right) L' + L \left(\frac{1}{L} L' \right)' \right\} dt = 0,$$

$$\int_R \left\{ \left((\ln L)' \right)^2 + (\ln L)'' \right\} L dt = 0,$$

$$M \left\{ \left((\ln L)' \right)^2 \right\} + M \{ (\ln L)'' \} = 0,$$

$$M \left\{ \left((\ln L)' \right)^2 \right\} = -M \{ (\ln L)'' \}.$$

3. Решение ищется для несмещенной оценки T функции $\tau(\theta)$ от параметра θ . Потребуем равенства математического ожидания статистики (оценки) T и функции от параметра $\tau(\theta)$, а также чтобы математическое ожидание расхождения величин $\{T - \tau(\theta)\}$ было минимальным:

$$\begin{aligned} M(T) &= M\{\tau(\theta)\} = \tau(\theta), \\ M\{T - \tau(\theta)\} &= \int_R (T - \tau(\theta)) L dt = 0. \end{aligned} \quad (1)$$

3.1. Неравенство Рао – Шварца – Крамера. Условия регулярности для функции правдоподобия L приводят к неравенству Рао – Шварца – Крамера для дисперсии оценки функции. Дифференцируя левую и правую части (1) с учетом условий регулярности и возведения в квадрат обеих частей, а также применив неравенство Коши – Шварца – Буняковского для математического ожидания случайных величин η_1, η_2 , $[M\{\eta_1 \eta_2\}]^2 \leq M\eta_1^2 M\eta_2^2$, получаем неравенство Рао – Шварца – Крамера относительно дисперсии статистики:

$$\begin{aligned} D(T) &= M\left\{(T - \tau(\theta))^2\right\} \geq \frac{\{-\tau'(\theta)\}^2}{M\left\{(\ln L)'\right\}^2}, \\ \Rightarrow D(T) &\geq \frac{\{-\tau'(\theta)\}^2}{M\left\{(\ln L)'\right\}^2} = -\frac{\{\tau'(\theta)\}^2}{M\{(\ln L)''\}}. \end{aligned} \quad (2)$$

3.2. Условие достижения нижней границы неравенства. В [4] определяется необходимое и достаточное условие для достижения нижней границы неравенства Рао – Шварца – Крамера как пропорциональность $(T - \tau(\theta))$ и $(\ln L)'_{\theta}$ на всем пространстве определения параметра:

$$(\ln L)' = \frac{\partial \ln L}{\partial \theta} = A(\theta)\{T - \tau(\theta)\}, \quad (3)$$

где A не зависит от наблюдений, но может быть функцией от θ , T – статистика для θ , $\tau(\theta)$ – некоторая функция только от θ . Выпишем дисперсию

$$D\left((\ln L)'\right) = M\left\{\left((\ln L)'\right)^2\right\} = \{A(\theta)\}^2 D(T).$$

Достижение нижней границы неравенства (2):

$$D(T) = \frac{\{\tau'(\theta)\}^2}{\{A(\theta)\}^2 D(T)}, \implies D(T) = \left| \frac{\tau'(\theta)}{A(\theta)} \right|.$$

3.3. Информация Фишера. Определяется количество информации, которое может дать выборка данных о неизвестном параметре. При оценивании самого параметра θ для несмещенной оценки будем иметь $\tau'(\theta) = 1$ и дисперсия

$$D(T) = \frac{\{\tau'(\theta) = 1\}^2}{M\left\{\left((\ln L)'\right)^2\right\}} = \frac{1}{M\left\{\left((\ln L)'\right)^2\right\}}.$$

Информационное количество Фишера

$$I_n(\hat{\theta}) = M\left\{\left((\ln L)'\right)^2\right\} = \{A(\theta)\}^2 D(T).$$

Также получаем эквивалентные функции для второй производной:

$$M\left\{(\ln L)''\right\} = -|A(\theta)|\tau'(\theta) = \{A(\theta)\}^2 D(T) = (\ln L)''.$$

Следовательно, статистика T будет оценкой функции $\tau(\theta)$ с минимально достижимой границей дисперсии неравенства Рао – Шварца – Крамера (Minimum Variance Unbiased Estimator MVUE). В силу достижения нижней границы эта статистика по [4] будет иметь свойство достаточной статистики для $\tau(\theta)$. Свойства оценки максимального правдоподобия для $\tau(\theta)$:

- статистика T является:
 - несмещенной оценкой для параметра θ ,
 - достаточной статистикой, так как является оценкой для параметра θ с минимально достижимой дисперсией;
- оценка $\hat{\theta}$ является:
 - эффективной оценкой, так как является несмещенной оценкой и достигается нижняя граница ее дисперсии в неравенстве Рао – Шварца – Крамера,
 - асимптотически состоятельной оценкой, так как увеличение точности обеспечивается с увеличением объема выборки.

4. Точность оценивания для параметра λ , тип 1. Представим $(\ln L)'_{\lambda}$ в виде (3), где выполнены условия достижения нижней границы неравенства для функции $\tau(\lambda) = \frac{1}{\lambda}$:

$$(\ln L)'_{\lambda} = A(\theta)\{T - \tau(\theta)\} = -n\left\{\bar{t} - \frac{1}{\lambda}\right\},$$

$$A(\theta) = -n, \quad T_{(\frac{1}{\lambda})} = \bar{t}, \quad \tau(\theta) = \tau(\lambda) = \frac{1}{\lambda},$$

$$\frac{\partial \tau(\lambda)}{\partial \lambda} = \tau'_{\lambda}(\lambda) = -\frac{1}{\lambda^2} = -\left(\frac{1}{\lambda}\right)^2.$$

Оценка параметра $\frac{1}{\lambda}$, ее дисперсия и информационное количество Фишера будут следующие:

$$(\ln L)'_{(\frac{1}{\lambda})} = -n\left\{\bar{t} - \frac{1}{\lambda}\right\} = 0, \quad \widehat{\left(\frac{1}{\lambda}\right)} = \bar{t},$$

$$D(T) = D(\bar{t})_{(\frac{1}{\lambda})} = \left| \frac{\tau'(\theta)}{A(\theta)} \right| = \left| -\frac{1}{n} \frac{d}{d\lambda} \frac{1}{\lambda} \right| = \frac{1}{n\lambda^2} = \frac{1}{n} \left(\frac{1}{\lambda}\right)^2,$$

$$D\left(\widehat{\left(\frac{1}{\lambda}\right)}\right) = \frac{1}{n}(\bar{t})^2,$$

$$I_n\left(\frac{1}{\lambda}\right) = M\left\{\left((\ln L)'_{(\frac{1}{\lambda})}\right)^2\right\} = \left\{\{A(\theta)\}^2 D(T)\right\} = n\left(\frac{1}{\lambda}\right)^2,$$

$$M\left\{(\ln L)''_{(\frac{1}{\lambda})}\right\} = (\ln L)''_{(\frac{1}{\lambda})} = -n\left(\frac{1}{\lambda}\right)^2.$$

Следовательно, дисперсия

$$D(\widehat{\lambda}) > D\left(\widehat{\left(\frac{1}{\lambda}\right)}\right) = \frac{(\bar{t})^2}{n}. \quad (4)$$

5. Точность оценивания параметра η , тип 2. Представим $(\ln L)'_{\eta}$ в виде (3) и решение уравнения правдоподобия:

$$(\ln L)'_{\eta} = \frac{n}{\eta^2}\{\bar{t} - \eta\} = 0, \quad A(\theta) = \frac{n}{\eta^2}, \quad T_{\eta} = \bar{t},$$

$$\tau(\theta) = \tau(\eta) = \eta, \quad \tau'_{\eta}(\eta) = 1.$$

$$S_\eta = (\ln L)'_\eta = -\frac{n}{\eta^2} \{\eta - \bar{t}\} = 0, \quad \hat{\eta} = \bar{t}.$$

Тогда $T_\eta = \bar{t}$ есть достаточная статистика для η и является оценкой с минимально достижимой дисперсией. Дисперсия $D(T_\eta)$ определена для η :

$$D(\bar{t}) = \left| \frac{\tau'(\eta)}{A(\eta)} \right| = \frac{\eta^2}{n},$$

$$D(\hat{\eta}) = \frac{(\bar{t})^2}{n}. \quad (5)$$

Определена информация Фишера $I_n(\hat{\eta})$, содержащаяся в выборке размером n , стандартное отклонение s_η и коэффициент вариации \hat{v}_η , которые характеризуют точность оценивания:

$$I_n(\hat{\eta}) = M\left\{((\ln L)')^2\right\} = \{A(\eta)\}^2 D(T) = \frac{n}{\eta^2},$$

$$M((\ln L)'') = M\left(\frac{\partial^2 \ln L}{\partial \eta^2}\right) = M\left(-\frac{n}{\eta^3} \{2\bar{t} - \eta\}\right) =$$

$$= -\frac{n}{\eta^3} \{2M(\bar{t}) - \eta\} = -\frac{n}{\eta^3} \{2\eta - \eta\} = -\frac{n}{\eta^2},$$

$$M\left\{((\ln L)')^2\right\} = -M((\ln L)'') = \frac{n}{\eta^2},$$

$$s_\eta = \sqrt{D(\hat{\eta})}, \quad v_\eta = \frac{s_\eta}{\hat{\eta}} = \frac{1}{\sqrt{n}}.$$

Таким образом, для второго типа выполнено необходимое и достаточное условие (3) для достижения нижней границы неравенства Рао–Шварца–Крамера. Выполнены требования по основным свойствам, которые предъявляются к оценкам параметра – состоятельность, несмещенность, эффективность. Второй тип распределения более предпочтителен.

6. Сравнение результатов для рассмотренных двух типов распределений. Имеем результаты (4) для функции $\tau(\lambda)$ от параметра λ и (5) для параметра η . Очевидна функциональная связь:

$$\hat{\eta} = \widehat{\left(\frac{1}{\lambda}\right)} = \bar{t}, \quad D(\hat{\eta}) = D\left(\widehat{\left(\frac{1}{\lambda}\right)}\right) = \frac{(\bar{t})^2}{n}.$$

Пример. Рассмотрена полная выборка из работы [5], приведенная в виде данных эксперимента по времени пробоя изолирующей жидкости (жидкий конденсатор) между электродами при напряжении 30 кВт. По теории распределение определяется как экспоненциальное. Выборка объема $n = 11$ содержит элементы: 7,74; 17,05; 20,46; 21,02; 22,66; 43,4; 47,3; 139,07; 144,12; 175,88; 194,9. Значения оценки, дисперсии, коэффициента вариации, и информации Фишера по выборке представлены в таблице.

Таблица. Результаты оценивания по выборке $n = 11$

$\hat{\eta}$	$D(\hat{\eta})$	v_{η}	$I_n(\hat{\eta})$
75,782	522,0804	0,3085	0,00191

7. Заключение. Из двух рассмотренных типов экспоненциального распределения установлено, что второй тип распределения с параметром масштаба η предпочтителен для использования. Достаточной статистикой является выборочное среднее, а $\hat{\eta}$ – несмещенной оценкой с минимально достижимой дисперсией. Для первого типа распределения с параметром λ не существует минимальной достижимой дисперсии оценки параметра, соответственно, $\hat{\lambda}$ – смещенная оценка. Параметры η и $\frac{1}{\lambda}$ двух типов распределений связаны функционально.

Литература

1. Кесечиоглу D. B. Reliability Engineering Handbook. Lancaster, Pennsylvania, Pa: DEStech Publications, 2002. 679 p.
2. McCool J. Using the Weibull Distribution: Reliability, Modeling, and Inference. Hoboken, New Jersey, Pa: John Wiley & Sons, 2012. 336 p.
3. Буре В. М., Парилина Е. М. Теория вероятностей и математическая статистика. СПб.: Лань, 2013. 416 с.
4. Кендалл М., Стьюарт А. Статистические выводы и связи. М.: Наука, 1973. 900 с.
5. Nelson W. Applied Life Data Analysis. New York, Pa: John Wiley & Sons, 1982. 634 p.

УДК 531.355

Кикенов А. А.

Математическая модель конвертоплана

Рекомендовано к публикации доцентом Шиманчуком Д. В.

1. Введение. Сейчас в технике как никогда актуальны вопросы управления [1,2], в частности, вопросы управления БПЛА [3,4]. Особый интерес вызывают БПЛА типа VTOOL. Для построения эффективного алгоритма управления такого летательного аппарата требуется правильно описать его математически. В данной работе выводится математическая модель конвертоплана.

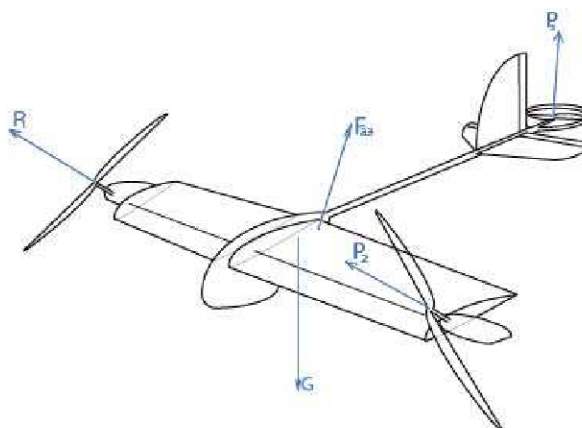


Рис. 1. Компонировка рассматриваемого в работе конвертоплана

2. Схема конвертоплана. Пусть БПЛА имеет трикоптерную схему (см. рис. 1) с двумя независимыми поворотными мотогондолами и одним задним двигателем, установленным так, чтобы его аэродинамический и реактивный моменты вокруг оси Oy связанной системы координат (СК) компенсировали друг друга. БПЛА спроектирован так, что поворот мотогондол не меняет его общий тензор

Кикенов Арманжан Александрович – магистрант, Санкт-Петербургский государственный университет; e-mail: st075803@student.spbu.ru, тел.: +7(981)892-85-14

инерции, а оси связанной СК параллельны главным осям его тензора инерции. Атмосферу считаем однородной и неподвижной.

3. Математическая модель. Уравнения будем записывать в связанной системе координат $Oxyz$ [5]. Углы будут использоваться самолетные: тангаж θ , крен γ и рысканье ψ . В самолетных углах кинематические уравнения Эйлера [6] запишутся в виде

$$\begin{cases} \dot{\psi} = \frac{\omega_y \cos \gamma - \omega_z \sin \gamma}{\cos \theta}, \\ \dot{\theta} = \omega_y \sin \gamma - \omega_z \cos \gamma, \\ \dot{\gamma} = \omega_x - \operatorname{tg}(\omega_y \cos \gamma - \omega_z \sin \gamma), \end{cases} \quad (1)$$

а движение БПЛА как материальной точки в виде

$$\begin{cases} m(\dot{v}_x + \omega_y v_z - \omega_z v_y) = P_1 \cos A_1 + P_2 \cos A_2 + X - G \sin \theta, \\ m(\dot{v}_y + \omega_z v_x - \omega_x v_z) = P_1 \sin A_1 + P_2 \sin A_2 + P_3 \cos B + \\ \quad + Y - G \cos \theta \cos \gamma, \\ m(\dot{v}_z + \omega_x v_y - \omega_y v_x) = P_3 \sin B + Z + G \cos \theta \sin \gamma, \end{cases} \quad (2)$$

где $P_1 = Q\Omega_1^2$, $P_2 = Q\Omega_2^2$, $P_3 = q\Omega_3^2$ – силы тяг двигателей, зависящие квадратично от угловых скоростей вращения винтов, A_1 , A_2 – углы поворота мотогондол, B – угол установки заднего двигателя. Проекции аэродинамических сил, действующих на БПЛА, на связанные оси запишутся в виде

$$X = c_x \frac{\rho V^2}{2} S, \quad Y = c_y \frac{\rho V^2}{2} S, \quad Z = c_z \frac{\rho V^2}{2} S,$$

где $c_{x,y,z}(\alpha, \beta, A_1, A_2)$ – аэродинамические коэффициенты, представленные в виде полиномов четырех переменных, где α – угол атаки, β – угол скольжения БПЛА, $G = mg$ – сила тяжести, ω_x , ω_y , ω_z – проекции угловых скоростей БПЛА на связанные оси, m – масса БПЛА.

Известно [7], что

$$\frac{d\mathbf{K}}{dt} = \mathbf{M}. \quad (3)$$

Для рассматриваемой модели

$$\mathbf{K} = \mathbf{K}_s + \mathbf{K}_1^o + \mathbf{K}_2^o + \mathbf{K}_1^p + \mathbf{K}_2^p + \mathbf{K}_3^p,$$

где \mathbf{K}_s – кинетический момент БПЛА с неподвижными гондолами и винтами, $\mathbf{K}_{1,2}^o$ – кинетические моменты гондол с неподвижными винтами, $\mathbf{K}_{1,2,3}^p$ – кинетические моменты винтов.

При наших допущениях \mathbf{K}_s и центр тяжести БПЛА не зависят от углов поворота мотогондол $A_{1,2}$. Соответствующие тензоры

$$I_s = \begin{pmatrix} I_x & 0 & 0 \\ 0 & I_y & 0 \\ 0 & 0 & I_z \end{pmatrix}, \quad I_o^{1,2} = \begin{pmatrix} I_1^o & 0 & 0 \\ 0 & I_1^o & 0 \\ 0 & 0 & I_2^o \end{pmatrix},$$

$$I_p^{1,2} = \begin{pmatrix} I_1^p & 0 & 0 \\ 0 & I_2^p & 0 \\ 0 & 0 & I_2^p \end{pmatrix}, \quad I_p^3 = \begin{pmatrix} I_1^d & 0 & 0 \\ 0 & I_2^d & 0 \\ 0 & 0 & I_1^d \end{pmatrix}.$$

Таким образом

$$\mathbf{K} = I_s \boldsymbol{\omega} + I^o (\dot{\mathbf{A}}_1 + \dot{\mathbf{A}}_2) + I_p^{1,2} (\boldsymbol{\Omega}_1 + \boldsymbol{\Omega}_2) + I_p^3 \boldsymbol{\Omega}_3,$$

где $\boldsymbol{\Omega}_{1,2,3}$ – угловые скорости вращения винтов. В связанной СК

$$\dot{\mathbf{A}}_{1,2} = \begin{pmatrix} 0 \\ 0 \\ \dot{A}_{1,2} \end{pmatrix}, \quad \boldsymbol{\Omega}_{1,2} = \begin{pmatrix} \Omega_{1,2} \cos A_{1,2} \\ \Omega_{1,2} \sin A_{1,2} \\ 0 \end{pmatrix}.$$

Тогда

$$\mathbf{K} = \begin{pmatrix} I_x \omega_x + I_1^p (\Omega_1 \cos A_1 + \Omega_2 \cos A_2) \\ I_y \omega_y + I_2^d \Omega_3 \cos B + I_2^p (\Omega_1 \sin A_1 + \Omega_2 \sin A_2) \\ I_z \omega_z + I_1^d \Omega_3 \sin B + I_2^o (\dot{A}_1 + \dot{A}_2) \end{pmatrix},$$

$$\dot{\mathbf{K}} = \begin{pmatrix} I_x \dot{\omega}_x + (I_z - I_y) \omega_y \omega_z + \omega_y I_1^d \Omega_3 \sin B - \omega_z I_2^d \Omega_3 \cos B \\ I_y \dot{\omega}_y + (I_y - I_z) \omega_z \omega_x + I_2^d \dot{\Omega}_3 \cos B - \omega_x I_1^d \Omega_3 \sin B \\ I_z \dot{\omega}_z + (I_z - I_x) \omega_x \omega_y + I_1^d \dot{\Omega}_3 \sin B + \omega_x I_2^d \Omega_3 \cos B \end{pmatrix} +$$

$$+ \begin{pmatrix} I_1^p (\dot{\Omega}_1 \cos A_1 + \dot{\Omega}_2 \cos A_2 - \Omega_1 \dot{A}_1 \sin A_1 - \Omega_2 \dot{A}_2 \sin A_2) \\ I_2^p (\dot{\Omega}_1 \sin A_1 + \dot{\Omega}_2 \sin A_2 + \Omega_1 \dot{A}_1 \cos A_1 + \Omega_2 \dot{A}_2 \cos A_2) \\ I_2^o (\ddot{A}_1 + \ddot{A}_2) \end{pmatrix} +$$

$$+ \begin{pmatrix} \omega_y I_2^o (\dot{A}_1 + \dot{A}_2) - \omega_z I_1^p (\Omega_1 \sin A_1 + \Omega_2 \sin A_2) \\ \omega_x I_2^o (\dot{A}_1 + \dot{A}_2) + \omega_z I_1^p (\Omega_1 \cos A_1 + \Omega_2 \cos A_2) \\ \omega_x I_2^p (\Omega_1 \sin A_1 + \Omega_2 \sin A_2) - \omega_y I_1^p (\Omega_1 \cos A_1 + \Omega_2 \cos A_2) \end{pmatrix}.$$

Момент внешних сил

$$\mathbf{M} = \sum_{i=1}^3 \left(\mathbf{M}_i^p + \mathbf{M}_i^\Gamma \right) + \mathbf{M}_E + \mathbf{M}_e + \mathbf{M}_\chi,$$

где $\mathbf{M}_{1,2,3}^p$ – реактивные моменты винтов, $\mathbf{M}_{1,2}^\Gamma = -L\Omega_{1,2}\mathbf{\Omega}_{1,2}$, $\mathbf{M}_3^\Gamma = -l\Omega_3\mathbf{\Omega}_3$ – аэродинамические моменты винтов, \mathbf{M}_E – аэродинамический момент всего конвертоплана, \mathbf{M}_e – момент элеронов, \mathbf{M}_χ – момент рулей высоты и направления.

$$\mathbf{M}_E = \frac{\rho LV^2}{2} S \begin{pmatrix} c_{mx}(\alpha, \beta, A_1, A_2) \\ c_{my}(\alpha, \beta, A_1, A_2) \\ c_{mz}(\alpha, \beta, A_1, A_2) \end{pmatrix},$$

$$\mathbf{M}_e = \frac{\rho LV^2}{2} S \begin{pmatrix} c_{mx}^e(\alpha, \delta_e) \\ c_{my}^e(\alpha, \delta_e) \\ 0 \end{pmatrix}, \quad \mathbf{M}_\chi = \frac{\rho LV^2}{2} S \begin{pmatrix} c_{mx}^{\chi_1}(\beta, \delta_{\chi_1}) \\ c_{my}^{\chi_1}(\beta, \delta_{\chi_1}) \\ c_{mz}^{\chi_2}(\alpha, \delta_{\chi_2}) \end{pmatrix}$$

где δ_e, χ_1, χ_2 – углы отклонения элеронов и рулей направления и высоты, соответственно.

$$\begin{aligned} \mathbf{M} = & \begin{pmatrix} M_1^\Gamma \cos A_1 - M_2^\Gamma \cos A_2 \\ M_1^\Gamma \sin A_1 - M_2^\Gamma \sin A_2 + M_3^\Gamma \cos B - x_3 P_3 \sin B \\ M_3^\Gamma \sin B + x_3 P_3 \cos B \end{pmatrix} + \\ & + \begin{pmatrix} z_o(P_2 \sin A_2 - P_1 \sin A_1) \\ z_o(P_1 \cos A_1 - P_2 \cos A_2) \\ x_1 P_1 \sin A_1 - y_1 P_1 \cos A_1 + x_2 P_2 \sin A_2 - y_2 P_2 \cos A_2 \end{pmatrix} + \\ & + \frac{\rho LV^2}{2} S \begin{pmatrix} c_{mx}(\alpha, \beta) + c_{mx}^e(\alpha, \delta_e) + c_{mx}^{\chi_1}(\beta, \delta_{\chi_1}) \\ c_{my}(\alpha, \beta) + c_{my}^e(\alpha, \delta_e) + c_{my}^{\chi_1}(\beta, \delta_{\chi_1}) \\ c_{mz}(\alpha, \beta) + c_{mz}^{\chi_2}(\alpha, \delta_{\chi_2}) \end{pmatrix}, \end{aligned}$$

где $[x_{1,2,3}, y_{1,2,3}, z_{1,2,3}]^T$ – координаты винтов.

Связь проекций скоростей в нормальной и связанной СК отражена в формуле

$$[\dot{X}_g, \dot{Y}_g, \dot{Z}_g]^T = \zeta [v_x, v_y, v_z]^T, \quad (4)$$

где $[X_g, Y_g, Z_g]$ – координаты центра масс БПЛА в нормальной СК, а матрица ζ имеет вид

$$\begin{pmatrix} \cos \psi \cos \theta & \sin \psi \sin \gamma - \cos \psi \sin \theta \cos \gamma & \sin \psi \sin \gamma + \cos \psi \sin \theta \sin \gamma \\ \sin \theta & \cos \theta \cos \gamma & -\cos \theta \sin \gamma \\ -\sin \psi \cos \theta & \cos \psi \sin \gamma + \sin \psi \sin \theta \cos \gamma & \cos \psi \cos \gamma - \sin \psi \sin \theta \sin \gamma \end{pmatrix}$$

Полученные 12 уравнений (1)–(4) могут быть представлены в виде системы из 19 уравнений в нормальной форме.

4. Заключение. Построена математическая модель БПЛА типа конвертоплан. Она представлена в виде системы из 19 дифференциальных уравнений в нормальной форме. Вектор фазового состояния имеет вид

$$[X_g, Y_g, Z_g, v_x, v_y, v_z, \psi, \theta, \gamma, \omega_x, \omega_y, \omega_z, A_1, \dot{A}_1, A_2, \dot{A}_2, \Omega_1, \Omega_2, \Omega_3]^T,$$

а вектор управляющих параметров – $[\ddot{A}_1, \ddot{A}_2, \dot{\Omega}_1, \dot{\Omega}_2, \dot{\Omega}_3, \delta_e, \delta_{\chi_1}, \delta_{\chi_2}]^T$.

Литература

1. Смирнова М. А., Смирнов М. Н., Смирнов Н. В. Система многоцелевого управления роботом-манипулятором // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2022. Т. 18. № 4. С. 621–630.
2. Жабко А. П., Жабко Н. А., Яковлев П. В. Метод оптимального демпфирования В. И. Зубова в задачеуправления одной гироскопической системой // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2022. Т. 18. № 2. С. 278–284.
3. Непочирук А. А., Лепихин Е. А., Щербинин А. В. Моделирование и управление квадрокоптером // Процессы управления и устойчивость. 2017. Т. 4. № 1. С. 445–451.
4. Баранов О. В., Жолобов Е. В. Моделирование аварийных управляемых движений квадрокоптера // Процессы управления и устойчивость. 2019. Т. 6. № 1. С. 213–217.
5. Микеладзе В. Г., Титов В. М. Основные геометрические и аэродинамические характеристики самолетов и ракет. М.: Машиностроение, 1982. 149 с.
6. Беспилотные летательные аппараты / под ред. Л. С. Чернобровкина. М.: Машиностроение, 1967. 440 с.
7. Бабаджанянц Л. К., Пупышев Ю. А., Пупышева Ю. Ю. Классическая механика. СПб.: Изд-во СПбГУ, 2011. 259 с.

УДК 539.3

Кондратьева А. Д.

Выпучивание под собственным весом нанокантилевера с упругим поворотом

Рекомендовано к публикации доцентом Бочкарёвым А. О.

1. Введение. Нанотехнологии широко используются при производстве различных датчиков, акселерометров и т. п. Несмотря на наноразмеры, моделирование их механического поведения не обходится без континуальной механики. Один из её подходов – учёт поверхностной энергии на границах наноструктуры, что с определённой долей достоверности позволяет описать размерный эффект. Работа посвящена моделированию выпучивания под действием собственного веса нанокантилевера с упругим поворотом.

2. Модели изгиба наностержня. В данной работе будем исходить из двух моделей изгиба стержня: Эйлера – Бернулли (с гипотезой прямых поперечных сечений, остающимися ортогональными изогнутой оси стержня) и Тимошенко [1] (поперечные сечения претерпевают сдвиг с осью стержня). Для описания поверхностных эффектов используем модель поверхностной упругости Гуртина – Мёрдоха [2, 3].

2.1. Кинематические и определяющие соотношения. Будем считать, что изгиб стержня длины l и прямоугольного поперечного сечения $[b \times c]$ происходит в плоскости xz , где x – ось, направленная вертикально вверх вдоль оси стержня. Сформулируем кинематические и определяющие соотношения:

- гипотезы Кирхгофа о смещениях:

$$u_x = u(x) - z\vartheta(x), \quad u_z = w(x);$$

- гипотезы Тимошенко о деформациях:

$$\varepsilon_{xx} = \frac{\partial u_x}{\partial x} = \varepsilon - z\vartheta'(x), \quad \gamma_{zx} = \frac{\partial u_z}{\partial x} + \frac{\partial u_x}{\partial z} = w'(x) - \vartheta(x);$$

Кондратьева Алёна Дмитриевна – студент, Санкт-Петербургский государственный университет; e-mail: st087638@student.spbu.ru, тел.: +7(996)940-75-70

$$\gamma_{zx} = \frac{\partial u_z}{\partial x} + \frac{\partial u_x}{\partial z} = w'(x) - \vartheta(x);$$

- обратный закон Гука для внутренних (объёмных) напряжений:

$$\sigma_{xx} = E\varepsilon_{xx}, \quad \sigma_{zx} = k\mu\gamma_{zx} = k\mu\gamma;$$

- поверхностные напряжения по модели Гуртина–Мёрдоха [2, 3] на гранях $z = \pm \frac{c}{2}$:

$$\tau_{xx} = \tau_0 + (2\mu_s + \lambda_s - \tau_0)\varepsilon_{xx} + \tau_0 \frac{\partial u_x}{\partial x}, \quad \tau_{zx} = \tau_0 \frac{\partial u_z}{\partial x};$$

- поверхностные напряжения по модели Гуртина–Мёрдоха [2, 3] на гранях $y = \pm \frac{b}{2}$:

$$\tau_{xx} = \tau_0 + (2\mu_s + \lambda_s - \tau_0)\varepsilon_{xx} + \tau_0 \frac{\partial u_x}{\partial x}, \quad \tau_{zx} = 2(\mu_s - \tau_0)\varepsilon_{zx} + \tau_0 \frac{\partial u_z}{\partial x},$$

где $w(x)$ – прогиб, $\vartheta(x)$ – угол поворота нормали сечения к оси стержня, E – модуль Юнга, k – сдвиговой коэффициент Тимошенко, μ – модуль сдвига, μ_s, λ_s – поверхностные константы, τ_0 – предварительное поверхностное натяжение.

2.2. Энергия деформации наностержня. Как показано в работе [4], при данных кинематических и определяющих соотношениях выражения полной энергии деформации наностержня по моделям Тимошенко и Эйлера–Бернулли имеют соответственно вид

$$U_{\text{Tim}} = \frac{1}{2} \int_0^l (p\tau_0(2 + 2\varepsilon + u'^2 + w'^2) + SE_t^* \varepsilon^2 + S\mu_b^* \gamma^2 + IE_b^* \vartheta'^2) dx, \quad (1)$$

$$U_{\text{EB}} = \frac{1}{2} \int_0^l (p\tau_0(2 + 2\varepsilon + u'^2 + w'^2) + SE_t^* \varepsilon^2 + IE_b^* w''^2) dx, \quad (2)$$

где введены эффективные модули при растяжении (t) и изгибе (b)

$$E_t^* = E + \frac{p}{S}(2\mu_s + \lambda_s), \quad E_b^* = E + \frac{c^2(3b + c)}{6I}(2\mu_s + \lambda_s),$$

$$\mu_b^* = \mu^* + \frac{2c}{S}(\mu_s - \tau_0), \quad \mu^* = k\mu = \frac{5}{6} \frac{E}{2(1 + \nu)},$$

$$I = \frac{bc^3}{12}, \quad S = bc, \quad p = 2(b + c).$$

3. Метод Ритца при выпучивании под собственным весом нанокантилевера с упругим поворотом. Рассмотрим нанокантилевер с опорным нижним концом, который может упруго поворачиваться согласно следующим граничным условиям:

$$\begin{aligned} w(0) &= 0, \\ M_{\text{Tim}}(0) &= \eta \vartheta(0), \\ M_{\text{EB}}(0) &= \eta w'(0), \end{aligned} \quad (3)$$

где η – коэффициент поворотной жёсткости. Стоит отметить, что при $\eta \rightarrow \infty$ решение нашей задачи должно совпасть со случаем классической жёсткой заделки.

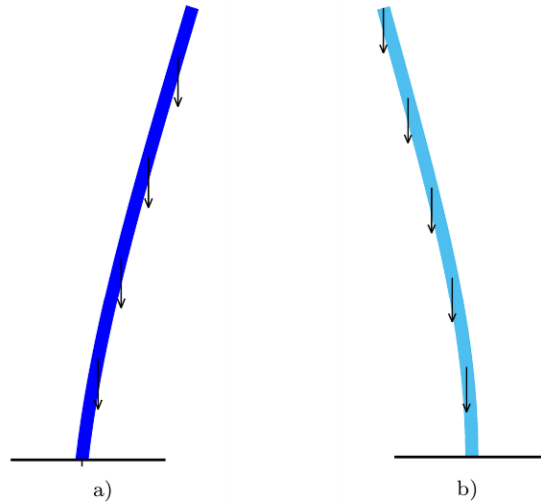


Рис. 1. Выпучивание нанокантилевера под собственным весом:
а) с упругим поворотом, б) с жёсткой заделкой

Вдоль оси стержня действует линейно распределенная сжимающая нагрузка, равная $-q(l-x)$, где q – удельный собственный вес на единицу длины. Решается задача о потери устойчивости прямолинейной формы равновесия нанокантилевера под действием собственного веса (см. рис. 1).

3.1. Метод Ритца по модели Тимошенко. Приращение энергии изгиба (1) по модели Тимошенко при выпучивании под собственным весом и с граничными условиями (3) имеет вид

$$\Delta U = \frac{1}{2}\eta\vartheta^2(0) + \frac{1}{2} \int_0^l \left((2p\tau_0 - q(l-x))w'^2 + S\mu_b^*(w' - \vartheta)^2 + IE_b^*\vartheta'^2 \right) dx. \quad (4)$$

Здесь $\eta\vartheta^2(0)$ – дополнительное слагаемое, описывающее энергию поворотной опоры.

Согласно методу Ритца представим прогиб и угол в виде линейной комбинации базовых функций, каждая из которых удовлетворяет кинематической части граничных условий (3):

$$w = \sum_n^N A_n w_n, \quad \vartheta = \sum_n^N B_n \vartheta_n. \quad (5)$$

Подставив (5) в (4) и рассматривая A и B как обобщённые координаты, условия равновесия можно записать в матричном виде:

$$\begin{aligned} \frac{\delta U}{\delta A} &= \left(\frac{(2p\tau_0 + S\mu_b^*)}{l} U_{11} - qU_{12} \right) A - S\mu_b^* U_{21} B = 0, \\ \frac{\delta U}{\delta B} &= \left(lS\mu_b^* U_{22} + \frac{IE_b^*}{l} U_{23} \right) B - S\mu_b^* U_{21}^T A = 0, \end{aligned} \quad (6)$$

где

$$\begin{aligned} \{U_{11}\}_{mn} &= \int_0^l w'_n w'_m d\bar{x}, & \{U_{12}\}_{mn} &= \int_0^l (1 - \bar{x}) w'_n w'_m d\bar{x}, \\ \{U_{21}\}_{mn} &= \int_0^l w'_n \vartheta'_m d\bar{x}, & \{U_{22}\}_{mn} &= \int_0^l (1 - \bar{x}) \vartheta'_n \vartheta'_m d\bar{x}, \\ \{U_{23}\}_{mn} &= K \vartheta_n(0) \vartheta'_m(0) + \int_0^l (1 - \bar{x}) \vartheta'_n \vartheta'_m d\bar{x}, & \bar{x} &= \frac{x}{l}, \quad K = \frac{\eta l}{E_b^* I}. \end{aligned}$$

Выражая из условий равновесия (6) вектор B через A , приходим к обобщённой задаче на отыскание собственных чисел Λ_{Tim} :

$$\begin{aligned} U_{12}A &= \Lambda_{\text{Tim}} \left(\Lambda_s U_{11} - U_{21} (\Lambda_b U_{22} + U_{23})^{-1} U_{21}^T \right) A, \\ \Lambda_b &= \frac{l^2 S \mu_b^*}{I E_b^*}, \quad \Lambda_s = \frac{(2p\tau_0 + S \mu_b^*) I E_b^*}{l^2 (S \mu_b^*)^2}, \quad \Lambda_{\text{Tim}} = \frac{l (S \mu_b^*)^2}{q I E_b^*}. \end{aligned}$$

Тогда критическая эйлерова нагрузка будет выражаться через максимальное Λ_{Tim} :

$$q_{cr} = \frac{l (S \mu_b^*)^2}{I E_b^* \max \Lambda_{\text{Tim}}}.$$

Кинематическим условиям нашей задачи удовлетворяют базовые функции

$$w = \sum_{n=1}^N A_n \left(\frac{x}{l} \right)^n, \quad \vartheta = \sum_{n=1}^N B_n \left(\frac{x}{l} \right)^{n-1}.$$

Вычислим элементы матриц

$$\begin{aligned} \{U_{11}\}_{mn} &= \frac{nm}{n+m-1}, & \{U_{12}\}_{mn} &= \frac{nm}{(n+m)(n+m-1)}, \\ \{U_{21}\}_{mn} &= \frac{m}{n+m-1}, & \{U_{22}\}_{mn} &= \frac{1}{(n+m-1)}, \\ \{U_{23}\}_{mn} &= \begin{cases} \frac{(n-1)(m-1)}{n+m-3}, & n \neq m, \\ K, & n = m = 1. \end{cases} \end{aligned}$$

3.2. Метод Ритца по модели Эйлера – Бернулли. Приращение энергии изгиба (2) по модели Эйлера – Бернулли при выпучивании под собственным весом и с граничными условиями (3) имеет вид

$$\Delta U(w) = \frac{1}{2}\eta w'^2(0) + \frac{1}{2} \int_0^l ((2p\tau_0 - q(l-x)) w'^2 + IE_b^* w'^2) dx. \quad (7)$$

Здесь энергию поворотной опоры описывает слагаемое $\eta w'^2(0)$.

Согласно методу Ритца представим прогиб в виде линейной комбинации базовых функций, каждая из которых удовлетворяет кинематической части граничных условий (3):

$$w = \sum_n^N A_n w_n. \quad (8)$$

Подставив (8) в (7) и рассматривая A как обобщённые координаты, условие равновесия можно записать в матричном виде:

$$\frac{\delta U}{\delta A} = \left(\frac{2p\tau_0}{l} U_{11} - q U_{12} + \frac{IE_b^*}{l^3} U_{13} \right) A = 0,$$

$$\bar{x} = \frac{x}{l} : \{U_{11}\}_{mn} = \int_0^l w'_n w'_m d\bar{x}, \quad \{U_{12}\}_{mn} = \int_0^l (1 - \bar{x}) w'_n w'_m d\bar{x},$$

$$\{U_{13}\}_{mn} = K w'_n(0) w'_m(0) + \int_0^l w''_n w''_m d\bar{x}, \quad K = \frac{\eta l}{E_b^* I}.$$

Приходим к обобщённой задаче на отыскание собственных чисел $\Lambda_{\text{ЕВ}}$:

$$U_{12} A = \Lambda_{\text{ЕВ}} (U_{13} + \Lambda_s U_{11}) A,$$

$$\Lambda_s = \frac{2l^2 p\tau_0}{IE_b^*}, \quad \Lambda_{\text{ЕВ}} = \frac{IE_b^*}{l^3 q}.$$

Критическая эйлерова нагрузка будет выражаться через максимальное Λ_{EB} :

$$q_{cr} = \frac{IE_b^*}{l^3 \max \Lambda_{\text{EB}}}.$$

Кинематическим условиям нашей задачи удовлетворяют базовые функции

$$w = \sum_{n=1}^N A_n \left(\frac{x}{l}\right)^n,$$

Вычислим элементы матриц

$$\begin{aligned} \{U_{11}\}_{mn} &= \frac{nm}{n+m-1}, & \{U_{12}\}_{mn} &= \frac{nm}{(n+m)(n+m-1)}, \\ \{U_{13}\}_{mn} &= \begin{cases} \frac{nm(n-1)(m-1)}{n+m-3}, & n \neq m, \\ K, & n = m = 1. \end{cases} \end{aligned}$$

Найдём связь между собственными числами моделей Тимошенко и Эйлера – Бернулли:

$$\begin{aligned} q_{cr} &= \frac{l(S\mu_b^*)^2}{IE_b^* \max \Lambda_{\text{Tim}}} = \left(\frac{l^2 S\mu_b^*}{IE_b^*}\right)^2 \frac{IE_b^*}{l^3 \max \Lambda_{\text{Tim}}} = \frac{IE_b^*}{l^3 \max \Lambda_{\text{EB}}}, \\ \frac{1}{\max \Lambda_{\text{EB}}} &= \left(\frac{l^2 S\mu_b^*}{IE_b^*}\right)^2 \frac{1}{\max \Lambda_{\text{Tim}}}. \end{aligned}$$

4. Результаты расчёта. Расчёты проводились для серебряного нановолокна с упругими характеристиками $E = 80$ ГПа, $\nu = 0,37$ при отношении длины к толщине $\frac{l}{c} = 10$ для квадратного поперечного сечения ($b = c$).

В таблицах 1–4 приведены результаты расчёта обратного максимального собственного числа критической эйлеровой нагрузки (для модели Тимошенко – приведенного к модели Эйлера – Бернулли) для двух наборов поверхностных параметров [5]. В таблицах 1, 2 показаны результаты в зависимости от толщины наностержня, а в таблицах 3, 4 – в зависимости от поворотной жёсткости K (значение 10^6 соответствует жёсткой заделке).

Таблица 1. Значения $\frac{1}{\max \Lambda_{\text{ЕВ}}}$ в зависимости от толщины стержня при $K = 10^6$

Поверхностные модули, Н/м	Толщина c							
	2нм	10нм	100нм	1мкм	10мкм	100мкм	1мм	
$\tau_0, \lambda_s + 2\mu_s$	87,804	27,951	10,106	8,067	7,860	7,840	7,838	
0,64, -1,597	56,214	29,318	10,916	8,160	7,870	7,841	7,838	
0,9, 28,29								

Таблица 2. Значения $\left(\frac{l^2 S \mu_b^*}{I E_b^*}\right)^2 \frac{1}{\max \Lambda_{\text{Тим}}}$ в зависимости от толщины стержня при $K = 10^6$

Поверхностные модули, Н/м	Толщина c							
	2нм	10нм	100нм	1мкм	10мкм	100мкм	1мм	
$\tau_0, \lambda_s + 2\mu_s$	86,546	27,644	10,025	8,000	7,794	7,774	7,772	
0,64, -1,597	54,170	28,888	10,827	8,092	7,804	7,775	7,772	
0,9, 28,29								

Таблица 3. Значения $\frac{1}{\max \Lambda_{\text{ЕВ}}}$ в зависимости от коэффициента поворотной жёсткости при $c = 10\text{нм}$

Поверхностные модули, Н/м	Коэффициент поворотной жёсткости K							
	0.1	0.5	1	2	10	50	10^6	
$\tau_0, \lambda_s + 2\mu_s$	13,351	14,480	15,717	17,713	23,893	27,006	27,951	
0,64, -1,597	14,255	15,406	16,669	18,714	25,096	28,334	29,318	
0,9, 28,29								

Таблица 4. Значения $\left(\frac{l^2 S \mu_b^*}{I E_b^*}\right)^2 \frac{1}{\max \Lambda_{\text{Тим}}}$ в зависимости от коэффициента поворотной жёсткости при $c = 10\text{нм}$

Поверхностные модули, Н/м	Коэффициент поворотной жёсткости K							
	0.1	0.5	1	2	10	50	10^6	
$\tau_0, \lambda_s + 2\mu_s$	13,302	14,415	15,630	17,586	23,633	26,695	27,627	
0,64, -1,597	14,185	15,313	16,546	18,535	24,727	27,889	28,853	
0,9, 28,29								

На рис. 2 показана зависимость относительной критической нагрузки к соответствующей макрозадаче для жёсткой заделки для двух наборов значений поверхностных параметров.

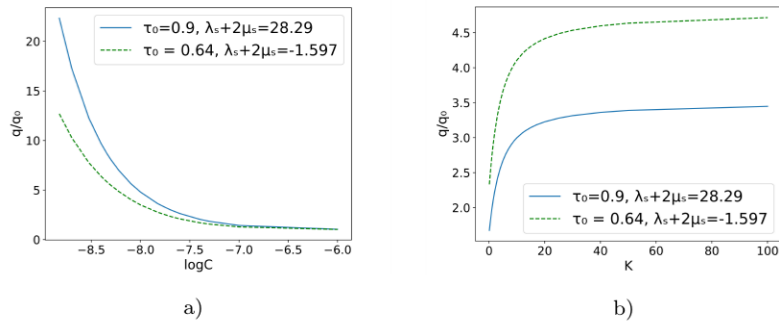


Рис. 2. Относительная (к жёсткой заделке в макрозадаче) критическая эйлерова нагрузка для двух наборов поверхностных параметров серебряного нановолокна: а) в зависимости от толщины, б) в зависимости от поворотной жёсткости

5. Выводы. Результаты расчёта показали, что балочные модели Эйлера–Бернулли и Тимошенко на наноразмерном уровне дают близкие результаты, а учёт поверхностной энергии приводит к увеличению жёсткостных характеристик нановолокна. При макро-размерах этот эффект затухает. Задавая неограниченно большие значения коэффициента жёсткости, приходим к результатам, полученным в [4] для случая жёсткого нанокантилевера.

Литература

1. Timoshenko S., Woinowsky-Krieger S. Theory of Plates and Shells. New York: McGraw-Hill, 1959. 591p.
2. Gurtin M. E., Murdoch A. I. A continuum theory of elastic material surfaces // Archive for Rational Mechanics and Analysis. 1975. Vol. 57. P. 291–323.
3. Gurtin M. E., Murdoch A. I. Surface stress in solids // International Journal of Solids and Structures. 1978. Vol. 14. P. 431–440.
4. Bochkarev A. O. Buckling of a nano-rod with taken into account of surface effect // ZAMM Zeitschrift für Angewandte Mathematik und Mechanik. 2023. Art. no e202300738.
5. Chhapadia P., Mohammadi P., Sharma P. Curvature-dependent surface energy and implications for nanostructures // J. Mech. Phys. Solids. 2011. Vol. 59. P. 2103–2115.

УДК 519.688

Набиев Т. К., Шиманчук Д. В.

Расчёт траекторий движения группы БПЛА в строю

1. Введение. В последнее время в связи с развитием электроники и увеличением мощностей вычислительной техники большой интерес специалистов как гражданской сферы, так и военной привлекает создание и управление беспилотными летательными аппаратами. Это приводит к необходимости создания алгоритмов для автоматического построения траектории, переводящей БПЛА из одной точки в другую. Для таких задач построенная траектория должна быть оптимальной как по времени полёта, так и по расходу топлива при полёте. Следовательно, существует потребность в разработке и программной реализации алгоритмов построения оптимальной траектории. Кроме того, из-за обширного распространения беспилотников появилась необходимость в управлении одновременно несколькими летательными аппаратами.

2. Анализ публикаций. Описание модели летательного аппарата, построение пространственной траектории движения и расчёт соответствующих управлений можно найти в [1]. Описание приёмов использования концепции обратных задач динамики представлена в [2, 3]. В [4] приведено описание построения движений в соответствии с используемой моделью, также показано, что такие движения являются допустимыми. Для получения представления об управлении группой роботов был использован труд [5], в котором представлены основные положения группового управления и некоторые алгоритмы управления. Подход к построению группового управления с использованием геометрических характеристик разработан в [6]. Идея геометрического подхода была использована для построения траекторий группы летательных аппаратов в настоящей работе.

Набиев Тимур Казымович – студент, Санкт-Петербургский государственный университет; e-mail: st102631@student.spbu.ru, тел.: +7(987)754-64-04

Шиманчук Дмитрий Викторович – доцент, Санкт-Петербургский государственный университет; e-mail: d.shimanchuk@spbu.ru, тел.: +7(906)240-51-19

Работа выполнена при финансовой поддержке РФФ, проект № 23-21-00027

Можно сказать, что большее распространение получили алгоритмы группового управления, построенные по принципу единой механической системы. Однако такой подход описывает общие положения, но ограничивает область и способ применения БПЛА на практике.

3. Постановка задачи. На сегодняшний день актуальной проблемой является разработка алгоритмов решения задачи группового управления БПЛА, удовлетворяющего следующим требованиям: использование геометрических характеристик, описывающих положение каждого аппарата относительно других; относительная простота алгоритма при программной реализации; получение характеристик физической модели каждого объекта из группы; универсальность.

Рассмотрим математическую модель БПЛА [1]

$$\begin{cases} \dot{H} = V \sin(\vartheta), \\ \dot{L} = V \cos(\vartheta) \cos(\psi), \\ \dot{Z} = -V \cos(\vartheta) \sin(\psi), \\ \dot{V} = g(n_x - \sin(\vartheta)), \\ \dot{\vartheta} = g \frac{n_y \cos(\gamma) - \cos(\vartheta)}{V}, \\ \dot{\psi} = -g \frac{n_y \sin(\gamma)}{V \cos(\vartheta)}, \end{cases} \quad (1)$$

где H – высота полёта, L – дальность полёта, Z – боковое смещение, V – путевая скорость, ϑ – угол наклона траектории, ψ – путевой угол, γ – угол крена, n_x – продольная перегрузка, n_y – поперечная перегрузка, g – ускорение свободного падения (см. рис 1).

Целью работы является разработка алгоритма, учитывающего ограничения реальных летательных аппаратов и обеспечивающего минимальное время перелёта. В качестве группы аппаратов будем рассматривать шеренгу из БПЛА, где каждый объект в любой момент времени находится на фиксированном расстоянии от соседних.

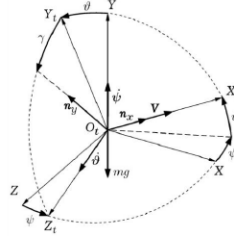


Рис. 1. Земная система координат и траекторная система координат

4. Задача терминального управления. Рассмотрим задачу терминального управления, которое приводило бы БПЛА из заданного начального состояния $(V_0, H_0, L_0, Z_0, \vartheta_0, \psi_0), (n_{x0}, n_{y0}, \gamma_0)$ при $t = 0$ в конечное состояние $(V_1, H_1, L_1, Z_1, \vartheta_1, \psi_1), (n_{x1}, n_{y1}, \gamma_1)$ при $t = \Delta t$. Пусть Δt известно заранее. Для того, чтобы определить траекторию, необходимо выразить L, Z и H , которые описывают положение БПЛА. Воспользуемся обратными методами динамики. Введём новые управления по формулам $u_1 = n_x, u_2 = n_y \cos(\gamma), u_3 = n_y \sin(\gamma)$ и новые переменные $x_1 = H, x_2 = L, x_3 = Z, x_4 = V \sin(\vartheta), x_5 = V \cos(\vartheta) \cos(\psi), x_6 = V \cos(\vartheta) \sin(\psi)$, которые определяют гладкую невырожденную замену переменных. Тогда система (1) запишется в виде

$$\begin{cases} \dot{x}_1 = x_4, \\ \dot{x}_4 = -g + u_1 g \sin(\vartheta) + u_2 g \cos(\vartheta), \\ \dot{x}_2 = x_5, \\ \dot{x}_5 = u_1 g \cos(\vartheta) \cos(\psi) - u_2 g \sin(\vartheta) \cos(\psi) + u_3 g \sin(\psi), \\ \dot{x}_3 = x_6, \\ \dot{x}_6 = -u_1 g \cos(\vartheta) \sin(\psi) + u_2 g \sin(\vartheta) \sin(\psi) + u_3 g \cos(\psi). \end{cases} \quad (2)$$

Систему (2) можно разрешить относительно новых управлений:

$$\begin{cases} u_1 = \frac{1}{g} \left((\ddot{x}_1 + g) \sin(\vartheta) + \ddot{x}_2 \cos(\vartheta) \cos(\psi) - \ddot{x}_3 \cos(\vartheta) \sin(\psi) \right), \\ u_2 = \frac{1}{g} \left((\ddot{x}_1 + g) \cos(\vartheta) - \ddot{x}_2 \sin(\vartheta) \cos(\psi) + \ddot{x}_3 \sin(\vartheta) \sin(\psi) \right), \\ u_3 = \frac{1}{g} \left(\ddot{x}_2 \sin(\psi) + \ddot{x}_3 \cos(\psi) \right). \end{cases} \quad (3)$$

Будем искать траекторию в виде многочленов пятой степени, в силу того что на функции x_1, x_2, x_3 накладывается по 6 ограничений. Получим

$$x_i(t) = \sum_{j=0}^2 \frac{x_{i0}^{(j)}}{j!} t^j + \sum_{j=1}^3 c_{ij} t^{2+j}, \quad i = \overline{1,3}, \quad (4)$$

где c_{ij} находятся из системы

$$\begin{cases} (\Delta t)^3 c_{i0} + (\Delta t)^4 c_{i1} + (\Delta t)^5 c_{i2} = x_{i1} - x_{i0} - \dot{x}_{i0}(\Delta t) - 0,5\ddot{x}_{i0}(\Delta t)^2, \\ 3(\Delta t)^3 c_{i0} + 4(\Delta t)^4 c_{i1} + 5(\Delta t)^5 c_{i2} = \dot{x}_{i1} - \dot{x}_{i0} - \ddot{x}_{i0}(\Delta t), \\ (\Delta t)^3 c_{i0} + (\Delta t)^4 c_{i1} + (\Delta t)^5 c_{i2} = \ddot{x}_{i1} - \ddot{x}_{i0}. \end{cases}$$

Из представлений (2)–(4), можно получить физические характеристики БПЛА.

Каждый летательный аппарат обладает набором характеристик, которые стоит учитывать при построении траектории. Каждый БПЛА имеет максимальную и минимальную скорость, ограничения высоты и т. д. Поэтому имеют место дополнительные ограничения на переменные состояния: $H_{\min} \leq H \leq H_{\max}$, $L_{\min} \leq L \leq L_{\max}$, $Z_{\min} \leq Z \leq Z_{\max}$, $V_{\min} \leq V \leq V_{\max}$, $\vartheta_{\min} \leq \vartheta \leq \vartheta_{\max}$, $\psi_{\min} \leq \psi \leq \psi_{\max}$, $n_{x \min} \leq n_x \leq n_{x \max}$, $n_{y \min} \leq n_y \leq n_{y \max}$, $\gamma_{\min} \leq \gamma \leq \gamma_{\max}$.

Поиск минимальной по времени движения БПЛА траектории с заданной точностью $\varepsilon > 0$ заключается в переборе по времени t (в порядке возрастания) траекторий, увеличивая время манёвра t на величину $\delta > 0$. Начальным значением t_0 будем считать минимальное время, за которое аппарат перейдёт из начального состояния в конечное, двигаясь прямолинейно с максимально возможной скоростью:

$$t_0 = \frac{\sqrt{(H_1 - H_0)^2 + (L_1 - L_0)^2 + (Z_1 - Z_0)^2}}{V_{\max}}, \quad (5)$$

конечным значением для перебора будем считать $t_{\max} = 15(t_0 + 5)$. Опытным путём было установлено, что искомое время лежит в промежутке $[t_0, t_{\max}]$.

Схема алгоритма поиска оптимальной траектории включает следующие шаги:

1. На каждом шаге осуществляется проверка, удовлетворяет ли траектория и реализующие её управления условиям, которые накладываются на решение максимальными и минимальными значениями переменных состояния и управлений.

2. Если найдена удовлетворяющая граничным условиям траектория, то необходимо уменьшить соответствующее значение времени t на величину δ , а потом уменьшить значение δ в 2 раза (проводить уменьшение необходимо пока $\delta \neq \varepsilon$).
3. Если найдена траектория, не удовлетворяющая наложенным на неё ограничениям, то значение t необходимо увеличить на соответствующую величину δ .
4. Первая траектория, которая удовлетворяет этим ограничениям при $\delta = \varepsilon$, и есть оптимальная по времени с точностью ε .

5. Формирование строя БПЛА. Рассмотрим задачу о построении траекторий для нескольких БПЛА при их движении в строю (параллельно друг другу). Решение задачи о формировании строя заключается в вычислении траекторий нескольких аппаратов по траектории первого. Рассмотрим гладкую кривую Γ на плоскости, в каждой её точке A проведём перпендикуляр и возьмем на этом перпендикуляре точки, находящиеся на некотором фиксированном расстоянии h от точки A . Полученные точки формируют две кривые, каждая из которых является эквидистантой кривой Γ . Эквидистанта позволяет построить кривые траекторий той же гладкости, что и исходная кривая.

Далее представим алгоритм определения траекторий для n роботов при расстоянии d между ними:

1. Выбор целевого положения Z_1, L_1, H_1 и определение робота, относительно которого формируется строй. Робот, оказавшийся ближе других к точке с радиус-вектором $\mathbf{r}_c = \frac{1}{n} \sum_{i=1}^n \mathbf{r}_i$, будет являться опорным, где r_i – радиус-вектор аппарата с номером i . Присваиваем ему номер 1.
2. Вычисление условно-оптимальной траектории опорного аппарата с конечной и начальной точками, определенными на шаге 1, с использованием представленной выше схемы алгоритма поиска оптимальной траектории.
3. Определение знака и величины расстояния относительно опорного аппарата по чётности номера робота, для чётных номеров $d_1 = d$, для нечётных $d_2 = -d$. Для каждого аппарата с номером $i = \overline{2, n}$, выполняем шаги 4–7, начиная с $i = 2$.
4. Вычисление точек эквидистанты для рассматриваемого робота с номером i относительно траектории опорного аппарата.

5. Аппроксимация функции по методу наименьших квадратов, с использованием данных, полученных на шаге 4.
6. Вычисление параметров физической модели i -го БПЛА.
7. Если номер аппарата чётный, то $d_1 = d_1 + d$, в противном случае $d_2 = d_2 - d$, $i = i + 1$.

6. Результаты компьютерного моделирования. Рассмотрим пример моделирования движения группы БПЛА при $n = 5$ и расстоянии $d = 5$ м. Исходные данные представлены в таблице 1.

Таблица 1. Начальные и конечные данные

	H , м	L , м	Z , м	V , км/ч	θ , °	ψ , °	n_x	n_y	γ , °
$t = 0$	0	0	0	120	0	0	0	1	0
$t = \Delta t$	900	500	200	110	0	-90	0	1	0

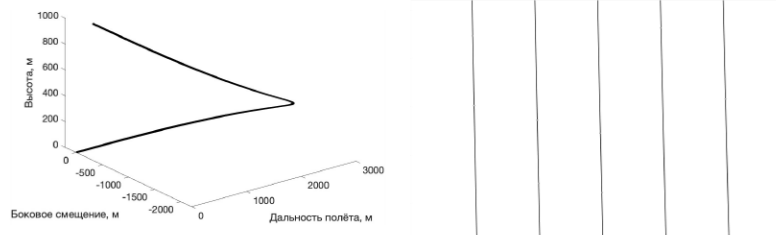


Рис. 2. Вид траектории в пространстве и в приближении

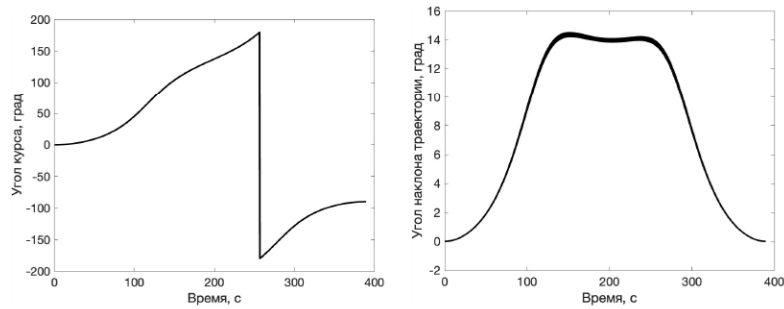


Рис. 3. График угла наклона траектории ϑ и график угла курса ψ

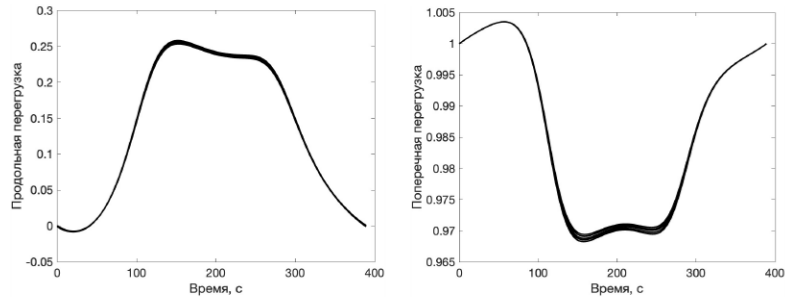


Рис. 4. График продольной перегрузки n_x и график поперечной перегрузки n_y

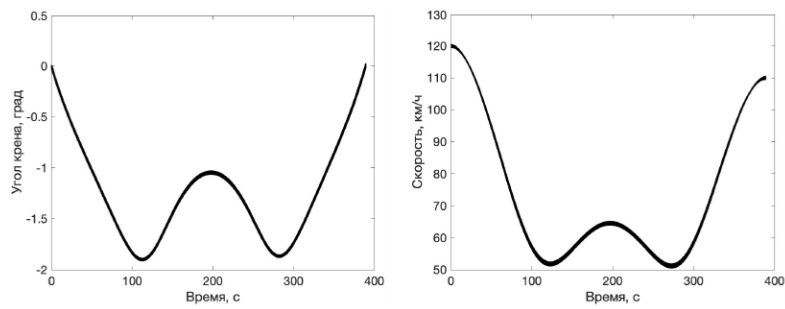


Рис. 5. График угла крена γ и график земной скорости БПЛА V

Из численного моделирования видно (см. рис 2), что при движении БПЛА заданная структура группы сохраняется. Однако проверке выполнения ограничений параметров модели БПЛА (см. рис 3–5) из группы необходимо проводить независимо, так как расчёт траекторий осуществляется согласно предложенному алгоритму на основании опорного БПЛА, для которого приняты ограничения (см. таблицу 2).

Таблица 2. Ограничения модели

	H , м	L , м	Z , м	V , км/ч	θ , °	ψ , °	n_x	n_y	γ , °
Min	0	-10000	-10000	50	-89	-179	-3	0,2	-60
Max	3000	10000	10000	180	89	179	3	6	60

7. Заключение. В статье представлен алгоритм для построения траектории для заданного количества БПЛА. Алгоритм позволяет строить программные управления и оценивать динамические характеристики на программном движении для каждой модели БПЛА из группы. Для синтеза управления с обратной связью с целью удержания группы беспилотников на программном движении можно рассмотреть подход из [7].

Литература

1. Канатников А. Н., Шмагина Е. А. Задача терминального управления движением летательного аппарата // *Нелинейная динамика и управление*, 2010. Вып. 7. С. 79–94.
2. Крутько П. Д. Обратные задачи динамики управляемых систем. *Нелинейные модели*. М.: Наука, 1988. 330 с.
3. Крутько П. Д. Обратные задачи динамики в теории автоматического управления. М.: Машиностроение, 2004. 576 с.
4. Канатников А. Н., Крищенко А. П., Ткачев С. Б. Допустимые пространственные траектории беспилотного летательного аппарата в вертикальной плоскости // *Наука и образование*, 2012. № 3. С. 1–17.
5. Пшихопов В. Х. Групповое управление подвижными объектами в неопределённых средах. М.: Физматлит, 2015. 307 с.
6. Иванов Д. Я. Решение строевой задачи в группе беспилотных квадрокоптеров // *Известия Южного федерального университета. Технические науки*. 2014. № 8. С. 138–147.
7. Смирнова М. А., Смирнов М. Н., Смирнов Н. В. Система многоцелевого управления роботом-манипулятором // *Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления*. 2022. Т. 18. Вып. 4. С. 621–630.

UDC 510

Belenok A. I.

Ritz method for building parametric representation of cost optimal trajectory

Recommended for publication by professor Majid Abbasov

1. Introduction. The problem of obtaining the most economically advantageous route for building a road between two points is studied. The problem is to determine the equation of the optimal trajectory, which corresponds to minimal construction costs. Using mathematical modeling tools, an integral cost functional is derived, which depends on the parametric functions representing the trajectory. The result is a problem in the calculus of variations, which is proposed to be solved by the Ritz method. The results of numerical experiments are presented.

2. Formulation of the problem. Let points A and B be given on the terrain, acting as the starting and ending points, respectively. The challenge is to connect these two points, which will be the most optimal in terms of construction costs. Let α be the cost of delivery per unit length of the road for the volume of materials required for the construction of its individual section. Since construction conditions vary from point to point due to various factors, the cost β of constructing a separate section of road at each point along the route is different.

As it was shown in [1], the functional for finding the optimal path can be written as

$$J(y) = \frac{\alpha}{2} \left(\int_{x_A}^{x_B} \sqrt{1 + y'(x)^2} dx \right)^2 + \int_{x_A}^{x_B} \beta(x, y) \sqrt{1 + y'(x)^2} dx,$$

where $\beta : \mathbb{R}^2 \rightarrow \mathbb{R}$ is a given function with continuous partial derivatives of the second order.

Belenok Anna Igorevna – undergraduate student, Saint Petersburg State University; e-mail: st087969@student.spbu.ru, phone: +7(900)657-33-89

The work was carried out with the financial support of the Russian Science Foundation, project No. 23-21-00027

The main goal of the current study is to find the optimal trajectory in a parametric form. The parametric form allows us to work with a wider set of trajectories than the explicit one, since a parametrically represented curve can have multiple intersections with a line parallel to the O_y axis, while this is not true for an explicitly defined curve.

Consider a parametric definition of the curve $r(t) = (x(t), y(t))$ such that $x(0) = y(0) = 0$ and $x(1) = l$, $y(1) = y_l$. The cost of the curve is the value of the following functional

$$J(r) = \frac{\alpha}{2} \left(\int_0^1 \sqrt{(x')^2 + (y')^2} dt \right)^2 + \int_0^1 \beta(x, y) \sqrt{(x')^2 + (y')^2} dt. \quad (1)$$

3. Application of the Ritz method. First, we briefly present the idea of the Ritz method in general form [2], and then we concretize it for the task under study.

Let's consider the task of minimizing the functionality

$$I = \int_0^1 F(x(t), y(t), x'(t), y'(t)),$$

where F is a continuous function of its arguments, and $y \in C^1[0, l]$ subject to

$$y(0) = x(0) = 0, \quad y(1) = x(1) = 0.$$

Let

$$\begin{aligned} \varphi(t, a_1, \dots, a_n), \quad n = 1, 2, \dots, \\ \psi(t, b_1, \dots, b_m), \quad m = 1, 2, \dots, \end{aligned}$$

be two sequences of n - and m -parametric families of functions, each of which is wider than the previous one by adding an additional parameter, and, moreover, boundary conditions are met for all parameter values. For each family, one can solve the problem of minimizing the function of $n + m$ arguments:

$$I(a_1, \dots, a_n, b_1, \dots, b_m) = \int_0^1 F\left(\varphi(t, a_1, \dots, a_n), \psi(t, b_1, \dots, b_m), \varphi'(t, a_1, \dots, a_n), \psi'(t, b_1, \dots, b_m)\right) dt,$$

which is reduced to solving a system of equations

$$\frac{\partial I}{\partial a_i} = 0, \quad i = 1, \dots, n, \quad \frac{\partial I}{\partial b_j} = 0, \quad j = 1, \dots, m.$$

The n -parametric family of functions can be introduced as a linear combination of the following system of functions

$$\mu_k(t) = \sin(k\pi t), \quad k = 1, 2, \dots,$$

or

$$\mu_k(x) = (1 - x)x^k, \quad k = 1, 2, \dots$$

This fact immediately follows from the Weierstrass approximation theorems.

We find a solution of the equations (1) in the form

$$x(t) = t + \sum_{k=1}^{\infty} a_k \mu_k(t), \quad y(t) = t + \sum_{k=1}^{\infty} b_k \mu_k(t).$$

4. The results of numerical experiments. Let us look at an example in which we assume that the ground surface on which the road should be built is described by a moderate altitude difference.

Example. Let $\alpha = 0.1$, $l = 1$, $y_l = 1$ and $\beta : \mathbb{R}^2 \rightarrow \mathbb{R}$ given by

$$\beta(x, y) = 1 + \sin 5x \cdot \sin y.$$

In order to get an explicitly given approximation of the optimal trajectory we substitute

$$y = \frac{y_l}{l}x + \sum_{k=1}^{\infty} a_k \phi_k(x)$$

to functional J and minimize it with respect to a_k , $k = 1, \dots, 5$. We use the Ritz method, which produces the following solution $a_1 = -0.31397$, $a_2 = 0.07367$, $a_3 = -0.03138$, $a_4 = 0.01212$, $a_5 = -0.00396$, for which the cost equals 1.279.

Now let us find the solution of the same problem in parametric form.

$$x(t) = t + \sum_{k=1}^3 a_k \sin(\pi kt),$$

$$y(t) = t + \sum_{k=1}^3 b_k \sin(\pi kt).$$

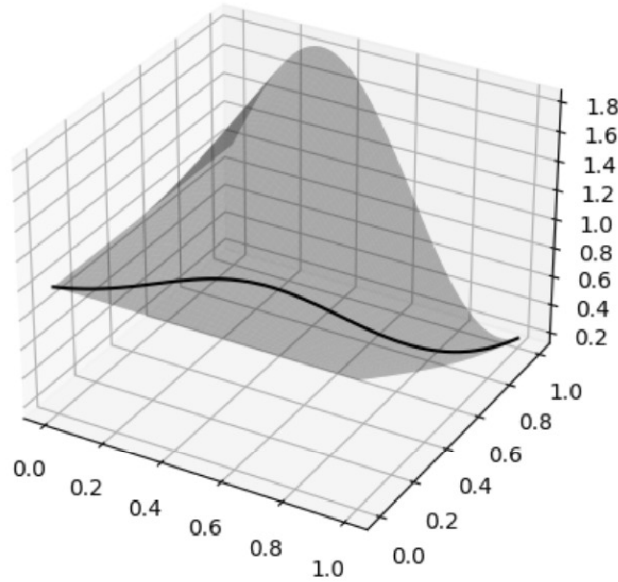


Fig. 1. A route given explicitly by a curve

Using Ritz method we get

$$\begin{aligned}
 a_1 &= 0.24047405, & a_2 &= -0.02688373, & a_3 &= -0.02494643, \\
 b_1 &= -0.04031987, & b_2 &= -0.08527058, & b_3 &= 0.00430315.
 \end{aligned}$$

The cost of the solution (i.e., the value of functional J on the solution produced) is 1.278.

5. Conclusion. We considered the problem of finding the cost-optimal path in parametric form. The proposed approach is based on the calculus of variations. Unlike commonly used heuristics, our method guarantees the quality of the results obtained. Thanks to the developed approach, we get the optimal solution and therefore can be sure that we do not waste resources when building roads and other transport networks such as waterways, pipelines and so on.

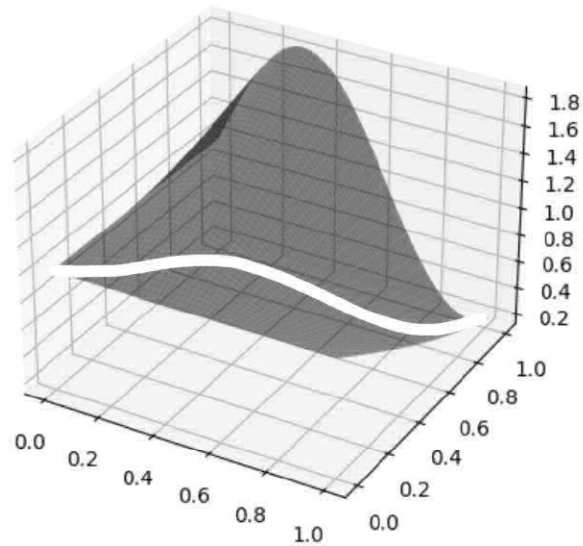


Fig. 2. Route defined by a curve in parametric form

References

1. Abbasov M.E., Sharlay A.S. Search for the cost-optimal road construction trajectory on the terrain // Vestnik of St. Petersburg University. Applied Mathematics. Computer science. Management processes. 2021. T. 17. No 1. P. 4–12.
2. Kantorovich L.V., Krylov V.I. Priblizhennye metody vysshego analiza. L.: Fizmatiz, 1962. 708 p. (In Russian)
3. Elsgolts L.E. Differential equations and calculus of variations. M.: Nauka, 1969. 424 p. (In Russian)

UDC 517.958, 539.3

Mullayanov T. O.

Investigating elastic response of nanostructured solid surface using boundary perturbation method

Recommended for publication by associate professor Kostyrko S. A.

1. Introduction. Surface properties play a critical role in determining the strength and failure mechanisms of materials as fractures often initiate from surface defects [1]. Traditionally, these defects are associated with local stress concentrations arising from surface roughness. However, as surface features shrink to the nanoscale, surface energy becomes increasingly influential. Consequently, there is a growing interest in developing theoretical models that accurately capture the behavior of nanostructured materials, particularly focusing on surface elasticity, driven by the rapid advancements in nanoscience and nanotechnology.

Understanding stress distribution on solid surfaces is essential due to its impact on various applications. Surface defects, such as roughness from manufacturing processes or corrosion-induced pits, can significantly affect stress concentrations. This understanding is crucial for predicting crack nucleation sites in fatigue and fracture damage. Additionally, in engineering problems involving surface reconstruction, surface stress conditions influence surface evolution due to diffusion and phase transition mechanisms [2]. While finite element method offers solutions for stress concentrations in complex geometries, it is time-consuming and require new calculations for each surface profile. Conversely, analytical solutions offer a more efficient approach to understand and predict stress concentration fields and address multi-physics problems.

2. Mathematical Framework. This research introduces a mathematical framework for analyzing the elastic behavior of a solid surface, incorporating both geometric and physico-mechanical properties, particularly in relation to nanosized surface asperities under plane strain conditions. The presented framework is built upon the Gurtin–Murdoch

Mullayanov Timur Olegovich – undergraduate student, Saint Petersburg State University; e-mail: st087024@student.spbu.ru, phone: +7(987)053-38-01

model of surface elasticity, treating the surface layer as a thin but significant component adhered to the bulk material with distinct elastic properties [3, 4]. Surface stresses are identified as the stress resultants in this layer, leading to the derivation of constitutive relations and equilibrium equations that accommodate the effects of surface elasticity and tension. The solid is subjected to the remote tensile loading T and surface traction q^s induced by surface stress and tension:

$$\begin{aligned} \sigma_{nn}(\zeta) + i\sigma_{nt}(\zeta) &= q^s(\zeta), \quad \zeta \in \Gamma, \\ q^s(\zeta) &= \kappa\sigma_0^s + \frac{\kappa}{2} \left[p_s \frac{\partial u}{\partial \zeta} + m_s \overline{\frac{\partial u}{\partial \zeta}} \right] - \\ &\quad - \frac{1}{2h} \left[p_s(i - \varepsilon f') \frac{\partial^2 u}{\partial \zeta^2} + m_s(i + \varepsilon f') \overline{\frac{\partial^2 u}{\partial \zeta^2}} \right], \end{aligned} \quad (1)$$

where σ_{nn} and σ_{nt} are the components of the bulk stress tensor defined in the local Cartesian coordinates (n, t) ; $\zeta = x_1 + i\varepsilon f(x_1)$, $f(x_1)$ is a function describing surface undulation; $A = \varepsilon a$ is a maximum deviation of the surface from a flat configuration (i.e. plane $x_2 = 0$); $p_s = M_s + \sigma_0^s$, $m_s = M_s - \sigma_0^s$, $M_s = \lambda_s + 2\mu_s$, λ_s and μ_s are the surface elastic constants, σ_0^s is a surface tension; κ and h are the local principal curvature and metric coefficient on the undulated boundary Γ , respectively.

The additional equation determines the continuity of displacements across the surface region:

$$u^s(\zeta) = u(\zeta), \quad \zeta \in \Gamma, \quad (2)$$

where $u^s = u_1^s + iu_2^s$ and $u = u_1 + iu_2$ are the displacements of the surface and bulk phases along the corresponding coordinate axes x_1 and x_2 .

Employing the Goursat–Kolosoov complex potentials and boundary perturbation method [5, 6], we derive the asymptotic expressions for stress and strain tensor components. The unknown functions are expanded as power series in a small parameter ε representing the amplitude-to-wavelength ratio of interface undulation. This approach reduces the formulated boundary value problem (1), (2) to a recurrent series of integral equations for anyorder approximation N :

$$2\mu u'_{(n)}(x_1) + \frac{\varkappa + 1}{2\pi i} \int_{-\infty}^{+\infty} \frac{q_{(n)}^s(t)}{x_1 - 1} dt + \frac{1 - \varkappa}{2} q_{(n)}^s(x_1) = F_n(x_1), \quad (3)$$

where $n = 1 \dots N$; $\varkappa = (\lambda + 3\mu)/(\lambda + \mu)$, λ and μ are the bulk elastic constants; $u'_{(n)}$ and $q^s_{(n)}$ are the coefficients of the power series expansions of the corresponding functions; F_n are the known functions derived at each iteration of boundary perturbation method using the complex potentials of the previous approximations. The solution for each iteration can be found in terms of a trigonometric series employing the properties of Cauchy type integrals. Subsequent numerical analysis allows us to explore the distribution of elastic stresses along the curved surface, considering the influence of both remote tensile loading T and surface tension σ_0^s .

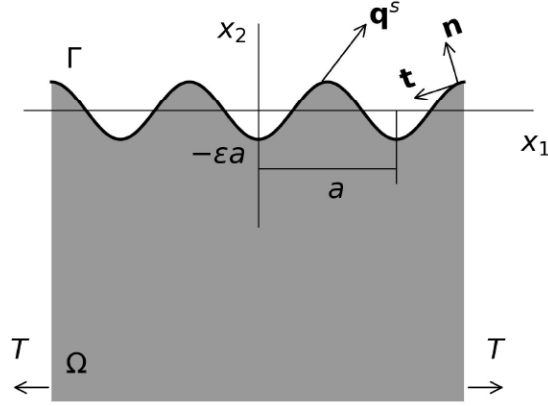


Figure 1. The model of a semi-infinite elastic solid with a nanosized surface asperities

3. Numerical Results. Within the developed framework, we examine the stress distribution in an aluminum specimen characterized by a surface profile defined as $f(x_1) = -a \cos(2\pi x_1/a)$. The elastic properties of both the surface and bulk are quantified by parameters $M_s = 6.099$ N/m, $\lambda = 58.17$ GPa, and $\mu = 26.13$ GPa. Figures 2 and 3 depict the variations of local stresses σ_{tt} and σ_{nn} , respectively, evaluated at the bottom of a surface valleys ($x_1/a = 0$), where the maximum stress concentration effect is observed. These figures illustrate the influence of undulation amplitude A for profiles with different undulation wavelengths $a = 10$ nm and $a = 100$ nm (Figures 2 and 3, respectively) under various loading conditions: $T = 0.1$ GPa, $\sigma_0^s = 1$ N/m $T = 0$,

$\sigma_0^s = 1 \text{ N/m}$ and $T = 0.1 \text{ GPa}$, $\sigma_0^s = 0$ (curves 1, 2, and 3, respectively). The dashed and solid lines represent the results of linear (i.e., $N = 1$) and nonlinear boundary perturbation solutions, where a 15-order approximation (i.e., $N = 15$) was selected based on convergence analysis.

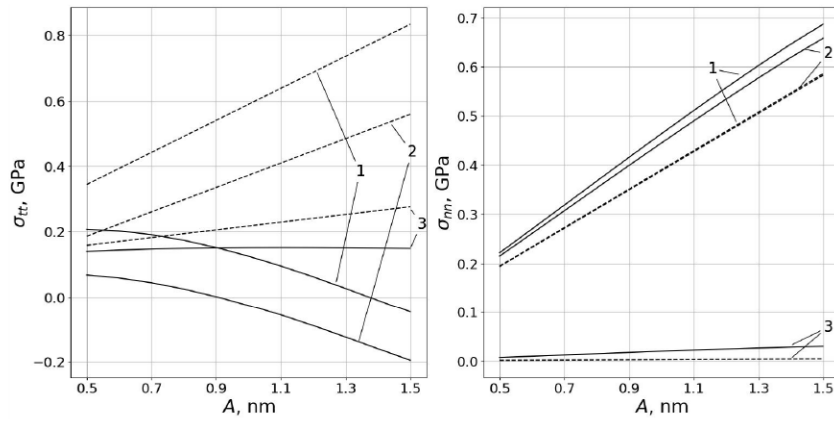


Figure 2. The dependence of the local stresses σ_{tt} and σ_{nn} on the undulation amplitude A for surface profile with undulation wavelength $a=10 \text{ nm}$

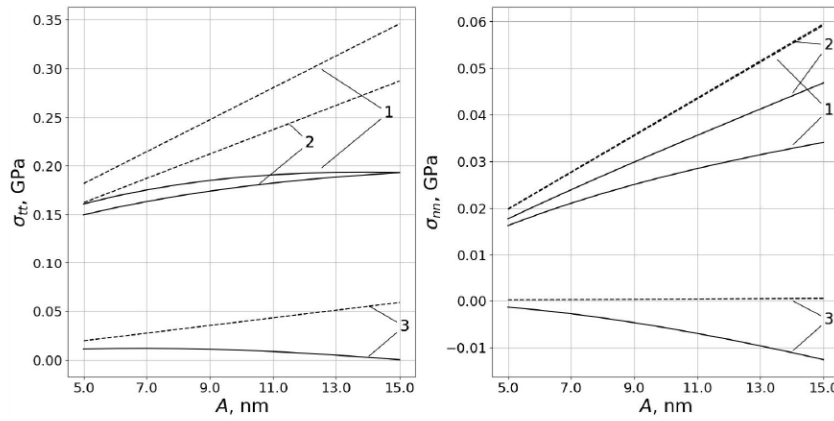


Figure 3. The dependence of the local stresses σ_{tt} and σ_{nn} on the undulation amplitude A for surface profile with undulation wavelength $a=100 \text{ nm}$

References

1. Freund L. B., Suresh S. Thin film materials: stress, defect formation and surface evolution. Cambridge University Press, 2004. 750 p.
2. Shuvalov G., Kostyrko S., Altenbach H. Interplay of surface and bulk elasticity in morphological stability of ultra-thin film coatings // Continuum Mechanics and Thermodynamics. 2024. Vol. 36. P. 503–523.
3. Gurtin M. E., Murdoch A. I. A continuum theory of elastic material surfaces // Archive of Rational Mechanics Analysis. 2014. Vol. 57. P. 291–323.
4. Gurtin M. E., Murdoch A. I. Surface stress in solids // International Journal of Solids Structures. 1978. Vol. 14. P. 431–440.
5. Kostyrko S., Grekov M., Altenbach H. Stress distribution at the wavy surface of a solid incorporating surface stresses and surface tension // Advances in Solid and Fracture Mechanics. Springer, Cham. 2022. P. 151–166.
6. Mullayanov T. O. The boundary perturbation method for solving a problem of nanoscale surface relief in a solid // Control Processes and Stability. 2023. Vol. 10. No 1. P. 127–131.

UDC 519.62

Xing J.

MS-stability comparison of Runge–Kutta–Chebyshev methods application to SDEs

Recommended for publication by associate professor Eremin A. S.

1. Introduction. Stochastic differential equations (SDEs) are widely used in many fields such as finance, neural networks and biology. However, it is often difficult to obtain their exact solutions for many SDEs, so it is very important to use suitable numerical methods to approximate the solution of SDEs. Abdulle and Li [1] proposed a class of explicit stochastic orthogonal Runge–Kutta–Chebyshev (S-ROCK) methods with large stability domains in the mean-square sense. The mean-square stability of these methods has been analyzed mainly for scalar linear SDEs.

In this paper, we propose a class of modified S-ROCK methods (mSROCK), analyze their mean-square stability, and compare them with the MS-stability of S-ROCK methods.

2. S-ROCK methods. We consider an SDE of It type given by

$$dX_t = f(t, X_t)dt + g(t, X_t)dW_t, \quad X(t_0) = X_0, \quad (1)$$

where X_t is a random variable, $f : [t_0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ is called the drift function, $g : [t_0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ is called the diffusion function, and W_t is a standard one-dimensional Wiener process.

For such equations, Abdulle and Li [1] proposed a class of explicit stabilized methods called S-ROCK (for stochastic orthogonal Runge–Kutta–Chebyshev). An S-ROCK method of strong order 1/2 and weak

Xing Juan – graduate student, Saint Petersburg State University; e-mail: st115836@student.spbu.ru, phone: +7(906)226-90-78

order 1, here denoted as SROCK, is given for s stages by

$$\begin{aligned}
\mathbf{K}_0 &= \mathbf{y}_n, \quad \mathbf{K}_1 = \mathbf{y}_n + h \frac{\omega_1}{\omega_0} \mathbf{f}(\mathbf{K}_0), \\
\mathbf{K}_i &= 2 \frac{T_{i-1}(\omega_0)}{T_i(\omega_0)} (h\omega_1 \mathbf{f}(\mathbf{K}_{i-1}) + \omega_0 \mathbf{K}_{i-1}) \\
&\quad - \frac{T_{i-2}(\omega_0)}{T_i(\omega_0)} \mathbf{K}_{i-2}, \quad i = 2, \dots, s, \\
\mathbf{y}_{n+1}^* &= \mathbf{K}_s, \quad \mathbf{y}_{n+1} = \mathbf{y}_{n+1}^* + g(\mathbf{y}_{n+1}^*) \Delta w_n,
\end{aligned} \tag{2}$$

where

$$\omega_0 = 1 + \frac{\eta}{s^2}, \quad \omega_1 = \frac{T_s(\omega_0)}{T_s'(\omega_0)}$$

and η is the damping parameter. $T_k(x)$ is the first kind Chebyshev polynomial defined by $T_k(\cos \theta) = \cos k\theta$ for $\theta \in [-1, 1]$ or for any $x \in \mathbb{R}$ by the recursion:

$$T_0(x) = 1, \quad T_1(x) = x, \quad T_k(x) = 2xT_{k-1}(x) - T_{k-2}(x), \quad k \geq 2.$$

In order to obtain a wider stability domain at optimal damping, we have modified an SROCK (getting an mSROCK). Its definition has a recursive relation similar to that of (2), except that the computation of the diffusion term is changed in the last stage:

$$\begin{aligned}
\mathbf{K}_0 &= \mathbf{y}_n, \quad \mathbf{K}_1 = \mathbf{y}_n + h \frac{\omega_1}{\omega_0} \mathbf{f}(\mathbf{K}_0), \\
\mathbf{K}_j &= 2 \frac{T_{j-1}(\omega_0)}{T_j(\omega_0)} (h\omega_1 \mathbf{f}(\mathbf{K}_{j-1}) + \omega_0 \mathbf{K}_{j-1}) \\
&\quad - \frac{T_{j-2}(\omega_0)}{T_j(\omega_0)} \mathbf{K}_{j-2}, \quad j = 2, \dots, s, \\
\mathbf{y}_{n+1}^* &= \mathbf{K}_s, \quad \mathbf{y}_{n+1} = \mathbf{y}_{n+1}^* + \sum_{i=1}^s b_i g(\mathbf{K}_i) \Delta W_n.
\end{aligned} \tag{3}$$

3. Mean-square stability analysis. To analyze the MS-stability of numerical methods, we consider a particular case of (1) — the scalar linear test equation [2]

$$dX_t = \lambda X_t dt + \mu X_t dW_t, \quad X(t_0) = X_0, \tag{4}$$

where $\lambda, \mu \in \mathbb{C}$.

The solution of (4) is said to be *mean-square stable* if and only if

$$\lim_{t \rightarrow \infty} \mathbb{E}(|X_t|^2) = 0 \Leftrightarrow 2\Re(\lambda) + |\mu|^2 < 0.$$

We denote the stability region of the test problem (4) by

$$S_{\text{MS}} = \{(\lambda, \mu) : 2\Re(\lambda) + |\mu|^2 < 0\}. \quad (5)$$

When a one-step method with constant step-size h is applied to the test problem, we can obtain the following expression

$$X_{n+1} = \phi(p, q, J_n)X_n \quad (6)$$

with some function ϕ , where $p = h\lambda$, $q = \mu\sqrt{h}$, $J_n = W(t_{n+1}) - W(t_n) \sim N(0, h)$. Squaring the modulus of (6), taking the expected value, we get

$$\mathbb{E}[|X_{n+1}|^2] = R_2(p, q)\mathbb{E}[|X_n|^2].$$

A numerical method is mean-square stable if and only if $R_2(p, q) < 1$. $R_2(p, q)$ is called the MS-stability function of the one-step method.

When we apply an mSROCK (3) to the test equation (4), we obtain

$$E[|X_{n+1}|^2] = R_m(p, q)E[|X_n|^2].$$

The stability function of the considered method is

$$R_m(p, q) = |R(p)|^2 + |\hat{R}(p)|^2 q^2,$$

where $R(p) = \frac{T_s(\omega_0 + \omega_1 p)}{T_s(\omega_0)}$ and $\hat{R}(p)$ is a polynomial about p and needs to be derived by performing s times on the drift term. The MS-stability region of (3) is given by

$$S = \{(p, q) : R_m(p, q) < 1\}. \quad (7)$$

We consider a method to be *more stable*, if its MS-stability region covers a larger part of (5) from $(0, 0)$ to negative p direction with all stable q values included. Figures 1–3 show the MS-stability regions for different values of η for $s = 2$. The darkened area is the exact solution MS-stability region (5) (in (λ, μ^2) axes scaled by h , which doesn't change

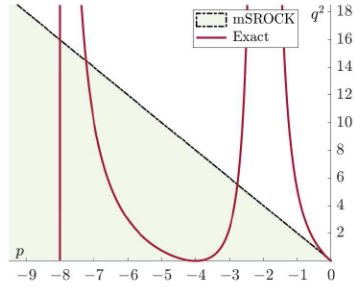


Fig. 1. 2-stage mSROCK
MS-stability with $\eta = 0$

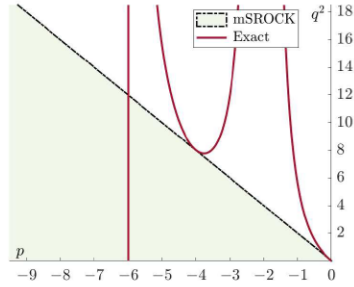


Fig. 2. 2-stage mSROCK
MS-stability with $\eta = 0.9$

it's form). The curve borders the MS-stability region (7) from above. As it can be seen, for $\eta = 0$ there is a gap in the middle which greatly restricts the stability area closest to the origin. We note that such gap does not enter the exact stability region when $\eta \geq 0.9$ and as η grows the width of the stability region is reduced. Thus we call $\eta = 0.9$ the *optimal* damping and the corresponding stability region is called the optimal stability domain.

In Figures 4–6 we respectively show the stability domain for the 2-, 3- and 4-staged SROCKs and mSROCKs at their *optimal* η values (different for the same s) [3]. There η_0 stands for the *optimal* damping for SROCK and η_1 – for mSROCK. We can see that mSROCKs have wider stability domains and smaller *optimal* damping parameter compared to SROCKs. Thus we can say that the proposed modification provides better stability characteristics compared to the original S-ROCK method.

4. Conclusion. A class of modified S-ROCK methods is proposed for stiff stochastic problems. Comparing the modified methods to the original S-ROCK on the basis of mean-square stability analysis, it is found that the modified S-ROCK method have wider stability domains with smaller *optimal* damping parameters. It should be mentioned though, that the implementation cost grows significantly, since now we need to calculate g term s times instead of only once. Still better mean-square stability may be worth it, so the proposed improvement should not be neglected and should be further investigated.

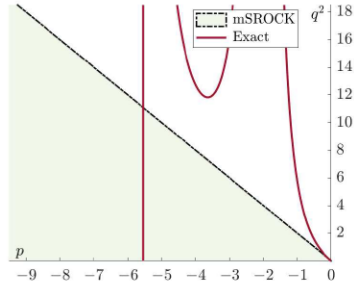


Fig. 3. 2-stage mSROCK MS-stability with $\eta = 1.36$

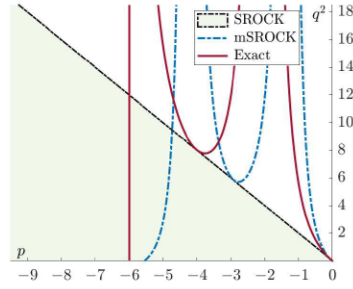


Fig. 4. 2-stage mSROCK vs SROCK with optimal η

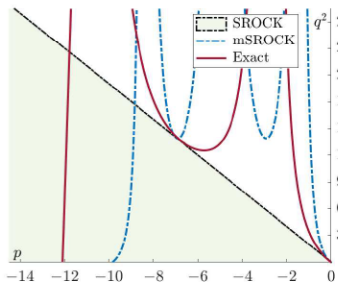


Fig. 5. 3-stage mSROCK vs SROCK with optimal η

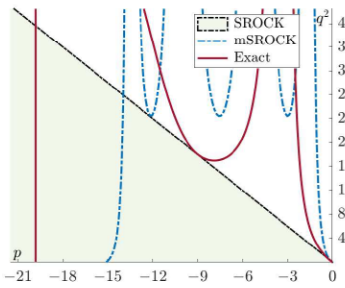
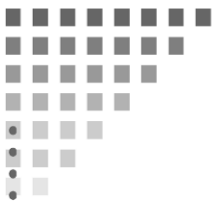


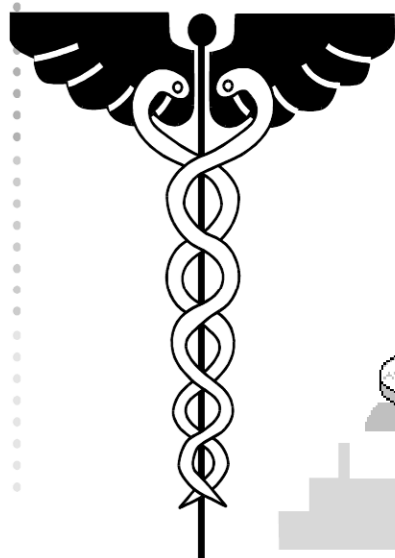
Fig. 6. 4-stage mSROCK vs SROCK with optimal η

References

1. Abdulle A., Li T. S-ROCK methods for stiff Ito SDEs // Communications in Mathematical Sciences. 2008. Vol. 6. P. 845–868.
2. Rathinasamy A., Balachandran K. Mean-square stability of second-order RungeKutta methods for multi-dimensional linear stochastic differential systems // Journal of Computational and Applied Mathematics. 2008. Vol. 219. P. 170–197.
3. Komori Y., Eremin A. S., Burrage K. S-ROCK methods for stochastic delay differential equations with one fixed delay // Journal of Computational and Applied Mathematics. 2019. Vol. 353. P. 345–354.



*3. Математические
модели медико-
биологических систем*



УДК 57.087.1

Давыденко Б. Н.

Прогнозирование результата коррекции гиперметропии с использованием CHAID анализа

Рекомендовано к публикации доцентом Корниковым В. В.

Актуальность. В настоящее время гиперметропия – одно из наиболее распространенных нарушений зрения. По некоторым оценкам [1] доля дальнозоркости среди прочих аномалий рефракции составляет от 27 до 40%. При этом биомеханика глаза способствует меньшей прогнозируемости результата коррекции этого заболевания, большей вероятности послеоперационного регресса по сравнению с коррекцией миопии [2]. Таким образом, статистические исследования в данной области остаются актуальными [3].

Данные и методы. В изучаемую выборку вошли данные о 169 рефракционных операциях по методу Femto LASIK. Среди этих данных были следующие измерения и исследования:

- НОЗ и КОЗ (некорригированная и корригированная острота зрения) с узким зрачком;
- субъективное измерение рефракции;
- рефрактометрия с узким и широким зрачком;
- кератометрия по рефрактометрии и кератотопографу;
- средний радиус кривизны роговицы;
- тонография и дифференциальная тонометрия (GlauTest-60);
- диаметр лимба по оптическому биометру;
- пахиметрия в центре и на расстоянии 3, 7 и 11 мм от него;
- ВГД (внутриглазное давление) по бесконтактному пневмотонометру;
- ВГД по Маклакову с тонометрами массой 15, 10, 7,5 и 5 г.

Давыденко Богдан Николаевич – студент, Санкт-Петербургский государственный университет; e-mail: st086788@student.spbu.ru, тел.: +7(900)569-15-87

Данные были предоставлены Чебоксарским филиалом ФГАУ НМИЦ «МНТК Микрохирургия глаза им. акад. С. Н. Федорова» Минздрава России. Для вычислений использовались программные пакеты MATLAB и SPSS.

В качестве предварительной обработки данных были исключены резко выделяющиеся значения. Исключение производилось на основании биомеханического значения. Для нормально распределенных данных (для проверки близости к нормальному распределению применялся критерий Колмогорова [4]) – на основании квантилей нормального распределения. Для параметров, распределение которых значительно отличалось от нормального, применялся фильтр Хэмпеля [5].

Для простоты интерпретации после снижения размерности параметры группировались по биомеханическому смыслу и степени корреляции. К каждой из таких групп по отдельности был применен метод главных компонент [6]. Далее рассматривались главные компоненты, объясняющие не менее 85% дисперсии исходных параметров. Некоторые переменные, такие как НОЗ, КОЗ и диаметр лимба были сохранены в изначальном виде. Таким образом, из исходных 34 переменных были получены 15 главных компонент и некоррелирующих переменных, над которыми производился весь последующий анализ. При наличии большого количества переменных-предикторов и множества пропущенных значений для построения прогностической модели подходит метод CHAID анализа [7]. В качестве прогнозируемых категорий приняты «благоприятный» (КОЗ при выписке более 0,7) и «неблагоприятный» (КОЗ при выписке не более 0,7) результаты операции. Такие категории были выбраны из соображений оценки результата и возможности обучения модели (в каждой из групп должно быть достаточное количество наблюдений). Для проверки качества прогнозирования исходная выборка делилась на две равные по объему – обучающую и тестовую.

Анализ результатов. На рис. 1 представлен результат CHAID анализа – дерево классификации. Основные переменные-предикторы КОЗ (CVA на рис. 1, 2) и SPH – главная сферическая компонента рефракции. Из метода главных компонент $SPH = 0,4805sph + 0,5920sph_n + 0,6471sph_w$, где sph , sph_n , sph_w – сферические компоненты рефракции по субъективному измерению рефракции и рефрактометрии с узким и широким зрачком соответственно. Можно

заметить, что обучающая выборка разделена на пять узлов: узлы 1 и 2 определяются исключительно по значению КОЗ, тогда как узлы 4-6, являющиеся дочерними для узла 3, – по КОЗ и SPH. Таким об-

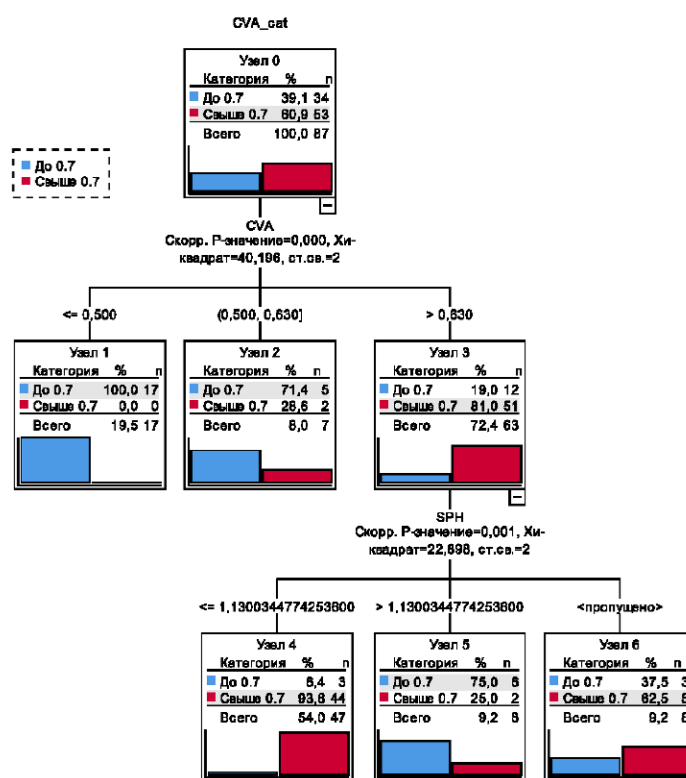


Рис. 1. Обучающая выборка

разом, для любого неклассифицированного наблюдения можно определить узел, которому оно принадлежит, и оценить, в соответствии с долями обучающих наблюдений, вероятность «благоприятного» или «неблагоприятного» исхода операции. Темным цветом отмечен «благоприятный» исход (послеоперационная КОЗ выше 0,7), соот-

ветственно светлым отмечен «неблагоприятный» исход.

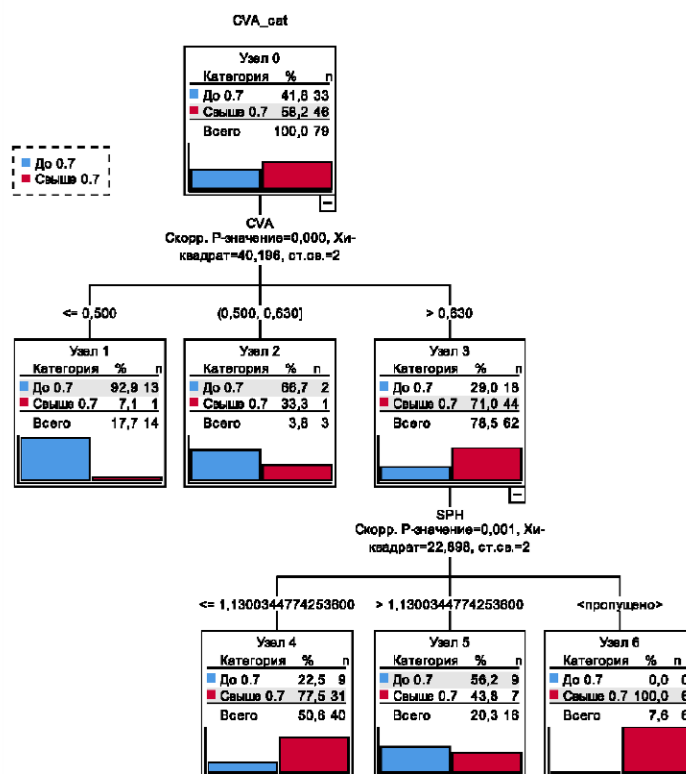


Рис. 2. Тестовая выборка

В соответствии с построенной моделью 88,5% наблюдений были классифицированы верно. Доля ошибочно классифицированных составила 0,115 со стандартной ошибкой 0,034. На рис. 2 представлено дерево классификации для тестовой выборки. Его структура соответствует дереву обучающей выборки. На диаграммах для каждого из узлов представлены доли наблюдений из тестовой выборки, попавших в соответствующий узел. Можно заметить, что доли на-

блюдений в каждом из узлов близки к долям в обучающей выборке, что свидетельствует в пользу качества прогностической модели. В соответствии с этим, в тестовой выборке доля ошибочно классифицированных данных – 0,228, ее стандартная ошибка – 0,047. Верно было классифицировано 77,2% наблюдений.

Заключение. Построена прогностическая модель, на основе которой возможно в определенной степени предсказать результат коррекции гиперметропии, оценить вероятность наилучшего исхода операции для пациента. Например, для пациентов с КОЗ не более 0,5 (узел 1) вероятность «благоприятного» исхода операции близка к 0,07.

Результат исследования и модель можно использовать на практике, а основные переменные-предикторы – для статистических исследований с применением иных методов.

Литература

1. Бауэр С. М., Венатовская Л. А., Авершина Л. А., Пикусова С. М. Оценка биомеханических характеристик на основе эластотонометрии после операций по коррекции гиперметропии // Вестник офтальмологии. 2020. № 5. С. 32–38.
2. Балашевич Л. И. Рефракционная хирургия. СПб.: СПбМАПО, 2002. 288 с.
3. Буре В. М., Щербакова А. А. Применение дискриминантного анализа и метода деревьев принятия решений для диагностики офтальмологических заболеваний // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2013. № 1. С. 70–76.
4. Большев Л. Н., Смирнов Н. В. Таблицы математической статистики. 3 изд. М.: Паука, 1983. 416 с.
5. Manoj K., Senthamarai K. K. Comparison of methods for detecting outliers // International Journal of Scientific & Engineering Research. 2013. No 9. P. 709–714.
6. Калинина В. Н., Соловьев В. И. Введение в многомерный статистический анализ: Учебное пособие. М.: ГУУ, 2003. 66 с.
7. Kass G. V. An exploratory technique for investigating large quantities of categorical data // Journal of the Royal Statistical Society. Series C (Applied Statistics). 1980. No 2. P. 119–127.

УДК 519.217

Легкова И. А.

Обнаружение отклонений и нарушений сердечного ритма в данных ЭКГ пациентов

Рекомендовано к публикации доцентом Гончаровой А. Б.

1. Введение. Электрокардиограмма (ЭКГ) является широко распространенным неинвазивным методом диагностики, который предоставляет ценную информацию о здоровье сердца. Анализ ЭКГ играет решающую роль в выявлении различных сердечных заболеваний, таких как блокада левой/правой ножки пучка Гиса, предсердная/желудочковая экстрасистолия.

Традиционные методы анализа ЭКГ полагаются на экспертные знания и ручную интерпретацию, что может быть подвержено субъективности и ошибкам. В последние годы динамические марковские модели (ДММ) получили широкое признание как мощный инструмент для обнаружения аномалий в данных. Данные подходы были использованы и при обнаружении аномалий в работе [1].

ДММ представляют собой статистические модели, которые описывают динамику последовательных наблюдений. Они учитывают как текущее состояние системы, так и ее историю, что делает их особенно подходящими для анализа временных рядов данных, таких как ЭКГ.

В этом исследовании предлагается использовать ДММ для обнаружения аномалий в данных ЭКГ. Предполагается, что аномальные сигналы ЭКГ могут быть смоделированы как отклонения от нормального поведения, которое можно захватить с помощью ДММ. Цель данной работы состоит в том, чтобы разработать модель, которая сможет точно идентифицировать аномалии в данных ЭКГ, тем самым повышая точность и объективность диагностики сердечных заболеваний.

2. Обзор литературы. За последнее время разработано ограниченное количество методов, способных эффективно решать задачу

Легкова Ирина Анатольевна – студент, Санкт-Петербургский государственный университет; e-mail: st088253@student.spbu.ru, тел.: +7(911)299-63-71

обнаружения аномалий во временных рядах, большинство из которых узкоспециализированы и адаптированы лишь под одну область применения. Например, модели на основе марковских цепей в основном используют для поиска сетевых аномалий.

В книге [2] дано понятное и развернутое объяснение многих аспектов, связанных с поиском аномалий, как точечных, так и во временных рядах. Достоинством данного источника является подробно рассмотренная теоретическая концепция построения скрытых марковских моделей (СММ), а также приведенные примеры использования СММ для обнаружения сетевых аномалий во временных рядах.

С помощью марковских цепей [3] и методов теории нечетких множеств [4] моделируется система постановки предварительного диагноза.

3. Теоретическая часть. Динамические марковские модели – это мощные вероятностные модели, которые могут использоваться для моделирования временных зависимостей и обнаружения аномалий во временных рядах. ДММ предполагают, что данные генерируются скрытым марковским процессом, где скрытые состояния представляют собой последовательность неизвестных переменных, которые управляют наблюдаемыми данными.

Преимуществом использования ДММ является экономия памяти, так как для их работы необходимо лишь знать значение на предыдущем шаге и в рассматриваемый момент времени.

4. Материалы и методы. База данных МПТ-ВН собиралась с 1975 по 1979 год и в сентябре 1980 года стала доступна для академических и промышленных групп. Всего в базе данных 99754 записей ЭКГ, размерность каждой записи 2000 единиц времени.

Записи ЭКГ были оцифрованы с частотой 360 Гц на канал с 11-битным разрешением в диапазоне 10 мВ. Два или более кардиолога независимо аннотировали каждую запись; разногласия были разрешены для получения компьютерочитаемых справочных аннотаций для каждой доли, включенных в базу данных.

Для реализации алгоритма построения динамических марковских моделей использовался язык программирования Python. В частности, такие модули как Numpy, Matplotlib, Pandas и Neurokit2.

5. Реализация. Прежде чем приступить к реализации алгоритма построения цепей Маркова [5], необходимо определить, какие ха-

рактические характеристики ЭКГ будут наиболее ценными в качестве состояний. Эти характеристики должны быть дискретными, наблюдаемыми и иметь четкие границы. Они также должны быть клинически значимыми и предоставлять информацию о состоянии сердца.

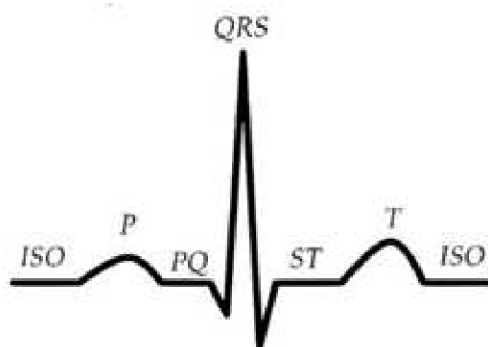


Рис. 1. Сегментация состояний ЭКГ

При погружении в медицинскую литературу [6], делаем вывод, что существенными характеристиками для исследования являются переходные состояния: P, PQ, QRS, ST и T (см. рис. [1]). Следовательно, поставленная задача сводится к анализу вероятностей попадания в эти состояния. Иными словами, требуется ответить на вопрос об устойчивости состояний. Для решения такой задачи применяются подходы, используемые при анализе ультразвуковых изображений нарушений в мягких тканях организма [7].

Использование цепей Маркова предусматривает, что какого-то состояния может не быть, что свидетельствует об аномалии.

Например, матрица переходных состояний для здорового пациента представлена в таблице.

Таблица. Переходная матрица состояний для здорового пациента

	P	PQ	QRS	ST	T
P	1	0	0	0	0
PQ	0	0,88	0,12	0	0
QRS	0	0	1	0	0
ST	0	0	0	0,9	0,1
T	0	0	0	0	1

Как видно из таблицы, вероятность остаться в текущем состоянии много больше, чем вероятность перехода в следующее состояние, что говорит об устойчивости рассматриваемой системы.

Переходные матрицы, вычисленные для здоровых пациентов, будут стремиться к единичной матрице.

6. Результаты. По результатам проведенной работы было обнаружено, что модель идентифицирует сигналы ЭКГ пациентов с заболеванием предсердной экстрасистолией с точностью 83%.

Вероятность обнаружения других заболеваний крайне мала, что говорит о недостатках построенной модели.

7. Заключение. Построена модель на основе ДММ, способная идентифицировать предсердную экстрасистолию с точностью 83%.

В дальнейшем алгоритм предполагается модифицировать для распознавания других сердечных заболеваний. В качестве модификации предусмотрено встраивание в модель учета состояний: положительный/отрицательный пик Т, длина интервалов RR, QT, PQ и комплекса QRS. Это позволит более тонко отслеживать аномальные сигналы ЭКГ и классифицировать их в пространстве заболеваний.

Литература

1. Легкова И. А. Обнаружение аномалий в медицинских данных // Современные методы прикладной математики, теории управления и компьютерных технологий (ПМТУКТ-2023). Сборник трудов Международной научной конференции. Воронеж, 2023. С. 51–52.
2. Шелухин О. И. Сетевые аномалии. Обнаружение, локализация, прогнозирование. М.: Горячая линия – Телеком, 2020. 448 с.
3. Кривополенова С. Д., Гончарова А. Б. Программная реализация системы постановки предварительного диагноза // Процессы управления и устойчивость. 2020. Т. 7. № 1. С. 153–157.
4. Гончарова А. Б. Постановка предварительного медицинского диагноза на основе теории нечетких множеств с использованием меры Сугено // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2019. Т. 15. № 4. С. 529–543.

5. Саакян Р. Р., Шпехт И. А., Петросян Г. А. Нахождение наличия заимствований в научных работах на основе марковских цепей // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2023. Т. 19. № 1. С. 43–50.
6. Ярцев С. С. Практическая электрокардиография. Справочное пособие для анализа ЭКГ. 3-е изд. М.: ГЭОТАР-Медиа, 2021. 144 с.
7. Гончарова А. Б., Бусько Е.А. Программная реализация системы принятия диагностических решений на основе мультипараметрических ультразвуковых показателей образований молочной железы // Сибирский журнал клинической и экспериментальной медицины. 2020. Т. 35. № 4. С. 137–142.

УДК 615.849

Халтобина М. М., Елизарова М. В.

Анализ выживаемости клеток в рамках RMR-модели

1. Введение. Цель математического моделирования в радиобиологии – возможность предсказать характер и степень воздействия ионизирующего излучения на биологические объекты, когда экспериментирование не представляется возможным [1–7]. Ионизирующее излучение вызывает не только явные повреждения клеток и тканей, но и скрытые повреждения, которые в некоторых структурах организма могут не обнаруживать себя на протяжении долгого времени, в частности, до тех пор, пока не потребуется задействовать их регенерационный потенциал. Количественное описание имеющихся экспериментальных данных в рамках математической модели может оказаться полезным не только для понимания механизмов реакции биологических объектов на облучение, но и для экстраполяции результатов в область отсутствующих экспериментальных данных.

2. RMR-модель. Модели выживания клеток, связанные исключительно с физикой и статистикой событий первоначального накопления энергии, пытаются соотнести конечное проявление биологических эффектов непосредственно с событиями ионизации [7]. С точки зрения выживаемости клеток наиболее важное последствие воздействия ионизирующих излучений – возникновение повреждений генетического материала. Эукариотические клетки обладают ферментативными механизмами репарации, которые способны устранять повреждения молекул ДНК и восстанавливать генетическую целостность. Процессы репарации, таким образом, на клеточном уровне изменяют результаты радиационного воздействия и влияют на выживаемость клеток.

В RMR-модели (repair-misrepair model) учитывается влияние ферментативной репарации на выживаемость клеток [1, 2]. Основ-

Халтобина Мария Михайловна – студент, Санкт-Петербургский государственный университет; e-mail: st085324@student.spbu.ru, тел.: +7(981)790-00-40

Елизарова Марина Владиславовна – младший научный сотрудник, Санкт-Петербургский государственный университет; e-mail: st809707@spbu.ru, тел.: +7(921)300-35-91

ными повреждениями, приводящими к гибели клеток, являются двунитевые разрывы ДНК. Если происходит взаимодействие сублетальных одиночных разрывов, то возможно увеличение числа летальных для клетки двойных разрывов. Также имеют место одиночные разрывы цепи ДНК, не принявшие участие в образовании двойных разрывов, которые с большой вероятностью могут быть отрепарированы и не приводят к летальному для клетки исходу.

Биохимические процессы восстановления начинаются, когда физическая и химическая фаза действия ионизирующего излучения завершены. На данном этапе происходит распознавание повреждений, мобилизация последовательности специфических ферментов, ресинтез поврежденных участков ДНК. Известные типы макромолекулярных повреждений – одно- и двунитевые разрывы ДНК, межнитевые поперечные связи, изменение оснований и димеризация. Скорость восстановления и последовательности ферментов для каждого типа поражения различны, а временной промежуток восстановления может варьироваться от минут до дней [8].

RMR-модель описывает количество соответствующих макромолекулярных поражений на клетку в зависимости от дозы D ; зависящие от времени t трансформации этих поражений; и зависящие от времени и дозы вероятности выживания S , летальности L и мутации. Величина U означает повреждения до того, как они подвергаются ферментативному восстановлению и модификации.

Репарационные состояния R представляют собой результат преобразований U состояния после ферментативной репарации. Клетка может в результате репарации точно восстановить последовательность ДНК такой же, какой она была до радиационного воздействия. Или же может быть реализован вариант неправильной репарации, в результате которой образуется мутация и изменение, которое может присутствовать в жизнеспособной клетке, вызвать гибель клетки или вызвать гибель ее потомков. В модели рассматриваются состояния, являющиеся результатом биохимической репарации, как в случае с полным сохранением генетической последовательности, так и в случае, когда восстановление последовательности является неполным.

Для однократной, быстро доставляемой дозы излучения с низкой линейной передачей энергии (LET, linear energy transfer) зависящее от времени поведение U описывается квадратичным дифференциальным уравнением первого порядка

$$\frac{dU}{dt} = -\lambda U(t) - kU^2(t) \quad (1)$$

с начальным условием $U(0) = U_0$. В этом уравнении λ – коэффициент линейной репарации, а k – коэффициент квадратичной репарации. При этом под линейной репарацией понимается исправление одноцепочечных разрывов ДНК, повреждений оснований, а также гомологичная репарация двунитевых разрывов. В свою очередь, квадратичная репарация подразумевает процесс негомологичной репарации двунитевых разрывов, которые можно рассматривать как результат взаимодействия пары одонитевых разрывов ДНК. Считается, что повреждения распределены равномерно по ядру клетки. Разделяя переменные в уравнении (1), найдем

$$\int -\frac{dU}{\lambda U + kU^2} = \int dt,$$

тогда

$$U = \frac{U_0 e^{-\lambda t}}{1 + \frac{U_0}{\varepsilon}(1 - e^{-\lambda t})}, \quad (2)$$

где U_0 – исходные повреждения на момент $t = 0$. Уравнение (2) описывает изменение количества повреждений U в результате репарации.

Рассматривая результаты линейной и квадратичной репарации, перейдем к рассмотрению выживаемости и ее зависимости от времени и от дозы излучения.

В каждый момент времени $U(t) + R_L(t) + R_Q(t) = U(0)$, где R_L и R_Q – результаты линейной и квадратичной репарации, соответственно. Проинтегрировав уравнение (1) с соответствующим начальным условием, получим

$$R_L(t) = \varepsilon \ln \left(1 + \frac{U_0}{\varepsilon}(1 - e^{-\lambda t}) \right), \quad (3)$$

$$R_Q(t) = \frac{U_0(1 + \frac{U_0}{\varepsilon})(1 - e^{-\lambda t})}{1 + \frac{U_0}{\varepsilon}(1 - e^{-\lambda t})} - \varepsilon \ln \left(1 + \frac{U_0}{\varepsilon}(1 - e^{-\lambda t}) \right). \quad (4)$$

Вероятность выживания клетки зависит от количества неправильно репарированных и нерепарированных повреждений на момент клеточного деления. Выживаемость обычно измеряется по репродуктивной целостности клеток-потомков после нескольких клеточных делений. Временную переменную можно ограничить, зафиксировав t_{\max} – допустимый интервал времени репарации. Если предположить, что любая линейная репарация является успешной и не вызывает гибель клетки, а все квадратичные репарации приводят к летальному исходу, получим простейшее уравнение выживаемости RMR-модели

$$S(t) = \exp(-R_Q - U) = e^{-U_0} \left[1 + \frac{U_0}{\varepsilon} (1 - e^{-\lambda t}) \right]^\varepsilon. \quad (5)$$

В этом случае все остаточные повреждения U считаются летальными, и в момент времени t любая клетка, имеющая результаты квадратичной репарации повреждения, погибнет, а выживут только клетки, у которых нет нерепарированных повреждений U или R_Q .

Доза не входит явно в вероятность выживания (5), однако U_0 зависит от дозы. Будем считать, что количество образовавшихся дефектов является линейной функцией от дозы $U_0 = \alpha D$, где α задает количество разрывов нитей на клетку в расчете на единицу дозы. Пропорциональность сохраняется независимо от используемого излучения, но величина α зависит от типа пучка.

Таким образом, простейшее уравнение выживаемости клеток в RMR-модели с параметрами α – количество повреждений U на единицу дозы, $\varepsilon = \frac{\lambda}{k}$ – коэффициент репарации, имеет вид

$$S = e^{-\alpha D} \left[1 + \frac{\alpha D}{\varepsilon} (1 - e^{-\lambda t}) \right]^\varepsilon. \quad (6)$$

3. Кривые выживаемости клеток и параметры модели. В рамках RMR-модели был проанализирован набор данных [6] по выживаемости для различных клеточных линий, облученных в одинаковых условиях.

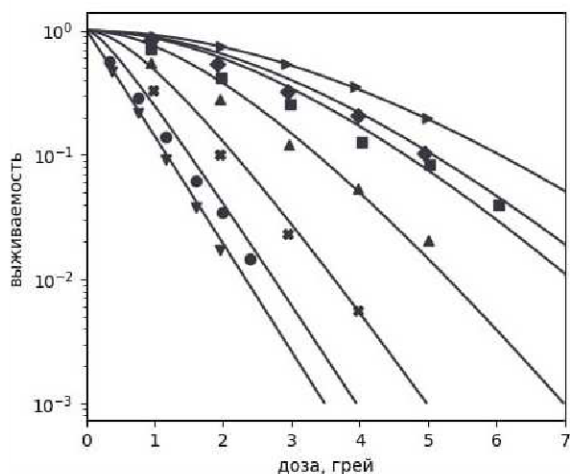


Рис. 1. Экспериментальные данные выживаемости различных клеточных линий (символы) [6] и расчетные кривые (сплошные линии) RMR-модели

На рис. 1 представлены экспериментальные данные [6] совместно с кривыми, рассчитанными в RMR-модели и демонстрирующими наилучшее согласие с экспериментом. Поскольку облучение проводилось в одинаковых условиях, параметр α , характеризующий количество повреждений на единицу дозы, был зафиксирован единым на значении, при котором согласие расчета и эксперимента достигалось для всех представленных клеточных линий. Для оценки динамики коэффициента репарации ε мы установили длительный допустимый интервал времени репарации, что соответствует значениям параметра λt , выше которых изменение этого параметра уже не влияет на форму кривых. Таким образом, изменение кривых выживаемости от наименее радиочувствительных линий опухолевых клеток к наиболее радиочувствительным в рамках RMR-модели может быть описано изменением одного параметра – снижением величины коэффициента репарации ε . Этот факт можно интерпретировать следующим образом: клетки, демонстрирующие более высокую радиорезистентность, более эффективно осуществляют репарацию

радиационных повреждений, лучше сохраняют клоногенную способность и демонстрируют выживаемость при более высоких значениях дозы облучения.

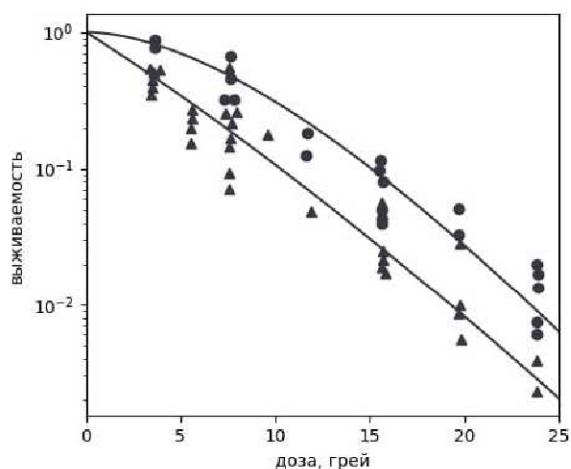


Рис. 2. Экспериментальные данные выживаемости опухолевых клеток (символы) [9] и расчетные кривые (сплошные линии) RMR-модели

Далее были проанализированы данные выживаемости мышей при солидном раке Эрлиха [9]. На рис. 2 представлены экспериментальные данные выживаемости клеток после облучения живых животных в обычных условиях и клеток, облученных в состоянии апоносии.

При подборе расчетных кривых учитывалось, что исследование гипоксии проводилось на одинаковых опухолевых клетках, т. е. коэффициент репарации ε должен быть одинаковым. Хотя облучение производилось одинаковым излучением, известно, что дефицит кислорода приводит к снижению количества повреждений в химической фазе действия ионизирующего излучения. Следовательно, нет оснований фиксировать для двух представленных наборов данных одинаковое значение параметра α . Кроме того, если предположить, что

способность клеток к репарации сохраняется и после смерти животных от асфиксии, неизвестны временные рамки такой репарации без восстановления снабжения клеток кислородом.

Изменение показателей выживаемости, а также изменение формы кривой для клеток с дефицитом кислорода очевидно, несмотря на разброс экспериментальных данных. Хотя такая динамика может быть описана расчетными кривыми в рамках RMR-модели, интерпретация изменений расчетных параметров в данном случае затруднена.

4. Заключение. RMR-модель использована для анализа данных выживаемости клеток после воздействия ионизирующего излучения. Методику можно применять для случаев фракционированного облучения, а также при использовании излучения с различной относительной биологической эффективностью. При интерпретации данных необходимо учитывать особенности постановки радиобиологического эксперимента.

Литература

1. Tobias C. A. The repair-misrepair model in radiobiology: comparison to other model // *Radiat. Res.* 1985. No 104. P. 77–92.
2. Pirruccello M. C., Tobias C. A. Repair-misrepair model of cell survival // Berkeley: California Univ., Lawrence Berkeley Lab, Nov 1980. P. 137–147.
3. Котина Е. Д. Математическое и компьютерное моделирование в ядерной медицине. СПб.: ВВМ, 2022. 143 с.
4. Котина Е. Д., Елизарова М. В. Введение в дозиметрическое планирование дистанционной лучевой терапии. СПб.: ВВМ, 2019. 60 с.
5. Гончарова А. Б., Колпак Е. П., Виль М. Ю., Абрамова А. В., Бусько Е. А. Математическое моделирование злокачественных опухолей яичников // *Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления.* 2022. Т. 18. № 1. С. 120–134.
6. Chapman J. D. Single-hit mechanism of tumour cell killing by radiation // *Int. J. Radiat. Biol.* 2003. Vol. 79. No 2. P. 71–81.

7. Kappos A., Pohlit W. A. A cybernetic model for radiation reaction in living cells. I. Sparsely-ionising radiation; stationary cells // *Int. J. Radiat. Biol.* 1972. Vol. 22. No 1. P. 51–65.
8. Cleaver J. E. DNA repair and its coupling to DNA replication in eukaryotic cells // *Biochim. Biophys. Acta.* 1978. Vol. 516. No 4. P. 489–516.
9. Ярмоненко С. П., Вайнсон А. А. Радиобиология человека и животных. М.: Высшая школа, 2004. 549 с.

УДК 004.8

Хоменко И. Е.

**Преобразование количественных данных
в бинарные с помощью адаптивной двусторонней
оценки для построения модели
выживаемости пациентов**

Рекомендовано к публикации доцентом Гончаровой А. Б.

1. Введение. Прогнозирование выживаемости пациентов с коронавирусной инфекцией играет важную роль в борьбе с заболеванием.

Прогнозирование выживаемости позволяет определить наиболее уязвимые группы населения, которым необходимо уделить особое внимание и медицинскую помощь. Оно позволяет оценивать не только общую выживаемость, но и прогнозировать, какие именно пациенты будут иметь лучшие шансы на выздоровление. Таким образом, медицинские учреждения могут адаптировать свои лечебные планы, сфокусировавшись на наиболее эффективных методах лечения для каждого пациента, уменьшая затраты и повышая качество медицинской помощи.

Без прогнозирования выживаемости медицинским работникам труднее принимать оптимальные решения и предоставлять наиболее качественную медицинскую помощь пациентам.

Поэтому задача анализа выживаемости пациентов с коронавирусной инфекцией является важной, а как следствие, важным является и улучшение качества прогностической модели. В настоящее время в медицинском анализе часто прибегают к переходу к дихотомическим данным, так как это упрощает работу медицинскому персоналу. Для перехода к дихотомическим данным чаще используют метод с применением индекса Юдена или же медианы [1]. Однако данные методы являются односторонними и при работе с медицинскими параметрами могут давать некачественный результат [2]. Поэтому построим вариант адаптивной двусторонней оценки.

Хоменко Ирина Евгеньевна – студент, Санкт-Петербургский государственный университет; e-mail: st087797@student.spbu.ru, тел.: +7(924)150-11-84

2. Цель исследования. Построение метода бинарной классификации медицинских параметров двусторонней оценкой на примере данных госпитализированных пациентов, поступивших в отделение реанимации и интенсивной терапии (ОРИТ). Сравнение полученных результатов с классическим односторонним методом перехода к бинарным данным через индекс Юдена.

3. Материалы и методы. База медицинских данных включает обезличенные медицинские данные 250 параметров 299 пациентов, поступивших в ОРИТ в 2020 году. База данных предоставлена ФГБУ «Федеральный научно-клинический центр специализированных видов медицинской помощи и медицинских технологий Федерального медико-биологического агентства» [3].

Предобработка данных, расчеты, а также построение модели анализа и его оценка проводилась с использованием высокоуровневого языка программирования Python с применением программных библиотек для работы с базами данных – Pandas и NumPy.

4. Разработка адаптивной двусторонней оценки параметров. Для задачи из 250 параметров отобраны непрерывные данные, которые имеют не более 25% пустых значений. Также при работе учитываются только данные, полученные за первые пять дней пребывания пациента в ОРИТ. Таким образом в дальнейшем рассматриваем 42 непрерывных параметра и результирующий параметр – «Выживаемость».

Для корректной работы с данными, так как у части пациентов отсутствуют некоторые критерии, произведем восстановление данных при помощи машинного обучения методом линейной регрессии на имеющихся параметрах. Проверка показала, что заполнение пустых ячеек не меняет общего поведения случайных значений.

Для построения адаптивных оценок используются квартили [4,5]. Для этого каждый показатель разбивается на два класса в зависимости от значения результирующего параметра и ищутся квартили данных распределений. Далее полученные значения подставляются в алгоритм (см. рис. 1), где x_1, x_2, x_3 – квартили большего по мощности множества, а y_1, y_2, y_3 – меньшего. Алгоритм построен за счет оценки работы задачи, решаемой при помощи машинного обучения, направленной на максимизацию значения чувствительности при классификации параметров. Уровни соотносятся с порядком расположения квартилей по возрастанию. Нижний уровень характе-

ризует границу разбиения по классам. В случае, если указано двойное значение, применяется двусторонняя оценка.

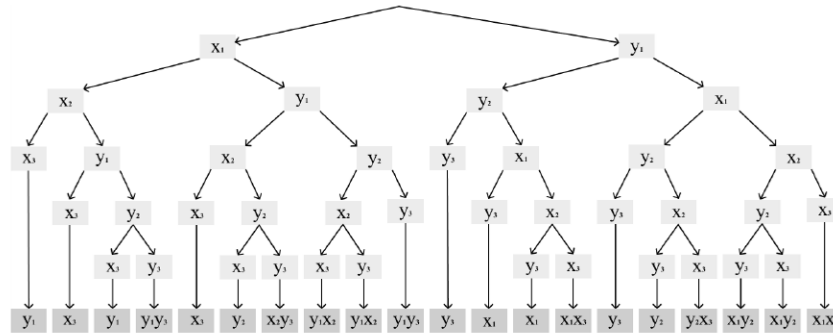


Рис. 1. Алгоритм выбора границ классификации

Таким образом получаем алгоритм, который позволяет без обучения на выборке производить перевод непрерывных данных в бинарные. Он адаптивный, дает возможность производить перевод любых данных без предварительного анализа необходимости использовать непосредственно двустороннюю оценку или одностороннюю. Алгоритм не сложен в реализации и не привязан к конкретным данным, может быть использован на стадии постановки предварительного диагноза [6].

Пример. Рассмотрим работу алгоритма на примере параметра «Гемоглобин при поступлении». В данном случае большим по мощности множеством являются анализы пациентов с летальным исходом, поэтому $x_1 = 116,6$; $x_2 = 131$; $x_3 = 148$; $y_1 = 118,25$; $y_2 = 136,15$; $y_3 = 144,775$. Расположим значения по возрастанию: $x_1 y_1 x_2 y_2 y_3 x_3$. Обращаясь к схеме алгоритма (см. рис. 1), получаем, что в качестве границ разбиения будут выступать $x_2 = 131,0$ и $y_3 = 144,775$. Данный вариант разбиения хорошо соотносится с реальными медицинскими оценками уровня гемоглобина.

Как видно из примера, алгоритм прост в применении. Оценим его эффективность, построив классификационную модель с его использованием.

5. Построение классификационной модели. Для построения классификационной модели используется биномиальная логи-

стическая регрессия [7]:

$$P = \frac{1}{1 + e^{-y}},$$

где $y = b_0 + \sum_{i=1}^n (b_i x_i)$, x_i – дихотомические переменные, b_i – коэффициенты модели, n – число переменных. Коэффициенты модели находятся при помощи обучения модели на обучающей выборке.

Таким образом строятся две модели, лишь с тем различием, что в первом случае для построения модели используется перевод непрерывных данных в бинарные с опорой на индекс Юдена, а в другом случае – при помощи адаптивной двусторонней оценки, методика которой описана в пункте 4.

Построение моделей важно производить на одинаковых по сбалансированности данных обучающей и тестовой выборках. Максимальный по численности показатель (в данном случае, летальный исход – 1) был урезан до минимального по численности (в данном случае, выздоровление пациента – 0), тем самым база содержит данные уже 284 пациентов. В нашем случае в качестве тестовой выборки отберем 20% данных, т. е. обезличенные данные 58 пациентов.

Таблица. Сравнение моделей

	Точность	Чувствительность	Специфичность
Индекс Юдена	0,741	0,793	0,689
Адаптивная двусторонняя оценка	0,948	0,966	0,931

6. Вывод. Построена регрессионная модель высокого качества, включающая только бинарные данные. По приведенному сравнению в таблице видно, что метод имеет преимущество перед классическим подходом, точность модели выросла более чем на 20%, чувствительность – на 17%, а специфичность – на 24%.

При этом метод сохранил основные преимущества использования методов бинарной классификации параметров: простоту хранения и обработки данных, а также устойчивость к незначительным ошибкам в измерении данных.

Литература

1. Аржаник А. А., Гончарова А. Б. Сравнение способов преобразования количественных данных в бинарные при предсказании рисков осложнения внебольничной пневмонии // Процессы управления и устойчивость. 2020. Т. 7. № 1. С. 148–152.
2. Гончарова А. Б., Бусько Е. А. Программная реализация системы принятия диагностических решений на основе мультипараметрических ультразвуковых показателей образований молочной железы // Сибирский журнал клинической и экспериментальной медицины. 2020. Т. 35. № 4. С. 137–142.
3. Легкова И. А., Хоменко И. Е. Применение логистической регрессии для анализа необходимости подключения больного коронавирусной инфекцией к аппарату ИВЛ // Процессы управления и устойчивость. 2023. Т. 10. № 1. С. 207–210.
4. Tayal K., Ravi V. Fuzzy association rule mining using binary particle swarm optimization: Application to cyber fraud analytics // 2015 IEEE International Conference on Computational Intelligence and Computing Research (ICCIIC). 2015. P. 1–5.
5. Большаков А. А., Егоров И. В., Лобанов В. В., Лачугин Д. В. Разработка стенда для оценки применимости транспортных протоколов в задачах обработки потоковой информации для создания адаптивной системы преобразования данных // Вестник ТГТУ. 2014. № 3. [Электронный ресурс]: URL:<https://clck.ru/39UmyZ> (дата обращения: 12.03.2024).
6. Кривополенова С. Д., Гончарова А. Б. Программная реализация системы постановки предварительного диагноза // Процессы управления и устойчивость. 2020. Т. 7. № 1. С. 153–157.
7. Zhang S., Yang J. Factors influencing TCM syndrome types of acute cerebral infarction: A binomial logistic regression analysis // Medicine. 2023. Vol. 102. No 46. P. 36–80.

UDC 519.6

Valieva N. L., Goncharova A. B.

Breast cancer classification in mammograms

1. Abstract. Breast Cancer causes millions of deaths every year. So, about 2.3 million women were diagnosed with this disease in 2020. Usually, breast cancer is detected by doctors using X-Ray, CT and MRI. Nevertheless, the development of Artificial Intelligence (AI) gave rise to the Computer-Aided Diagnosis (CAD) systems to assist doctors in their routine work. The need for interpretable and trustworthy AI is now key aim in most fields of AI research. Therefore, the aim of this research is to classify breast cancer images using different approaches and try to interpret them automatically.

2. Introduction. According to the research of World Health Organisation (WHO) the breast cancer caused 685,000 deaths in 2020 with about 2.3 million women diagnosed with this disease. Approximately in half of the cases the women have no risks of having breast cancer excluding such factors as sex and age. In breast cancer classification 2 classes are identified: benign and malignant tumors. Masses and microcalcifications are the mean indicators of breast cancer. Some breast tissues sometimes look similar to masses, so mammography is sometimes not enough to detect malignant tumor. Moreover, false positive cases put healthy patients in emotional and physical risk. CAD can be the solution to the described problems. For the past few years with the rapid development of AI CAD became widely used in many research areas in medicine: colonoscopy [1], leukaemia [2] etc.

Deep Learning (DL) is widely used in medical systems. According to [3], the State-of-the-Art (SOTA) classification, segmentation and detection models are AlexNet, GoogleNet, RetinaNet, Resnet and YOLO. Abunasser et al. [4] state, that their fine-tuned Xception DL model managed to get high results of classification of breast images into 8 classes, the f1-score is 97.58%. Despite of using DL, classical Machine Learning (ML) approaches also show high results in detection. Ensemble is

Valieva Nafisa Lenisovna – undergraduate student, Saint Petersburg State University; e-mail: st117021@student.spbu.ru, phone: +7(900)321-12-07

Goncharova Anastasia Borisovna – associate professor, Saint Petersburg State University; e-mail: a.goncharova@spbu.ru, phone: +7(812)428-47-26

an approach, when results of several models are combined together to get a higher result at the end. Das et al. [5] use this method and high accuracy in detection and classification of breast tumors. Dewangan et al. [6] proposed a new optimisation method to get the accuracy of 99.6% in breast cancer detection at early stage using MRI images.

Deep Learning methods seem as an opportunity to trust routine work to CAD systems, but the more and more ethical questions are raising with the bigger usage of AI. Interpretable AI models allow doctors to make decisions based on imaging without using invasive methods which might be harmful for the patient's life. Zohaib et al. [7] observe modern interpretability methods commonly used in medical AI: concept learning methods, case-based methods, counterfactual explanation etc.

In this work we will use Resnet-18 for breast images classification with Grad-CAM interpretability and try Topological Data Analysis (TDA) approach for breast cancer classification.

3. Dataset. This section describes the characteristics of the dataset containing mammograms used for breast cancer classification in medical images.

3.1. MosMedData. MosMedData-MMG-BREASTCR-type I-v 3 [8] is an online public dataset of results of mammography researches calibrated for services based on AI. The dataset contains 100 records, overall 402 DICOM-files. The dataset is used for binary (pathology and no-pathology) and multi-label classification (BI-RADS 0-2). The labels in all mammograms are recognized and annotated by expert radiologists.

3.2. CBIS-DDSM. Curated Breast Imaging Dataset (CBIS) of DDSM Dataset [9] is an updated version of Digital Database of Screening Mammography. The CBIS-DDSM consists of 6775 studies and also has 6 files. In this research only 1318 were selected and used for fine-tuning the pre-trained model for binary image classification: the mass file contains 3 classes for pathology: "BENIGN", "MALIGNANT", "BENIGN WITHOUT CALLBACK". For the binary classification case the "BENIGN" was changed to "No pathology" class and "MALIGNANT", "BENIGN WITHOUT CALLBACK" to "Pathology".

4. Dataset preparation. Two datasets presented in the section Dataset will be used in the model training for binary image classification. Overall, the dataset contains 1720 images of breast with labels, where 1 means an indicated pathology, and 0 – no pathology. Dataset

preparation is one of the most important parts of model training. This number of images may be insufficient for the model to get enough features for classification, therefore we make several experiments using data augmentations and without augmentations. Data augmentation is the process of enriching dataset by applying geometrical and morphological transformations to the existing dataset. Al-Antari et al. [10] use several types of transformation for data augmentation: 8 types of rotation with angles of 45 degrees. Other works such as [11], [12] and [13] also use rotational transformations, so in this work we will also use rotation of images with angles of 90 degrees: 0, 90, 180, 270.

5. Model Description. In this work we used pre-trained ResNet-18 model with some minor changes, the classification metric is accuracy:

1. Change in the 1st Convolution layer: As we have .dcm images with only one channel, the `in_channels` parameter should be equal to one;
2. Changes in the last Fully-Connected layer's output parameters for binary classification;
3. In evaluation part of the models adding `forward_hook` and `backward_hook` to extract the features that are important for classification and interpretation of the model.

ResNet-18 became a SOTA model in a wide range of tasks related to image-processing: object detection, classification, and segmentation. Its use of residual blocks and skip connections make it a powerful and efficient choice for medical tasks.

The overall accuracy we got using this model is 0.6657. We also experimented with Topological Data Analysis techniques for computer vision. 73 topological features we generated for each image in the dataset and further Random Forest Classifier was used to classify the tabular data of features: Cubical Persistence, Amplitude of Betti curve, Persistence Entropy, Betti Curve, Binarizer, Height Filtration, Erosion Filtration, Dilation Filtration. The `f1_score` got from this experiment was 0.55, which is not significant to make decisions about the malignancy of found tumor in the image.

6. Limitations. The work has several limitations to discuss. Firstly, the number of images might be not enough to make generalisable models

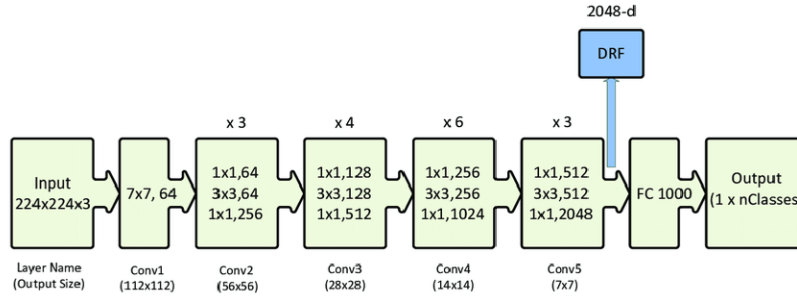


Figure 1. Schematic of convolutional neural network image model

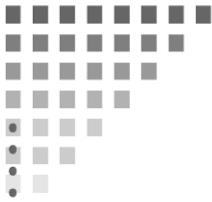
for different kinds of architecture. Secondly, the more classification models could be observed to make a more full research of DL possibilities. These limitations should be taken into account in further researches.

7. Conclusion. Breast Cancer causes lots of deaths worldwide and lots of people aim to find cure for the disease. The aim of the study was to study current approaches towards classifying breast cancer images using DL methods and try implementing other ways of classification. The dataset was collected from MosMed data and CDIS-DDSM with several minor changes. The ResNet-18 and TDA with Random Forest were used for classification with 0.6657 and 0.55 accuracy respectively. In further researches we will aim to improve accuracy with adding more diverse data and using other models for classification.

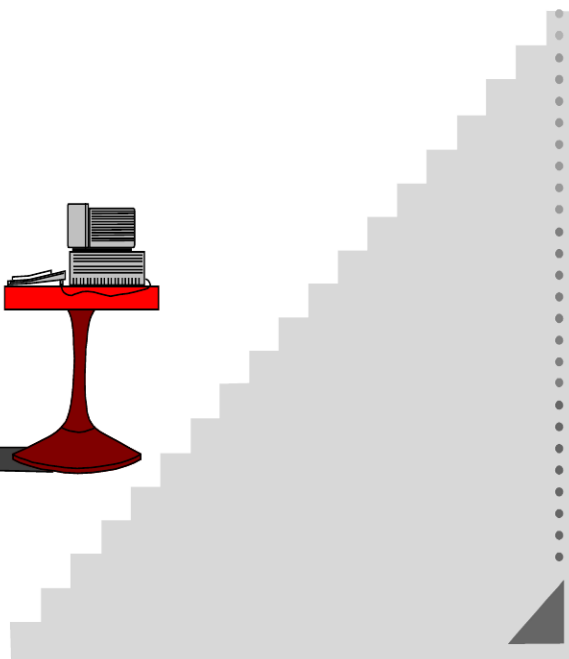
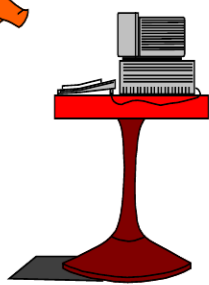
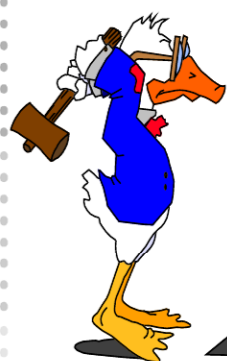
References

1. Mori Y. et al. Computer-aided diagnosis for colonoscopy // Endoscopy. 2017. Vol. 49. No 8. P. 813–819.
2. Shafique S., Tehsin S. Computer-aided diagnosis of acute lymphoblastic leukaemia // Computational and Mathematical Methods in Medicine. 2018. Art. no 6125289.
3. Hamed G., Marey M. A. E. R. et al. Deep learning in breast cancer detection and classification // Conference AICV2020. 2020. P. 322–333.

4. Abunasser S. et al. Convolution neural network for breast cancer detection and classification using deep learning // Asian Pacific Journal of Cancer Prevention. 2023. Vol. 24. No 2. P. 531–544.
5. Das A. et al. Breast cancer detection using an ensemble deep learning method // Biomedical Signal Processing and Control. 2021. Vol. 70. Art. no 103009.
6. Dewangan K. K. et al. Breast cancer diagnosis in an early stage using novel deep learning with hybrid optimization technique // Multimed Tools Appl. 2022. Vol. 81. No 10. P. 13935–13960.
7. Salahuddin Z. et al. Transparency of deep neural networks for medical image analysis: A review of interpretability methods // Computers in Biology and Medicine. 2022. Vol. 140. Art. no 105111.
8. Morozov S. P. et al. Mammography studies results for calibrating artificial intelligence services. Certificate of state reg. of the database 2020621741 [Internet resource]: URL:<https://rb.gy/6afqen> (date: 18.03.2024).
9. Sawyer-Lee R. et al. Curated breast imaging subset of DDSM [Internet resource]: URL:<https://pubmed.ncbi.nlm.nih.gov/29257132/> (date: 11.03.2024).
10. Al-Antari A. et al. A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification // Int J Inform. 2018. Vol. 117. P. 44–54.
11. Al-Masni A. et al. Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system // Comput Methods Programs Biomed. 2018. Vol. 157. P. 85–94.
12. Levy D., Arzav J. Breast mass classification from mammograms using deep convolutional neural networks // 30th Conference on Neural Information Processing Systems. 2016. P. 1–6.
13. Yuan Z.-W., Jun Z. Feature extraction and image retrieval based on AlexNet // Eighth International Conference on Digital Image Processing (ICDIP 2016). 2016. Vol. 10033. Art. no 100330E.



*4. Информационные
и компьютерные
технологии*



УДК 519.6

Антропова Е. Г.

Решение оптимизационных задач при помощи нейронных сетей

Рекомендовано к публикации профессором Смирновым Н. В.

1. Введение. Решение задач оптимизации является актуальной темой, так как эти задачи имеют множество приложений в различных областях науки, техники и экономики. Известно, например, что такие задачи встречаются и играют важную роль в алгоритмах машинного обучения, использующих нейронные сети [1]. В рамках таких задач обучение нейронной сети есть по сути решение задачи оптимизации параметров сети для минимизации функции ошибки, чтобы обученная нейронная сеть могла наилучшим образом аппроксимировать зависимость между входными и выходными данными.

В данной работе на основе статьи [2] исследуется подход к решению задач оптимизации с помощью нейронных сетей. Идея данного подхода состоит в том, чтобы, используя архитектуру многослойного перцептрона, аппроксимировать нелинейной регрессией исходную сложную целевую функцию, возникающую в задаче оптимизации. При этом производная полученной аппроксимированной целевой функции будет иметь полиномиальный вид, что позволяет легче вычислить решение для задачи оптимизации.

2. Многослойный перцептрон. В данной работе с учетом результатов [2] для вычисления приближенной целевой функции будет использоваться одна из простейших классических архитектур – многослойный перцептрон с одним скрытым слоем – которая тем не менее обладает хорошими аппроксимационными свойствами [3].

Для многослойного перцептрона правило прямого распространения от входного слоя сети к j -му нейрону в скрытом слое есть взвешенная сумма значений выходов $x_i \in \mathbb{R}$ нейронов входного слоя с

Антропова Екатерина Геннадьевна – студент, Санкт-Петербургский государственный университет; e-mail: st093661@student.spbu.ru; тел.: +7(917)428-44-15

весами $w_{ij} \in \mathbb{R}$, взятая со смещением $\theta_j \in \mathbb{R}$:

$$\sum_{i=1}^n w_{ij} x_i + \theta_j,$$

где n – количество нейронов во входном слое.

После вычисления правил распространения применяется функция активации. В качестве активационной функции нейронов будет использоваться функция $f(\cdot) = \operatorname{arctg}(\cdot)$, имеющая дробно-рациональную производную $[\operatorname{arctg}(x)]' = \frac{x'}{1+x^2}$, $x \in \mathbb{R}$.

Определив активационную функцию нейронов скрытого слоя, по аналогии записано правило прямого распространения от всех m нейронов скрытого слоя к k -му нейрону выходного слоя:

$$y_k(t) = \sum_{j=1}^m \bar{w}_{jk} f\left(\sum_{i=1}^n w_{ij} x_i + \theta_j\right) + \bar{\theta}_k. \quad (1)$$

Пусть нужно решить следующую задачу оптимизации:

$$\begin{aligned} F(x_1, \dots, x_n) &\rightarrow \max, \\ r_1(x_1, \dots, x_n) &\leq 0, \quad \dots, \quad r_m(x_1, \dots, x_n) \leq 0, \end{aligned} \quad (2)$$

где F – исходная целевая функция, а r_i – ограничения. Тогда, сгенерировав данные для обучения из области параметров, заданных ограничениями исходной задачи, можно аппроксимировать целевую функцию F с помощью нейронной сети (1) с одним нейроном в выходном слое (т. е. при $k = 1$) (см. рис. 1):

$$\begin{aligned} F(x_1, \dots, x_n) &\approx \sum_{j=1}^m \bar{w}_{j1} f\left(\sum_{i=1}^n w_{ij} x_i + \theta_j\right) + \bar{\theta}_1, \\ r_1(x_1, \dots, x_n) &\leq 0, \quad \dots, \quad r_m(x_1, \dots, x_n) \leq 0. \end{aligned}$$

Полученную новую задачу оптимизации возможно теперь решить, например, используя теорему Куна – Таккера [4]:

$$\begin{aligned} \frac{\partial F(x_1, \dots, x_n)}{\partial x_j} \pm \sum_{i=1}^m \lambda_i \frac{\partial r_i(x_1, \dots, x_n)}{\partial x_j} &\leq 0, \quad j = 1, \dots, n, \\ \lambda_1 r_1(x_1, \dots, x_n) = 0, \quad \dots, \quad \lambda_m r_m(x_1, \dots, x_n) &= 0. \end{aligned} \quad (3)$$

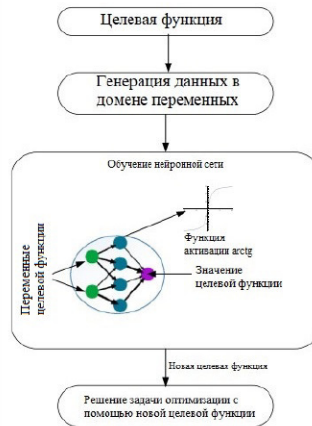


Рис. 1. Схема алгоритма аппроксимации целевой функции с помощью нейронной сети

В случае задачи (2) с ограничениями в форме равенств также можно использовать метод множителей Лагранжа, записать соответствующий функционал Лагранжа

$$L(x_1, \dots, x_n, \lambda_1, \dots, \lambda_m) = \sum_{j=1}^m \bar{w}_{j1} f\left(\sum_{i=1}^n w_{ij} x_i + \theta_j\right) + \bar{\theta}_1 + \sum_{i=1}^m \lambda_i r_i$$

и решить систему

$$\begin{cases} \frac{\partial L}{\partial x_i} = 0 & \forall i \in \{1, \dots, n\}, \\ \frac{\partial L}{\partial \lambda_j} = 0 & \forall j \in \{1, \dots, m\}. \end{cases} \quad (4)$$

Вид выражений для систем (3) и (4) подчеркивает важность использования функции активации $f(\cdot) = \text{arctg}(\cdot)$, чтобы упростить вычисление производных и искомого решения.

3. Примеры. Описанный метод опробуем на двух задачах. Программы написаны на Python, запускались на персональном компьютере, использовались библиотеки Tensorflow, Keras. В среднем на обучение затрачивалось 70 мс.

3.1. Первая задача. Рассматривается пример простой линейной задачи оптимизации

$$15x + 10y \rightarrow \max$$

с условиями

$$\begin{aligned} \frac{1}{3000}x + \frac{1}{6000}y &\leq \frac{2}{25}, \\ \frac{1}{3000}x + \frac{1}{2000}y &\leq \frac{1}{100}, \\ x &\geq 0, \quad y \geq 0. \end{aligned}$$

Реализована архитектура многослойного перцептрона с двумя нейронами во входном слое, пятью нейронами в скрытом слое, с шагом разбиения области параметров 0,125 и функцией активации $f(\cdot) = \arctg(\cdot)$. Результат, полученный точными методами:

$$x = 30, \quad y = 0, \quad \max(15x + 10y) = 450.$$

Решение, полученное с помощью нейронной сети:

$$x = 30, \quad y = 0, \quad \max(15x + 10y) = 447,9375.$$

Таким образом, получено хорошее приближение решения, которое можно улучшить, выбрав меньший шаг разбиения при генерации данных или увеличив количество эпох обучения. Увеличение количества эпох обучения, с одной стороны, улучшает качество аппроксимации целевой функции для данных по обучающей выборке, но, с другой стороны, из-за переобучения может привести к падению точности на этапе тестирования при вычислении экстремума аппроксимированной целевой функции.

3.2. Вторая задача. Описанный подход применен к построению регулятора в дискретной задаче линейного квадратичного оптимального управления (LQR) [5]. Рассмотрим систему вида

$$x_{k+1} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} x_k + \begin{pmatrix} 1 \\ 0 \end{pmatrix} u_k, \quad y_k = (1 \quad 0) x_k, \quad x_0 = (1 \quad 0)^T,$$

где минимизируется критерий оптимальности

$$J = x_N^T P x_N + \sum_{k=0}^{N-1} (x_k^T Q x_k + u_k^T R u_k).$$

В данном примере $N = 3$ – временной горизонт, $n = 2$ – размерность вектора состояний системы, $m = 1$ – размерность вектора управления, и весовые матрицы

$$Q = P = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad R = 0,3.$$

Задача LQR сводится к задаче оптимизации [5], и новая задача решается предложенным подходом. Для данного двумерного примера подход позволил построить оптимальное управление, утверждение об оптимальности найденного решения основано на сравнении с точным решением, которое получено решением уравнения Риккати.

Отмечается, что размерность исходной оптимизационной задачи, к которой сводится LQR, равна $(N + 1)n + Nm$, таким образом, размерность оптимизационной задачи растет с ростом числа уравнений, описывающих динамическую систему, что затрудняет применение предложенного подхода для многомерных задач LQR и задач с управлением на длинных временных горизонтах.

4. Заключение. В данной работе рассмотрен и реализован метод решения оптимизационных задач при помощи нейронных сетей. Главным достоинством данного подхода является упрощение решения задачи оптимизации со сложной целевой функцией с помощью аппроксимации нелинейной регрессией, используя обученные веса нейронной сети, с полиномиальной производной, которая позволяет проще решить задачу, применяя такие способы, как метод множителей Лагранжа и метод Куна – Таккера. Нейронная сеть обучается находить значение целевой функции для значений параметров на основе данных, полученных при помощи применения ограничений и подстановки значений параметров в исходную целевую функцию задачи оптимизации.

Недостатком данного подхода является возможная погрешность при численном нахождении значения целевой функции, которая появляется из-за погрешности обучения нейронной сети. Можно отметить, что для каждой задачи строится своя нейронная сеть, поэтому архитектура нейронной сети подбирается отдельно, и обучение происходит для каждой задачи также отдельно.

В статье [2] представлены результаты применения подхода на нескольких задачах оптимизации. Одна из этих задач была рассмотрена в данной работе, результат применения метода отличается от результата, полученного в [2], в лучшую сторону.

Вторым примером для применения метода стала дискретная задача оптимального управления (LQR). Стандартный метод решения таких задач – численное решение алгебраического уравнения Риккати. Это вычислительно трудозатратная процедура, особенно для задач большой размерности и с большим горизонтом, поэтому имеет смысл исследовать подход с многослойным перцептроном на маломерной задаче LQR, которая сводится к виду задачи оптимизации [5]. Сделан вывод, что такой подход не применим для задач LQR большей размерности и с длинным временным горизонтом.

Литература

1. Haykin S. Neural networks: a comprehensive foundation. Upper Saddle River: Prentice Hall PTR, 1998. 842 p.
2. Villarrubia G., De Paz J. F., Chamoso P., De la Prieta F. Artificial neural networks used in optimization problems // Neurocomputing. 2018. Vol. 272. P. 10–16.
3. Schmidt-Hieber J. The Kolmogorov–Arnold representation theorem revisited // Neural networks. 2021. Vol. 137. P. 119–126.
4. Kuhn H. W., Tucker A. W. Nonlinear programming // Giorgi G., Kjeldsen T. (eds). Traces and emergence of nonlinear programming. Basel: Birkhauser, 2014. P. 247–258.
5. Moore K. L., Naidu S. Linear quadratic regulation using neural networks // IJCNN-91-Seattle International Joint Conference on Neural Networks. 1991. IEEE. P. 735–739.

УДК 004.932.4

Бабкина А. В.

Очистка медицинских ультразвуковых изображений от спекл-шума

Рекомендовано к публикации доцентом Козыньченко В. А.

1. Введение. На сегодняшний день наблюдается стремительное увеличение количества исследований, связанных с нейросетевыми технологиями. Разработки, основанные на искусственном интеллекте и машинном обучении используются во многих областях, в том числе в медицине. Одной из задач данной области исследований является обработка медицинских изображений с целью очистки от шума, поиска медицинских артефактов, решения задач распознавания и т. д.

Ультразвуковое исследование (УЗИ) – одно из востребованных и наиболее безопасных способов исследования организма человека. В основе УЗИ лежит прием отраженных от внутренних органов ультразвуковых сигналов. Применение УЗИ дает возможность получить как 2D, так и 3D изображение с информацией о мягких тканях, состоянии сосудов, пульсации и скорости кровотока, движении плода в реальном времени. Однако полученные изображения содержат помехи в виде спекл-шума, возникающего вследствие интерференции отраженных от тканей когерентных эхо-сигналов, суммируемых со случайной фазой. Данный шум визуально представляет набор вытянутых пятен, образующих случайный узор, и является неотъемлемой частью ультразвуковых изображений. Спекл-шум затрудняет оценку итогового изображения, ухудшает контрастность, смазывает границы и скрывает медицинские артефакты, поэтому необходима разработка алгоритмов, способных снизить или вовсе устранить данный шум. Стоит отметить, что спекл-шум возникает не только на медицинских изображениях. Его можно обнаружить и на радиолокационных изображениях с синтезированной апертурой SAR (Synthetic Aperture Radar) [1], и в лазерных интерферометрических системах [2].

2. Модель спекл-шума. Датчик медицинского прибора сканирует

Бабкина Анна Владимировна – магистрант, Санкт-Петербургский государственный университет; e-mail: st076184@student.spbu.ru, тел.: +7(904)331-72-88

тело человека ультразвуковыми волнами с частотой от 2 до 20 МГц. Отражаясь от внутренних тканей, эхо-сигналы формируют изображение, которое в конечном итоге выводится на экран прибора. Из-за неровностей отражающей поверхности происходит суммирование когерентных сигналов с произвольным запаздыванием, что ведет к паразитной амплитудной модуляции и появлению пятен спекл-шума на итоговом изображении. Исходя из физической природы, спекл-шум преимущественно мультипликативен и является неотъемлемой частью ультразвукового сканирования. Пример изображения со спекл-шумом приведен на рис. 1.

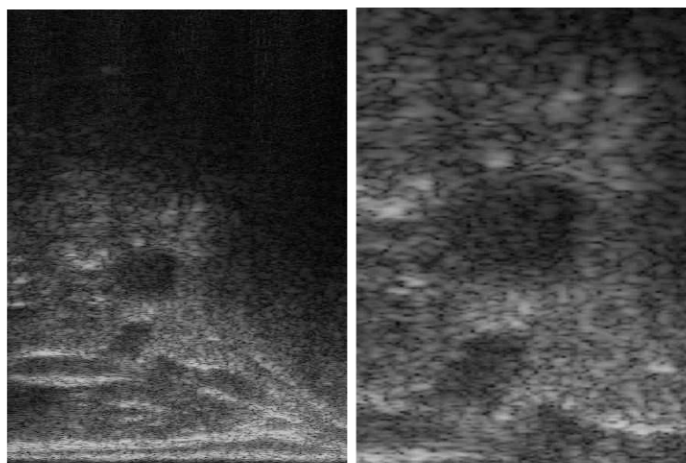


Рис. 1. Пример УЗИ-изображения со спекл-шумом (исходный и увеличенный вариант)

Представим изображение с мультипликативным шумом следующим образом: $g(x, y) = f(x, y)u(x, y) + w(x, y)$ [3], где $f(x, y)$ – исходное изображение (без шума); $u(x, y)$ – мультипликативная шумовая составляющая; $w(x, y)$ – аддитивная шумовая составляющая.

Поскольку на практике влияние аддитивной составляющей мало, можно рассматривать упрощенную модель: $g(x, y) = f(x, y)u(x, y)$.

Известны следующие варианты моделирования спекл-шума:

1. В виде мультипликативного шума с распределением Релея. Данная модель описана в статье [1].

2. В виде модели, соответствующей формуле:

$$g(x, y) = f(x, y) + f^\gamma(x, y)n(x, y), \quad (1)$$

где $n(x, y)$ – гауссова шумовая составляющая с нулевым средним, выбрано $\gamma = 0,5$. В статье [4] показано, что после обработки спекл-шума внутренними функциями аппарата УЗИ он перестает быть мультипликативным и может быть представлен формулой (1). Также отмечается гибкость модели и лучшее соответствие статистике изображения.

3. В виде статистической модели [5], полученной в рамках предположений о монохромности и идеальной поляризации лучей прибора.

Также существуют готовые функции, моделирующие спекл-шум. Например, в среде MATLAB функция *imnoise()*, добавляющая на изображение мультипликативный спекл-шум со стандартным отклонением, а также программные реализации, например, Field II Simulation Program, позволяющие моделировать наложение эхосигналов, полученных с использованием различных схем расположения отражающих поверхностей (см. рис. 2).

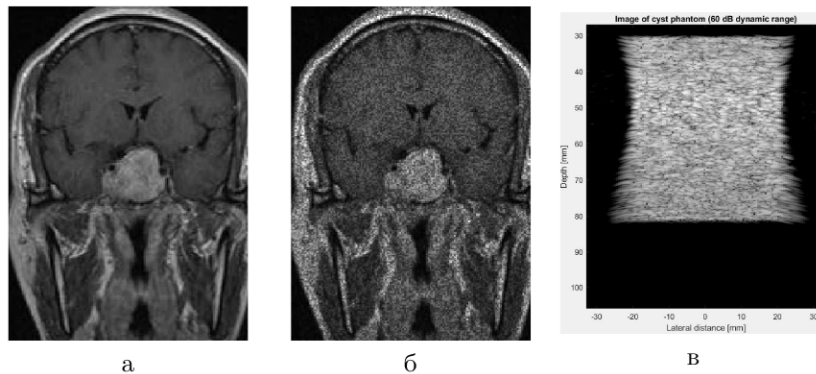


Рис. 2. Наложение спекл-шума в MATLAB: а) оригинальное изображение, б) функция *imnoise* (среднеквадратичное отклонение 0,05), в) программа Field II

3. Обучающие данные. Одной из основных трудностей применения нейронных сетей в задачах шумоочистки медицинских изображений является создание обучающих наборов данных. Проблема заключается в отсутствии образцовых изображений без шума. Для обучения необходимо иметь идентичные пары с шумом и без, а медицинские приборы, во-первых, не могут выдать изображение без спекл-шума, а во-вторых, они не позволяют одновременно записывать исходное шумовое и полученное после внутренних преобразований очищенное изображение. Возможны следующие способы создания обучающего набора данных:

1. Создание синтетического набора данных путем наложения спекл-шума на исходный набор не зашумленных изображений. Трудность данного подхода заключается в выборе наиболее реалистичной шумовой модели.
2. Записи изображений одних и тех же участков тканей с обработкой, без нее и дальнейшее совмещение полученных изображений по характерным точкам. Трудность данного подхода заключается в совмещении изображений, а также в наличии доступа к подходящим приборам.
3. Создание каким-либо способом очищенного изображения по зашумленному. Имеются в виду алгоритмы, дающие хороший результат шумоочистки, но не подходящие для промышленной реализации, например, из-за длительного времени обработки одного изображения.

Для дальнейшей работы выбран третий способ. Образцовые (очищенные) изображения были получены с помощью модели SCUNet [6]. В качестве входных данных использовался набор ультразвуковых изображений [7], состоящий из 520 снимков. В тренировочное множество вошли 350 изображений, в тестовое – 170. Пример пары исходного и очищенного изображений представлен на рис. 3.

4. Построение нейронной сети. Для первоначального решения задачи была выбрана архитектура автоэнкодера, представленная на рис. 4. На вход сети подавалось зашумленное изображение, увеличенное в два раза до размера (1008×752) , на выходе получалось изображение того же размера, после чего оно уменьшалось до исходного размера (504×376) . Сеть состояла из шести сверточных слоев, двух слоев Max pooling (mp) и двух слоев Upsampling (ups). В

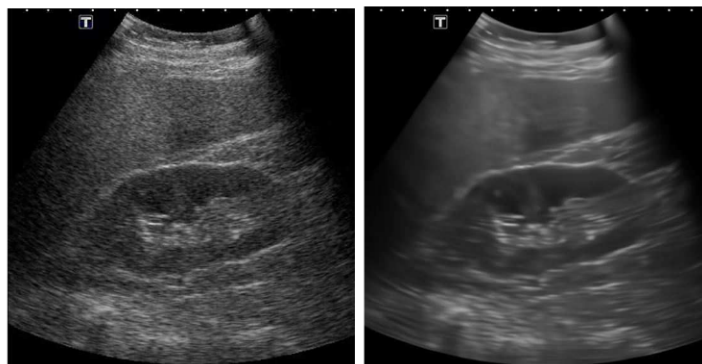


Рис. 3. Исходное УЗИ-изображение со спекл-шумом и очищенное с помощью SCUNet

качестве функции активации использовалась Leaky ReLU, как менее подверженная обнулению весов при обучении. Всего сеть содержала 743 тыс. параметров.

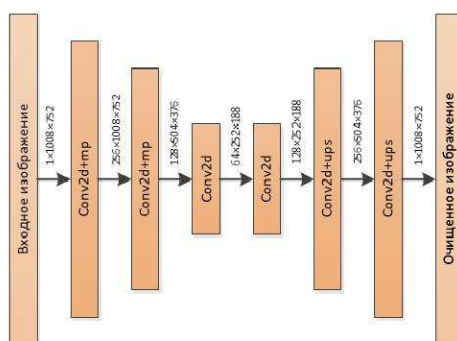


Рис. 4. Архитектура нейронной сети автоэнкодера

В качестве метрик качества использовались следующие [8]:

1. Среднеквадратичная ошибка $RMSE = \sqrt{\frac{\sum (g(i,j) - f(i,j))^2}{m \times n}}$, где $f(i, j)$ – изображение без шума, $g(i, j)$ – изображение с шумом, $m \times n$ – размер изображений, $i = 1, \dots, m, j = 1, \dots, n$.
2. Пиковое соотношение сигнал-шум $PSNR = 20 \log_{10}(\frac{I_{\max}}{RMSE})$, где I_{\max} – максимально возможное значение белого (например, для

8-битного изображения $I_{\max} = 255$). Чем больше значение, тем лучше произведена шумоочистка.

3. Индекс структурного сходства [9]

$$\text{SSIM} = \frac{(2\mu_f\mu_g + C_1)}{(\mu_f^2 + \mu_g^2 + C_1)} \frac{(2\sigma_{f,g} + C_2)}{(\sigma_f^2 + \sigma_g^2 + C_2)},$$

где μ – средняя интенсивность, σ^2 – дисперсия изображений, $\sigma_{f,g}$ – коэффициент ковариации, C_1, C_2 – малые константы, позволяющие избежать близкого к нулю значения в знаменателе ($C_i = (I_{\max} \cdot k_i)^2, k_1 = 0,01, k_2 = 0,03$). Чем ближе SSIM к единице, тем лучше сохранена структура исходного изображения.

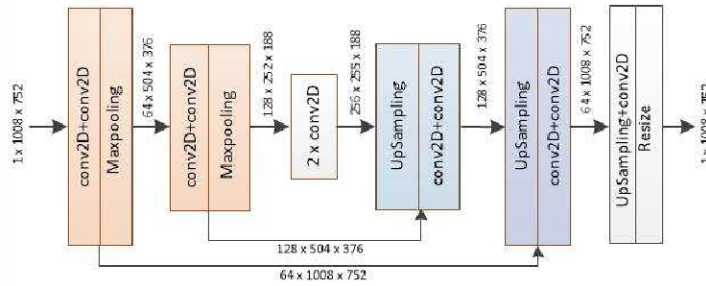


Рис. 5. Нейронная сеть на основе упрощенной модели UNet

Для сравнения использовалась вторая сеть – small UNet, построенная на основе упрощенной модели UNet [10]. Сеть состояла из последовательного набора блоков понижения и увеличения размерности, каждый из которых включает в себя два сверточных слоя и слой Max pooling или Upsampling, а также горизонтальные связи (всего в сети 11 сверточных слоев и 1,9 млн коэффициентов). Блок-схема сети

приведена на рис. 5. Выходы обеих сетей представлены на рис. 6.

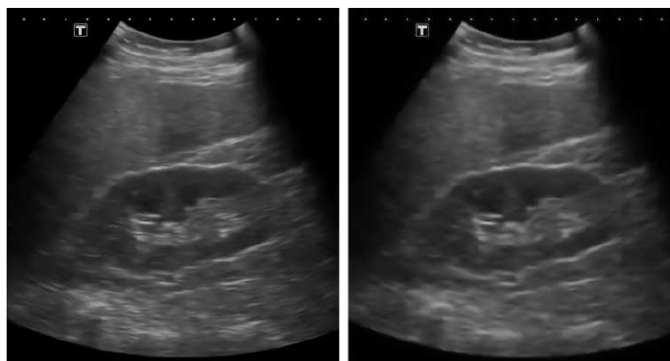


Рис. 6. Очистка от спекл-шума медицинского изображения с помощью двух моделей (слева – автоэнкодер, справа – small UNet)

В таблице 1 приведены результаты тестирования этих сетей. В качестве итоговых были взяты средние значения метрик по набору тестовых данных (170 изображений).

Таблица 1. Сравнение архитектур

	RMSE	PSNR	SSIM
без обработки	0,029	30,7	0,7712
autoencoder	0,023	32,8	0,9382
small UNet	0,010	39,8	0,9897

5. Заключение. В ходе эксперимента реализованы две нейронные сети с архитектурой автоэнкодера для борьбы со спекл-шумом. При сравнении выходов сетей (см. рис. 6 и таблицу 1) можно сделать вывод о том, что обе сети справляются с шумоподавлением. Однако при этом наблюдается некоторое снижение четкости контуров, что приводит к визуальному размытию изображения. В дальнейшем предполагается решить задачу подавления спекл-шумов с сохранением первоначальной четкости изображения. В зависимости от вычислительной мощности конечного устройства можно использовать одну из приведенных сетей при реализации прибора УЗИ.

Литература

1. Singh P., Shree R. Speckle noise: modelling and implementation // International Journal of Control Theory and Applications. 2016. Vol. 9. No 17. P. 8717–8727.
2. Васильев В. Н., Гуров И. П. Компьютерная обработка сигналов в приложении к интерферометрическим системам. СПб.: БХВ-Санкт-Петербург, 1998. 237 с.
3. Krithiga R., Lakshmi C. Reduction of speckle-noise in ultrasound images using MMM filter // International Journal of Pure and Applied Mathematics. 2017. Vol. 113. P. 84–95.
4. Coup P., Hellier P. et al. Nonlocal means-based speckle filtering for ultrasound images // IEEE transactions on image processing. 2009. Vol. 18. No 10. P. 2221–2229.
5. Goodman J. W. Some fundamental properties of speckle // JOSA. 1976. Vol. 66. No 11. P. 1145–1150.
6. Zhang K., Li Y., Liang J. et al. Practical blind image denoising via Swin-Conv-UNet and data synthesis // Machine Intelligence Research. 2023. Vol. 20. No 6. P. 822–836.
7. Breast Ultrasound Images Dataset [Электронный ресурс]: URL:<https://www.kaggle.com/datasets/aryashah2k/breast-ultrasound-images-dataset> (дата обращения 28.02.24).
8. Rosa R., Monteiro F. C. Performance analysis of speckle ultrasound image filtering // Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization. 2016. Vol. 4. No 3-4. P. 193–201.
9. Wang Z., Bovik A. C. et al. Image quality assessment: from error visibility to structural similarity // IEEE transactions on image processing. 2004. Vol. 13. No 4. P. 600–612.
10. Ronneberger O., Fischer P., Brox T. U-net: Convolutional networks for biomedical image segmentation // Medical Image Computing and Computer-Assisted Intervention–MICCAI: 18th International Conference. 2015. P. 234–241.

УДК 004.89

Горбунов В. И., Салимов Т. А.

Проблемы практического применения результатов Data Science и подходы к их решению

Рекомендовано к публикации доцентом Блекановым И. С.

1. Введение. Использование методов науки о данных, в особенности машинного обучения, плотно проникло в разные области исследований. Возможность обрабатывать большие объемы данных, извлекая из них новую информацию, позволяет проводить более масштабные исследования и за меньшее время. Однако уже давно идут разговоры о кризисе воспроизводимости, и предполагается, что методы, применяемые в науке о данных, позволят решить кризис воспроизводимости, либо снизить его влияние, так как используют современные инструменты, позволяющие сохранять намного больше информации, такие как использование программного языка для проведения вычислительных экспериментов и инструменты хранения и версионирования данных, программного кода и документации по проведенным исследованиям.

И, как будет показано в этой работе, кризис воспроизводимости с приходом методов науки о данных не стал меньшей проблемой. Будут обозначены ключевые проблемы и подняты вопросы, требующие решения.

2. Обзор литературы. Рассмотрим исследования последних лет, в которых поднимаются вопросы практического применения результатов, полученных методами науки о данных, а также проблемы воспроизводимости.

2.1. Опрос «Rexer 2023 Data Science Survey». Согласно этому опросу [1], в котором приняли участие 328 специалистов по данным из 49 стран, только 22% специалистов сообщают, что инициативы с использованием методов машинного обучения доходят до

Горбунов Владислав Игоревич – аспирант, Санкт-Петербургский государственный университет; e-mail: gorbunov.v93@gmail.com, тел.: +7(911)900-41-62

Салимов Тимур Альфредович – магистрант, Санкт-Петербургский государственный университет; e-mail: st054614@mail.ru, тел.: +7(951)671-75-96

интеграции полученных моделей. Примерно 43% специалистов высказались о том, что 80% решений и более не удалось развернуть. Из всех видов проектов машинного обучения, включая обновление моделей для существующих решений, только 32% специалистов сказали, что их модели обычно разворачиваются.

Выделяются следующие проблемы:

- отсутствие четко заданных целей проекта до старта работ. Только 34% специалистов по данным говорят, что цели проекта обычно хорошо определены до того, как проект начнется;
- недостаточная прозрачность в работе полученных моделей. Только 49% специалистов подтвердили, что лица, принимающие решение о развертывании модели, в целом достаточно хорошо осведомлены, чтобы принять такое решение.

Респондентам также было предложено описать причины такой ситуации в открытой форме. Чаще всего назывались следующие причины:

- заинтересованные стороны не доверяли работе моделей, считая что важнее личный опыт;
- неясная связь между полученными результатами и бизнес целями компании;
- проблемы коммуникации с другими командами, ответственными за внедрение.

В большей части специалисты выделяют организационный характер проблемы, а не технический.

2.2. Опрос «Algorithmia 2020 State of Enterprise Machine Learning». Согласно опросу, в котором приняли участие 745 компаний, использующих анализ данных и машинное обучение в своих решениях, основными проблемами, с которыми сотрудники сталкиваются при разработке решений на основе машинного обучения, являются: масштабирование, контроль версий, воспроизводимость модели и согласование применения моделей от заинтересованных сторон [2].

В 2020 году начал отслеживаться новый показатель – время, необходимое организации для развертывания одной модели машинного обучения. Около половины представителей опрошенных компаний утверждают, что тратят от 8 до 90 дней на развертывание одной модели. А 18% компаний тратят на развертывание более 90 дней, и в 5% случаев это может превышать год.

В независимости от отрасли и размера компании, 58% респондентов заявили, что усилия по машинному обучению успешны, если они обеспечивают рентабельность инвестиций, сокращают отток клиентов, помогают во внедрении продуктов и/или способствуют повышению качества бренда. Еще 58% респондентов заявили, что усилия по machine learning успешны, когда модель достигает удовлетворительных технических показателей. Важно отметить, что респондентам предлагалось выбрать более одного варианта ответа, потому в сумме может получиться больше 100%.

Таким образом, в данном опросе помимо технических проблем были также зафиксированы организационные трудности, связанные с интеграцией полученных при помощи анализа данных и машинного обучения решений. Вновь заметно расхождение в понимании ключевых целей и метрик успешности проекта.

2.3. Исследование воспроизводимости в лингвистике.

Помимо организационных проблем стоит также обратить внимание на технические проблемы, о которых в том числе говорили респонденты в рассмотренных выше опросах. Согласно исследованию, проведенному в 2008–2009 и затем повторно 2018–2019 годах, выявленный кризис воспроизводимости в области лингвистики не был преодолен [3]. Исследование проводилось на 600 случайно выбранных статьях из рецензируемых изданий, в которых использовались методы анализа данных.

По результатам данного исследования лишь 35% статей были опубликованы в открытом доступе, и при этом уровень предоставления материалов, данных и протоколов экспериментов были ниже 10%. Из всей выборки лишь 1% статей сообщал о воспроизведении результатов.

2.4. Вычислительная воспроизводимость Jupyter тетрадей. В исследовании, проведенном в 2022 году, проверяли воспроизводимость 8160 Jupyter тетрадей, с использованием языка программирования Python, из 1117 репозиториях GitHub, связанных с 1419 публикациями в рецензируемом издании PubMed Central [4].

Результаты показали, что из 8160 тетрадей лишь 4169 содержали указание о зависимостях. Удалось успешно установить зависимости в 2686 из них, а запустить тетради без ошибок только в 396 случаях. В итоге, лишь 245 Jupyter тетрадями дали результаты, идентичные тем, о которых сообщалось в оригинальном исследовании.

2.5. Исследование доступности данных. Целью исследования 2022 года стало изучение соблюдения авторами статей заявлений о готовности предоставить данные, использующиеся в исследовании [5]. Были отобраны 333 журнала с открытым доступом, в которых были требования к обязательной доступности данных.

Исследователи связались с авторами 3556 статей и попросили их предоставить данные для проверки их работы, в результате:

- 42% авторов указали, что готовы предоставить данные лишь по обоснованному запросу;
- среди 1792 статей, в которых авторы указали, что готовы поделиться данными, 1669 (93%) авторов либо не ответили, либо отказались;
- среди 254 (14%) из 1792 авторов, ответивших на запрос о предоставлении данных, только 123 (6,8%) предоставили запрошенные данные.

Несмотря на то, что научные журналы в последние года стали либо устанавливать обязательное требование о доступности данных, либо отдельно положительно отмечать статьи, в которых авторы выразили готовность предоставить данные, проблема воспроизводимости все также присутствует для большей части работ.

3. Обсуждение основных проблем. В первую очередь рассмотрим проблемы, связанные с воспроизводимостью. Кризис воспроизводимости не является проблемой только науки о данных, и, как было показано в пункте 2, применение инженерных практик из науки о данных не решило кризис воспроизводимости в других областях.

Так как есть путаница при применении термина «воспроизводимость», была предложена новая терминология, в которой выделяют три уровня воспроизводимости [6]:

- воспроизводимость методов – возможность получения одинаковых результатов, используя предоставленные данные и методы оригинального исследования;
- воспроизводимость результатов – возможность получения подтверждающих результатов в новом независимом исследовании с использованием тех же экспериментальных методов;
- воспроизводимость выводов – возможность получения качественно аналогичных выводов либо в результате независимого повторения исследования, либо в результате повторного анализа оригинального исследования.

Воспроизводимость методов – это только первый шаг к общей достоверности исследований. Например, исследование может обладать воспроизводимостью методов, но иметь аналитические проблемы, которые делают результаты недействительными. Если возможно воспроизвести исследование хотя бы на уровне методов, то, по крайней мере, эти аналитические проблемы можно будет выявить и исправить.

На данный момент отсутствуют единые стандарты и требования о воспроизводимости к исследованиям, использующим методы науки о данных в качестве инструмента анализа и экспериментирования, даже в научных изданиях. Необходимо в рамках каждой области проработать методики обеспечения воспроизводимости методов, а также для каждого исследования задавать погрешность при воспроизведении результатов [6].

Проблемы с воспроизводимостью в том числе влияют на другую группу проблем – недостаточная прозрачность процесса проведения исследований с использованием методов науки о данных. Невозможно говорить о прозрачности в исследованиях, результаты которых в большинстве случаев сложно, или даже невозможно повторить. Однако возможность обеспечения для исследований хотя бы воспроизводимости методов не является решением данной проблемы. Требуется выстроить четко установленный процесс, понятный всем заинтересованным лицам.

4. Потенциальные пути решения проблем. В существующих работах по естественным, социальным и формальным наукам невозможно выделить общие метрики и правила по обеспечению воспроизводимости методов. После изучения часто возникающих проблем, были выделены следующие ключевые аспекты, требующие улучшений, и подготовлены общие рекомендации для повышения уровня воспроизводимости методов в области науки о данных:

1. Версионирование и отслеживание изменений:
 - использование контроля версий кода, данных и отчетов об исследованиях;
 - сохранение точных версий всех использованных внешних зависимостей;
 - фиксация случайного зерна для алгоритмов, использующих случайность.
2. Автоматизация и устранение ручного управления:
 - устранение этапов ручного управления данными или рабочим процессом работы с данными.
3. Стандартизация хранения данных:
 - введение стандарта хранения промежуточных данных;
 - хранение всех промежуточных данных и вычислений для получения итогового результата по экспериментам.
4. Документирование и прозрачность:
 - фиксация в отчетах по экспериментам связи между инфографикой, данными и кодом для получения инфографики;
 - фиксация не только интерпретации наблюдения в текстовом виде, но и самого наблюдения.
5. Доступность данных и результатов:
 - обеспечение доступности результатов экспериментов, данных и исследований.

Процесс проведения исследований в каждой доменной области, где используются методы науки о данных, специализирован, поэтому невозможно выделить единый процесс работы, который обеспечит решение проблемы прозрачности. Для повышения прозрачности процессов можно адаптировать существующие методологии проведения исследований на данных в организациях, интегрируя элементы научного метода:

- характеристики предмета исследования – наблюдения, определения и измерения предмета исследований;
- гипотезы и теории – формулирование, сохранение и ранжирование;
- прогнозы – формулирование и сохранение;
- эксперименты – дизайн, процедура проведения и интерпретация результатов экспериментов.

Таким образом на основе элементов научного метода получится выделить каркас для дальнейшего совершенствования прозрачности процессов без необходимости создания единой методологии или отдельной проработки вопроса прозрачности в каждой доменной области.

5. Выводы. В данной работе проведено обзорное исследование проблем воспроизводимости в области науки о данных, где воспроизводимость является важным критерием достоверности результатов научной работы. Выделены ключевые проблемы, выраженные в недостаточной прозрачности процессов проведения исследований и отсутствии стандартов воспроизводимости методов. Представлен набор правил для повышения уровня воспроизводимости методов в исследованиях на данных. Представлен подход к повышению прозрачности процессов проведения исследований на основе адаптации методологий исследований с интеграцией элементов научного метода. Без решения данных проблем сложно говорить об успешном практическом применении результатов анализа данных и построения моделей машинного обучения для достижения бизнес-целей в организациях.

Литература

1. Rexer Analytics Data Science Survey 2023 [Электронный ресурс]: URL:<https://www.rexeranalytics.com/data-science-survey> (дата обращения: 15.03.2024).
2. Algorithmia 2020 state of enterprise machine learning [Электронный ресурс]: URL:https://cdn2.hubspot.net/hubfs/2631050/0284%20CDA0%20FS/Algorithmia_2020_State_of_Enterprise_ML.pdf (дата обращения: 15.03.2024).
3. Bochynska A., Keeble L., Halfacre C. et al. Reproducible research practices and transparency across linguistics // *Glossa Psycholinguistics*. 2023. Iss. 2(1). P. 1–36.
4. Samuel S., Mietchen D. Computational reproducibility of Jupyter notebooks from biomedical publications [Электронный ресурс]: URL:<https://arxiv.org/abs/2209.04308> (дата обращения: 15.03.2024).
5. Gabelica M., Bojic R., Puljak L. Many researchers were not compliant with their published data sharing statement: a mixed-methods study // *Journal of Clinical Epidemiology*. 2022. Vol. 150. P. 33–41.
6. Goodman S. N., Fanelli D., Ioannidis J. P. A. What does research reproducibility mean? // *Science Translational Medicine*. 2016. Vol. 8. No 341. P. 341–353.

УДК 51-37

Гусева М. Ю.

Расчет аналитического решения трехмерной системы обыкновенных дифференциальных уравнений с релейным гистерезисом и возмущением

Рекомендовано к публикации доцентом Евстафьевой В. В.

Введение. Исследуется трехмерная система обыкновенных дифференциальных уравнений, правая часть которой содержит постоянную матрицу, релейную гистерезисную нелинейность с параметром и периодическую функцию возмущения. В статье [1] такая система рассмотрена с рассчитанными значениями параметров. Найдены условия, при которых существует периодическое решение с двумя точками переключения за период и фазовой траекторией, склеенной из двух кусков в точках переключения. Получены формулы для расчета параметров системы и решения. Такое решение представляется важным в теории автоматических систем [2].

Цель работы – численная реализация теоретических результатов, полученных в [1], и проведение математического моделирования на основе вычислительных экспериментов. Настоящая статья является продолжением работы [3].

1. Постановка задачи. В качестве модели автоматической релейной системы в [1] предложена следующая трехмерная система обыкновенных дифференциальных уравнений:

$$\dot{X} = A_0 X + B_0 F(\sigma) + K_0 f(t), \quad \sigma = (\Gamma, X), \quad (1)$$

где $X = (x_1, x_2, x_3)^T$ – вектор состояний системы; A_0 – диагональная матрица системы с собственными числами $\lambda_1 = 0, 1$, $\lambda_2 = -1$ и $\lambda_3 = -2$; $B_0 = (b_1^0, b_2^0, b_3^0)^T = (1, 1, 1)^T$ – вещественный вектор; $F(\sigma)$ – нелинейность, заданная следующим образом: $F(\sigma) = m_1 = -1, 11$

Гусева Мария Юрьевна – студент, Санкт-Петербургский государственный университет; e-mail: st089778@student.spbu.ru, тел.: +7(921)899-97-62

Исследование выполнено за счет гранта Российского научного фонда № 23-21-00069, <https://rscf.ru/project/23-21-00069/>

при $\sigma < \ell_2$ (ℓ_2 – параметр) и $F(\sigma) = m_2 = 5,30$ при $\sigma > \ell_1 = 0,75$; $K_0 = (k_1^0, k_2^0, k_3^0)^T = (-2; 1; 0,5)^T$ – вещественный вектор; $f(t) = f_0 + f_1 \sin(\omega t + \varphi_1) + f_2 \sin(2\omega t + \varphi_2) = 1 + 2 \sin(t + \pi/3) + 5 \sin(2t)$ – функция возмущения с периодом $T = 2\pi$; $\Gamma = (\gamma_1, 0, 0)^T = (-0,14; 0; 0)^T$ – вектор обратной связи.

Под *точкой переключения* понимается состояние системы, при котором σ принимает пороговое значение ℓ_1 или ℓ_2 , а релейная нелинейность $F(\sigma)$ при этом меняет значение выхода m_2 на m_1 , или наоборот. Обозначим через L_μ , $\mu = 1, 2$, *гиперплоскость переключения* [1], заданная уравнением $\sigma = \gamma_1 x_1 = \ell_\mu$. В точках переключения на этих гиперплоскостях происходит «сшивание» фазовых траекторий, заданных системой (1) при $F(\sigma) = m_1$ и $F(\sigma) = m_2$.

Решением системы (1) является непрерывная T -периодическая вектор-функция для $t \geq 0$ в форме Коши

$$X(t) = e^{A_0(t-t_0)} X(t_0) + \int_{t_0}^t e^{-A_0(\tau-t)} (B_0 m_\mu + K_0 f(\tau)) d\tau, \quad (2)$$

где t_0 – начальный момент времени. Поведение изображающей точки решения следующее. Движение начинается в точке $X^1 = X(t_0)$ на L_1 в момент времени $t_0 = 0$ и в момент времени t_1 попадает на L_2 в точку $X^2 = X(t_1)$ в силу системы (1) при $F(\sigma) = m_1$. Затем возвращается в момент времени T в точку $X^1 = X(T)$ в силу системы (1) при $F(\sigma) = m_2$. Под параметрами решения понимаем точки переключения $X^1 = (x_1^1, x_2^1, x_3^1)^T \in L_1$, $X^2 = (x_1^2, x_2^2, x_3^2)^T \in L_2$ и моменты переключения t_1, T .

Поставим задачу: разработать программу для расчета аналитического решения системы (1) и построения его траектории.

2. Теоретические результаты. Приведем формулы из [1]

$$t_1 = T + \frac{1}{\lambda_1} \ln \frac{m_2 - m_1}{(e^{\lambda_1 T} - 1)L + m_2 - m_1 e^{\lambda_1 T}}, \quad (3)$$

$$\ell_2 = (\ell_1 + H) e^{\lambda_1 t_1} + \lambda_1^{-1} \gamma_1 (m_1 + k_1^0 f_0) (e^{\lambda_1 t_1} - 1) - H(\omega t_1), \quad (4)$$

$$x_1^\mu = \ell_\mu / \gamma_1, \quad \mu = 1, 2, \quad x_j^1 = \frac{e^{\lambda_j T}}{1 - e^{\lambda_j T}} \left(m_1 \int_0^{t_1} e^{-\lambda_j \tau} d\tau + m_2 \int_{t_1}^T e^{-\lambda_j \tau} d\tau + k_j^0 \int_0^T e^{-\lambda_j \tau} f(\tau) d\tau \right),$$

$$x_j^2 = \frac{e^{\lambda_j t_1}}{1 - e^{\lambda_j T}} \left(\int_{t_1}^T e^{\lambda_j(T-\tau)} (m_2 + k_j^0 f(\tau)) d\tau + \int_0^{t_1} e^{-\lambda_j \tau} (m_1 + k_j^0 f(\tau)) d\tau \right), \quad j = 2, 3, \quad (5)$$

$$H(\omega t) = \gamma_1 k_1^0 \left(\frac{f_1 \sin(\omega t + \varphi_1 + \delta_1)}{\sqrt{\lambda_1^2 + \omega^2}} + \frac{f_2 \sin(2\omega t + \varphi_2 + \delta_2)}{\sqrt{\lambda_1^2 + 4\omega^2}} \right),$$

$$L = - \left(\frac{\lambda_1 \ell_1}{\gamma_1} + k_1^0 f_0 \right) + \frac{\lambda_1 (H(\omega T) - H e^{\lambda_1 T})}{\gamma_1 (e^{\lambda_1 T} - 1)}, \quad H = H(0),$$

$$\delta_\xi = \text{arctg}(\xi\omega/\lambda_1), \quad \xi = 1, 2, \quad t \geq 0.$$

Если в некоторый момент t_0 изображающая точка решения принадлежит гиперплоскости L_μ , $\mu = 1, 2$, то наименьший момент времени $t^\mu > t_0$, в который изображающая точка принадлежит $L_{3-\mu}$, называется *моментом первой встречи изображающей точки с гиперплоскостью* $L_{3-\mu}$ [1]. Выпишем условия переключения с $t^1 = t_1$ и $t^2 = T$

$$\ell_2 - \gamma_1 x_1(t) > 0, \quad t \in [0, t_1), \quad \ell_2 - \gamma_1 x_1(t_1) = 0, \quad (6)$$

$$\ell_1 - \gamma_1 x_1(t) < 0, \quad t \in [t_1, T), \quad \ell_1 - \gamma_1 x_1(T) = 0. \quad (7)$$

3. Расчет значений параметров и аналитического решения. Разработан программный модуль для расчета значений параметров в формате двойной точности (16 цифр). Ниже представлены результаты расчета по формулам (3)–(5) в формате с тремя десятичными знаками без округления. Итак, $\ell_2 = 3,998$, $t_1 = 1,601$, $T = 6,283$, $X^1 = (-5,357; 4,618; 2,521)^T$, $X^2 = (-28,563; 4,462; 0,776)^T$.

Программный модуль также предусматривает расчет траектории по формуле (2) при $m_\mu = m_1$ с начальной точкой $X(0) = X^1$ на интервале $[0, t_1)$ и при $m_\mu = m_2$ с начальной точкой $X(t_1) = X^2$ на интервале $[t_1, T)$.

4. Результаты моделирования траекторий. Разработан программный модуль для проверки условий переключения с $t^1 = t_1$,

$t^2 = T$ и построения траекторий, соответствующих работе автоматической системы. Проведены вычислительные эксперименты. Результаты двух экспериментов представлены на рис. 1, 2 и в таблице, где $\Delta_\mu = \ell_\mu - \gamma_1 x_1(t)$, $\mu = 1, 2$.

Эксперимент 1 с шагом $h = 0,1$ и точностью до 10^{-3} : условия переключения (6), (7) выполняются. Параметры решения имеют значения, приведенные в пункте 3. Фазовая траектория замкнутая. Результаты эксперимента согласуются с результатами работы [1].

Таблица. Анализ выполнения условий переключения

t	$\gamma_1 x_1(t)$	$\Delta_2(t)$	t	$\gamma_1 x_1(t)$	$\Delta_1(t)$
$h = 0,1 \quad \Delta_\mu(t) < 10^{-3}$					
0,000	0,750	3,248	1,601	3,998	-3,248
0,100	0,865	3,133	1,701	3,994	-3,244
...
1,600	3,997	0,001	6,201	0,752	-0,002
1,601	3,998	0,000	6,283	0,750	0,000
$h = 0,001 \quad \Delta_\mu(t) < 10^{-4}$					
0,0000	0,7500	3,2489	1,6013	3,9989	-3,2489
...
1,6010	3,9985	0,0003	6,2183	0,7501	-0,0001
1,6013	3,9989	0,0000	6,2193	0,7500	-0,0000
		
			6,2823	0,7499	0,0000
			6,2831	0,7500	0,0000

Эксперимент 2 с шагом $h = 0,001$ и точностью до 10^{-4} : условие переключения (6) выполняется, (7) нет. Параметры решения имеют следующие значения: $t^1 = t_1 = 1,6013$, $t^2 = 6,2193 < T = 6,2831$. В этом эксперименте значения представлены с четырьмя десятичными знаками без округления. Таким образом, X^1 совпадает со значением, приведенным в пункте 3, а X^2 не достигается. Фазовая траектория не является замкнутой, так как изображающая точка решения возвращается на гиперплоскость L_1 в точку $X(t^2) = (-5,3571; 4,4176; 2,4229)^T$, находящуюся от начальной точки $X(0)$ на расстоянии, равном 0,0505 (см. рис. 2). Таким образом, найденные значения параметров не соответствуют траектории решения, которая построена с точностью выше 10^{-3} .

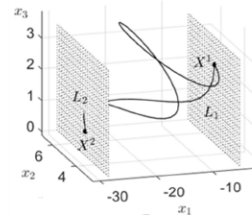


Рис. 1. Фазовая траектория (эксперимент 1)

Закключение. На основе теоретических результатов из [1] разработана программа в среде компьютерной алгебры для расчета параметров системы и решения, расчета траектории аналитического решения и проверки условий переключения с заданной точностью.

Проведены вычислительные эксперименты по моделированию режима работы автоматической релейной системы с переключениями в момент достижения пороговых значений. Эксперименты показали, что в автоматической системе существует режим T – периодических колебаний при регистрации значений с шагом $h = 0,1$ и точностью измерений до 10^{-3} . При уменьшении шага до $h = 0,001$ и увеличении точности до 10^{-4} второй момент переключения наступает раньше, отчего нарушается замкнутость фазовой траектории. Для обеспечения работы в этом режиме следует изменить значения параметров системы, которые в условиях задачи подлежат настройке.

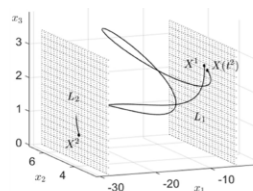


Рис. 2. Фазовая траектория (эксперимент 2)

Литература

1. Евстафьева В. В. О существовании двухточечно-колебательных решений возмущенной релейной системы с гистерезисом // Дифференциальные уравнения. 2021. Т. 57. № 2. С. 169–178.
2. Евстафьева В. В., Камачкин А. М., Потапов Д. К. Периодические режимы в системе автоматического управления с трехпозиционным гистерезисным реле // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2022. Т. 18. № 4. С. 596–607.
3. Евстафьева В. В., Гусева М. Ю. Математическое моделирование периодического режима в неавтономной системе с неидеальным реле // Математическая физика, механика и их приложения: материалы Международной научной конференции, посвященной 75-летию академика АН РБ В. Ш. Шагапова (г. Стерлитамак, 21–24 июня 2023 г.) / отв. ред. К. Б. Сабитов. Уфа: Уфимский университет науки и технологий, 2023. С. 206–208.

УДК 519.652

Ковалев А. Ю., Калинин П. Е.

Альтернативный алгоритм построения шума Перлина

Рекомендовано к публикации доцентом Лежниной Е. А.

Введение. Процедурная генерация – это способ автоматического создания контента с помощью различных вычислительных алгоритмов. Данный способ на основе заранее заданных параметров позволяет создавать различные модели, объекты или графику: ландшафт, текстуры, эффекты и т. п.

Основным преимуществом процедурной генерации является создание множества уникального контента при низких трудозатратах. Однако в сравнении с ручным созданием контента недостатком может служить низкое качество или нереалистичный вид получаемого контента.

На сегодняшний день существует несколько наиболее популярных алгоритмов процедурного создания карт высот: шум Перлина, симплекс шум, diamond-square, шум Ворлея. В данной работе предлагается альтернативный вариант построения шумовой карты на основе шума Перлина.

Обзор методов генерации шума и их модификаций. Одним из самых популярных методов генерации численного шума для процедурной генерации шумовых карт является шум Перлина. В 2002 году, Кен Перлин представил модификацию собственного шума, которая используется и по сей день [1]. Подробное описание самого алгоритма изложено в статье [2].

Альтернативным и не менее популярным алгоритмом генерации численного шума является алгоритм diamond-square [3].

В работе [4] представлены процедурные методы генерации ландшафтов, делается акцент на преимуществах использования процедурных методов:

Ковалев Александр Юрьевич – студент, Санкт-Петербургский государственный университет; e-mail: Fullerene@yandex.ru, тел.: +7(922)044-36-75

Калинин Павел Евгеньевич – студент, Санкт-Петербургский государственный университет; e-mail: st087365@student.spbu.ru, тел.: +7(981)958-75-13

- значительное снижение трудозатрат в сравнении с классическими «ручными» методами;
- возможность быстро создавать различные изображения, указывая лишь несколько входных параметров для работы алгоритма;
- возможность видоизменения и усовершенствования алгоритма для получения принципиально других шумовых картин.

Во многих работах, например [2, 4, 5], авторы отмечают особый потенциал развития области процедурной генерации и указывают на возможность модификаций некоторых алгоритмов для увеличения их реалистичности и разнообразия.

Генерация шума Перлина на четырех узловых точках.

Классический алгоритм генерации двумерного шума, предложенный К. Перлином, включает в себя несколько этапов:

1. Разбиение координатной плоскости на единичные квадраты.
2. Генерация в узлах полученной сетки случайных двумерных единичных векторов.
3. Последовательный перебор всех квадратов сетки с перебором всех точек внутри квадрата.
4. Для точки (x, y) внутри выбранного квадрата строится четыре скалярных произведения (между случайными узловыми векторами и векторами, направленными от вершин квадрата к фиксированной точке).
5. Для скалярных произведений, вычисленных для пары нижних узлов и пары верхних узлов квадрата, проводится линейная интерполяция [6] по координате x выбранной точки.
6. Для полученной пары значений проводится линейная интерполяция по координате y выбранной точки.

Полученное значение – это значение шума в точке с координатами (x, y) .

Генерация шума Перлина на множестве узловых точек.

Классический алгоритм основан на использовании четырех узлов на одном квадрате сетки, т. е. на стороне квадрата находятся две узловые точки. Рассмотрим n равноудаленных узловых точек на стороне квадрата, тогда i -я точка имеет координату $\frac{i-1}{n-1}$ (см. рис. 1).

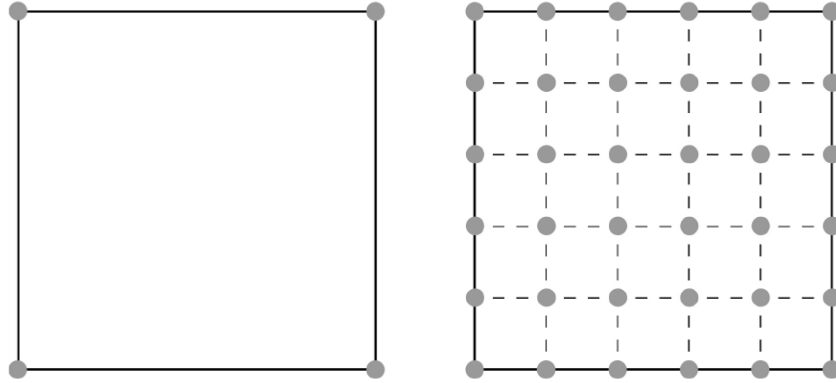


Рис. 1. Пример добавления узловых точек на квадрат сетки

Тогда пункты 5 и 6 из классического алгоритма видоизменяются: линейная интерполяция уже не может быть использована, поскольку в общем случае нельзя провести прямую, проходящую через n точек при $n > 2$. Рассмотрим два подхода:

- построение интерполяционного полинома, проходящего через все n точек, в форме Лагранжа или в форме Ньютона;
- интерполяция полученных значений скалярных произведений при помощи сплайнов.

Тогда интерполяция скалярных произведений, соответствующих узловым точкам, лежащим на i -й горизонтали, в случае использования полинома Лагранжа будет проводиться по формуле

$$L_{n-1}^i(x) = \sum_{k=1}^n l_k(x) f(x_k),$$

где $f(x_k)$ – значение скалярного произведения, соответствующего k -му узлу на i -й горизонтали, а множитель Лагранжа

$$l_k(x) = \frac{(x - x_1) \dots (x - x_{k-1})(x - x_{k+1}) \dots (x - x_n)}{(x_k - x_1) \dots (x_k - x_{k-1})(x_k - x_{k+1}) \dots (x_k - x_n)}.$$

Далее, после интерполяции по всем горизонталям проводится интерполяция по вертикали для получения значения шума в точке с

координатами (x, y) :

$$F(x, y) = \sum_{k=1}^n l_k(y) L_{n-1}^k(x).$$

В реализации второго подхода использовались кубические сплайны второй гладкости – $s_{3,2}(x)$. На рис. 2 показаны результаты сравнения генерации шума на четырех узлах и на множестве узлов.

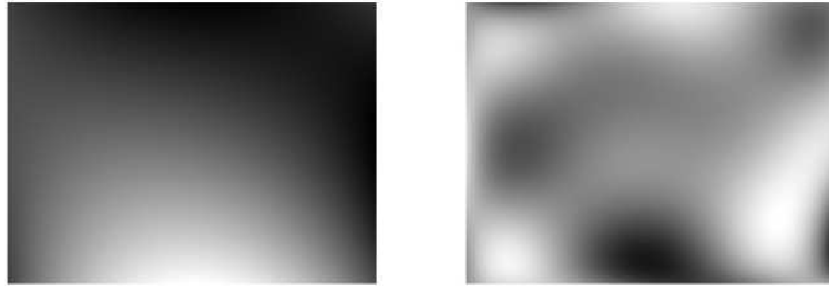


Рис. 2. Сравнение шума на четырех узлах и на множестве узлов

К недостаткам классического шума Перлина относится однообразие получаемых картин: высоты и низины шумовых значений однотипны и похожи друг на друга. В любом из двух предлагаемых подходов похожие части шумовой карты встречаются намного реже. Преимуществом предлагаемых подходов является возможность генерации карт с более сложным рельефом при сохранении гладкости шумовой функции.

Однако предлагаемый метод обладает и несколькими недостатками. Во-первых, из-за нелинейной интерполяции скалярных произведений теряется возможность динамической генерации шума – добавления новых квадратов в сетку. Во-вторых, экспериментальным путем было установлено, что при использовании большого количества узловых точек, получаемая шумовая карта имеет множество значений, близких к нулю. Это происходит по причине интерполяции значений через множество узловых точек, имеющих нулевые скалярные произведения.

Закключение. Рассмотрены альтернативные подходы к построению шума Перлина, приведены результаты генерации шумовых карт

на множестве узловых точек, проведено сравнение используемых алгоритмов с классическим способом. Предложенный метод может быть использован для создания небольших карт высот с гладкими элементами сложной формы.

Литература

1. Perlin K. An image synthesizer // SIGGRAPH. 1985. Vol. 19. No 3. P. 287–296.
2. Барсуков А. В. Исследование популярных алгоритмов процедурной генерации ландшафтов на возможность усовершенствования // Альманах научных работ молодых учёных университета ИТМО. Материалы XLIX научной и учебно-методической конференции. 2020. Т. 3. С. 30–35.
3. Fournier A., Fussell D., Carpenter L. Computer rendering of stochastic models // SIGGRAPH. 1982. Vol. 25. No 6. P. 371–384.
4. Воронкина А. Ю., Шляга В. С. Моделирование процедурных ландшафтов // Вестник МГУП имени Ивана Федорова. 2015. № 1. С. 34–42.
5. Дрешпан Н. В., Шепляков И. А. Моделирование случайного ландшафта // Международная научно-техническая конференция молодых учёных БГТУ им. В.Г. Шухова, посвящённая 300-летию Российской Академии Наук. 2022. Т. 13. С. 130–134.
6. Утешев А. Ю., Боровой И. И. Решение задачи рациональной интерполяции с использованием ганкелевых полиномов // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2016. Т. 12. № 4 С. 31–43.

УДК 004

Колесов В. В.

Алгоритмы для классификации жанров фильма

Рекомендовано к публикации доцентом Блекановым И. С.

1. Введение. Задача определения жанров фильма довольно нетривиальна. У фильма может быть более одного жанра. Также может возникнуть ситуация, в которой помимо данных о сюжете фильма известен ещё и год выпуска, режиссер и другие признаки. По сравнению с более ранними статьями [1, 2], уже существуют более современные методы [3, 4] с более высоким качеством в задаче текста. Здесь будут описаны как те методы, что встречались в предыдущих работах, так и те, что ещё не применяли.

2. Постановка задачи. Задача данной работы заключается в классификации жанров фильмов. Постановка задачи классификации заключается в том, чтобы, имея обучающую выборку, получить категориальный ответ, количество которых ограничено. Обучающая выборка – это некие исходные данные, прямые результаты измерений, которые обычно представляются набором некоторых чисел. Ответ алгоритма также часто называют целевым вектором.

Пусть имеется обучающая выборка $X = \{x_i\}_{i=1}^n$, что характеризуется вектором вещественных признаков $x_i = (x_{i,1}, \dots, x_{i,d})$. В качестве ответа x фигурирует переменная, принимающая конечное множество значений из множества $T = \{1, \dots, l\}$. Задача заключается в том, чтобы построить алгоритм (классификатор), который по вектору признаков x вернул бы метку класса $t \in T$ или вектор оценок принадлежности к каждому из классов $\{p(s|x)_{s=1}^l\}$.

3. Алгоритмы для классификации сюжета фильма. Рассмотрим сначала алгоритмы, которые можно использовать для классификации сюжета.

3.1. Наивный байесовский классификатор. Этот подход основан на двух ключевых предположениях: порядок слов не имеет значения и предположение об условной независимости признаков.

Колесов Владислав Владимирович – магистрант, Санкт-Петербургский государственный университет; e-mail: st068048@student.spbu.ru, тел.: +7(960)892-93-74

Для документа $d = w_1 \dots w_N$, где w_i – слово в документе, N – размер словаря, и класса $c \in C$, вероятности $P(w_i|c)$ – независимы. Это означает, что

$$P(d|c) = P(w_1, w_2, w_n|c) = \prod_{i=1}^N P(w_i|c). \quad (1)$$

Таким образом, уравнение для класса, выбранного наивным байесовским классификатором, имеет вид

$$c_N B = \arg \max_{c \in C} \frac{p(c) \prod_{i=1}^N P(w_i|c)}{P(d)} = \arg \max_{c \in C} p(c) \prod_{i=1}^N P(w_i|c).$$

В данной работе рассматривается подход, заключающийся в построении модели полиномиального наивного байесовского классификатора (mNB). Для каждого текста сюжета d , mNB может оценить апостериорную вероятность $P(c|d)$ в правой части уравнения (1). Модель предсказывает, что фильм относится к жанру c , если $P(c|d) > P(c)$, где $P(c)$ – априорная вероятность жанра c . Однако в ситуации, когда слово w_i не содержится в обучающей выборке, вероятность $P(w_i|c)$ не задана для любого $c \in C$, чего хотелось бы избежать. Для решения этой проблемы используют сглаживание, основная идея которого – увеличить вероятности всех биграмм в уравнении на некоторую величину α , чтобы вероятности стали ненулевыми. Таким образом, для слова w^* , не содержавшегося в обучающей выборке, можно показать, используя теорему Байеса, что значение

$$P(w^*|c) = \frac{P(c)P(c|w^*)}{P(w^*)}$$

не определено, так как $P(w^*) = 0$. После сглаживания получим:

$$P(w^*|c) = \frac{P(c)P(c|w^*) + \alpha}{P(w^*) + \alpha N},$$

где α является гиперпараметром модели.

3.2. Метод SVM с применением TF-IDF для векторизации текста. Одна из проблем с наивным байесовским классификатором заключается в том, что функция набора слов, а также предположение об условной независимости нереалистичны и могут

не отражать смысл предложений в тексте. Поэтому в следующем алгоритме используются представления слов с помощью TF-IDF (TF – term frequency, IDF – inverse document frequency). Самый простой способ закодировать текст в вектор – использовать словарь, состоящий из всех уникальных слов в рассматриваемом корпусе текстов, на позиции, соответствующей номеру слова в словаре, задать значение, равное количеству таких слов в отдельном тексте. Однако у такого метода есть недостаток: любые слова, часто встречающиеся во всех рассматриваемых текстах будут иметь высокую частоту, но не нести информации. Метод TF-IDF решает эту проблему.

В нём считается отношение частоты употребления слова в отдельном тексте к частоте употребления слова во всех документах. Таким образом, большой вес получают слова с высокой частотой в пределах конкретного документа и с низкой частотой употреблений в других документах. Вектор в данном случае формируется по такому же принципу, за исключением того, что на позиции, соответствующей номеру слова будет значение, равное весу TF-IDF. Полученное векторное представление сюжета может быть использовано алгоритмами машинного обучения, один из которых – метод опорных векторов.

Метод опорных векторов (SVM) – линейная модель для задач классификации и регрессии. Идея алгоритма состоит в создании линии или гиперплоскости, которая разделяет данные на классы. Предположим, что каждая заданная точка данных x_i (в нашем случае – вектор значений TF-IDF весов описания сюжета i -го фильма) принадлежит к одному из двух классов, 1 или -1 . Точка данных рассматривается как p -мерный вектор, и требуется выяснить, можем ли мы разделить такие точки с помощью $(p - 1)$ -мерной гиперплоскости. Есть много гиперплоскостей, которые могут классифицировать данные. Разумный выбор в качестве лучшей гиперплоскости – та, которая представляет наибольшее расстояние или границу между двумя классами. Поэтому выберем гиперплоскость так, чтобы расстояние от нее до ближайшей точки данных с каждой стороны было максимальным.

Если такая гиперплоскость существует, то множества точек классов линейно разделимы и задача оптимизации имеет вид

$$\begin{cases} \|w\|^2 \rightarrow \min, \\ c_i(wx_i - b) \geq 1, \quad 1 \leq i \leq n, \end{cases}$$

где c_i принимает значение 1 или -1 , в зависимости от того, какому классу принадлежит точка x_i . Уравнение разделяющей гиперплоскости имеет вид $wx - b = 0$.

Однако такой гиперплоскости может не существовать (случаем линейной неразделимости). Чтобы алгоритм мог работать в данном случае, позволим ему допускать ошибки на обучающей выборке. Введём набор дополнительных переменных ξ_i , характеризующих величину ошибки на объектах x_i :

$$\begin{cases} \frac{1}{2}\|w\|^2 + C \sum_{i=1}^n \xi_i \rightarrow \min, \\ c_i(wx_i - b) \geq 1 - \xi_i, \quad 1 \leq i \leq n, \\ \xi_i \geq 0, \quad 1 \leq i \leq n, \end{cases}$$

где коэффициент C — параметр настройки метода, который позволяет регулировать отношение между максимизацией ширины разделяющей полосы и минимизацией суммарной ошибки.

3.3. Нейронные сети класса трансформеров (BERT, RoBERTa, XLNet и прочие). На сегодняшний момент в числе state-of-the-art моделей по решению задачи классификации текста с наилучшими векторными представлениями есть множество моделей, основанных на «механизме внимания». Опишем лишь один из основных, часто используемых, и на основе архитектуры которого построено большинство остальных алгоритмов класса трансформеров — BERT.

BERT относится к классу трансформеров, т. е. это модель глубокого обучения, которая использует механизм самоконтроля, дифференциально взвешивая значимость (attention) каждой части входных токенов.

Для любой задачи распознавания языка, в том числе классификации текста, обучение нейронной сети состоит из двух шагов. Сначала, в стадии так называемого «предобучения» BERT обучается на решении задачи моделирования замаскированного языка и задачи предсказания следующего предложения.

Следующий шаг — этап «точной настройки». Добавление одного слоя позволяет использовать нейросеть для конкретной задачи. Для классификации текста — это обычный линейный слой размером, совпадающим с количеством классов. После применяется сигмоидальная функция для каждого выхода, в результате получаем «уверенность» модели в принадлежности к каждому из классов.

4. Алгоритмы для классификации признаков директора фильма, актеров и других. В нашей задаче также есть признаки, которые не являются текстовыми, относятся к категориальным или числовым: например, директор фильма или год выхода фильма. В этом случае подойдут следующие алгоритмы.

4.1. Алгоритм Random forest. Это ансамблевый метод обучения для классификации и регрессии, который работает путем построения множества деревьев решений во время обучения. Для задач классификации выход случайного леса – класс, выбранный большинством деревьев.

Деревья в лесу строятся независимо по следующей схеме:

1. Выбирается подвыборка размера N и по ней строится дерево, причем у каждого дерева своя подвыборка.
2. Для построения расщепления в каждом дереве просматриваем свое множество признаков (для каждого нового разбиения свои случайные признаки).
3. По заранее заданному критерию выбираем наилучший признак и разбиение по нему. Дерево строится до исчерпания выборки (пока в листьях не останутся представители одного класса), или пока не нарушается ограничение высоты дерева и/или числа объектов в подвыборке.

4.2. Алгоритм LightGBM. Это метод градиентного бустинга, который основан на деревьях решений. LightGBM оптимизирован для высокой производительности в распределенных системах и создает деревья решений, которые растут по листам. Это означает, что при заданном условии разделяется только один лист, в зависимости от значения прироста информации.

LightGBM использует метод, основанный на гистограмме, в котором данные распределяются по ячейкам с использованием гистограммы распределения для вычисления коэффициента усиления и разделения данных.

Для выборки набора данных в LightGBM используется односторонняя выборка на основе градиента (GOSS). Метод GOSS присваивает точкам данных с большими градиентами при расчете усиления более высокий вес. В этом методе экземпляры, которые недостаточно использовались для обучения, вносят больший вклад. Точки данных

с меньшими градиентами удаляются случайным образом, некоторые из них сохраняются для поддержания точности.

5. Результаты применения алгоритмов. Описанные выше алгоритмы были применены на датасете Wiki movie plots [5]. Была произведена обработка текстов (исправление опечаток, удаление фильмов, встречающихся реже, чем в 1% выборки) и разбиение данных на тестовое, валидационное и обучающее множество в соотношении 0,2/0,2/0,6 соответственно. Полученные результаты приведены в таблице 1.

Таблица 1. Результаты работы алгоритмов с использованием только сюжета

Название метода	ROC AUC
multinomialNB	0,831
Tf-Idf + SVM	0,872
RoBERTa	0,8967
XLNet	0,8959

Использовалась метрика (ROC) AUC [6] (среднее её значение по всем жанрам), поскольку является независимой от порога классификации и масштаба данных. Значение метрики в 0,5 соответствует случайному классификатору, 1,0 – идеальному.

На валидационной выборке получены оптимальные параметры: для SVM $C = 0,5$ (шаг поиска 0,1); для multinomialNB $\alpha = 0,01$ (шаг поиска 0,001); для RoBERTa – 3 эпохи, скорость обучения $2 \cdot 10^{-5}$; для XLNet – 2 эпохи, скорость обучения $2 \cdot 10^{-5}$.

Таблица 2. Результаты работы алгоритмов с использованием категориальных признаков

Название метода	ROC AUC
RandomForest	0,7764
LightGBM	0,765

Оптимальные значения параметров алгоритмов из таблицы 2 подбирались на валидации с использованием Optuna [7].

Также производилось ансамблирование методов, обученных только на сюжете и методов, обученных только с использованием категориальных признаков. Результаты представлены в таблице 3.

Таблица 3. Результат ансамблирования алгоритмов

Название метода	ROC AUC
RoBERTa + Random Forest	0,9047
RoBERTa + LightGBM	0,9062
RoBERTa + Random Forest + LightGBM	0,9071
RoBERTa + XLNet + Random Forest + LightGBM	0,9138

6. Заключение. В данной работе рассмотрены алгоритмы для классификации жанров фильма как для случая с использованием только сюжета, так и с использованием категориальных признаков (директор фильма, актерский состав и прочее).

Метод RoBERTa оказался лучшим с точки зрения метрики ROC AUC. При использовании только категориальных признаков лучшее качество показал алгоритм LightGBM. Лучшего же результата удалось достичь ансамблированием методов RoBERTa, XLNet, Random Forest, LightGBM.

Литература

1. Predicting Movie Genres Based on Plot Summaries [Электронный ресурс]: URL:<https://arxiv.org/pdf/1801.04813.pdf> (дата обращения 29.03.23).
2. Ertugrul A. M., Karagoz P. Movie genre classification from plot summaries using bidirectional LSTM // IEEE 12th International Conference on Semantic Computing (ICSC), 2018. P. 248–251.
3. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding [Электронный ресурс]: URL:<https://arxiv.org/abs/1810.04805> (дата обращения 29.03.23).
4. XLNet: Generalized Autoregressive Pretraining for Language Understanding [Электронный ресурс]: URL:<https://arxiv.org/abs/1906.08237> (дата обращения 29.03.23).
5. Wikipedia Movie Plots [Электронный ресурс]: URL:<https://www.kaggle.com/datasets/jrobischon/wikipedia-movie-plots> (дата обращения 08.02.23).
6. Classification: ROC Curve and AUC [Электронный ресурс]: URL:<https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc> (дата обращения 25.02.23).
7. Optuna: A Next-generation Hyperparameter Optimization Framework [Электронный ресурс]: URL:<https://arxiv.org/abs/1907.10902> (дата обращения 29.03.23).

УДК 004.021

Кукло Р. А.

Применение метода кластеризации по времени и географическим данным для уменьшения объема передаваемых данных

Рекомендовано к публикации доцентом Корховым В. В.

1. Введение. Беспроводная сенсорная сеть (Wireless sensor network, WSN) – это сеть небольших сенсорных узлов, цель которых фиксировать изменение данных в своем окружении и пересылать их на принимающий узел. После этого принимающий узел обрабатывает полученные данные и пересылает результаты в центр мониторинга [1]. Крупномасштабные беспроводные сенсорные сети все чаще развертываются в различных классах приложений для точного мониторинга и в приложениях интернета вещей.

Однако передача информации от каждого узла к узлу-приемнику требует больших затрат энергии [2]. Кроме того, при передаче большого объема информации возникают проблемы, такие как лимит на объем передаваемых сообщений и потеря сигнала сенсора.

Высокая плотность узлов позволяет сделать предположение, что у некоторых узлов информация коррелирует как в пространстве, так и во времени [3]. Это позволяет объединить сенсоры с коррелируемой информацией в кластер (группу «родственных» объектов) и передавать лишь данные с одного сенсора каждого кластера. Данный метод позволяет сократить объем передаваемого сообщения [4].

Большинство методов основаны либо на пространственной корреляции, либо на временной. Однако более эффективным методом является пространственно-временная корреляция. Она позволяет сократить общий объем информации в десятки раз [5]. Большинство статей посвящены схемам сжатия данных [1] и методам восстановления сжатой информации с высокой точностью [5]. Кроме того, во многих методах данные кластеризуются, как если бы мы имели доступ к полной информации со всех сенсоров сразу. Такой подход почти не реализуем на практике, поэтому предложим практический

Кукло Роман Алексеевич – студент, Санкт-Петербургский государственный университет; e-mail: st097551@student.spbu.ru, тел.: +7(921)031-90-99

способ обхода сенсоров для последовательного сравнения данных. Существуют и более точные методы кластеризации и сокращения объема передаваемых данных, такие как Adaptive Prediction Strategy with Clustering (APSCT) [6]. Однако использование подобных алгоритмов сопряжено с дополнительными вычислительными мощностями.

В данной работе рассмотрены в отдельности пространственная и временная корреляции, а также пространственно-временная корреляция. Продемонстрирован способ обхода узлов для объединения в кластер.

2. Постановка задачи. Пусть имеются n сенсоров, расположенных на плоскости. Каждому сенсору присвоены координаты (x, y) , обозначающие координаты по оси абсцисс и ординат соответственно.

Каждый сенсор собирает одинаковый объем данных за промежуток времени T . Пусть за данное время имеется дискретный набор данных с количеством значений 500 для каждого сенсора. Эти данные каждый сенсор передает узлу-приемнику.

Задача состоит в уменьшении объема передаваемых данных путем объединения сенсоров в кластеры. Объединение в кластеры будет происходить путем последовательного обхода сенсоров и сравнения их данных. Необходимо рассмотреть различные способы корреляции данных: временную, пространственную и пространственно-временную и оценить их эффективность.

3. Начальные данные. Пусть имеются 4 различных функции от переменной t , где $t \in (0, 10)$, с шагом 0,02:

$$l_1 = \cos(2t), \quad (1)$$

$$l_2 = \sin(3t), \quad (2)$$

$$l_3 = \cos(0,05t), \quad (3)$$

$$l_4 = \sin(0,01t). \quad (4)$$

Будем полагать, что каждый из десяти рассматриваемых сенсоров с номерами от 0 до 9 будет обладать данными, равными значениям одной из четырех вышеперечисленных функций (1)–(4).

Для наглядности представим данные в датафрейме, где каждой колонке соответствуют значения для определенного сенсора. Количество строк в такой таблице будет равно количеству значений, т. е.

500. Весь объем данных (типа float) будет занимать 41640 байт. В дальнейшем полученные данные будут носить название Пример 1.

Для полученных данных с помощью встроенных функций легко построить матрицу корреляции (см. рис. 1), где значения ячеек соответствуют коэффициенту корреляции между значениями сенсоров.

Из матрицы корреляции видно, что многие сенсоры имеют близкие значения (коэффициент корреляции близок к единице). Следовательно, при передаче всех данных от каждого сенсора по отдельности будет передаваться одна и та же информацию несколько раз.

	sensor_0	sensor_1	sensor_2	sensor_3	sensor_4	sensor_5
sensor_0	1.000000	0.181136	0.181136	0.084865	0.181136	1.000000
sensor_1	0.181136	1.000000	1.000000	-0.099667	1.000000	0.181136
sensor_2	0.181136	1.000000	1.000000	-0.099667	1.000000	0.181136
sensor_3	0.084865	-0.099667	-0.099667	1.000000	-0.099667	0.084865
sensor_4	0.181136	1.000000	1.000000	-0.099667	1.000000	0.181136
sensor_5	1.000000	0.181136	0.181136	0.084865	0.181136	1.000000

Рис. 1. Пример матрицы корреляции данных первых шести сенсоров

4. Временная корреляция. На практике значения чаще всего повторяются с какой-то периодичностью. Например, в рассматриваемом Примере 1 все функции периодичны. Следовательно, можно сократить объем передаваемых данных.

Учитывая специфику конкретного Примера 1, реализуем алгоритм, способный найти промежутки с коррелируемыми значениями. Рассмотрим алгоритм на примере сенсора с номером 0. Всего значений 500. Будем брать последовательно по 25 значений. Рассмотрим значения сенсора 0 в нашем датафрейме со строки 0 по 24 строку (первый интервал). Возьмем следующий промежуток с 25 по 49 строку (второй интервал). И так далее, всего получим 20 промежутков. Составим из них новый датафрейм, где каждый столбец соответствует промежутку с номером от 0 до 19. Составим для нового датафрейма матрицу корреляций для столбцов (см. рис. 2).

Следовательно, можно объединять коррелирующие промежутки в кластеры и вместо всех данных передавать сообщение, сформированное следующим образом: для каждого кластера будем передавать номера промежутков (типа int64), входящих в этот кластер, все значения первого промежутка в кластере (25 значений типа float), его коэффициент корреляции с каждым последующим промежутком

	0	1	2	3	4	5
0	1.000000	0.973162	0.917540	-0.994081	-0.977330	-0.932891
1	0.973162	1.000000	0.984421	-0.942400	-0.999822	-0.990734
2	0.917540	0.984421	1.000000	-0.868907	-0.980929	-0.999180
3	-0.994081	-0.942400	-0.868907	1.000000	0.948542	0.888239
4	-0.977330	-0.999822	-0.980929	0.948542	1.000000	0.987996
5	-0.932891	-0.990734	-0.999180	0.888239	0.987996	1.000000

Рис. 2. Пример матрицы корреляции для первых шести временных промежутков сенсора 0

этого кластера (типа float), а также первое значение каждого промежутка этого кластера (типа float). Если данную процедуру проделать со значениями каждого сенсора для Примера 1, то объем таких переданных данных будет равен примерно 10000 байт.

5. Пространственная корреляция. Присвоим каждому сенсору координаты (x, y) , исходя из того, что коррелирующие сенсоры находятся близко друг к другу. Иначе данная задача не имела бы смысла. Пусть сенсоры на плоскости расположены как на рис. 3.

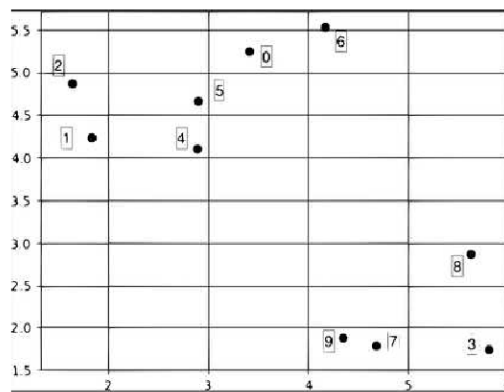


Рис. 3. Пример расположения сенсоров на плоскости для Примера 1

Алгоритм обхода всех сенсоров:

1. Выбираем любой начальный сенсор, например, сенсор 0.
2. Располагаем оставшиеся сенсоры в порядке удаленности от начального вектора. Далее обходим сенсоры таким образом.

3. Если сенсор коррелирует с начальным сенсором, добавляем сенсор в кластер, идем к следующему сенсору по удаленности и снова повторяем пункт 3. Иначе создаем новый кластер, где начальным сенсором будет текущий, идем в пункт 2.

Алгоритм выполняется, пока все сенсоры не войдут в какой-то кластер. Для каждого кластера будем передавать номера сенсоров (типа `int64`), входящих в него, все значения начального сенсора кластера (типа `float`) и первые значения каждого сенсора (типа `float`). Таким образом, для Примера 1 объем переданного сообщения получится равным примерно 16000 байт и получены следующие кластеры: [0, 5, 6], [1, 2, 4], [7, 9], [3, 8].

6. Пространственно-временная корреляция. Наиболее эффективный вид корреляции включает в себя два этапа, которые выполняются последовательно друг за другом:

1. Пространственная корреляция.
2. Временная корреляция.

В результате пространственной корреляции получается массив значений начальных сенсоров каждого кластера. Затем к этому массиву применяется метод временной корреляции. Для Примера 1 размер итогового сообщения равен примерно 4000 байт, что приблизительно в 10 раз меньше начального объема.

7. Заключение. В данной работе рассмотрены различные виды корреляции данных. Выяснилось, что пространственно-временная корреляция приводит к наилучшим результатам и максимальному сокращению объема передаваемого сообщения.

Литература

1. Chen S. et al. Layered adaptive compression design for efficient data collection in industrial wireless sensor networks // Journal of Network and Computer Applications. 2019. Vol. 129. P. 37–45.
2. Guo Z., Peng J. et al. Minimizing redundant sensing data transmissions in energy-harvesting sensor networks via exploring spatial data correlations // IEEE Internet of Things Journal. 2020. Vol. 8. No 1. P. 512–527.

3. Yoon S.H., Shahabi C. Exploiting spatial correlation towards an energy efficient clustered aggregation technique (cag) [wireless sensor network applications] // IEEE International Conference on Communications. 2005. Vol. 5. P. 3307–3313.
4. Глазырин А.Г. Численное моделирование и оценка характеристик плотности распределения модельных кластеров на евклидовой плоскости // Процессы управления и устойчивость. 2020. Т. 7. №. 1. С. 225–229.
5. Villas L. A., Guidoni D. L. et al. An energy-aware spatio-temporal correlation mechanism to perform efficient data collection in wireless sensor networks // Computer Communications. 2013. Vol. 36. No 9. P. 1054–1066.
6. Kumar R., Jain V. et al. An adaptive prediction strategy with clustering in wireless sensor network // International Journal of Wireless Information Networks. 2020. Vol. 27. P. 575–587.

УДК 004.75

Мосина М. А.

Исследование методов распределения нагрузки на основе муравьиного алгоритма

Рекомендовано к публикации доцентом Корховым В. В.

1. Введение. С развитием технологий увеличивается спрос на облачные вычисления. Они позволяют быстро реагировать на изменения в вычислительных потребностях и значительно увеличивают скорость вычислений и обработки данных. Системам облачных вычислений приходится взаимодействовать с большим количеством пользовательских задач и большим объемом данных. Балансировка нагрузки необходима для повышения производительности виртуальных машин и оптимального использования облачных ресурсов.

Система состоит из заранее заданного числа виртуальных машин. Количество заданий (независимых или зависимых) и их сложность задаются до начала работы алгоритма. Задачей алгоритма является оптимизация времени выполнения всех заданий.

В данной работе рассмотрены методы, основанные на муравьином алгоритме, предложенном в статье [1]. Идея муравьиного алгоритма основана на принципе поиска пищи колонии муравьев. Муравьи ориентируются в пространстве с помощью феромонного следа, оставленного каждым муравьем. Самый короткий путь является самым популярным.

2. Муравьиный алгоритм и его модификации. Муравьиный алгоритм – это случайный оптимизационный поиск, который используется для распределения поступающих заданий по виртуальным машинам. Предшественниками муравьиного алгоритма в задаче оптимизации нагрузки являются алгоритмы FCFS и Round-Robin. Результаты экспериментов в статье [2] показали, что оптимизация муравьиной колонии превосходит оба эти алгоритма.

Выполняемые системой задания могут быть представлены в виде m суперузлов. Каждый суперузел имеет n узлов – виртуальных машин, выполняющих задание. Муравей путешествует от одного узла

Мосина Мария Александровна – магистрант, Санкт-Петербургский государственный университет; e-mail: st076887@student.spbu.ru, тел.: +7(981)762-29-71

к другому. На каждой итерации искусственные муравьи перемещаются по графу.

Обновление феромонного следа может происходить глобально или локально. Локальное обновление случается, когда муравей выбрал следующий узел для посещения. Глобальное обновление значения феромона происходит после завершения тура всеми муравьями.

Идеи последующих алгоритмов описаны в статьях [1, 3, 4].

2.1. Базовый алгоритм Ant System (AS). Изначально муравьи распределяются по случайно выбранным узлам. На начальной итерации они пойдут по случайно выбранным путям. На каждом этапе путешествия применяется вероятностное правило для определения следующего узла:

$$p_{ij}^k = \frac{[\tau_{ij}]^\alpha [\eta_{ij}]^\beta}{\sum_{l \in N_i^k} [\tau_{il}]^\alpha [\eta_{il}]^\beta}, \quad j \in N_i^k, \quad (1)$$

где τ_{ij} – значение феромона на пути от узла i к узлу j , η_{ij} – величина, обратно пропорциональная времени выполнения задачи j , а N_i^k – множество допустимых узлов, которые муравей k может посетить после узла i , β и α – параметры, которые определяют степень воздействия времени выполнения задачи и значения феромонного следа на выбор следующего узла.

После завершения тура всеми муравьями феромонные следы обновляются по формулам

$$\tau_{ij} \leftarrow (1 - \rho) \tau_{ij} + \sum_{k=1}^m \Delta \tau_{ij}^k, \quad \forall (i, l) \in E,$$

$$\Delta \tau_{ij}^k = \begin{cases} 1/C^k, & \text{если путь } (i, j) \text{ принадлежит } T^k, \\ 0, & \text{в противном случае.} \end{cases}$$

Здесь ρ , $0 \leq \rho \leq 1$, – скорость испарения феромона, $\Delta \tau_{ij}^k$ отвечает за количество феромона, оставленного муравьем k на пути между узлами i и j , C^k – время полного путешествия муравья k по всем суперузлам, T^k – полный тур k -го муравья, E – множество всех возможных путей в системе.

2.2. Elitist Ant System (EAS). В данном методе дополнительно при глобальном обновлении феромона добавляется феромон на путях, которые принадлежат туру T^{bs} , время выполнения которого минимально.

2.3. Система муравьиной колонии (Ant Colony System, ACS). Алгоритм системы муравьиной колонии позволяет муравьям исследовать новые маршруты. Глобальное обновление феромона доступно только на лучшем маршруте и некоторое количество феромона испаряется при переходе от задачи i к задаче j после прохождения муравья по пути.

Итак, муравей выберет лучший путь по вероятностному правилу (1) с некоторой установленной вероятностью q . С вероятностью $1 - q$ муравей пойдет по неисследованному пути.

Глобальное обновление феромонного следа происходит только на лучшем маршруте. Локальное обновление включает в себя также испарение феромона с заданным коэффициентом $(1 - \xi)$.

2.4. Load Balancing Ant Colony Optimization (LBACO). Алгоритм основан на предположении, что при выборе узла должны учитываться параметры виртуальных машин системы.

Изначальное количество феромона на путях задается формулой

$$\tau_{ij}(0) = \text{re_num}_l \times \text{re_mips}_l + \text{vm_bw}_l, \quad i = \overline{1, m}.$$

Здесь re_num_l – количество процессоров в виртуальной машине vm_l , на которой будет выполнено задание j , re_mips_l – MIPS (количество миллионов инструкций в секунду) каждого процессора виртуальной машины vm_l , а vm_bw_l связана с пропускной способностью сети, с помощью которой этой виртуальной машине передаются данные.

Выбор следующего пункта назначения происходит с помощью вероятностного правила

$$p_{ij}^k = \frac{[\tau_{ij}]^\alpha [EV_l]^\beta [LB_l]^\gamma}{\sum_{e \in N_i^k} [\tau_{ie}]^\alpha [EV_l]^\beta [LB_l]^\gamma}, \quad j \in N_i^k, \quad (2)$$

где τ_{ij} – значение феромона на рассматриваемом ребре, EV_l – вычислительная мощность виртуальной машины vm_l , которая будет использоваться для решения задания j , а LB_l – коэффициент балансировки нагрузки виртуальной машины vm_l .

Вычислительная мощность определяется по формуле

$$EV_l = \text{re_num}_l \times \text{re_mips}_l + \text{vm_bw}_l,$$

где re_num_l – это количество процессоров в виртуальной машине vm_l , re_mips_l – MIPS (количество миллионов инструкций в секунду)

каждого процессора виртуальной машины vm_l , а vm_bw_l связана с пропускной способностью сети, связанной с виртуальной машиной vm_l .

Коэффициент балансировки определяется по формуле

$$LB_j = 1 - \frac{res_l - lastAver_res}{res_l + lastAver_res},$$

где $lastAver_res$ – это среднее время выполнения заданий на оптимальном пути последней итерации, а ожидаемое время выполнения задания на виртуальной машине l

$$res_l = \frac{total_tasklength}{EV_l} + \frac{InputFilesize}{vm_bw_l}.$$

Здесь $total_tasklength$ – это общая длина задачи, т. е. количество операций, которое должно быть выполнено, а $InputFilesize$ – размер входных данных задачи.

В формуле (2) α, β и γ – параметры, определяющие относительный вес феромонного следа, вычислительной мощности виртуальных машин и коэффициента балансировки нагрузки соответственно.

Глобальное обновление феромона происходит только для посещенных путей по формулам

$$\tau_{ij} \leftarrow (1 - \rho) \tau_{ij} + \Delta\tau_{ij},$$

$$\Delta\tau_{ij} = 1/C^k.$$

Здесь ρ , $0 \leq \rho \leq 1$, – коэффициент затухания феромонного следа, C^k – время пути муравья k . Кроме того, дополнительно поощряется оптимальный маршрут

$$\Delta\tau = D/C^{bs},$$

где C^{bs} – время путешествия муравья по оптимальному маршруту T^{bs} , а D – коэффициент поощрения.

3. Эксперименты. Большинство существующих на данный момент работ [2, 4–9] содержит в себе сравнительный анализ одного конкретно взятого алгоритма с базовыми алгоритмами, например, Round-Robin (RR), анализ эффективности работы алгоритмов проведен на независимых задачах. В связи с этим, в ходе работы была разработана модификация рассматриваемых алгоритмов для зависимых задач.

3.1. Независимые задания. Рассмотрим ситуации, в которых система не имеет приоритетных для выполнения заданий, т. е. задания могут быть выполнены в любом порядке. В ситуациях с различными сложностями заданий, которые измеряются в миллионах инструкций (mi), и мощностями виртуальных машин, измеряемых количеством выполняемых инструкций в секунду (mips), для сравнительного анализа будет приведен также алгоритм Round-Robin. Для оценки алгоритмов рассматривается среднее арифметическое времени путешествия муравьев в трех последних итерациях алгоритмов. Для подсчета среднего значения и оценки каждый эксперимент проводился пять раз.

3.1.1. Сложность заданий. Рассмотрены три возможные ситуации, при которых в систему поступает различное количество заданий с низкой и высокой сложностями. Для всех рассматриваемых случаев закрепим остальные параметры: система состоит из 7 виртуальных машин (ВМ), каждая производительностью 100 mips (по 4 процессора в каждой ВМ), и 7 виртуальных машин, каждая производительностью 400 mips (по 4 процессора в каждой ВМ). Количество муравьев равно 15. Количество заданий и их сложность представлены в таблице 1.

Таблица 1. Количество и сложность заданий

Ситуация	Количество заданий	Сложность заданий, mi
1	12	300
	3	5000
2	3	300
	12	5000
3	15	1000

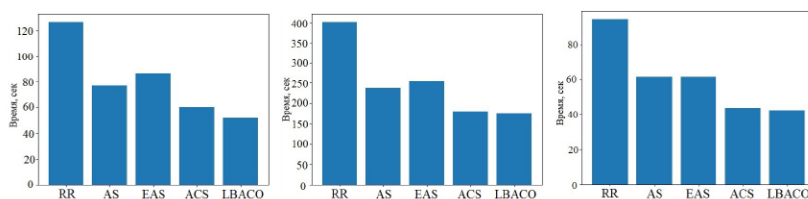


Рис. 1. Среднее время работы алгоритмов в ситуациях 1–3

Исходя из рис. 1, алгоритмы AS и EAS во всех ситуациях возвращают результат хуже, чем алгоритмы ACS и LBACO, которые получают близкие по времени результаты.

3.1.2. Мощности виртуальных машин. Рассмотрим случаи с различным процентным соотношением ВМ с высокой и низкой мощностями (таблица 2). Для обеих ситуаций параметры системы таковы: 15 заданий, сложность каждого составляет 1000 m_i , 15 муравьев.

Таблица 2. Количество и мощность ВМ

Ситуация	Количество ВМ	Мощность ВМ
1	12	1000
	3	100
2	3	1000
	12	100

Из рис. 2 видно, что в системе, где преобладают машины с небольшой мощностью, алгоритмы ACS и LBACO получают результаты более, чем в два раза лучше, чем алгоритмы AS и EAS, которые значительно более чувствительны к вычислительным мощностям виртуальных машин.

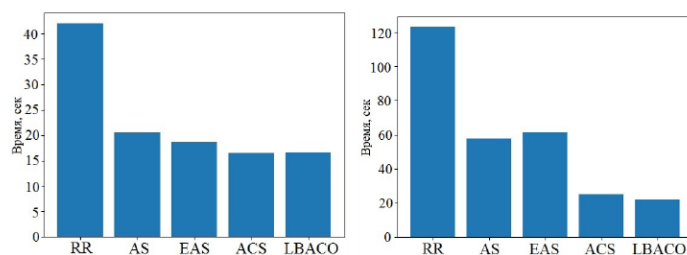


Рис. 2. Среднее время работы алгоритмов в ситуациях 1, 2

3.2. Зависимые задания. Зависимые задания – это задания, которые не могут быть выполнены до выполнения другого задания, с которым они связаны. Для всех ситуаций установлено 14 заданий со сложностью 1000 m_i . В системе 7 виртуальных машин мощностью 100 $mips$, 7 виртуальных машин мощностью 400 $mips$, 14 муравьев.

Ситуация 1. Все задания можно разделить на пары, где одно – зависимое, а другое – то, от которого оно зависит.

Ситуация 2. Среди заданий присутствуют цепочки вложенных зависимостей. Например, задание 1 зависит от задания 2, задание 2 зависит от задания 3 и т. д.

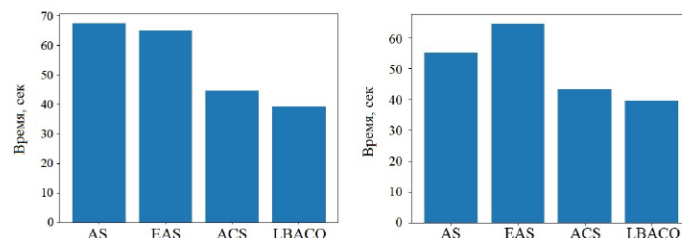


Рис. 3. Среднее время работы алгоритмов в ситуациях 1, 2

Из рис. 3 видно, что в ситуации 1 значительных отклонений от поведения алгоритмов при независимых заданиях выявлено не было. При невложенных зависимостях, как и при независимых заданиях, алгоритмы AS и EAS уступают алгоритмам ACS и LBACO. В ситуации 2 при вложенных зависимостях алгоритм AS получает результат лучше, чем алгоритм EAS, но по-прежнему не достигает результатов алгоритмов ACS и LBACO.

4. Выводы. В большинстве рассмотренных случаев результаты AS и EAS были схожи, поэтому нельзя точно утверждать, какой алгоритм является предпочтительнее для дальнейшего использования. Кроме того, ACS и LBACO показывали значительно лучшие результаты.

Сравнивая алгоритмы ACS и LBACO, приходим к выводу, что в большинстве рассмотренных случаев LBACO работает быстрее. Тем не менее, бывают ситуации, когда LBACO уступает по времени алгоритму ACS. Такое поведение может быть объяснено как зависимостью методов от определенных параметров системы, так и случайной расстановкой муравьев в начальный момент времени.

Результаты алгоритм EAS в ситуации с последовательными вложенными задачами значительно хуже, чем результаты остальных рассматриваемых алгоритмов. Исходя из этого, приходим к выводу, что алгоритм EAS стоит использовать лишь в системах, на вход которым поступают только независимые задачи.

Литература

1. Dorigo M., Gambardella L. M. Ant colony system: a cooperative learning approach to the traveling salesman problem // IEEE Transactions on Evolutionary Computation. 1997. No 1. P. 53–66.

2. Tawfeek M. A., El-Sisi A., Keshk A. E., Torkey F. A. Cloud task scheduling based on ant colony optimization // International Arab Journal of Information Technology. 2013. Vol. 12. No 2. P. 64–69.
3. Chaturvedi R., Banka H. Modified ant colony optimization algorithm for travelling salesman problem // International Journal of Computer Applications. 2014. Vol. 97. No 10. P. 20–24.
4. Li K., Xu G., Zhao G., Dong Y., Wang D. Cloud task scheduling based on load balancing ant colony optimization // Sixth Annual China Grid Conference. 2011. P. 3–9.
5. Ali A. D., Belal M. A., Al-Zoubi M. B. Load balancing of distributed systems based on multiple ant colonies optimization // American Journal of Applied Sciences. 2010. Vol. 7. No 3. P. 428–433.
6. Keskinurk T., Yildirim M. B., Barut M. An ant colony optimization algorithm for load balancing in parallel machines with sequence-dependent setup times // Computers & Operations Research. 2012. Vol. 39. Iss. 6. P. 1225–1235.
7. Guo Q. Task scheduling based on ant colony optimization in cloud environment // AIP Conference Proceedings. 2017. Vol. 1834. Art. no 040039.
8. Umarani S., Senthilprakash T. Parallel asynchronous particle swarm optimization for job scheduling in grid environment // International Journal on Recent and Innovation Trends in Computing and Communication. 2014. Vol. 2. Iss. 8. P. 2384–2389.
9. Чабанец А. В., Власова Т. В. Сравнительный анализ муравьиного алгоритма и метода имитации отжига для задачи маршрутизации с ограничением на грузоподъёмность // Процессы управления и устойчивость. 2020. Т. 7. № 1. С. 428–432.

УДК 004.93

Набиев Т. Р.

Применение методов компьютерного зрения для обнаружения и отслеживания дронов на видео

Рекомендовано к публикации доцентом Головкиной А.Г.

1. Введение. Реконструкция траектории движения объектов по видеопотоку становится все более важной, особенно в контексте распространения беспилотных летательных аппаратов (БПЛА). Способность дронов маневрировать в трехмерном пространстве, их небольшие размеры и возможность совершения резких и непредсказуемых движений создают сложности в их обнаружении и отслеживании.

Отслеживание объектов – это задача компьютерного зрения, которая заключается в автоматическом обнаружении и слежении за объектами в видеопоследовательностях. В процессе решения данной задачи возникают дополнительные сложности, связанные с возможностью перекрытия объектов, изменения их внешнего вида и исчезновения из кадра.

Целью данного исследования является сравнение эффективности различных алгоритмов для отслеживания дронов в видеопотоке. Рассматриваются алгоритмы на основе машинного обучения, так как на данный момент они являются самыми эффективными для решения поставленной задачи. Оценка будет проводиться на основе метрики НОТА (Higher Order Tracking Accuracy) [1], объединяющей показатели точности обнаружения и отслеживания объектов. Для оценки методов был использован размеченный набор данных [2], который представляет собой набор из 77 видео общей продолжительностью более часа, с записями полетов 1–5 дронов (рис. 1) и размеченные ограничительные рамки для каждого кадра видео. Общее количество последовательностей кадров составляет около 100 тысяч.

2. Алгоритмы отслеживания объектов. Существуют традиционные методы отслеживания, которые основаны на точках интереса во времени и пространстве, а также на низкоуровневых характеристиках, таких как изменение интенсивности цвета пикселей.

Набиев Тимур Ришатович – студент, Санкт-Петербургский государственный университет; e-mail: st088226@student.spbu.ru, тел.: +7(986)703-42-15

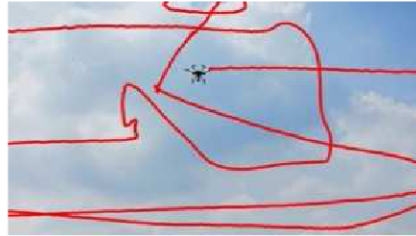


Рис. 1. Траектория дрона на видео из используемого набора данных

Однако в настоящее время методы машинного обучения и глубокие нейронные сети показывают лучшие результаты по сравнению с традиционными алгоритмами [3]. Рассмотрим наилучшие из известных алгоритмов:

2.1. Алгоритм DeepSORT. Deep Simple Online and Realtime Tracking [4] – это алгоритм отслеживания для видео, расширение функциями глубокого обучения алгоритма SORT [5], который использует фильтр Калмана для предсказания движения объектов и Венгерский алгоритм для ассоциации. DeepSORT объединяет глубокое представление признаков, предсказание движения и эффективное объединение данных для точного отслеживания объектов. Интеграция функций глубокого обучения повышает способность алгоритма справляться со сложными сценариями отслеживания, такими как окклюзия и изменения внешнего вида.

2.2. Алгоритм ByteTrack. Это многообъектный метод отслеживания, который основан на ассоциации всех обнаружений [6]. В отличие от других подходов, ByteTrack объединяет все обнаружения в пары, используя критерии пространственного и временного совпадения, и рассчитывает центр и площадь для каждого обнаружения. Затем идет расчет и обновление весов ассоциаций для каждого обнаружения. Таким образом улучшается точность ассоциаций объектов.

2.3. Алгоритм OC-SORT. Он объединяет основные принципы системы отслеживания объектов (SORT) с дополнительными функциями центрирования вокруг наблюдений [7]. На каждом кадре видеопотока применяется система SORT для обнаружения и отслеживания объектов с использованием алгоритмов детекции объектов и фильтра Калмана. Для каждого обнаруженного объекта вычисляется его центр и сохраняется в памяти для последующего использо-

вания. Алгоритм сопоставляет обнаруженные объекты с сохраненными центрами предыдущих кадров, используя различные метрики сходства и дополнительные функции центрирования. Это повышает устойчивость алгоритма к изменениям внешних условий съемки и обеспечивает более точное и устойчивое отслеживание объектов по сравнению с базовым методом SORT.

3. Метрики. Рассмотрим составляющие метрики НОТА для оценки эффективности отслеживания объектов.

3.1. Локализация. Локализация измеряет пространственное соответствие между предсказанным и истинным обнаружением. Показатель IoU локализации (Loc-IoU) [8] рассчитывается как отношение площади перекрытия (пересечения) между двумя обнаружениями к общей площади, покрываемой ими обоими (объединение). Измеряется общая точность локализации (LocA) путем усреднения Loc-IoU по всем парам совпадающих предсказанных и истинных обнаружений во всем наборе данных по формуле:

$$\text{LocA} = \frac{1}{|\text{TP}|} \sum_{c \in \text{TP}} \text{Loc-IoU}(c).$$

3.2. Обнаружение. Обнаружение измеряет соответствие между предсказанными и истинными обнаружениями через показатель Detection IoU (Det-IoU), который определяется по формуле:

$$\text{Det-IoU} = \frac{|\text{TP}|}{|\text{TP}| + |\text{FN}| + |\text{FP}|}.$$

Det-IoU использует порог локализации (например, $\text{Loc-IoU} > 0,5$) для определения пересечения двух обнаружений, что делает их истинно положительными (TP). Обнаружения, не совпадающие с истинными, являются ложноположительными (FP), а истинные обнаружения, которые не были предсказаны, – ложноотрицательными (FN).

3.3. Ассоциация. Ассоциация измеряет, насколько хорошо треки связывают обнаружения во времени с одинаковыми идентификаторами (ID) за счет сравнения предсказанных и истинных ID. Истинно положительные совпадения между двумя треками являются истинно положительными ассоциациями True Positive Associations (TPA). Аналогично определяются False Positive Associations (FPA)

и False Negative Associations (FNA). Ассоциативный коэффициент можно рассчитать по формуле:

$$\text{Ass-IoU} = \frac{|\text{TPA}|}{|\text{TPA}| + |\text{FNA}| + |\text{FPA}|}.$$

4. Сравнение. Результаты оценки на наборе данных рассмотренных алгоритмов представлены в таблице. На графике (рис. 2) можно видеть, что по метрике LocA лучше всего справился алгоритм DeepSORT. Метод OC-SORT лучше всех по метрике DetA, однако очень плохо показал себя в точности ассоциации. По метрике AssA и объединяющей их всех метрике HOTA метод ByteTrack лучше всего справился с задачей отслеживания дрона на видео. Так как при потере объекта интереса и последующем его повторном обнаружении ByteTrack продолжал соответствующую этому объекту траекторию, и по остальным параметрам не сильно уступает другим алгоритмам.

Таблица. Результаты оценки

Модель	HOTA	DetA	AssA	LocA
deepocsort	20,277	31,740	13,315	74,563
bytetrack	21,048	32,871	14,000	73,866
ocsort	20,707	35,453	12,591	73,908

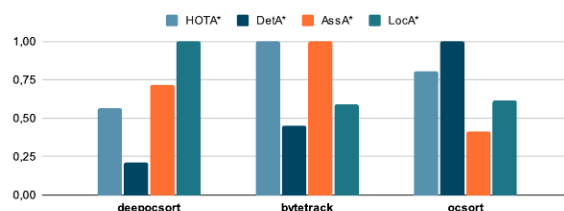


Рис. 2. Нормированные значения метрик

По результатам проведенного сравнительного анализа можно утверждать, что алгоритм ByteTrack является самым эффективным в решении задачи отслеживания дрона по видео. Это обосновано тем, что предложенный метод успешно объединяет все обнаружения в пары, используя критерии пространственного и временного совпадения и рассчитывая для этих ассоциаций веса, что позволяет повысить точность ассоциаций и обнаружений.

Отметим, что повышение эффективности слежения за дронами возможно при помощи настройки гиперпараметров нейронных сетей для детекции дронов, а также за счет комбинации нескольких методов. Данные подходы предполагается рассмотреть в дальнейшем.

Литература

1. Luiten J., Aljosa O. et al Hota: A higher order metric for evaluating multi-object tracking // International Journal of Computer Vision. 2021. No129. P. 548–578.
2. Drone-vs-Bird Detection Challenge // [Электронный ресурс] URL:<https://github.com/wosdetc/challenge> (дата обращения: 19.03.2024).
3. Raghava N., Gupta K. et al. An experimental comparison of different object tracking algorithms // 2020 International Conference on Communication and Signal Processing (ICCSP). 2020. P. 0726–0730.
4. Wojke N., Bewley A., Paulus D. Simple online and realtime tracking with a deep association metric //2017 IEEE International Conference on Image Processing (ICIP). 2017. P. 3645–3649.
5. Bewley A., Ge Z. et al. Simple online and realtime tracking // 2016 IEEE International Conference on Image Processing (ICIP). 2016. P. 3464–3468.
6. Zhang Y., Sun P. et al. Bytetrack: Multi-object tracking by associating every detection box // European Conference on Computer Vision. 2022. P. 1–21.
7. Cao J., Pang J. et al. Observation-centric sort: Rethinking sort for robust multi-object tracking // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023. P. 9686–9696.
8. Rahman M. A., Wang Y. Optimizing intersection-over-union in deep neural networks for image segmentation // International Symposium on Visual Computing. 2016. P. 234–244.

УДК 004.912

Неронов Р. М.

Решение задачи разведочного анализа текстов через кластеризацию и аппроксимацию числа кластеров

Рекомендовано к публикации доцентом Блекановым И. С.

1. Введение. Многие компании работают с огромными потоками текстовой информации, большая часть которых представляет собой неразмеченные и неструктурированные данные. Поэтому специалистов Data Science часто просят провести разведочный анализ коллекции текстов, который является вспомогательным этапом, помогающий бизнесу сформулировать конкретную задачу.

В данной статье рассмотрены план и методы, которые были использованы в программной реализации решения задачи разведочного анализа, разработанной автором статьи.

2. Постановка задачи. За основу разведочного анализа текстов взята задача кластеризации. Затем на основании полученных разбиений осуществляется расчет аналитики. Под аналитикой для группы текстов понимаются ключевые слова для всего класса в целом, а также типичные представители кластера.

Рассмотрим план всего цикла анализа документов, который был взят за основу для решения задачи:

1. Загрузка корпуса текстов в одном из доступных форматов.
2. Обработка набора данных.
3. Вычисление векторного представления для каждого текста.
4. Нахождение оптимального числа кластеров для первичной кластеризации.
5. Проведение первичной кластеризации.
6. Формирование списка ключевых слов для каждого кластера.
7. Интерактивное взаимодействие с кластерами одним из предложенных способов: разделение или объединение.

Неронов Роман Михайлович – студент, Санкт-Петербургский государственный университет; e-mail: roman.neronov@mail.ru, тел.: +7(929)128-44-54

8. Если полученный результат устраивает пользователя, он может выгрузить свой корпус с метками кластеров, а также догрузить дополнительный корпус документов, который будет классифицирован на основе меток, полученных при кластеризации.

3. Этапы разведочного анализа текстов. Рассмотрим основные этапы цикла анализа документов.

3.1. Загрузка текстов. Для загрузки корпуса текстов доступен ряд широко используемых форматов файлов, включая JSON, XLSX, Parquet, CSV и NPY. После их загрузки осуществляется трансформация данных в формат DataFrame.

3.2. Обработка текстов. Производится нормализация текстов путем трансформации всех символов к нижнему регистру и исключения из текста избыточных специальных символов и пробелов. Далее, с целью обеспечения мультиязычной совместимости обработки данных, осуществляется определение языка текста с использованием библиотеки langdetect.

3.3. Векторные представления для текстов. Для получения векторного представления текстов была использована предобученная модель BERT: distiluse-base-multilingual-cased-v2 с максимальной длиной последовательности токенов – 128 и выходным вектором размером 512. Данная модель является мультиязычной и подходит для 50 языков, среди которых присутствуют английский и русский.

Длина входного текста может превосходить 128 токенов, тогда возникает проблема потери информации об остальной части документа. Чтобы решить данную проблему, для каждого текста высчитывались первые 2048 токенов, затем данная последовательность разбивалась на подпоследовательности по 128 токенов и уже для каждой подчасти высчитывался числовой вектор. Потом все векторы усреднялись и получался вектор, который отражал большую часть текста.

3.4. Оптимальное число кластеров. Для использования алгоритмов кластеризации, которым нужно предоставлять число кластеров, требуется найти оптимальное значение для данного гиперпараметра с помощью стороннего алгоритма.

Для решения этой проблемы был выбран метод снижения размерности представлений UMAP [1], с последующей кластеризации

ей. Данный подход работает быстрее других алгоритмов (Truncated SVD, PCA, t-SNE), решающих задачу понижения размерности, и лучше сохраняет структуры данных. Более того, UMAP дает теоретические гарантии качества представления данных.

Таким образом, сначала с помощью UMAP понижаем векторные представления до представлений низкой размерности, а потом применяем алгоритм кластеризации на различных значениях количества кластеров и находим наиболее оптимальные значения, минимизирую определенную метрику.

В качестве алгоритма для определения оптимального числа кластеров использовался модифицированный алгоритм k -means – Bisecting k -means, который основан на идее рекурсивного двоичного разделения кластеров.

Для определения наилучшего разбиения использовалась метрика на основе индекса Калинского – Харабаза.

3.5. Первичная кластеризация текстов. Первичная кластеризация после получения оптимального числа кластеров проводится с помощью k -means в исходном пространстве признаков.

3.6. Ключевые слова для кластеров. Для выделения ключевых слов для каждого кластера использовался keyBERT [2]. Рассмотрим этапы работы данного алгоритма:

1. Выделяем список кандидатов-слов. Удаляем стоп-слова.
2. Вычисляем векторные представления для текста и для каждого кандидата.
3. Выбираем из кандидатов наиболее схожие с текстом и формируем список топ-слов. Для расчета сходства между кандидатами и документом используется косинусное расстояние между векторами.
4. Для повышения различия между ключевыми словами используется алгоритм максимального суммарного сходства: максимизируется сходство кандидата с документом и одновременно минимизируется сходство между кандидатами.

Чтобы найти ключевые слова для всего кластера в целом, сначала находим топ шесть слов для каждого текста отдельно, затем подсчитываем частоты для каждого слова среди всех ключевых слов для текстов из одного кластера и выбираем десять наиболее часто встречающихся слов. При этом, чтобы кластеры не имели схожих

ключевых слов, удалим те, которые встречаются в топе слов более чем у 40% кластеров.

3.7. Взаимодействие с кластерами и обучение модели. У пользователя есть два способа взаимодействия с кластерами: разделение кластеров (пользователь задает количество новых кластеров) и объединение кластеров (пользователь выбирает кластеры, которые хочет объединить).

Для новых кластеров пересчитываются списки ключевых слов и их центры на основе входящих в них документов и их векторных представлений.

4. Программная реализация. Практической реализацией плана является программа «ETA» [3,4], которая предназначена для решения задачи разведочного анализа текстов с применением модели Supporting Clustering with Contrastive Learning [5] для кластеризации. Она обеспечивает удобный и интуитивно понятный интерфейс, способствующий упрощению процесса анализа данных и интерпретации результатов.

К преимуществам данной программы (по отношению к известным) относятся:

1. Интерактивная работа с данными: пользователь может редактировать кластерную структуру, которая была получена в ходе обработки.
2. Дообучение модели кластеризации по требованию пользователя на основании получившегося разбиения текстов.
3. Отсутствие необходимости ручного подбора параметров для функций кластеризации, подбора ключевых слов и аппроксимации числа кластеров.
4. Инкапсуляция всех этапов загрузки, обработки и постобработки текстовой информации в виде единого интерфейса.
5. Поддержка Metal Performance Shaders backend.

5. Заключение. В данной работе рассмотрен план программы для ЭВМ, предназначенной для выполнения разведочного анализа текстов, который покрывает ключевые аспекты данной задачи. В дальнейшем, чтобы расширить получаемую аналитику по данным, нужно конкретизировать предметную область для исследуемых документов, что позволит определить специфические методы, нацеленные на решение конкретных прикладных задач.

Литература

1. McInnes L. et al. UMAP: Uniform Manifold Approximation and Projection // Journal of Open Source Software. 2018. Vol. 3. No 29. Art. no 861.
2. Priyanshu A., Vijay S. AdaptKeyBERT: An Attention-Based approach towards Few-Shot & Zero-Shot Domain Adaptation of KeyBERT [Электронный ресурс]: URL:<https://arxiv.org/abs/2211.07499> (дата обращения: 17.03.2024).
3. Неронов Р. М. Программа для решения задачи разведочного анализа текстов на основе кластеризации и аппроксимации числа кластеров (ETA) / пат. 2024612451 РФ № 2024611258, 01.02.2024.
4. Библиотека AnaText [Электронный ресурс]: URL:<https://github.com/RomiconEZ/AnaText> (дата обращения: 17.03.2024).
5. Zhang D. et al. Supporting clustering with contrastive learning // Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2021. P. 5419–5430.

УДК 004.932.1

Никитин А. В.

Применение нейросетевых методов для сегментации болезней культурных растений по цветным снимкам в точном земледелии

Рекомендовано к публикации доцентом Блекановым И. С.

1. Введение. В эпоху стремительного роста населения нашей планеты вопрос обеспечения продовольственной безопасности становится более актуальным, чем когда-либо. Потребность в качественной и здоровой пище многократно возросла, и для многих стран сельское хозяйство остаётся краеугольной отраслью, поддерживающей жизнеспособность экономики. Так, например, по состоянию на 2023 год Россия занимает 3-е место в рейтинге мировых лидеров по производству зерна [1] и 1-е место по экспорту зерна [2]. В условиях такого динамичного изменения сельскохозяйственного ландшафта встает вопрос не только о количестве произведенных продуктов, но и об их качестве. Для достижения высокого качества урожая становится неотложной задачей раннее выявление заболеваний растений. Однако этот процесс усложняется тем, что специалистам часто бывает сложно своевременно определить текущее состояние растений. В этих условиях бурно развивается направление на стыке информатики и сельского хозяйства в части автоматизации процесса ухода за культурами (см. напр. [3]). В данной работе решается задача сегментации поражённых участков культурных растений по цветным снимкам с помощью нейросетевых подходов. Проводится сравнительный анализ различных архитектур. Показано, что модели, основанные на архитектуре Transformer [4], решают эту задачу с отличной точностью.

2. Постановка задачи. Нейросетевая модель получает на вход цветные изображения сельскохозяйственных культур. Нужно обучить модель верно предсказывать поражённые области растения. В качестве ответа модель выдаёт сегментационную маску, на которой выделены нездоровые участки растений.

Никитин Александр Викторович – студент, Санкт-Петербургский государственный университет; e-mail: st087322@student.spbu.ru, тел.: +7(958)139-78-13

3. Набор данных. Для обучения модели был выбран открытый набор данных [5], состоящий из 2352 изображений листьев различных сельскохозяйственных растений, поражённых некоторым заболеванием, и сегментационных масок к ним. Примеры изображений представлены на рис. 1.



Рис. 1. Примеры изображений и их масок

4. Протестированные модели. Несмотря на то, что для семантической сегментации изображений могут использоваться модели, основанные на операции свёртки [6], для оценки эффективности были протестированы три различные архитектуры, основанные на механизме внимания: SegFormer [7], MobileViT [8] и UPerNet [9]. Модель MobileViT представляет собой облегчённую версию архитектуры ViT [10]. В качестве основы у UPerNet выбрана модель Swin [11]. Из семейства моделей SegFormer было протестировано шесть: SegFormer-b0 – SegFormer-b5. Предварительно обученные версии этих моделей взяты с платформы HuggingFace. После этого проведена процедура fine-tuning на выбранном датасете, а именно, исследованы два метода дообучения моделей. В первом случае все веса модели были подвергнуты дообучению на найденном наборе данных. Во втором случае использована техника Parameter Efficient Fine-Tuning, а именно метод LoRA (Low Rank Adapter) [12]. Суть алгоритма LoRA заключается в замораживании всех слоев модели, при

этом к некоторым из них добавляются обучаемые параметры, представляющие разложение матрицы весов слоя на произведение двух матриц более низкого ранга. Эта стратегия сокращает количество обучаемых параметров, что приводит к сокращению времени обучения и успешной адаптации крупной предобученной модели к узкому домену задачи сегментации болезней культурных растений.

5. Результаты экспериментов. Для экспериментов набор данных был разделён на тренировочную, валидационную и тестовую части в соотношении 0,75 : 0,1 : 0,15 соответственно. Все модели обучались на тренировочном наборе данных, после чего тестировались на валидационном наборе. Результаты оценивались по трём метрикам: Intersection over Union (IoU), Dice coefficient и Cohen’s kappa.

5.1. Дообучение всех весов. При дообучении всех весов был использован оптимизатор AdamW, так как он показывает более стабильные результаты

Гиперпараметры обучения с оптимизатором AdamW представлены в таблице 1.

Таблица 1. Гиперпараметры обучения с AdamW

Количество эпох	Скорость обучения	Размер батча	β_1	β_2	Уменьшение весов
10	0,00006	5	0,9	0,999	0,01

Результаты обучения представлены в таблице 2. Можно сделать вывод, что семейство моделей SegFormer справляется с задачей лучше всего. Чем глубже модель, тем лучший результат она показывает. Наилучший результат показала модель SegFormer-b5.

Таблица 2. Сравнение моделей

Модель	IoU	Dice	Кappa
SegFormer-b0	0,921	0,959	0,892
SegFormer-b1	0,928	0,962	0,907
SegFormer-b2	0,943	0,971	0,914
SegFormer-b3	0,945	0,971	0,926
SegFormer-b4	0,947	0,973	0,926
SegFormer-b5	0,948	0,973	0,928
Mobile-ViT	0,727	0,842	0,703
UPerNet-Swin	0,931	0,964	0,922

5.2. Метод LoRA. С помощью техники LoRA были дообучены модели SegFormer-b0 и SegFormer-b2. Гиперпараметры обучения представлены в таблице 3, а результаты обучения – в таблице 4.

Таблица 3. Гиперпараметры LoRA

r	alpha	target modules	dropout	bias
32	32	0,1	0,1	0,2

Таблица 4. Результаты дообучения с помощью LoRA

Модель	Доля обучаемых параметров	IoU
SegFormer-b0	0,14	0,848
SegFormer-b2	0,14	0,904

Как видно, вариант дообучения с помощью LoRA показывает более низкие результаты, в результате чего было принято решение о нецелесообразности использования этого метода для остальных, относительно небольших моделей. Однако эта техника может быть полезной при дообучении более крупных моделей, таких например, как Swin или SAM, сокращая время обучения.

6. Заключение. В данной работе применены новейшие архитектуры нейронных сетей для решения задачи сегментации болезней культурных растений. Эксперименты показали отличные результаты, в связи с чем нейросетевые подходы могут активно внедряться в сельское хозяйство, улучшая урожайность полей и экономя как человеческие, так и материальные ресурсы.

Литература

1. Top 10 Wheat Producing Countries 2023 [Электронный ресурс]: URL:<https://www.geeksforgeeks.org/top-wheat-producing-countries/> (дата обращения: 7.03.24).
2. Россия в 2023 г. восьмой год подряд осталась крупнейшим в мире экспортером пшеницы [Электронный ресурс]: URL:<https://dzen.ru/a/ZeVurLMsSQc7oigW> (дата обращения: 7.03.24).
3. Молин А. Е., Блеканов И. С., Митрофанов Е. П., Митрофанова О. А. Методы генерации синтетических данных для обучения нейросетей в задаче сегментации уровня азотного режима растений на снимках беспилотных летательных аппаратов на сельскохозяйственном поле // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2024. Т. 20. № 1. С. 20–33.

4. Vaswani A., Shazeer N., Parmar N. et al. Attention is all you need // Advances in Neural Information Processing Systems – NeurIPS Proceedings. 2017. Vol. 30. P. 5998–6008.
5. Leaf Disease Segmentation Dataset [Электронный ресурс]: URL:<https://www.kaggle.com/datasets/fakhrealam9537/leaf-disease-segmentation-dataset> (дата обращения: 1.12.23).
6. Сугоняев А. Д. Применение сверточных нейросетей в задаче семантической сегментации фотографий // Процессы управления и устойчивость. 2019. Т. 6. № 1. С. 363–367.
7. Xie E., Wang W., Yu Z. et al. SegFormer: simple and efficient design for semantic segmentation with transformers // Advances in Neural Information Processing Systems – NeurIPS Proceedings. 2021. Vol. 34. P. 12077–12090.
8. Mehta S., Rastegari M. MobileViT: light-weight, general-purpose, and mobile-friendly vision transformer [Электронный ресурс]: URL:<https://arxiv.org/abs/2110.02178v2> (дата обращения: 2.12.23).
9. Xiao T., Liu Y., Zhou B. et al. Unified perceptual parsing for scene understanding // Computer Vision – ECCV 2018. Lecture Notes in Computer Science. Springer, Cham. 2018. Vol. 11209. P. 418–434.
10. Dosovitskiy A., Beyer L., Kolesnikov A. et al. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale [Электронный ресурс]: URL:<https://arxiv.org/abs/2010.11929> (дата обращения: 2.12.23).
11. Liu Z., Lin Y., Cao Y. et al. Swin transformer: hierarchical vision transformer using shifted windows // 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada, 2021. P. 9992–10002.
12. Hu E., Shen Y., Wallis P. et al. LoRA: Low-Rank Adaptation of Large Language Models [Электронный ресурс]: URL:<https://arxiv.org/abs/2106.09685> (дата обращения: 3.12.23).

УДК 004.8

Павлов А. В.

Методы точной настройки: RAG vs LoRA

Рекомендовано к публикации доцентом Блекановым И. С.

1. Введение. Большие языковые модели в последние годы получили широкое распространение и популярность благодаря возможности генерировать естественные диалоги. Однако их функциональные возможности имеют определенные ограничения, связанные с объемом и характером данных, используемых для обучения.

Адаптация больших языковых моделей для работы в произвольных предметных областях является сложной и нетривиальной задачей. В данной работе рассматриваются два подхода к ее решению – RAG и LoRA. Отдельное внимание уделяется их сравнительному анализу с точки зрения применимости в различных предметных областях.

2. Методы точной настройки и их реализация. Приведем необходимую теорию для понимания используемых методов точной настройки. Под точной настройкой понимается процесс адаптации предварительно обученной большой языковой модели к предметной области или конкретной задаче путем ее дальнейшего обучения на новом, значительно меньшем наборе данных.

Традиционные методы точной настройки подразумевают обновление всех весовых коэффициентов предварительно обученной модели. Этот подход может быть дорогостоящим с точки зрения вычислений и памяти [1], особенно для архитектур с миллиардами параметров. Помимо этого, при использовании такого подхода может произойти ситуация, когда модель потеряет способность хорошо выполнять первоначальную задачу при адаптации к новой.

Во избежание этих проблем были предложены различные методы точной настройки. Особенно выделяются генерация с расширенным поиском (RAG) и адаптация низкого ранга (LoRA), на которые было сосредоточено значительное внимание.

Павлов Алексей Владимирович – студент, Санкт-Петербургский государственный университет; e-mail: st086869@student.spbu.ru, тел.: +7(921)439-64-06

2.1. Система RAG. Она состоит из двух компонент, предварительно обученной языковой модели и поисковой системы. Языковая модель отвечает за генерацию текста, в то время как поисковая система используется для извлечения релевантной к запросу информации из внешнего источника знаний. Генерация с расширенным поиском работает следующим образом:

1. *Обработка входных данных.* Входной текст обрабатывается поисковой системой, которая использует такие методы, как разреженное векторное сходство или плотное векторное сходство для поиска наиболее релевантных документов или отрывков из источника знаний. Построение внешнего источника знаний показано на рис. 1.

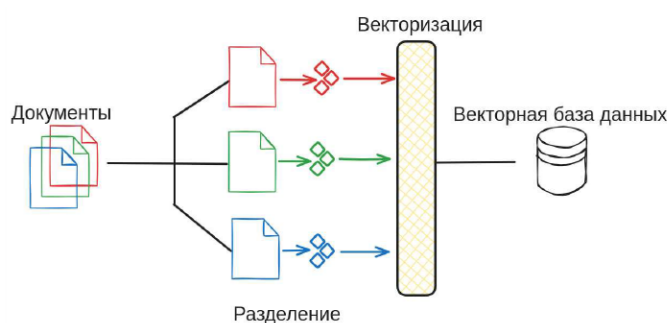


Рис. 1. Построение внешнего источника знаний

2. *Извлечение.* Извлекаются наиболее релевантные документы или отрывки из них на основе сходства с входным текстом. Эти извлеченные документы используются в качестве дополнительного контекста для языковой модели.
3. *Генерация.* Предварительно обученная языковая модель принимает входной текст и извлеченные документы в качестве входных данных и генерирует окончательный выходной текст. В дополнение к этому языковая модель также может быть обучена для наилучшего обусловливания своей генерации извлеченным контекстом, что позволит ей включать наиболее релевантную информацию из внешнего источника знаний. На рис. 2 показана схема генерации.

Для оценки эффективности подхода использован «утопичный» подход. Для проверки концепции сделано предположение, что поисковая часть работает идеально и возвращает наиболее релевантный контекст.



Рис. 2. Схема генерации с расширенным поиском

2.2. Метод LoRA. Это один из эффективных по параметрам методов точной настройки [1]. Его основная идея заключается во введении небольшого количества обучаемых параметров, менее 0,1% от исходного размера модели, которые добавляются к существующим весам модели во время точной настройки. Такой подход позволяет проводить эффективную точную настройку, достигая при этом конкурентоспособных характеристик по сравнению с полной точной настройкой модели. Процесс на основе LoRA состоит из следующих этапов:

1. *Построение матрицы низкого ранга.* Для каждого выбранного слоя в предварительно обученной модели LoRA добавляет две матрицы низкого ранга: матрицу весов и матрицу смещений. На рис. 3 представлена схема с матрицами низкого ранга.
2. *Точная настройка.* Во время точной настройки обновляются только матрицы низкого ранга, в то время как предварительно обученные веса модели остаются замороженными. Матрицы низкого ранга действуют как адаптеры для конкретной задачи, позволяя модели адаптироваться к новой задаче или предметной области без изменения исходных предварительно обученных параметров.

Согласно работе [2] количество обучаемых параметров можно уменьшить в 10000 раз, а требования к памяти графических ускорителей

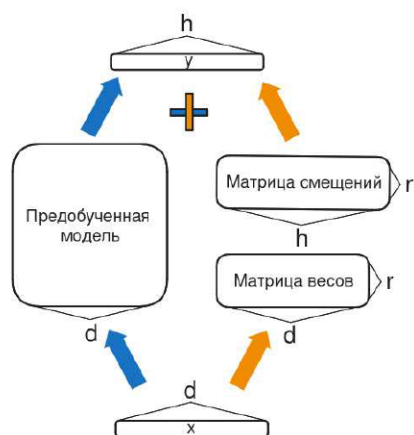


Рис. 3. Матрицы низкого ранга

второе. Для построения адаптации низкого ранга был использован подход QLoRA [3], который добавляет квантизацию. Найдены и выбраны параметры:

- r – ранг матриц низкого ранга. В результате экспериментов с различными значениями r было обнаружено, что $r = 64$ обеспечивает наилучший компромисс между точностью и эффективностью;
- *dropout* – коэффициент отсева для матриц низкого ранга. Был использован коэффициент отсева, равный 0,1, что помогло предотвратить переобучение.

3. Эксперимент. В работе поставлен эксперимент по оценке качества и сравнительному анализу метода генерации с расширенным поиском (RAG) и метода на основе адаптации низкого ранга (LoRA). Для данных методов в качестве базовой была использована языковая модель “falcon” [4]. Для обучения моделей и оценки был выбран набор данных [5], содержащий вопросы и ответы, а также их контекст.

В эксперименте качество рассматриваемых подходов оценивалось с помощью восьми метрик [6, 7]: BLEU, ROUGE-1, ROUGE-2,

ROUGE-L, METEOR, Precision, Recall, F1. В таблице 1 представлены результаты оценки качества.

Таблица 1. Результаты оценки качества трех моделей точной настройки

Метрики	Базовая модель falcon	LoRA модель	RAG модель
BLEU	0,3879	0,5242	0,6543
ROUGE-1	0,6553	0,7504	0,8582
ROUGE-2	0,5259	0,6754	0,7613
ROUGE-L	0,6259	0,7447	0,8300
METEOR	0,6718	0,8014	0,8835
Precision	0,9261	0,9416	0,9691
Recall	0,9412	0,9639	0,9759
F1	0,9334	0,9525	0,9724

Из таблицы 1 видно, что наивысшие значения по всем качественным метрикам имеет метод точной настройки на основе подхода RAG. В частности, данный метод в среднем показывает прирост качества по восьми метрикам на 27% по сравнению с базовой большой языковой моделью и на 10% по сравнению с подходом на основе LoRA.

Таблица 2. Пример работы методов точной настройки

Контекст запроса	...book, Ed Wood's Sleaze Paperbacks, which is now available...
Вопрос из датасета	What is the name of the companion book released by the curators of Ed Wood's paperback books exhibition?
Экспертный ответ	The name of the companion book is "Ed Wood's Sleaze Paperbacks"
Ответ falcon	The companion book released by the curators of Ed Wood's paperback books exhibition is called "Ed Wood: The Life and Works of a Hollywood Legend".
Ответ LoRA	The name of the companion book is "Ed Wood's Paperback Book Exhibition". It is a book of essays, interviews, and illustrations by the curators of Ed Wood's paperback books exhibition. It includes essays by the curators, interviews with Ed Wood's family and friends, and illustrations by the author.
Ответ RAG	The name of the companion book released by the curators of Ed Wood's paperback books exhibition is "Ed Wood's Sleaze Paperbacks".

В таблице 2 на реальном сценарии из датасета [5] приведен пример генерации текста тремя методами точной настройки.

4. Заключение. В данной работе продемонстрировано, что применение метода генерации с расширенным поиском является более эффективным, чем метод генерации на основе LoRA.

В дальнейшем планируется совместить эти два подхода, использовать адаптацию низкого ранга в качестве адаптера для тренировки модели ради более лучшего обуславливания генерации контекстом.

Литература

1. Ning D., Qin Y., Guang Y. et al. Parameter-efficient fine-tuning of large-scale pre-trained language models // *Nature Machine Intelligence*. 2023. Vol. 5. P. 220–235.
2. Hu J. E., Shen Y., Wallis P. et al. LoRA: low-rank adaptation of large language models // *arXiv:2106.09685*. 2021. P. 1–26.
3. Dettmers T., Pagnoni A., Holtzman A., Zettlemoyer L. QLoRA: efficient finetuning of quantized LLMs // *Advances in Neural Information Processing Systems*. 2023. Vol. 36. P. 1–26.
4. HuggingFace [Электронный ресурс]: URL:<https://huggingface.co/ericzzz/falcon-rw-1b-instruct-openorca> (дата обращения: 01.03.2024).
5. HuggingFace [Электронный ресурс]: URL:<https://huggingface.co/datasets/neural-bridge/rag-dataset-12000> (дата обращения: 01.03.2024).
6. Yang A., Liu K., Liu J. et al. Adaptations of ROUGE and BLEU to better evaluate Machine Reading comprehension task // *In Proceedings of the Workshop on Machine Reading for Question Answering*. ACL. 2023. P. 98–104.
7. Evtikhiev M., Bogomolov E., Sokolov Y., Bryksin T. Out of the BLEU: how should we assess quality of the Code Generation models? // *Journal of Systems and Software*. 2023. Vol. 203. Art. no. 111741.

УДК 004.89

Павлов М. В.

Нейросетевые методы тестирования игрового процесса на основе обучения с подкреплением в разработке компьютерных игр

Рекомендовано к публикации доцентом Блекановым И. С.

1. Введение. В настоящее время разработка компьютерных игр (*gamedev*) является быстроразвивающейся индустрией. При этом научные исследования последних лет показывают, что методологии и достижения в разработке компьютерных игр находят применение не только в сфере развлечений, но и в других областях:

- медицине (например, в терапии расстройств аутистического спектра) [1];
- образовании (обучение программированию и другим дисциплинам) [2];
- бизнесе (геймификация корпоративных систем) [3].

При разработке игр команды нередко сталкиваются с рядом глобальных проблем. К ним можно отнести как технические ошибки (недостаточная оптимизация, баги), так и ошибки в дизайне (неясное видение дизайна игры, отсутствие «фактора удовольствия») [4].

Все большее количество проблем исследователи пытаются решить с помощью использования искусственного интеллекта (ИИ). Основными сферами применения ИИ являются создание правдоподобных неигровых персонажей (*non-playable character*, NPC), процедурная генерация контента, моделирование опыта игроков, помощь с геймдизайном [5].

Одним из факторов возникновения проблем является ручное тестирование, преобладающее в разработке видеоигр [6], которое не всегда позволяет выявить возможные ошибки. Оно полностью полагается на *ad-hoc* мышление игровых тестировщиков, играющих в одну и ту же игру снова и снова. Все чаще игры выходят с ошибками, которые разработчики вынуждены исправлять в последующих

Павлов Михаил Вячеславович – студент, Санкт-Петербургский государственный университет; e-mail: st084445@student.spbu.ru, тел.: +7(921)851-16-67

обновлениях [7]. В этой связи актуальной задачей является автоматизирование тестирования игрового процесса, отличающегося от «традиционного» процесса разработки ПО и слабо поддающегося устоявшимся методам автоматизации [4].

Данная работа посвящена тестированию игрового процесса с помощью нейросетевых подходов на основе обучения с подкреплением (*reinforcement learning*, RL) с целью повышения качества разработки, а также минимизации затрачиваемых времени и ресурсов на тестирование.

2. Нейросетевые методы и технологии тестирования игр.

Условно выделяются три метода обучения игровых агентов с помощью RL: обучение по пикселям, по компактным представлениям состояний, а также обучение на основе моделирования игрового процесса (через мультиагентное обучение с подкреплением) [4].

- Метод *обучения по пикселям*, использующий выходные видеоданные игры, хорошо показал себя в играх для консоли *Atari 2600*, разрешение которой составляло 210×160 пикселей [8]. Применение такого подхода в современных играх, ориентированных на более высокие разрешения, не представляется целесообразным;
- обучение игровых агентов на основе *компактного представления состояний* (таких как положение игрока, расстояние до соседних объектов или до пункта назначения и т. д.) применяется для минимизации информации о пространстве состояний на каждом временном шаге [9];
- метод на основе *моделирования игрового процесса* для многопользовательских игр позволяет тестировать игровой процесс путем самостоятельной игры игровых агентов. Данная методология лежит в основе программ *AlphaZero* и *AlphaGo Zero*, разработанных для игры в шахматы, го и сёги [10].

Описанные выше методологии обучения с подкреплением имеют следующие преимущества [11]:

- RL-агенты способны обучаться на основе взаимодействия с игровой средой. В результате поведение агентов становится более похожим на поведение игроков-людей, что повышает вероятность обнаружения багов и ошибок, связанных с устройством самой среды;

- нейросетевой агент более устойчив к изменениям и обновлениям игровой среды: он может быть переобучен или доработан с минимальными изменениями;
- поведение агента может контролироваться путем изменения получаемой награды. Таким образом можно симулировать поведение игроков с разным игровым опытом и стилем игры.

3. Предлагаемое решение. В данной работе реализован подход к тестированию игрового процесса по компактным представлениям состояний, основанный на нейросетевой архитектуре *Proximal Policy Optimization (PPO)*.

3.1. Архитектура предлагаемого решения. Команды агенту (класс `CarController`) могут передаваться как напрямую с помощью выбранного устройства ввода, так и через класс `CarAgent`, выступающий в качестве обертки для нейросети. Архитектура решения представлена на рис. 1.

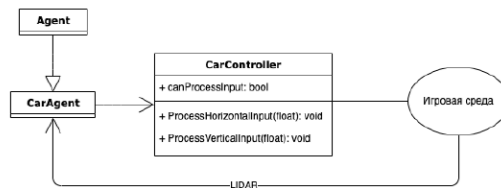


Рис. 1. Схематическое изображение основной архитектуры проекта

3.2. Эксперимент. В работе был поставлен эксперимент по оценке качества предложенного метода тестирования на основе PPO, в котором для наглядности разработана эмуляционная игровая среда *Race Against the Machine* – аркадная гоночная игра на движке *Unity 2021.3.14f1* (см. рис. 2). Игроком (агентом) в *Race Against the Machine* выступает электромобиль. Цель игры – пройти заданное количество кругов за минимально возможное время. Результаты загружаются на сервер и становятся доступными другим игрокам через меню «Rankings».

Этап тренировки агента организован следующим образом:

- в течение 10 тысяч шагов агент исследовал игровую среду, ему присуждалась награда в случае первичного нахождения локации вида *out of bounds* (локация, находящаяся за пределами



Рис. 2. Кадр геймплея эмуляционной игровой среды

предполагаемых границ игры) либо локации с числом кадров в секунду меньше 10, т. е. ниже округленного медианного значения в условиях тренировки;

- по прошествии эпохи позиция и угол поворота агента сбрасывались и инициализировались заново, а веса модели обновлялись с использованием собранных наблюдений.

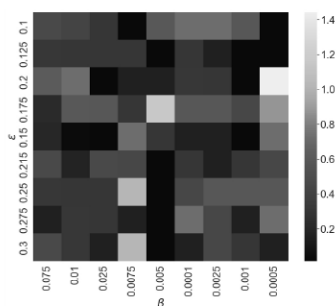


Рис. 3. Тепловая карта значений гиперпараметров ϵ , β и суммарной награды на момент остановки обучения

В рамках эксперимента также был проведен поиск оптимальных значений гиперпараметров ϵ и β , используемых в алгоритме PPO. Результат оптимизации путем поиска по решетке приведен на рис. 3.

В общем случае нейросетевой агент тренировался 100 эпох (10^6 шагов), и этап тренировки был повторен 10 раз для получения усредненных значений. Результаты, приведенные на рис. 4, показывают, что получаемое агентом за эпоху среднее значение награды стабилизируется в процессе обучения.

4. Заключение. Разработан метод тестирования на основе архитектуры PPO, с помощью которого в эмуляционной игровой среде было найдено 293 ошибок. Точки проявления ошибок формируют

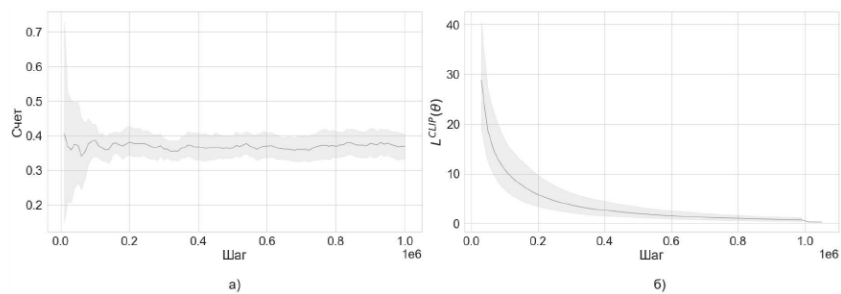


Рис. 4. а) кривая обучения модели, б) график функции потерь в десяти экспериментах. Для каждой эпохи приведены выборочное среднее и 95%-й доверительный интервал

кластеры, явно выделяя «проблемные» области. Данные результаты показывают состоятельность и эффективность предложенного подхода.

В дальнейшем планируется расширить эксперимент и сравнить предлагаемое решение с нейросетевым подходом на основе архитектуры *Soft Actor-Critic* (SAC).

Литература

1. Hassan A., Pinkwart N., Shafi M. Serious games to improve social and emotional intelligence in children with autism // *Entertainment Computing*. 2021. Vol. 38. Art. no 100417.
2. Mayer R. E. Computer games in education // *Annual review of psychology*. 2019. Vol. 70. P. 531–549.
3. Augustin K., Thiebes S. et al. Are we playing yet? A review of gamified enterprise systems // *PACIS 2016 Proceedings*. 2016.
4. Politowski C., Petrillo F., Guéhéneuc Y. G. A survey of video game testing // *2021 IEEE/ACM International Conference on Automation of Software Test (AST)*. 2021. P. 90–99.

5. Xia B., Ye X., Abuassba A. O. M. Recent research on AI in games // 2020 International Wireless Communications and Mobile Computing (IWCMC). 2020. P. 505–510.
6. Santos R. E. S., Magalhães C. V. C. et al. Computer games are serious business and so is their quality: particularities of software testing in game development from the perspective of practitioners // Proceedings of the 12th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement. 2018 P. 1–10.
7. Truelove A., de Almeida E. S., Ahmed I. We'll fix it in post: what do bug fixes in video game update notes tell us? // 2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE). 2021. P. 736–747.
8. Mnih V., Kavukcuoglu K. et al. Playing Atari with deep reinforcement learning // arXiv preprint arXiv:1312.5602. 2013.
9. Borovikov I., Zhao Y. et al. Winning isn't everything: training agents to playtest modern games // AAAI Workshop on Reinforcement Learning in Games. 2019.
10. Silver D., Hubert T. et al. Mastering chess and shogi by self-play with a general reinforcement learning algorithm // arXiv preprint arXiv:1712.01815. 2017.
11. Bergdahl J., Gordillo C. et al. Augmenting automated game testing with deep reinforcement learning // 2020 IEEE Conference on Games (CoG). 2020. P. 600–603.

УДК 519.175.3

Пономарев Ю. Ю.

Вопросы точных оценок характеристик деревьев при моделировании метаданных в информационных системах общего назначения

Введение. Современные информационные системы порождают термы определенной глубины вложенности [1]. Частным случаем порожденных термов являются метаданные любой природы. Общие постановки анализа и синтеза метаданных привели к созданию не только богатых результатами разделов системного анализа, теории хранилищ и структур данных в СУБД, но и выдвинули ряд ранее не обусловленных задач работы в информационных системах большого объема метаданных для анализа и синтеза деревьев. Классические результаты описаны в монографии Ф. Харари и Э. Палмера [2], однако заметим, что при создании информационных систем (ИС) важны прогнозные показатели роста объема данных исходя из метаданных. В работе [3] приведен ряд эффективных теорем о системе DOWN-UP кодов, дающих решение множества практических задач, но последний блок гипотез для метаданных необходимо исследовать более тщательно.

Описание предметной области. Далее будем следовать определениям понятий предметной области, опираясь на статьи [1, 3].

Предметная область \mathbf{D} спецификации $(c_0^0, c_1^{m_1}, \dots, c_n^{m_n})$ – это совокупность термов, получаемых из данного количества $c_0 \neq 0$ знаков переменных x_1, \dots, x_{c_0} и данных количеств c_i m_i -местных функциональных знаков

$$f_1^{(m_1)}, \dots, f_{c_i}^{(m_i)},$$

где $i = 1, \dots, n$, соединяемых с помощью композиции.

Терм, как обычно, определяется индуктивно: 1) переменная есть терм, 2) если $f_{(i)}$ – i -местный функциональный знак и a_1, \dots, a_i –

Пономарев Юрий Юрьевич – эксперт, ООО «Автоматизированные системы, консалтинг, информационные технологии»; e-mail: yura_graph@mail.ru, тел.: +7(903)537-46-11

термы, то $f_{(i)}(a_1, \dots, a_i)$ – терм, 3) никаких других термов не существует.

Глубиной вложенности данного вхождения терма $a \in \mathbf{D}$, как подтерма в терм $b \in \mathbf{D}$, назовем число термов, содержащих данное вхождение. Максимальную из глубин вложенности всех вхождений переменных в терм назовем глубиной терма. На языке функциональных деревьев, представляющих структуру терма, глубина терма – высота соответствующего функционального дерева.

Постановка задачи. Для метаданных предметных областей требуется провести анализ DOWN-UP кодов. По возможности, уточнить уже имеющиеся результаты [3] с целью более точного описания функционального дерева. Провести сравнение системы кодирования с существующими фреймворками.

Основные результаты. Сформулируем основные результаты в виде следующих утверждений.

Теорема 1. *Точная оценка средней высоты (изоморфных) корневых деревьев $h(n)$ с учетом результатов для числа вершин $f(A)$ по множеству DOWN-UP кодов A для заданного числа вершин n обладает средним значением $h(n)$ по n в два раза меньше, чем указано в работе [1].*

Доказательство. Суммарная высота деревьев по максимальной оценке $4^{(n-2)}$ для корневых деревьев числом вершин $n = 3, \dots, 16$ – это последовательность $4, 16, 64, \dots, 4^{(n-2)}$. При прямом построении, согласно [4], имеем: 2, 13, 60, 251, 1018, 4089, 16376, 65527, 262134, 1048565, 4194292, 16777203, 67108850, 268435441. Сравниваем число кодов по формуле Волошина о средней высоте корневых деревьев из [1] и при прямом построении [4] – они совпадают с большой точностью для $n < 100$. Однако, если найти значение средней высоты по формуле Волошина и при прямом построении, то разница будет ровно два раза. Это означает, что в формуле Волошина имеется опечатка или ошибка ровно в два раза. Рассмотрим ряд Волошина для средней высоты [1]: 4, 6, 9, 12, 15, 19, 22, 26, 31, 35, 40, 45, 50, 55. Ряд по построению [4]: 2, 3, 4, 6, 7, 9, 11, 13, 15, 17, 20, 22, 25, 27. Оба ряда округлены до целых отбрасыванием значений дробной части.

Теорема 2. *Основная работа с деревьями ведется только в очень узком классе так называемых «критических кодов», т. е. таких кодов, где четные элементы принимают максимально возможное значение.*

Доказательство. Будем использовать определение критического кода из работ [4, 5]. Все элементы на четных местах имеют максимально возможные значения. Рассмотрим любой стандарт кодирования XML информации в России, например [6]. Он имеет общий вид: $\langle \text{rows} \rangle \langle \text{row} \rangle \dots \langle / \text{row} \rangle \langle \text{row} \rangle \dots \langle / \text{row} \rangle \langle \text{row} \rangle \dots \langle / \text{row} \rangle \langle \text{row} \rangle \dots \langle / \text{row} \rangle \langle / \text{rows} \rangle$. Переводя на язык DOWN-UP кода получаем, что код такого корневого дерева будет содержать столько поддеревьев, исходящих из общего корня, сколько записей рассматривается, а точкой ветвления будет тот уровень, на котором стоят $\langle \text{row} \rangle \dots \langle / \text{row} \rangle$ тэги. Так как обычно информационная составляющая XML ровно на один уровень глубже точки ветвления, то порождаемый код — критический. Оценка числа критических кодов — часть таблицы с убывающим после максимума числом кодов из [3].

Теорема 3. *Создание алгоритмов, известных в литературе как “Rete” [7], и их уточнение для разных классов графов может осуществляться в терминах оптимального кодирования объектов.*

Доказательство. Алгоритмы типа Rete строятся на базисах из бизнес-правил и сопоставления их с образцами вывода требуемой структуры и информационного наполнения. Отсюда получаем, что естественными основами для такого сопоставления служат как алгоритмы естественного анализа графа (нахождения подграфов, автоморфизмов, изоморфизмов, раскрасок, критических путей), так и отдельное направление — создание системы оптимального кодирования графов для этих задач. В статьях [3–5] рассматривается только один из оптимальных способов кодирования, для которого получено достаточно много хороших свойств и оптимальных в своем классе алгоритмов.

Гипотеза 1. *Анализ реальных данных на покрытие всех их реляционных отношений силами любой ИС в динамике ее развития — почти не разрешенная проблема.*

Для конкретизации материала напомним определение DOWN-UP кодов, которые определяются рекурсивно.

Здесь и далее, для простоты изложения, положим $A = a_1, a_1 = 1$, либо $A = 1$ — граничное условие на код, единичный код. В общем виде код есть последовательность из элементов с четными и нечетными индексами: $A = \{ a_i, i \in (2n + 1), a_i, n \in N, a_i, i \in (2n), n \in N, a_i \in N \cup 0 \}$, при этом элементы, имеющие четные индексы строго меньше обеих соседствующих с ними элементов с нечетными индексами, $|A|$ — число элементов в коде или размерность кода, число

всегда нечетное, $|A| \geq 3$.

Как показано в [8], существует биекция между множеством корневых деревьев и DOWN-UP кодов, что позволяет распространять результаты, полученные для DOWN-UP кодов, на все множество упорядоченных корневых деревьев без потери общности сформулированных в данной статье утверждений. Для удобства перечисления деревьев, DOWN-UP коды создаются с минимальными возможными нечетными элементами (и, соответственно, минимальными четными). Критическими кодами назовем такие коды, где элементы с четными индексами принимают максимально возможные значения.

В статьях [3–5] представлено несколько теорем и следствий из них, которые утверждают, что многие особенности DOWN-UP кода могут быть установлены в виде описания его элементов на всем множестве изоморфных деревьев любой природы, причем есть прямые формулы для связи сложности и глубины дерева с количеством его конечных вершин.

Алгоритм Rete – эффективный алгоритм сопоставления с образцом для продукционных систем, экспертных систем и баз знаний, созданный Чарльзом Форги из Университета Карнеги – Меллона. Этот алгоритм впервые был описан им в рабочем документе 1974 г., затем в докторской диссертации 1979 г. и в статье [7] 1982 г. Rete стал основой многих популярных экспертных систем, включая CLIPS, Jess, Drools, BizTalk Rules Engine и Soag. К сожалению, последняя версия Rete-NT запатентована и закрыта.

Для управления метаданными существует ряд специализированных пакетов и фреймворков, ряд из которых требуют особого перечисления.

1. Общее управление данными: Ataccama ONE.
2. Для сложных ландшафтов данных: Collibra Data Governance.
3. Соблюдение нормативных требований: IBM Data Governance.
4. Просмотр корпоративных данных: erwin by Quest.
5. Аналитика на данных: Precisely Data360 Govern.
6. Цифровая трансформация: Informatica Data Governance.
7. Управление информационными рисками: OneTrust Data Governance.

8. Централизованное управление данными: SAP Master Data Governance.
9. Управление требованиями в продуктах BI & FMW Oracle: Oracle Enterprise Metadata Management.
10. Система автоматизированной передачи данных: Atlan.
11. Автоматизация документооборота по управлению данными: Alation.

Каждый из них – лидер в той или иной области управления метаданными.

Как показано в статье [9], наибольшую сложность мы обычно испытываем на перекрестке трех возможных путей создания ИС: прямого анализа информации для декомпозиции реляционных отношений, соотнесения метаданных об объектах, моделируемых в ИС, создания процедур очистки и поточного диагностирования данных для внесения в уже построенную и смоделированную (иногда и запрограммированную) систему пользовательских форм.

К сожалению, продуктов, которые работали эффективно в рамках существующих словарей данных для обычных объектно-реляционных, либо хранилищных моделей типа факт-измерение, почти не осталось. Корпорация Oracle (США) пытается связать моделирование, словари данных, типы архитектур воедино в ряде своих продуктов. Требуется заметить, что фрагментация знаний о метаданных, находящихся в СУБД, характеризуют только продукты 18 и 19 версий, что говорит о запоздалой реакции корпорации на рыночную среду.

Российские исследователи метаданных продолжают подвергать разумной критике принимаемые законы о персональных данных гражданина. Приведем несколько цитат из Федерального закона № 168-ФЗ «О едином федеральном информационном регистре, содержащем сведения о населении Российской Федерации».

Предлагается по конкурсным процедурам создать систему для обработки сведений о физическом лице:

- фамилия, имя и отчество (при наличии) и в случае их изменения – иные фамилия, имя и отчество (при наличии);
- дата рождения;
- дата смерти;
- место рождения;

- место смерти;
- пол и в случае его изменения иной пол;
- сведения о семейном положении физического лица (записи актов о заключении и расторжении брака);
- гражданство РФ и (или) гражданство (подданство) иностранного государства/иностраннных государств;
- сведения о наличии у гражданина РФ документа на право постоянного проживания в иностранном государстве (при наличии);
- сведения об обращении гражданина РФ в полномочный орган иностранного государства о его выходе из гражданства данного государства или об отказе от имеющегося у него документа на право постоянного проживания в иностранном государстве (при наличии).

Создать систему для обработки сведений об идентификаторах:

- записи акта о рождении;
- записи акта о смерти;
- документа, удостоверяющего личность физического лица, включая вид, номер и иные сведения о таком документе;
- документов или отметок в документах, удостоверяющих личность, подтверждающих право иностранного гражданина и лица без гражданства на пребывание (проживание) в РФ;
- сведений о регистрационном учете гражданина РФ и миграционном учете иностранного гражданина и лица без гражданства в РФ;
- сведений о принятом решении по вопросам гражданства РФ;
- сведений о постановке на учет в налоговом органе, в том числе в качестве налогоплательщика налога на профессиональный доход;
- сведений о регистрации физического лица в качестве индивидуального предпринимателя;
- сведений о постановке на воинский учет граждан РФ, обязанных на нем состоять;
- сведений о регистрации в системах обязательного пенсионного, медицинского и социального страхования;
- сведений о постановке на учет в органах службы занятости;

- документа об образовании и (или) о квалификации, документа об обучении, включая виды, номера и иные сведения о таких документах, сведений о присуждении, лишении, восстановлении ученой степени, ученого звания;
- учетной записи физического лица в федеральной государственной информационной системе «Единая система идентификации и аутентификации в инфраструктуре, обеспечивающей информационно-технологическое взаимодействие информационных систем, используемых для предоставления государственных и муниципальных услуг в электронной форме»;
- записей федерального регистра сведений о населении, о физических лицах, являющихся родителями физического лица, супругом (супругой) физического лица, ребенком (детьми) физического лица.

Заметим, что даже беглый анализ приведенных метаданных дает повод для углубленного размышления над многими законодательно неурегулированными моментами, приводящими к проблемам в госуправлении. Например, сколько у человека может быть идентификаторов за его жизнь? Будет ли законодательно устанавливаться равнозначность электронного и физического документа при потере того или иного идентификатора? Как бороться с кражей и подделкой цифрового следа физического лица? Почему юридические лица финансового сектора наделяют себя теми или иными полномочиями обработки метаданных и их интерпретации вне воли граждан РФ. Будет ли уравновешена правовая практика интерпретации е и ё в физическом и электронном представлении документов о личности гражданина? Насколько подлежат защите данные и метаданные о физическом лице, если они противоречиво обрабатываются и распечатываются по инструкциям и внутренним актам различных ведомств? Что есть гражданство как общеюридический термин?

При создании киберфизических систем для упрощения трансформаций уже в начале этого века была предложена платформа Eclipse VIATRA™ (Visual Automated model TRAnsformations) – фреймворк общего назначения для полного жизненного цикла моделей, включая в себя создание спецификаций, проектирование, исполнение, валидацию и конструирование трансформаций между различными языками создания моделей и предметных областей.

Фактически нужно признать, что на 2020 год в мире было до 50 миллиардов киберфизических систем типа «интернет вещей», «умный дом» и иных систем, от которых могут зависеть наша жизнь и здоровье, но хорошего и надежного, желательно быстрого, языка общения между ними и конструирования трансформаций порождаемого цифрового ландшафта так и не изобрели.

Таким образом, если мы сформулируем предметную область, то современные программные средства могут взять данные в виде XML (или сформировать термы по описаниям XML) и решить задачу в этой предметной области без участия программистов.

Но до сих пор нет достаточного инструментария, чтобы человек мог сравнить постановку задачи в терминах и сравнить ее с решением, построенным механическим исполнителем.

Полученные автором формулы оценок дают обозримый способ оценки постановки задачи и качества ее решения для некоторых предметных областей в терминах DOWN-UP кодирования [3–5].

Заключение. Результаты данной работы ранее обсуждались на семинарах компании «РДТЕХ» при проведении анализа и синтеза декомпозиции метаданных при построении ИС различной практической направленности, а также при создании моделей организационных систем.

Благодарности. Автор благодарит с.н.с. Пономареву А. С., в.н.с. Сухарева В. Г. (ИИФ г. Серпухов), г.н.с., заведующего 57 лабораторией ИПУ РАН Чхартишвили А. Г., за помощь в работе и обсуждение результатов.

Литература

1. Волошин Ю. М. Перечисление термов предметных областей по глубине вложенности // Докл. АН СССР. 1974. Т. 219. № 6. С. 1294–1297.
2. Харари Ф., Палмер Э. Перечисление графов / Перевод с англ. Г. П. Гаврилова. М.: Мир, 1977. 324 с.
3. Пономарев Ю. Ю. Вопросы точных оценок характеристик деревьев при моделировании термов предметных областей, нейросетей, XML данных // Процессы управления и устойчивость. 2023. Т. 10. № 1. С. 309–317.

4. Пономарев Ю. Ю. DOWN-UP последовательности и корневые деревья // Процессы управления и устойчивость: Труды 30-й научной конференции / под ред. В. Н. Старкова. СПб.: НИИ Химии СПбГУ, 1999. С. 374–380.
5. Пономарев Ю. Ю. Кодирование корневых деревьев DOWN-UP последовательностями // Процессы управления и устойчивость: Труды 29-й научной конференции / под ред. В. Н. Старкова. СПб.: НИИ Химии СПбГУ, 1998. С. 257–259.
6. Стандарты CommerceML [Электронный ресурс]: URL:<https://v8.1c.ru/tekhnologii/obmen-dannymi-i-integratsiya/standarty-i-formaty/standarty-commerceml/> (дата обращения: 01.04.24).
7. Forgy C. L. Rete: a fast algorithm for the many pattern/many object pattern match problem // Artificial Intelligence. 1982. Vol. 19. Iss. 1. P. 17–37.
8. Волошин Ю. М. Последовательности «ап-даун» и упорядоченные деревья // Докл. РАН. 1993. Т. 329. № 2. С. 125–127.
9. Случайные базы данных. Oracle Enterprise Data Quality – щит и меч корпоративного хранилища [Электронный ресурс]: URL: <https://habr.com/ru/companies/rdtex/articles/444700/> (дата обращения: 12.05.24).

УДК 519.254, 519.688

Потешкин Е. П.

Выбор параметров в методе Монте-Карло SSA

Рекомендовано к публикации доцентом Голяндиной Н.Э.

1. Введение. Обнаружение сигнала в красном шуме (красным шумом называется процесс авторегрессии первого порядка с положительным коэффициентом) с целью его дальнейшего выделения является важной задачей, например, для анализа поведения течения Эль-Ниньо [1] или глобального потепления [2]. Одним из методов, используемым для этой цели, является метод Monte Carlo SSA (MC-SSA) [3], который проверяет гипотезу об отсутствии сигнала, а в случае если она отвергается, позволяет приближенно выделить этот сигнал. При применении метода MC-SSA возникают следующие проблемы. Самый распространенный на практике вариант критерия MC-SSA является радикальным, поэтому для его использования необходимы поправки. В [4] предлагается поправка на множественное тестирование, частично исправляющая радикальность. Однако применяемая на практике модификация критерия все еще остается радикальной. Эту радикальность сложно исправить теоретически, поэтому используется метод эмпирической поправки критерия, основанный на оцененных ошибках первого рода в зависимости от уровня значимости. Например, в [5] этот подход используется в контексте метода MC-SSA.

Используемый в той же работе [5] подход через построение ROC-кривых позволяет сравнивать критерии по мощности. Соответственно возникает вопрос о выборе параметра метода MC-SSA, называемого длиной окна, дающего оптимальную мощность. Сложность состоит в том, что если исходно критерий сильно радикальный, то поправку, исправляющую радикальность, очень трудно или даже невозможно сделать.

Данная статья посвящена построению подхода к поиску длины окна, при которой критерий наиболее мощный среди ограниченно радикальных критериев.

Потешкин Егор Павлович – студент, Санкт-Петербургский государственный университет; email: egor.poteshkin@yandex.ru, тел.: +7(921)560-32-02

2. Вспомогательные результаты. Введем обозначения и приведем известные результаты.

2.1. Метод SSA [6]. Пусть $X = (x_1, \dots, x_N)$, $x_i \in \mathbb{R}$, – временной ряд длины N . Зафиксируем параметр L , $1 < L < N$, называемый длиной окна и построим так называемую траекторную матрицу $\mathbf{X} = [X_1 : \dots : X_K]$, состоящую из $K = N - L + 1$ векторов вложения $X_i = (x_i, \dots, x_{i+L-1})^T \in \mathbb{R}^L$.

Следующий шаг – разложение в сумму матриц единичного ранга $\mathbf{X} = \sum_{i=1}^d \mathbf{X}_i$. В базовом SSA используется сингулярное разложение матрицы X .

Далее компоненты полученного матричного разложения группируются, и каждая сгруппированная матрица преобразуется во временной ряд. Таким образом, результатом SSA является разложение временного ряда.

2.2. Метод Toeplitz SSA [6]. Этот метод является модификацией базового SSA и использует вместо сингулярного разложения разложение матрицы \mathbf{X} , основанное на собственных векторах $\{P_i\}_{i=1}^L$ тейплицевой матрицы \mathbf{C} с элементами

$$c_{ij} = \frac{1}{N - |i - j|} \sum_{m=1}^{N - |i - j|} x_m x_{m + |i - j|}, \quad 1 \leq i, j \leq L. \quad (1)$$

Тогда

$$\mathbf{X} = \sum_{i=1}^L \sigma_i P_i Q_i^T = \mathbf{X}_1 + \dots + \mathbf{X}_L,$$

где $Q_i = \mathbf{X}^T P_i / \sigma_i$, $\sigma_i = \|\mathbf{X}^T P_i\|$. Такое разложение представляется более естественным для стационарных временных рядов с нулевым средним, когда \mathbf{C} является оценкой автоковариационной матрицы.

2.3. Метод Monte Carlo SSA [3]. Рассмотрим задачу поиска сигнала во временном ряде. Модель временного ряда имеет вид

$$X = S + \xi,$$

где S – сигнал, ξ – красный шум с параметрами φ и δ . Тогда нулевая гипотеза $H_0 : S = 0$ и альтернатива $H_1 : S \neq 0$.

Зафиксируем длину окна L и обозначим траекторную матрицу ряда ξ как Ξ . Рассмотрим вектор $W \in \mathbb{R}^L$ единичной длины, называемый проекционным вектором. Введем величину

$$p = \|\Xi^T W\|^2.$$

Статистикой критерия является величина

$$\hat{p} = \|\mathbf{X}^T W\|^2.$$

Распределение статистики критерия оценивается с помощью моделирования согласно нулевой гипотезе, отсюда и название метода.

Если вектор W – синусоида с частотой ω , то \hat{p} отражает вклад частоты ω в исходный ряд. Так как частота ожидаемого сигнала неизвестна, то необходимо рассматривать несколько векторов W_k , $k = 1, \dots, H$. Решение возникающей при этом проблемы множественного тестирования рассматривается в [4]. Гипотеза об отсутствии сигнала отвергается, если хотя бы для одного вектора $W = W_k$ значение \hat{p} оказывается значимым.

Еще одним параметром MC-SSA является способ выбора векторов W_k . В данной работе в качестве векторов для проекции берутся собственные векторы матрицы \mathbf{C} , определенной в (1). Такой способ выбора самый распространенный, поскольку, если есть значимые векторы, можно восстановить сигнал с помощью SSA на их основе. Но этот вариант, вообще говоря, дает радикальный критерий, поскольку W_k зависят от ряда X , в котором ищется сигнал. С помощью эмпирической поправки неточных критериев можно бороться с этой проблемой.

2.4. Поправка неточных критериев. Данный алгоритм позволяет преобразовывать радикальные и консервативные статистические критерии в точные.

Зафиксируем нулевую гипотезу H_0 , уровень значимости α^* , количество выборок M для оценки $\alpha_I(\alpha)$ и их объем N (в случае временных рядов N – длина ряда).

Сначала моделируется M выборок объема N при верной H_0 . Затем по моделированным данным строится зависимость ошибки первого рода от уровня значимости $\alpha_I(\alpha)$. Результатом работы алгоритма является формальный уровень значимости $\tilde{\alpha}^* = \alpha_I^{-1}(\alpha^*)$. Критерий с таким уровнем значимости является асимптотически точным при $M \rightarrow \infty$.

Заметим, что если критерий сильно радикальный, то функция $\alpha_I(\alpha)$ имеет большую производную в нуле, что существенно затрудняет оценку $\alpha_I^{-1}(\alpha^*)$.

2.5. ROC-кривая. Это кривая, задаваемая параметрически:

$$\begin{cases} x = \alpha_I(\alpha), \\ y = \beta(\alpha), \end{cases} \quad \alpha \in [0, 1],$$

где $\alpha_I(\alpha)$ – функция зависимости ошибки первого рода α_I от уровня значимости α , $\beta(\alpha)$ – функция зависимости мощности β от уровня значимости α .

С помощью ROC-кривых можно сравнивать по потенциальной мощности неточные критерии. Отметим, что для точного (в частности, поправленного) критерия ROC-кривая совпадает с графиком мощности, так как $\alpha_I(\alpha) = \alpha$.

3. Зависимость радикальности и мощности MC-SSA от параметра L . Поскольку рассматриваемый вариант критерия MC-SSA является радикальным, существует проблема выбора такой длины окна L , которая дает максимально мощный критерий, но при этом не слишком радикальный, чтобы можно было применить поправку. Однако, в зависимости от длины ряда N и параметров красного шума ξ наблюдаются разные зависимости мощности от L .

Рассмотрим несколько примеров. Пусть дана модель

$$X = S + \xi,$$

где $S = \{A \cos(2\pi\omega n)\}_{n=1}^N$, а ξ – красный шум с параметрами φ и $\delta = 1$. Рассмотрим следующие нулевую гипотезу и альтернативу: $H_0 : A = 0$, $H_1 : A \neq 0$. В работе [5] показано, что оценка параметров модели почти не искажает критерий после применения алгоритма поправки, поэтому будем предполагать, что параметры красного шума известны. В первых трех примерах рассмотрим частоту сигнала $\omega = 0,075$.

Пример 1. Пусть $\varphi = 0,7$, $N = 100$. По графику ошибок первого рода на рис. 1 видно, что чем больше L , тем более радикальным становится критерий. На рис. 2 изображены ROC-кривые критериев, наибольшую мощность дает критерий с $L = 90$. На этом примере видно, что самым мощным является самый радикальный критерий.

Пример 2. Пусть $\varphi = 0,3$, $N = 100$. На рис. 3 изображен график ошибок первого рода. По нему видно, что, как и в примере 1, чем больше L , тем больше радикальность критерия. Если взглянуть на ROC-кривые на рис. 4, то видно, что с уменьшением параметра

φ уменьшается разброс мощности критериев после поправки в зависимости от длины окна. Лучшей из рассмотренных в этом случае является $L = 10$, хотя разница с $L = 50$ совсем небольшая, а для $L = 90$ для небольших ошибок I рода поправку сделать не удалось из-за радикальности.

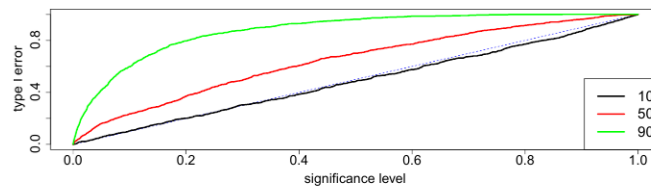


Рис. 1. Ошибка I рода ($\varphi = 0,7, N = 100$)

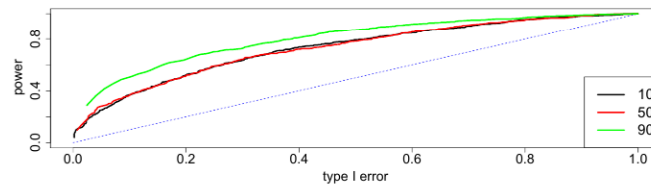


Рис. 2. ROC-кривая ($\varphi = 0,7, N = 100, \omega = 0,075$)

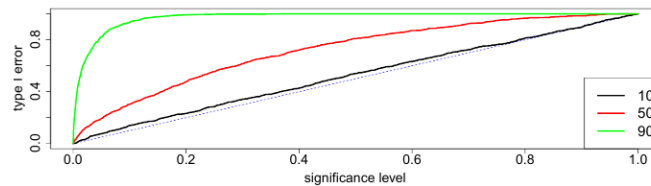


Рис. 3. Ошибка I рода ($\varphi = 0,3, N = 100$)

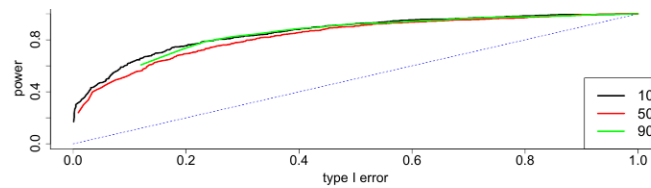


Рис. 4. ROC-кривая ($\varphi = 0,3, N = 100, \omega = 0,075$)

Пример 3. Теперь увеличим длину ряда до $N = 400$ и посмотрим на ROC-кривые для примеров 1, 2. На рис. 5, 6 видим, что для

обоих примеров с увеличением длины ряда уменьшается различие в мощности после поправки в зависимости от длины окна.

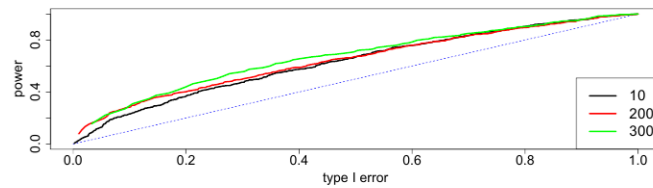


Рис. 5. ROC-кривая ($\varphi = 0,7$, $N = 400$, $\omega = 0,075$)

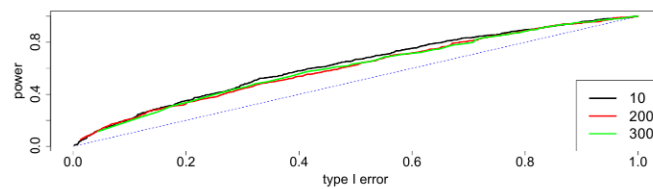


Рис. 6. ROC-кривая ($\varphi = 0,3$, $N = 400$, $\omega = 0,075$)

Пример 4. В условиях примера 1 рассмотрим разные частоты ω сигнала S и зависимость упорядоченности критериев по мощности от L . На рис. 7, 8 изображены ROC-кривые критериев при разных альтернативах. Видно, что упорядоченность L нарушается при маленьких частотах сигнала. Если упорядочить рис. 7, рис. 2 и рис. 8 по частоте ω , то видна динамика по соотношению ROC-кривых для $L = 10$ и $L = 50$.

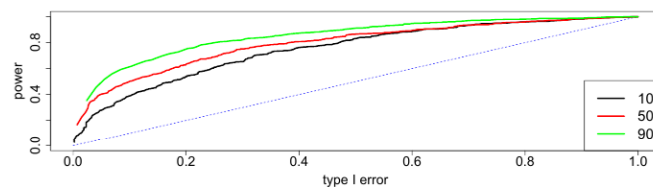


Рис. 7. ROC-кривая ($\varphi = 0,7$, $N = 100$, $\omega = 0,175$)

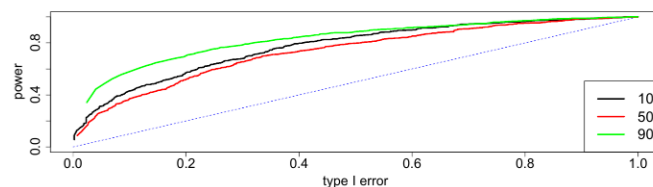


Рис. 8. ROC-кривая ($\varphi = 0,7$, $N = 100$, $\omega = 0,025$)

4. Выбор длины окна. Численные эксперименты показали, что длина окна L , дающая максимальную мощность критерия после поправки, зависит от параметров шума, длины ряда и, главное, от частоты сигнала в альтернативной гипотезе. Поэтому при выборе длины окна возможны следующие варианты:

1. Использовать поправленный критерий MC-SSA с маленькой длиной окна, например, $L = 10$ при длине ряда $N = 100$. Это нетрудозатратно, а также критерий не является сильно радикальным (в рассмотренных примерах он близок к точному). Недостаток состоит в том, что такой выбор L может являться неоптимальным, т. е. возможна некоторая потеря в мощности.
2. В поведении оптимальной по мощности длины окна L в зависимости от параметров ряда наблюдается некоторая регулярность. Поэтому можно было бы построить зависимость оптимальной длины окна от параметров ряда с помощью численного моделирования, оценив параметры красного шума. Однако было показано, что упорядоченность критериев по мощности зависит от частоты сигнала в альтернативе, эта рекомендация имеет практический смысл, только если есть дополнительная информация о диапазоне возможных частот в альтернативе.

5. Заключение. В работе выработан подход к выбору параметра «длина окна» в методе MC-SSA, позволяющем обнаруживать сигнал в красном шуме. Оказалось, что из-за того, что возможная частота сигнала, вообще говоря, неизвестна, использование численного моделирования не является решением задачи выбора оптимальной длины окна. В целом, численное моделирование позволяет дать рекомендации по выбору L при уточнении диапазона возможной частоты сигнала. В любом случае, численное моделирование является весьма трудоемким, поэтому, вполне возможно, что наиболее разумной альтернативой является выбор маленькой длины окна (например, $L = 10$ при длине ряда $N = 100$). Как показано на примерах, MC-SSA с такой длиной окна не всегда дает максимальную мощность, однако обладает небольшой трудоемкостью и малой радикальностью.

Литература

1. Jevrejeva S., Moore J.C., Grinsted A. Oceanic and atmospheric transport of multiyear El Niño–Southern Oscillation (ENSO) signatures to the polar regions // *Geophysical Research Letters*. 2004. Vol. 31. Art. no L24210.
2. Селиверстова К. А. Применение статистических методов для оценки глобального потепления // *Процессы управления и устойчивость*. 2023. Т. 10. № 1. С. 195–199.
3. Allen M., Smith L. Monte Carlo SSA: detecting irregular oscillations in the presence of coloured noise // *Journal of Climate*. 1996. Vol. 9. P. 3373–3404.
4. Golyandina N. Detection of signals by Monte Carlo singular spectrum analysis: multiple testing // *Statistics and Its Interface*. 2023. Vol. 16. No 1. P. 147–157.
5. Ларин Е. С. Применимость исправления статистических критериев на примере задачи обнаружения сигнала во временных рядах // *Процессы управления и устойчивость*. 2022. Т. 9. № 1. С. 267–271.
6. Golyandina N., Nekrutkin V., Zhigljavsky A. *Analysis of Time Series Structure: SSA and Related Techniques*. Boca Raton: Chapman and Hall/CRC, 2001. 320 p.

УДК 519.6

Рудович К. В.

Библиотека для создания и имплементации генетических алгоритмов на языке C++

Рекомендовано к публикации доцентом Никифоровым К. А.

1. Введение. Дискретная оптимизация [1] представляет собой раздел математического программирования, где исследуются и решаются экстремальные задачи на целочисленных решетках и конечных множествах. Задачи дискретной оптимизации имеют высокую вычислительную сложность из-за ограничений на значения переменных, которые могут принимать только дискретные (целочисленные) значения. Поэтому процесс поиска оптимального решения в общем случае требует рассмотрения всех возможных комбинаций значений переменных. Решения дискретных задач оптимизации усложняются из-за невозможности применения градиентных методов [2], которые широко используются в непрерывной оптимизации, так как они требуют наличия гладкой целевой функции, что не выполняется в случае целочисленных задач. Одной из ключевых ценностей дискретной оптимизации является ее применение в решении реальных проблем, таких как задача размещения [3], задача маршрутизации [4], задача о рюкзаке [5] и другие. Эти методы находят широкое применение в управлении проектами, логистике, транспорте, производстве и других областях, где требуется эффективное распределение ресурсов и оптимальное принятие решений. Для решения задач дискретной оптимизации существует широкий спектр подходов [6–8], все они делятся на точные методы и приближенные.

1.1. Генетический алгоритм. Один из приближенных методов решения – генетический алгоритм [9, 10]. Это эвристический алгоритм [11] поиска, который использует механизмы, напоминающие биологическую эволюцию.

Рудович Кирилл Владимирович – студент, Санкт-Петербургский государственный университет; e-mail: Rudovichkirill@gmail.com, тел.: +7(950)552-10-20
Работа выполнена при финансовой поддержке РФФ, грант № 23-29-00937

Такой подход к решению задачи по своей сути является оптимизированным перебором и активно применяется в робототехнике, компьютерных играх, обучении нейронных сетей, создании моделей, составлении расписаний, оптимизации запросов к базам данных, поиске оптимальных маршрутов и т. д.

1.2. Проблематика. Генетический алгоритм достаточно абстрактен, определить конкретную реализацию всех операций можно произвольным образом. Например, отбор для скрещивания можно производить на основе приспособленности или близости, детерминировано или вероятностно, а само скрещивание может быть одноточечным, двухточечным, ассоциативным [12] и т. д. Также для каждого такого оператора или для всего алгоритма в целом можно варьировать немалое количество гиперпараметров. Зачастую вопрос об оптимальной реализации генетического алгоритма можно решить только экспериментально и только под конкретную задачу. Автор такого алгоритма, как правило, может руководствоваться только спецификой задачи.

На данный момент существует множество программных решений по использованию данного алгоритма, например, PyGAD [13], ortlib [14], GAlib [15], Jenetics [16].

Большинство решений имеют достаточно узкую специализацию. Как правило, они реализованы под конкретную задачу, логику обучения алгоритма в них нельзя изменить, а единственный способ влиять на процесс – варьировать параметры и целевую функцию. В остальных решениях, которые позволяют в разной мере влиять на логику обучения, подразумевается детерминированное обучение одной популяции по заданным правилам, что не позволяет изменять правила обучения динамически или удобно проектировать каскады из нескольких различных реализаций генетического алгоритма.

В связи с этим, для решения возникающих задач появляется необходимость в инструменте, который избавляет пользователя от лишних действий во время создания генетического алгоритма, но при этом не ограничивает его конкретными реализациями операторов.

2. Описание библиотеки. Создание любого генетического алгоритма (ГА) начинается с определения постановки задачи в терминах ГА. Для этого в библиотеке реализован ряд сущностей, которые представляют исходные данные задачи: **Gen**, **Chromosome**, **Generation**, **Population**. Они используются в операторах ГА для

обучения.

Для того чтобы иметь возможность встраивать произвольную функциональность внутрь операторов используются интерфейсы генетического алгоритма (ИГА) – абстрактные классы, задающие каркас операторов: входные/выходные параметры и сущности. Таким образом пользователю достаточно на основе ИГА реализовать собственную логику операторов для работы алгоритма. Подобный подход требует самостоятельной реализации работы всех операторов, но никак не ограничивает функциональность. В ИГА можно передавать произвольные данные (как общие, так и локальные), а используя возможности C++, можно производить новые ИГА на основе существующих, группируя схожую логику.

Основные ИГА: `FitnessFunctionWrapper`, `CrossingoverWrapper`, `ConditionsForStoppingWrapper`, `SelectionWrapper`, `AnyWrapper`, `MutationWrapper`.

Единственное строгое ограничение на обучение популяции – это порядок вызовов операторов ГА. Для популяции, которая уже имеет ненулевое количество поколений порядок действий следующий:

1. Вычисление функции приспособленности для всех особей в последнем поколении G , для которых она не вычислена.
2. Проверка условий остановки алгоритма.
3. Формирование нового поколения G' с помощью оператора скрещивания.
4. Вызов оператора мутации для нового поколения G' .
5. Вычисление функции приспособленности для всех особей в поколении G' , для которых она не вычислена.
6. Вызов оператора селекции для нового поколения G' .
7. Добавление поколения G' в популяцию.
8. Вызов произвольного пользовательского оператора для всего поколения.
9. Повторение шагов 2–8.

Такой алгоритм является базовым процессом обучения, далее будем его называть алгоритм эволюции популяции (АЭП). Для его работы от пользователя требуется лишь реализовать логику работы операторов с помощью интерфейсов.

2.1. Дерево эволюции. Для решения реальных задач зачастую одного АЭП бывает недостаточно. Алгоритмы такого рода очень непредсказуемы: в зависимости от реализации и гиперпараметров они могут сходиться очень медленно или вовсе застрять в «локальном оптимуме». Чтобы решить эту проблему можно использовать различные каскады, состоящие из АЭП. Для этого была реализована концепция дерева эволюции (ДЭ).

Основными сущностями в ДЭ, как и в любом дереве, являются вершины. Любая вершина представляет собой набор данных для АЭП и некоторую дополнительную информацию, различающуюся для каждого типа вершин. Результатом работы любой вершины является популяция после обучения АЭП.

`Node` – базовый абстрактный класс любой вершины ДЭ, который помимо АЭП включает в себя следующие интерфейсы: `StartNode`, `EndNode`, `StartNodeLog`, `NewGenerationLog`, `EndNodeLog`. Как следует из названия, интерфейсы логирования служат для формирования и сохранения любой необходимой информации в процессе работы вершины. Интерфейсы `StartNode` и `EndNode` могут выполнять любую дополнительную внешнюю логику по отношению к АЭП (например, динамически задавать правила обучения).

Вершина `PopulationNode` включает в себя интерфейс формирования стартового поколения; является листом любого ДЭ.

Вершина `UnaryNode` включает в себя указатель на другую вершину.

Вершина `K_Node` включает в себя K указателей на другие вершины и интерфейс объединения K популяций в одну.

Каждая конкретная вершина использует собственные операторы, что позволяет гибко настроить процесс эволюции. ДЭ расширяет поле потенциальных экспериментов и решений, например, с помощью `UnaryNode` можно последовательно запустить разные АЭП с разными условиями, а с помощью `K_Node` можно эффективно избавляться от локальных оптимумов задачи, объединяя популяции, которые находятся в разных экстремумах. Любая вершина может быть корнем дерева, а отдельных корней может быть произвольное количество, если требуется провести независимые процессы.

2.2. Дополнительные возможности. Библиотека поддерживает распараллеливание процесса обучения ДЭ на множество потоков, так как обучение каждого поддерева для `K_Node` происходит независимо. Таким образом, если сверху ограничить время рабо-

ты каждой вершины дерева, то теоретическая временная сложность обучения всего ДЭ (при неограниченном количестве ядер процессора) будет зависеть линейно не от количества его вершин, а от его высоты (в случае сбалансированного дерева зависимость будет логарифмической от количества вершин). Стоит отметить, что АЭП в общем случае не поддается параллелизму, так как в нем задан строгий порядок выполнения, а на ИГА не накладывается ограничение в независимости.

Также, помимо встраиваемых пользователем операторов логирования, предусмотрено нативное логирование.

Offline логирование сохраняет подробную информацию работы ДЭ (время работы, максимальные/минимальные/средние показатели целевой функции и т. д.) для каждой вершины, чтобы после обучения была возможность понять, как шел процесс.

Online логирование во время работы ДЭ отрисовывает с определенным интервалом все дерево в терминале только с самой необходимой информацией (состояние вершины, время работы, максимальный показатель целевой функции и т. д.), чтобы оценить общее время работы и убедиться, что обучение идет корректно.

3. Пример использования. Применим данную библиотеку для решения задачи о 0-1 рюкзаке, зададим размер рюкзака $N = 1000$, вместительность $W = 100000$, веса предметов сгенерируем из нормального распределения с математическим ожиданием 3000 и дисперсией 10000, а ценности с математическим ожиданием 100 и дисперсией 900.

В качестве особи будет бинарная строка x , где $x_i = 1$, если предмет i берется в рюкзак, иначе $x_i = 0$.

Оператор `StartPopulation` будет генерировать одно поколение из $L = 200$ особей, заполненных нулями.

Оператор `Selection` будет заполнять новое поколение на 10% наиболее приспособленными особями, на 10% нулевыми особями, на 80% особями, которые были отобраны с вероятностями пропорциональными их приспособленностям с повторениями.

Оператор `Mutation` для каждого гена с вероятностью $m = 0,004$ будет менять значение на противоположное.

Оператор `PoolingPopulations` из $K = 2$ популяций будет отбирать $\frac{L}{K}$ наиболее приспособленных особей в одну популяцию.

Оператор `ConditionsForStopping` будет останавливать алгоритм,

если за $M = 200$ итераций приспособленность не увеличивалась.

Оператор `FitnessFunction` будет вычислять приспособленность особи по формуле:

$$\mu(X) = \begin{cases} \sum_{i=1}^n v_i x_i, & \sum_{i=1}^n w_i x_i \leq W, \\ 0, & \sum_{i=1}^n w_i x_i > W. \end{cases}$$

Все остальные интерфейсы использовать не будем.

ДЭ сделаем бинарным деревом высотой 3, в каждой вершине зададим одинаковый АЭП, меняя только вероятность мутации в зависимости от высоты вершины: $m_i = m/2^j$, где j – высота вершины i . Результат работы программы представлен на рисунке 1.

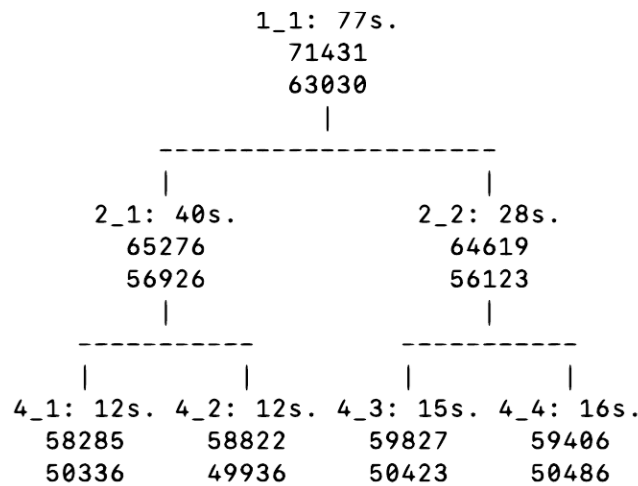


Рис. 1. Результат работы (точное решение: 76178)

Более подробно с примерами работы, исходными кодами, документацией и другими возможностями библиотеки можно ознакомиться в [17].

4. Заключение. В результате работы описан и предложен инструмент для создания и имплементации генетического алгоритма на языке C++. Данную библиотеку можно использовать как готовое

решение для проектирования генетических алгоритмов, или как модуль программного комплекса. В частности, в рамках создания цифровой модели автоэмиссионной системы с катодно-модуляторным узлом [18, 19] разрабатывается модуль оптимизации на основе библиотеки.

Литература

1. Ватутин Э. И., Титов В. С., Емельянов С. Г. Основы дискретной комбинаторной оптимизации. М.: Аргмак-Медиа, 2016. 270 с.
2. Polak E. Optimization: Algorithms and Consistent Approximations. New York: Springer-Verlag, 1997. 782 p.
3. Fowler R. J., Paterson M. S., Tanimoto S. L. Optimal packing and covering in the plane are NP-complete // Information Processing Letters. 1981. Vol. 12. No 3. P. 133–137.
4. Dantzig G. B., Ramser J. H. The Truck Dispatching Problem // Management Science. 1959. Vol. 6. No 1. P. 80–91.
5. Mathews G. B. On the partition of numbers // Proceedings of the London Mathematical Society. 1897. P. 486–490.
6. Land A. H., Doig A. G. An automatic method of solving discrete programming problems // Econometrica. 1960. Vol. 28. No 3. P. 497–520.
7. Bellman R. The theory of dynamic programming // Bulletin of the American Mathematical Society. 1954. Vol. 60. No 6. P. 503–516.
8. Kirkpatrick S., Gelatt C. D., Vecchi M. P. Optimization by Simulated Annealing // Science. 1983. Vol. 220. No 4598. P. 671–680.
9. Fraser A. Simulation of genetic systems by automatic digital computers. I. Introduction // Aust. J. Biol. Sci. 1957. Vol. 10. No 4. P. 484–491.
10. Whitley D. A genetic algorithm tutorial // Statistics and Computing. 1994. Vol. 4. No 2. P. 65–85.
11. Pearl J. Heuristics: Intelligent Search Strategies for Computer Problem Solving. United States: Addison-Wesley, 1984. 401 p.

12. Батищев Д. И., Неймарк Е. А., Старостин Н. В. Применение генетических алгоритмов к решению задач дискретной оптимизации: учебно-методическое пособие. Н. Новгород: ФГАОУ ВО «ННГУ им. Н.И. Лобачевского», 2007. 85 с.
13. PyGAD [Электронный ресурс]: URL:<https://github.com/ahmedfgad/GeneticAlgorithmPythonm> (дата обращения: 20.03.2024).
14. Optlib [Электронный ресурс]: URL:<https://github.com/dbrojas/optlib> (дата обращения: 20.03.2024).
15. GAlib [Электронный ресурс]: URL:<https://github.com/remryan/GAlib> (дата обращения: 20.03.2024).
16. Jenetics [Электронный ресурс]: URL:<https://github.com/jenetics/jenetics> (дата обращения: 20.03.2024).
17. Код проекта [Электронный ресурс]: URL:<https://github.com/Rudovich1/VKR> (дата обращения: 20.03.2024).
18. Соколов И. А., Никифоров К. А. Моделирование автоэмиссионного катодно-модуляторного узла // Процессы управления и устойчивость. 2023. Т. 10. № 1. С. 156–160.
19. Egorov N. V., Vinogradova E. M., Kurbatova G. I., Klemeshev V. A. Field emissionsystem with two emitters mathematical modeling // Vestnik of Saint Petersburg University. AppliedMathematics. Computer Science. Control Processes. 2023. Vol. 19, No 4. P. 443–448.

УДК 519.254

Хромов Н. А., Голяндина Н. Э.

High-order MSSA для выделения сигнала

1. Введение. SSA (Singular Spectrum Analysis) [1] является пространственным методом анализа и прогноза временных рядов. Этот метод используется, в частности, для выделения сигнала и выделения тренда и периодических компонент из временного ряда. В последнее время получили распространение методы машинного обучения. Например, в [2] рассматривается задача прогноза с предобработкой временного ряда методом SSA перед применением нейронных сетей для прогноза. Однако задача анализа одного нестационарного временного ряда не теряет своей актуальности. Ряд методов рассматривает при этом конкретные модели для сигнала/тренда. Например, в [3] рассматривается полиномиальная регрессия. Метод SSA не требует задания параметрической модели, поэтому он более универсальный. Однако, при этом метод SSA хорошо работает и с моделью в виде суммы комплексных экспонент или в виде суммы экспоненциально-модулированных гармоник в вещественном случае, что позволяет оценивать параметры в этих моделях.

Метод SSA основан на сингулярном разложении особой матрицы, построенной по временному ряду; она называется траекторной. Обобщение метода SSA на случай многомерных сигналов называется MSSA [4]. Для оценки параметров временного ряда в параметрической модели методов SSA и MSSA используется метод ESPRIT (Estimation of signal parameters via rotational invariance techniques) [5], который также основан на сингулярном разложении траекторной матрицы и относится к классу методов, основанных на подпространстве сигнала.

В работе [6] предлагается модификация метода ESPRIT, основанная на переходе от траекторной матрицы и матричного сингулярного разложения к траекторному тензору и тензорному сингулярному разложению, и демонстрируется преимущество такой модификации

Хромов Никита Андреевич – студент, Санкт-Петербургский государственный университет; e-mail: hromovn@mail.ru, тел.: +7(981)509-25-94

Голяндина Нина Эдуардовна – доцент, Санкт-Петербургский государственный университет; e-mail: n.golyandina@spbu.ru, тел.: +7(911)957-20-08

Работа выполнена при финансовой поддержке РФФИ, проект № 23-21-00222

при оценке параметров, в частности, частот. При этом для случая многомерного временного ряда особой структуры преимущество наиболее очевидно. В данной работе рассмотрим аналогичную тензорную модификацию метода MSSA (HO-MSSA, где HO – сокращение от high order) для выделения сигналов из временных рядов. Целью работы является сравнение метода MSSA с его тензорной модификацией на временных рядах, аналогичных тем, что были рассмотрены в [6].

2. Модель сигнала. Опишем рассматриваемые многомерные комплексные сигналы. Пусть дан одномерный временной ряд, состоящий из N комплексных значений s_n , $n = 0, \dots, N - 1$, и пусть этот ряд представим в виде конечной суммы R экспоненциально-модулированных комплексных гармоник

$$s_n = \sum_{j=1}^R a_j e^{-\alpha_j n} e^{i(2\pi\omega_j n + \varphi_j)}. \quad (1)$$

Параметрами модели являются амплитуды $a_j \in \mathbb{R} \setminus \{0\}$, фазы $\varphi_j \in [0, 2\pi)$, частоты $\omega_j \in [0, 1/2]$ и степени затухания $\alpha_j \in \mathbb{R}$.

Модель многомерного сигнала имеет вид набора P одномерных сигналов вида (1), т. е.

$$s_n^{(p)} = \sum_{j=1}^{R(p)} a_j^{(p)} e^{-\alpha_j^{(p)} n} e^{i(2\pi\omega_j^{(p)} n + \varphi_j^{(p)})}, \quad (2)$$

где $n = 0, \dots, N - 1$, а p ($1 \leq p \leq P$) отвечает за номер одномерного временного ряда. Далее будем рассматривать частный случай такой модели, при котором параметры $R(p)$, $\omega_j^{(p)}$ и $\alpha_j^{(p)}$ не зависят от номера ряда p :

$$s_n^{(p)} = \sum_{j=1}^R a_j^{(p)} e^{-\alpha_j n} e^{i(2\pi\omega_j n + \varphi_j^{(p)})}. \quad (3)$$

Это обосновано тем, что такая модель применяется в спектроскопии ядерного магнитного резонанса [7]. Кроме того, в работе [6] используется именно этот частный случай модели.

Замечание. В общем случае можно рассматривать модели временных рядов вида (1), в которых амплитуды являются ненулевыми многочленами.

3. Ряды конечного ранга и тензорные ранги. Одними из ключевых объектов в теории SSA являются *ряды конечного ранга*. В данном разделе изложим теорию, касающуюся рядов конечного ранга, а затем рассмотрим теорию тензорных рангов, и приведём утверждения, связывающие ранги рядов с рангами определённых тензоров.

3.1. Ряды конечного ранга. Введём несколько вспомогательных определений.

Определение 1. *Оператором вложения с длиной окна L* будем называть отображение, преобразующее временной ряд (возможно, комплексный) (x_0, \dots, x_{N-1}) в ганкелеву $(L \times K)$ -матрицу, $K = N - L + 1$, следующим образом:

$$\mathbb{H}_L((x_0, \dots, x_{N-1})) = \begin{pmatrix} x_0 & x_1 & x_2 & \dots & x_{K-1} \\ x_1 & x_2 & \dots & \dots & \vdots \\ x_2 & \dots & \dots & \dots & \vdots \\ \dots & \vdots & \vdots & \vdots & x_{N-2} \\ x_{L-1} & \dots & \dots & x_{N-2} & x_{N-1} \end{pmatrix}.$$

Определение 2. *Траекторной матрицей одномерного временного ряда (x_0, \dots, x_{N-1}) с длиной окна L* называется матрица \mathbf{H} , полученная применением оператора вложения с длиной окна L к данному временному ряду.

Определение 3. Ряд (x_0, \dots, x_{N-1}) называется *рядом конечного ранга d* ($d < N/2$), если для любой длины окна $L < N$ такой, что $\min(L, K) \geq d$, траекторная матрица этого ряда, построенная по данной длине окна, имеет ранг d .

Пример 1. Временной ряд вида (1) имеет ранг, равный количеству уникальных пар (α_j, ω_j) по всем j [8].

Рассмотрим многомерный временной ряд

$$\left(s_0^{(p)}, \dots, s_{N-1}^{(p)} \right), \quad p = 1, \dots, P.$$

Построим для каждого одномерного ряда $\left(s_n^{(p)} \right)_{n=0}^{N-1}$ траекторную матрицу с длиной окна L :

$$\mathbf{H}_p = \mathbb{H}_L \left(s_0^{(p)}, \dots, s_{N-1}^{(p)} \right).$$

Определение 4. Траекторной матрицей данного многомерного ряда с длиной окна L называется матрица \mathbf{H} , полученная приписыванием матриц \mathbf{H}_p друг за другом по столбцам:

$$\mathbf{H} = [\mathbf{H}_1 : \dots : \mathbf{H}_P]. \quad (4)$$

Замечание. Ранг многомерного ряда определяется аналогично рангу одномерного ряда с заменой траекторной матрицы одномерного ряда на траекторную матрицу многомерного ряда.

Пример 2. Многомерный временной ряд вида (2) имеет ранг, равный количеству уникальных пар $(\alpha_j^{(p)}, \omega_j^{(p)})$ по всем j и p [8].

Пример 3. Так как

$$\cos(2\pi\omega n + \varphi_n) = (e^{2\pi i\omega n + i\varphi_n} + e^{-2\pi i\omega n - i\varphi_n}) / 2,$$

то ранг ряда

$$s_n = Ae^{\alpha n} \cos(2\pi\omega n + \varphi)$$

с $A \neq 0$ равен 2, если $\omega \in (0, 1/2)$ и равен 1, если $\omega \in \{0, 1/2\}$.

Пример 4. Пусть ряд представляется в виде суммы M экспоненциально-модулированных гармоник

$$s_n = \sum_{j=1}^M A_j e^{\alpha_j n} \cos(2\pi\omega_j n + \varphi_j), \quad (5)$$

где $A_j \neq 0$. Обозначим

$$r(\omega) = \begin{cases} 1, & \omega \in \{0, 1/2\}, \\ 2, & \text{иначе.} \end{cases}$$

Тогда ранг ряда (5) равен $\sum_{(\omega, \alpha) \in \Omega} r(\omega)$, где Ω – множество уникальных пар (ω_j, α_j) по всем j .

3.2. Тензорные ранги. Введём несколько вспомогательных определений.

Определение 5. Векторами тензора $\mathcal{A} \in \mathbb{C}^{I_1 \times \dots \times I_M}$ по измерению n называются векторы вида $\left((a_{i_1 \dots i_{n-1} k i_{n+1} \dots i_M})_{k=1}^{I_n} \right)^T$, где индексы i_j фиксированы.

Определение 6. n -рангом тензора \mathcal{A} называют размерность линейной оболочки, построенной по векторам этого тензора по измерению n . Обозначим его $\text{rank}_n(\mathcal{A})$.

Определение 7. *Траекторным тензором многомерного временного ряда*

$$\left(x_0^{(p)}, \dots, x_{N-1}^{(p)}\right), \quad p = 1, \dots, P,$$

с длиной окна L называется тензор $\mathcal{H} \in \mathbb{C}^{L \times K \times P}$ ($K = N - L + 1$), элемент h_{lkp} которого равен $x_{l+k-2}^{(p)}$.

Таким образом, слой траекторного тензора с номером p по третьему измерению получается как траекторная матрица, построенная с длиной окна L по ряду с номером p :

$$\mathcal{H}_{..p} = \mathbb{H}_L \left(\left(x_0^{(p)}, \dots, x_{N-1}^{(p)} \right) \right). \quad (6)$$

Основываясь на виде траекторного тензора многомерного ряда, можно доказать следующее утверждение, связывающее понятие многомерного ряда конечного ранга и понятие n -рангов траекторного тензора этого ряда.

Утверждение (Об n -рангах траекторного тензора многомерного ряда конечного ранга). *Пусть временной ряд*

$$\left(x_0^{(p)}, \dots, x_{N-1}^{(p)}\right), \quad p = 1, \dots, P$$

имеет ранг d в терминах MSSA. Тогда для траекторного тензора \mathcal{H} , построенного по любой длине окна $L < N$ такой, что $\min(L, K) \geq d$, выполняется $\text{rank}_1(\mathcal{H}) = \text{rank}_2(\mathcal{H}) = d$, а 3-ранг этого тензора равен рангу матрицы, в строках которой записаны заданные одномерные ряды.

4. Методы для оценки сигнала. Пусть дан некоторый многомерный сигнал вида (3). В работе [6] рассматривается задача оценки параметров ω_j и α_j такого ряда. Мы же рассмотрим задачу оценки значения самого сигнала по наблюдаемому зашумлённому сигналу. В качестве шума можно рассматривать нерегулярные колебания вокруг нуля. Мы рассмотрим реализации независимого гауссовского шума с нулевым математическим ожиданием и одинаковой дисперсией для разных рядов.

4.1. MSSA. Приведём алгоритм MSSA для решения задачи оценки многомерного сигнала по зашумлённому ряду.

На первом шаге выбирается длина окна L и по ней строится траекторная матрица \mathbf{H} (4) данного многомерного ряда. Эту матрицу

можно представить в виде её сингулярного разложения (SVD)

$$\mathbf{H} = \sum_{j=1}^d \sigma_j U_j V_j^H,$$

где $d = \text{rank}(\mathbf{H}) \geq \min(L, K)$, σ_j – сингулярные числа матрицы \mathbf{H} , верхний индекс \mathbf{H} обозначает эрмитово сопряжение, а матрицы $\mathbf{U} = [U_1 : \dots : U_d] \in \mathbb{C}^{L \times d}$ и $\mathbf{V} = [V_1 : \dots : V_d] \in \mathbb{C}^{K \times d}$ унитарные.

Если на данный сигнал воздействует шум, то траекторная матрица будет матрицей полного ранга и тогда рассматривается наилучшее приближение траекторной матрицы матрицей заданного ранга $r \leq d$:

$$\hat{\mathbf{H}} = \sum_{j=1}^r \sigma_j U_j V_j^H.$$

Последний шаг алгоритма – восстановление, заключается в усреднении каждой матрицы $\hat{\mathbf{H}}_p$ по побочным диагоналям, т. е.

$$\tilde{s}_n^{(p)} = \frac{1}{|M_n|} \sum_{i,j \in M_n} \hat{h}_{ij}^{(p)}, \quad n = 0, \dots, N-1,$$

где $M_n = \{(i, j) : i + j - 2 = n, 1 \leq i \leq L, 1 \leq j \leq K\}$, а $\hat{h}_{ij}^{(p)}$ – элементы матрицы $\hat{\mathbf{H}}_p$. Результат такого усреднения $\tilde{s}_n^{(p)}$ считается оценкой сигнала $s_n^{(p)}$.

4.2. High-Order MSSA. В работе [6] предложена тензорная модификация алгоритма ESPRIT для оценки параметров многомерных сигналов. В этом разделе мы приведём расширение этого алгоритма для решения задачи выделения сигнала из многомерных рядов.

Рассмотрим многомерный временной ряд с сигналом вида (3). На первом шаге выбирается длина окна $1 < L < N$ и по данному ряду строится траекторный тензор многомерного ряда \mathcal{H} (6). Затем траекторный тензор представляется в виде своего HO-SVD [9]

$$\mathcal{H} = \sum_{l=1}^{d_1} \sum_{k=1}^{d_2} \sum_{p=1}^{d_3} c_{lkp} U_l^{(1)} \circ U_k^{(2)} \circ U_p^{(3)}.$$

Из примера 2 и утверждения 1 следует, что при некоторых условиях на N и L 1- и 2-ранги траекторного тензора будут равны количеству

уникальных пар $(\alpha_j^{(p)}, \omega_j^{(p)})$ по всем j и p , а 3-ранг будет равен рангу матрицы, в строках которой записаны заданные одномерные ряды. Если во временном ряде присутствует шум, то у траекторного тензора будут максимальные n -ранги, и тогда выбираются параметры r и r_3 и строится приближение траекторного тензора с 1- и 2-рангами, равными r , и 3-рангом, равным r_3 :

$$\hat{\mathcal{H}} = \sum_{l=1}^r \sum_{k=1}^r \sum_{p=1}^{r_3} \sigma_{lkp} U_l^{(1)} \circ U_k^{(2)} \circ U_p^{(3)}.$$

Последний шаг алгоритма аналогичен по смыслу шагу восстановления в алгоритме MSSA и заключается в усреднении матриц-слоёв $\hat{\mathcal{H}}_{..p}$ полученного приближения вдоль побочных диагоналей $l + k = \text{const}$.

5. Численные сравнения. В данном разделе приведем результаты численных сравнений метода MSSA с его тензорным аналогом НО-MSSA. Точность выделения сигнала сравнивалась с помощью RMSE оценки сигнала по 1000 реализациям шума. Методы сравнивались на одних и тех же реализациях шума, все различия значимы при уровне значимости $\alpha = 0,05$.

5.1. Сигнал с одинаковыми фазами. Пусть временной ряд имеет вид

$$x_n^{(p)} = c_1^{(p)} e^{-0,01n} \cos(2\pi 0,2n) + c_2^{(p)} e^{-0,02n} \cos(2\pi 0,22n) + \varepsilon_n^{(p)}, \quad (7)$$

где $n = 0, \dots, 24$, $p = 1, \dots, 12$, $c_k^{(p)}$ – независимые нормальные случайные величины с нулевым средним и стандартным отклонением 1, $\varepsilon_n^{(p)}$ – последовательность независимых нормальных случайных величин с нулевым средним и стандартным отклонением $\sigma = 0,02$.

Будем сравнивать точность выделения сигнала при выборе оптимальных для каждого метода длин окна (для MSSA $L = 22$, для НО-MSSA $L = 20$). Ранг сигнала в обоих случаях равен 4, а 3-ранг траекторного тензора равен 2. Соответственно, число компонент для оценки сигнала в MSSA будем выбирать равным 4, число компонент по первым двум измерениям в НО-MSSA тоже равным 4, а по третьему измерению – 2. В таблице 1 в первой строке приведены значения RMSE оценок многомерного сигнала (7) методами MSSA и НО-MSSA. Видно, что метод НО-MSSA отделил сигнал с бóльшей

точностью, чем метод MSSA, что аналогично результатам, полученным в работе [6] для точности оценок параметров сигнала.

5.2. Сигнал с линейно меняющимися фазами. Пусть временной ряд имеет вид

$$x_n^{(p)} = c_1^{(p)} e^{-0,01n} \cos(2\pi 0,2n + p\pi/6) + c_2^{(p)} e^{-0,02n} \cos(2\pi 0,22n + p\pi/9) + \varepsilon_n^{(p)}, \quad (8)$$

где $n = 0, \dots, 24$, $p = 1, \dots, 12$, $c_k^{(p)}$ – независимые нормальные случайные величины с нулевым средним и стандартным отклонением 1, $\varepsilon_n^{(p)}$ – последовательность независимых нормальных случайных величин с нулевым средним и стандартным отклонением $\sigma = 0,02$.

Будем сравнивать точность выделения сигнала при выборе оптимальных для каждого метода длин окна (для обоих методов оптимальное $L = 21$). Ранг сигнала в обоих случаях равен 4, а 3-ранг траекторного тензора равен 4. Соответственно, число компонент для оценки сигнала в MSSA будем выбирать равным 4 и число компонент по всем измерениям в HO-MSSA тоже равным 4.

Таблица. RMSE оценки многомерного сигнала

	MSSA	HO-MSSA
сигнал (7)	0,0093	0,0070
сигнал (8)	0,0092	0,0091

Во второй строке таблицы приведены значения RMSE оценок многомерного сигнала (8) методами MSSA и HO-MSSA. Видно, что метод HO-MSSA отделил сигнал с большей точностью, чем метод MSSA, однако это преимущество уменьшилось по сравнению со случаем равных фаз.

6. Заключение. В работе показано, что тензорный вариант HO-MSSA для выделения многомерного сигнала дал точность выше, чем обычный MSSA. Соответственно, в следующих исследованиях имеет смысл развивать теорию метода HO-MSSA для увеличения точности оценивания сигнала и его компонент.

Литература

1. Golyandina N., Nekrutkin V., Zhigljavsky A. A. Analysis of time series structure – SSA and related techniques. Boca Raton: Chapman and Hall/CRC, 2001. 320 p.

2. Ежов Ф. В. Исследование гибридных моделей нейронных сетей с применением SSA на примере реальных данных // Процессы управления и устойчивость. 2022. Т. 9. № 1. С. 223–231.
3. Головкина А. Г., Козынченко В. А., Клименко И. С. Метод последовательных приближений для построения модели динамической полиномиальной регрессии // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2023. Т. 18. № 4. С. 487–500.
4. Golyandina N., Zhigljavsky A. Singular Spectrum Analysis for Time Series. Heidelberg: Springer, 2020. 146 p.
5. Roy R., Kailath T. ESPRIT-estimation of signal parameters via rotational invariance techniques // IEEE Transactions on Acoustics, Speech, and Signal Processing. 1989. Vol. 37. No 7. P. 984–995.
6. Papy J. M., De Lathauwer L., Van Huffel S. Exponential data fitting using multilinear algebra: the single-channel and multi-channel case // Numerical Linear Algebra with Applications. 2005. Vol. 12. No 8. P. 809–826.
7. Van Huffel S., Chen H., Decanniere C., Van Hecke P. Algorithm for time-domain NMR data fitting based on total least squares // Journal of Magnetic Resonance. 1994. Vol. 110. No 2. P. 228–237.
8. Степанов Д. В., Голяндина Н. Э. Варианты метода «Гусеница»–SSA для прогноза многомерных временных рядов // Труды IV Международной конференции «Идентификация систем и задачи управления» SICPRO'05. 2005. С. 1831–1848.
9. De Lathauwer L., De Moor B., Vandewalle J. A Multilinear singular value decomposition // SIAM Journal on Matrix Analysis and Applications. 2000. Vol. 21. No 4. P. 1253–1278.

УДК 004.852

Чайкин Г. А.

Разработка инструментария для трекинга экспериментов для задачи отслеживания знаний при непрерывном обучении

Рекомендовано к публикации доцентом Блекановым И. С.

1. Введение. В задачах машинного обучения, помимо непосредственной тренировки модели, возникает и вопрос организации экспериментов, так как при увеличении выборки становится тяжелее вручную следить за предсказательной способностью модели или выводить визуальную информацию. В открытом доступе далеко не всегда находится код для обучения вместе с решением для сохранения результатов обучения модели, или решение находится, но хранение происходит лишь на локальной машине.

С помощью современного программного обеспечения можно решить проблему трекинга экспериментов по машинному обучению и автоматизировать остальные рутинные процессы, а эксперименты с увеличивающимся объемом данных можно автоматизировать, например, путем последовательного дообучения модели при поступлении новой выборки.

Данная работа посвящена вопросу автоматизации экспериментов для задачи отслеживания знаний, для которой определены последовательные данные пользовательской активности. Со временем они постепенно накапливаются.

2. Задача отслеживания знаний. Задача отслеживания знаний является задачей обучения с учителем, где в качестве данных используется пользовательская активность учащихся. Пользовательскую активность можно представить в виде последовательности ответов на задания. Эти последовательности и используются в качестве признаков. Меткой для предсказания является ответ на следующий вопрос.

Пусть задано множество последовательностей $X = \{(x_t)_{t=1}^{T_i} : T_i \in N, i \in N\}$, где $x_t = (q_t, a_t)$ – это взаимодействие учащегося

Чайкин Георгий Александрович – аспирант, Санкт-Петербургский государственный университет; e-mail: st061320@student.spbu.ru, тел.: +7(981)755-59-34

с платформой, q_t – это вопрос для учащегося на шаге t , а a_t – это ответ учащегося на вопрос q_t . Также задано множество допустимых ответов Y , и существует целевая функция $y^* : X \rightarrow Y$, значения которой $y_i = y^* \left((x_t)_{t=1}^{T_i} \right)$ известны не только на конечном подмножестве объектов $X_l = \left\{ (x_t)_{t=1}^{T_i} \right\}, i = \overline{1, l}$, но и на всем множестве X . В задаче отслеживания знаний $y_i \in Y$ – это некоторый аспект следующего взаимодействия x_{t+1} . Конкретно в данной задаче $y = [0, 1]$, где $y \in Y$ – это условная вероятность правильного ответа на следующем вопросе $p(a_{t+1} = 1 | q_{t+1}, X)$.

Полученные результаты можно использовать для разработки интеллектуальных систем обучения, для рекомендации учебных материалов [1] или для определения начальных данных для задачи распределения заданий, определенной в [2].

3. Непрерывное обучение. Задача непрерывного обучения заключается в последовательном обучении на выборках, которые поступают со временем. Одним из свойств, которое пытаются сохранить при решении задачи, является сохранение обобщающей способности на прошлых выборках [3]. Преимуществом непрерывного обучения является то, что для обучения используется лишь часть данных, а не вся накопленная за все время выборка, обработка которой может занять большое время.

4. Программная реализация. Всю автоматизацию экспериментов можно разделить на несколько разделов с учетом введенных задач и требования к сборке проекта в единый пакет.

4.1. Трекинг экспериментов через wandb. В источниках с открытым исходным кодом далеко не всегда есть код для регистрации экспериментов вместе с обучением модели машинного обучения. В решении проблемы трекинга запусков понадобится специальный инструментарий для работы с экспериментами машинного обучения, и wandb [4] является как раз одним из таких.

С использованием wandb, хватит нескольких функций, чтобы выполнять трекинг, логирование и версионирование различных объектов вроде моделей и выборок. Например, для логирования метрик во время обучения понадобится вызов функции `wandb.log`. В качестве аргумента она принимает словарь с любыми метриками. По мере работы кода метрики автоматически добавляются на сервер, где развернут wandb, и по окончании обучения будет предоставлена пол-

ная визуализация графиков с метриками или график со значениями функции потерь. Для версионирования моделей и разметки поможет `wandb.log_artifact`, принимающий на вход объект типа `Artifact` и содержащий в себе информацию о хранимом файле и информацию о версии файла, которая хранится на сервере с развернутым сервисом `wandb`. Правильно комбинируя функции, можно зарегистрировать все параметры запуска кода для повторения эксперимента.

4.2. Poetry и параметры для запуска. Весь инструментарий представляет собой код, который можно использовать не только для запуска в виде скрипта, но и в виде отдельного пакета на Python. Для разрешения конфликтов между зависимостями используется специальный инструмент Poetry [5]. Через него можно собирать пакеты, разрешать конфликты между зависимостями и отдельными командами запускать скрипты. Запуск скрипта для обучения и регистрации эксперимента выполняется следующим образом:

```
poetry install --without dev
# run_cl - команда, специально указанная для Poetry.
# Она содержит путь к функции, а не просто к .py файлу.
poetry run run_cl
```

Сборка пакета происходит следующим образом:

```
poetry build
```

Далее можно использовать пакет отдельно от кода проекта. Например, его можно установить через стандартный менеджер пакетов `pip`.

4.3. Реализация сценария непрерывного обучения. Код для реализации модели машинного обучения для решения задачи отслеживания знаний взят из [6] и далее модифицирован для работы со сценарием непрерывного обучения, который в свою очередь реализовывается через библиотеку `Avalanche` [7].

`Avalanche` является библиотекой на Python для реализации непрерывного обучения. Во многом она предназначена для задач компьютерного зрения или обработки естественных языков, но есть несколько методов, которые подходят для других задач обучения с учителем.

5. Численные эксперименты и визуализация результатов.

На основе подготовленного кода реализован метод для регистрации эксперимента и сбора информации по нему. Эксперимент состоял из обучения модели [8] на последовательных подвыборках на основе выборки из [9] с использованием нескольких подходов непрерывного обучения.

Выборка сначала была разделена на тренировочную и тестовую. Затем тренировочная разделена на 10 выборок, а после каждая выборка была разбита на тренировочную и тестовую. В итоге через wandb будет собираться по 11 метрик Area Under Curve (AUC). На рис. 1 представлен пример визуализации графиков для метрики.

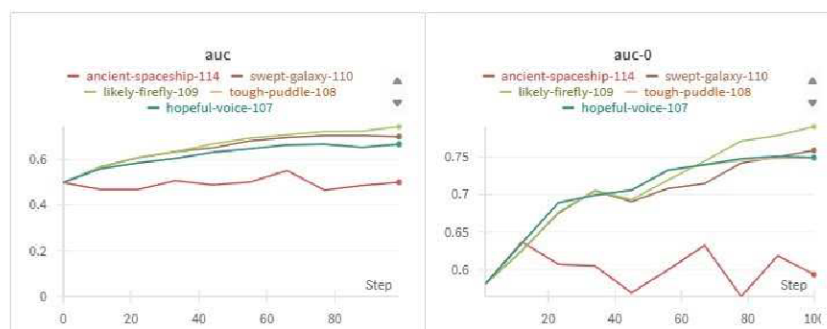


Рис. 1. Пример визуализации графиков

Графики визуализированы с помощью интерфейса wandb. Каждая точка на графиках обозначает значение метрики для i -й тестовой подвыборки. Можно подробнее изучить информацию о каждом запуске, перейдя по нужным вкладкам в интерфейсе. В частности, для запуска можно записывать аргументы для консольного скрипта, что позволяет воспроизвести эксперименты ещё раз.

Также для каждого запуска и зарегистрированного значения метрики можно сохранять версии модели. Конкретную версию модели можно загрузить как вручную, так и сделать это посредством кода. Таким образом, можно выгружать конкретные версии моделей для дальнейшей работы с ними.

Таблица. Метрики AUC для непрерывного обучения

	FromScratch	Naive	EWC	Replay	Cumulative
AUC	0,49	0,66	0,66	0,69	0,74
AUC 0	0,59	0,74	0,74	0,75	0,79
AUC 1	0,59	0,71	0,70	0,73	0,72
AUC 2	0,53	0,72	0,72	0,77	0,78
AUC 3	0,62	0,67	0,67	0,68	0,70
AUC 4	0,56	0,77	0,77	0,78	0,78
AUC 5	0,59	0,83	0,83	0,82	0,84
AUC 6	0,59	0,83	0,83	0,82	0,84
AUC 7	0,56	0,70	0,71	0,74	0,71
AUC 8	0,65	0,79	0,79	0,76	0,79
AUC 9	0,49	0,66	0,66	0,69	0,74

Для дополнительной визуализации подготовлена таблица с последними значениями метрики AUC для всех подготовленных способов непрерывного обучения для задачи отслеживания знаний. Лучше всего показывают себя методы Cumulative и Replay, так как они основаны на использовании данных с прошлых выборок. Кроме того, Cumulative и вообще на каждой эпохе работает со всей доступной выборкой, а не только с ее частью. Метод Naive представляет собой процесс последовательного дообучения на подвыборках, а EWC применяет подобие регуляризации для борьбы с потерей предсказательной способности на предыдущих подвыборках. Метод FromScratch на каждой эпохе обучает модель на подвыборках с заново инициализированными весами, поэтому метрики для него такие низкие по сравнению с остальными.

6. Заключение. В работе описана разработка инструментария для регистрации экспериментов для задачи отслеживания знаний с реализацией сценария непрерывного обучения. Выполнена автоматизация экспериментов так, что их можно однозначно отделить друг от друга, и более того воспроизвести заново результаты.

В дальнейшем планируется развивать репозиторий с кодом [10], чтобы кроме консольных скриптов был предоставлен полноценный пакет на Python, который можно использовать для эффективного решения задачи отслеживания знаний в разных сценариях.

Литература

1. Abdelrahman G., Wang Q., Nunes B. Knowledge tracing: a survey // ACM Computing Surveys. 2023. Vol. 55. No 11. P. 1–37.
2. Чайкин Г. А. Распределение проблемных тематик между учащимися // Процессы управления и устойчивость. 2023. Т. 10. № 1. С. 490–495.
3. Wang L., Zhang X. et al. A comprehensive survey of continual learning: theory, method and application // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2024. P. 1–20.
4. Weights & Biases [Электронный ресурс]: URL:<https://wandb.ai/> (дата обращения: 20.03.24).
5. Poetry – Python dependency management and packaging made easy [Электронный ресурс]: URL:<https://python-poetry.org/> (дата обращения: 20.03.24).
6. A collection of Knowledge Tracing model implementations with PyTorch [Электронный ресурс]: URL:<https://github.com/hcnoh/knowledge-tracing-collection-pytorch> (дата обращения: 20.03.24).
7. Avalanche: an End-to-End Library for Continual Learning [Электронный ресурс]: URL:<https://avalanche.continualai.org/> (дата обращения: 20.03.24).
8. Zeng J., Zhang Q. et al. Application of deep self-attention in knowledge tracing //arXiv preprint arXiv:2105.07909. 2021.
9. ASSISTmentsData – 2009-2010 ASSISTment Data [Электронный ресурс]: URL:<https://sites.google.com/site/assistmentsdata/home/2009-2010-assistment-data?authuser=0> (дата обращения: 20.03.24).
10. Continual Learning with knowledge tracing and sequential data [Электронный ресурс]: URL:<https://github.com/georgechaikin/clkt> (дата обращения: 20.03.24).

UDC 004

Huang L.

Deep residual neural network based on feature extraction for ultrasound image classification of focal liver lesions

Recommended for publication by associate professor Goncharova A. B.

1. Introduction. Liver cancer has been a disease of worldwide concern, and according to the latest statistics from the International Agency for Research on Cancer (IARC/WHO) of the World Health Group, 905,700 people will be diagnosed with liver cancer and 830,200 people will die from liver cancer in 2020, making it one of the top three causes of cancer deaths in 46 countries worldwide; it is expected that the number of new diagnoses of liver cancer worldwide in 2040 and deaths could increase by 55% [1]. Wang et al. [2] newly developed deep learning radiomics (DLRE) for elastography significantly outperformed other methods in assessing the staging of liver fibrosis. Yasaka et al. [3] demonstrated that the diagnostic performance of using a convolutional neural network (CNN) deep learning approach to differentiate liver masses in dynamic contrast-enhanced computed tomography (CT) outperforms the conventional approach. Wu et al. [4] applied a 3-layer DBN to temporal intensity curves (TICs) extracted from contrast-enhanced ultrasound (CEUS) video sequences to achieve benign and malignant classification of focal liver lesions, and they obtained the highest accuracy rate of 86.36%. Sato [5] proposed by the joint training of B-mode ultrasound images and other information from medical records and blood data an Artificial Intelligence model for benign and malignant tumor identification with an accuracy of 96.3% and an AUROC of 0.994. Tarek et al. [6] proposed a stacked sparse autoencoder feature representation based on deep learning techniques, which achieves an overall classification accuracy of 97.2% and outperforms (Multiple Support Vector Machines, K-Nearest Neighbors and Naive Bayes) and other techniques. However, most of these machine learning methods use shallow architectures. Given the powerful features and learning capabilities of deep neural networks, deep

Huang Lin – graduate student, Saint Petersburg State University; e-mail: st115840@student.spbu.ru, phone: +7(965)786-39-16

convolutional neural networks based on feature extraction for liver ultrasound image classification is a topic worthy of in-depth research. In this paper, we propose a deep residual neural network (ResNet) based on feature extraction for ultrasound image classification of liver lesions, which improves the accuracy of focal liver lesion classification.

2. Data Acquisition and Processing. The dataset used in this study was provided by The N. N. Petrov National Medicine Research Center of oncology. The collected video data were obtained by the ultrasound department through a uniform image acquisition standard, and all were verified by clinicopathological examination. Since the standard orientation and annotation of ultrasound images are very important for the clear transmission of the information presented in the static images. The information presented in static images includes transabdominal images, where the transducer is orientated in the longitudinal direction (sagittal plane); and planar abdominal images, where the transducer is in the transverse direction (axial). The videos of the cases involved contain both orientations (sagittal plane scan, subcostal scan). A total of 92 ultrasound videos were retrospectively collected from August 2022 to December 2023, including (17 Hemangioma videos, 17 Fatty Liver videos, 16 Liver Metastases videos, 25 Liver Cyst videos, and 17 Normal Liver videos). As the collected liver ultrasound is a real time video generated using high frequency sound waves. In this regard, image processing techniques were applied to segment the conventional ultrasound images of liver lesions as a basis for training and testing of classification algorithms for further feature extraction and lesion classification. A total of 3368 liver ultrasound images were obtained, including (651 Hemangioma images, 608 fatty liver images, 536 Metastases images, 1085 Cyst ultrasound images of liver cysts and 488 Normal liver images).

3. Classification methods and results.

3.1. Model. CNN have a strong representation learning capability and can extract features from images through convolution, pooling, activation, and full connectivity. A small number of shallow convolution operations can only capture simple features such as contours and lines of an image, while more complex feature representations can be captured with an increase in the number of convolutions and the size of the convolution kernel. As shown in Figure 3.2, a schematic diagram of liver ultrasound lesion classification using convolutional neural network is shown, i.e., the ultrasound image of liver lesion is input to the con-

volutional neural network to compute discriminative deep convolutional features, which are then inputted into a classifier to output the class attributes of the lesion.

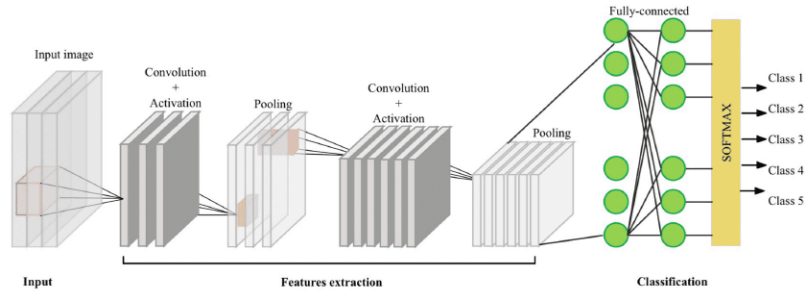


Figure 1. Schematic of convolutional neural network image classification model

Meanwhile, to improve the generalization performance of neural networks, various neural networks with different structures have emerged. For example, LetNet, AlexNet, VGGNet, GoogLeNet and so on. Among them, AlexNet introduces the Rectified Linear Unit (ReLU) activation function for the first time, which replaces the traditional Sigmoid or Tanh activation function, and the use of ReLU accelerates the training process and solves the problem of gradient disappearance. AlexNet has 8 layers to be trained, does not contain Local Response Normalization (LRN) layer and pooling layer, the first 5 layers are convolutional layers, the last 3 layers are fully connected, and the last fully connected layer is replaced by a SoftMax layer. The ResNet_18 network contains 17 convolutional layers (convolution) and 1 fully connected layer (Fully connected). Each convolutional layer uses a 3x3 convolutional kernel and a ReLU activation function to extract local features of the image.

3.2. Model testing and analysis. To ensure the objectivity and credibility of the model, the data were randomly divided into two parts: the training set (80%) and the test set (20%), of which 2694 were the training set and 674 were the test set, and the Epoch was set to 100 during the training process, considering the performance of the computer. In the training process, considering the performance of the computer, the Epoch is set to 100, and the initial learning rate is set to 0.0001. The training and testing accuracy rates and the loss

rate are calculated as shown in Figure 2. In addition, we compared the classification results with those of two models, AlexNet and VGGNet. The experimental results are shown in Table 1. The experimental results show that our model achieves an accuracy of 93.99% on the test set with an average accuracy of 92.67%.

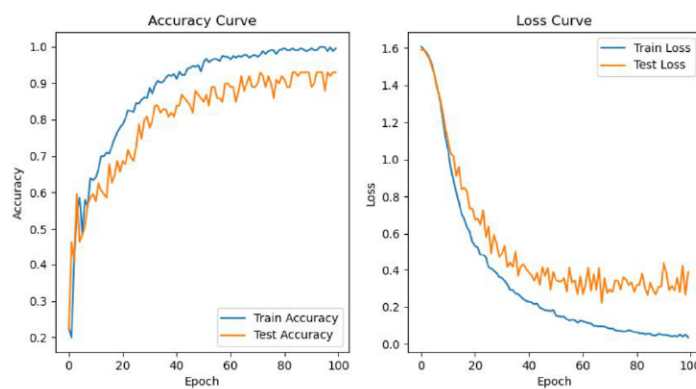


Figure 2. Model training accuracy

Table 1. Comparison of training effectiveness of different network models

	Train acc (%)	Val acc (%)
ALexNet	81.96	74.20
VGG	87.08	83.96
ResNet	93.99	90.01


This indicates that our model can effectively distinguish different types of liver lesions with high classification accuracy. Further analysis shows that compared to the two models, AlexNet and VGGNet, the ResNet_18 model has the fastest convergence speed and the best training effect on the training set and exhibits good classification performance. Through experiments, ResNet_18 based on feature extraction trained impressive results, proving the feasibility of classifying liver ultrasound images based on CNN.

4. Summary and outlook. Deep learning network models represented by CNN are often used in the field of medical image classification, and the classification training tests were conducted on three CNN models

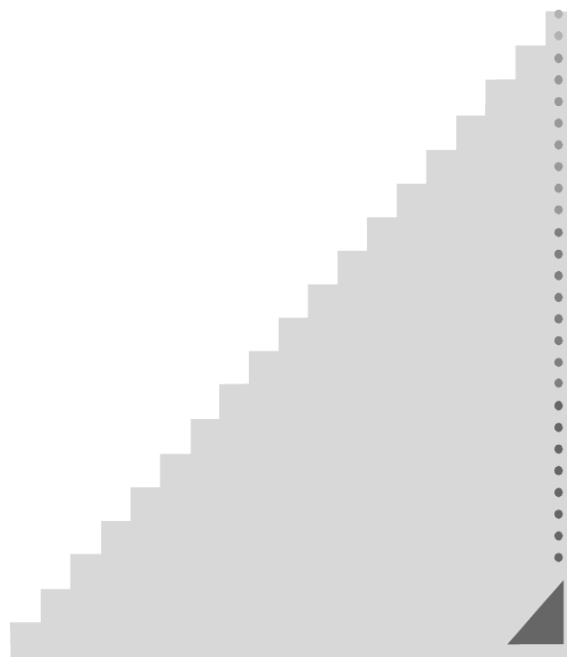
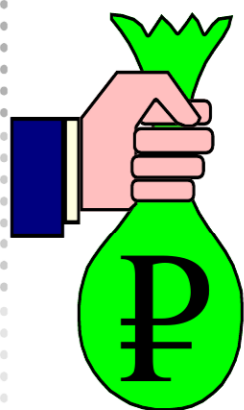
with different network depths, structures and functions, and the comprehensive view showed that the classification effect of the ResNet_18 model was better, and the correct rate was higher. The results of this study show that deep residual neural networks have potential applications in medical image classification tasks, especially in the field of liver disease diagnosis.

References

1. Runggay H., Arnold M., et al. Global burden of primary liver cancer in 2020 and predictions to 2040 // *Journal of hepatology*. 2022. Vol 77. No 6. P. 1598–1606.
2. Wang K., Lu X., et al. Deep learning Radiomics of shear wave elastography significantly improved diagnostic performance for assessing liver fibrosis in chronic hepatitis B: a prospective multicentre study // *Gut*. 2019. Vol. 68. No 4. P. 729–741.
3. Yasaka K., Akai H., et al. Deep learning with convolutional neural network for differentiation of liver masses at dynamic contrast-enhanced CT: A preliminary study // *Radiology*. 2018. Vol. 286. No 3. P. 887–896.
4. Wu K., Chen X., Ding M. Deep learning based classification of focal liver lesions with contrast-enhanced ultrasound // *Optik*. 2014. Vol. 125. No 15. P. 4057–4063.
5. Sato M., Kobayashi T., et al. Development of novel deep multimodal representation learning-based model for the differentiation of liver tumors on B-mode ultrasound images // *Journal of gastroenterology and hepatology*. 2022. Vol. 37. No 4. P. 678–684.
6. Hassan T. M., Elmogy M., Sallam E-S. Diagnosis of focal liver diseases based on deep learning technique for ultrasound images // *Arabian Journal for Science and Engineering*. 2017. Vol. 42. P. 3127–3140.



*5. Управление
социально-
экономическими
системами*



УДК 519.246.8

Аникушин В. А., Редько А. Ю.

Применение модели ARIMA для прогнозирования численности населения Земли

Рекомендовано к публикации профессором Захаровым В. В.

1. Введение. На протяжении всей истории человечества прогнозирование численности населения Земли всегда занимало важное место в научных исследованиях, обуславливая разработку стратегий социального, экономического и экологического развития общества. Поэтому важно внимательно отслеживать демографические тенденции и принимать меры для их коррекции и улучшения. В данной работе рассматривается задача прогнозирования численности населения Земли с использованием авторегрессионной модели скользящего среднего ARIMA, а также проведено исследование по выбору оптимальных параметров модели ARIMA(p,d,q), для чего были использованы два подхода: минимизация информационных критериев AIC и BIC, визуальный анализ графиков автокорреляции и частичной автокорреляции. Эти методологии позволяют не только точно прогнозировать численность населения на глобальном уровне, но и выявлять ключевые тренды и закономерности в изменениях демографической ситуации, что имеет важное значение для стратегического планирования в различных областях жизнедеятельности человека.

2. Постановка задачи. Временной ряд – это последовательность упорядоченных во времени числовых показателей, характеризующих уровень состояния и изменения изучаемого явления, $\{e_1, \dots, e_i, \dots, e_n\}$, где e_i – событие в момент времени t_i , n – общее количество событий.

Сформулируем задачу построения прогноза по временному ряду. Пусть дан временной ряд $\{e_1, \dots, e_i, \dots, e_n\}$. Требуется на его основании определить значение e_{n+k} при $k > 0$.

Редько Анна Юрьевна – студент, Санкт-Петербургский государственный университет; e-mail: st090114@student.spbu.ru, тел.: +7(989)801-17-35

Аникушин Василий Александрович – студент, Санкт-Петербургский государственный университет; e-mail: st087566@student.spbu.ru, тел.: +7(952)373-07-21

2.1. Модель АRIMA. Одним из статистических методов прогнозирования в демографии является одномерное АRIMA-моделирование. В основе этого метода лежит подход «чистого временного ряда»: текущие значения демографического показателя связаны с его прошлыми значениями, а для прогнозирования будущих значений используется информация о прошлых значениях изучаемого показателя.

Модель АRIMA универсальна и популярна при анализе одномерных временных рядов. Модель АRIMA сочетает в себе три процесса: авторегрессию (AR), скользящее среднее (MA) и процесс дифференцирования.

1. Модель авторегрессии порядка p (AR(p)) – регрессия ряда на собственные значения в прошлом.

$$y_t = \alpha + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t,$$

где y_t – стационарный ряд, ε_t – гауссов белый шум с нулевым средним и постоянной дисперсией σ_ε^2 .

2. Модель скользящего среднего порядка q (MA(q))

$$y_t = \alpha + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}.$$

3. Модель ARMA(p, q)

$$y_t = \alpha + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}.$$

Другой способ записи модели ARMA(p, q)

$$\phi(B)y_t = \theta(B)\varepsilon_t,$$

где B – разностный оператор, $By_t = y_{t-1}$.

4. Ряд описывается моделью АRIMA(p, d, q), если ряд его разностей

$$\nabla^d y_t = (1 - B)^d y_t$$

описывается моделью АRMA(p, q)

$$\phi(B)\nabla^d y_t = \theta(B)y_t.$$

2.2. Уравнение демографического баланса. Методология прогнозирования численности населения основана на анализе процессов притока и оттока в системе народонаселения и опирается на методы исследования временных рядов статистических данных о таких процессах [1–6]. К множеству таких процессов можно отнести не только рост населения Земли, но и, например, распространение вирусных заболеваний.

Пусть t_0 – начальный год имеющейся статистики, T – конечный год, $D(t)$ – число умерших людей в год t , $B(t)$ – число родившихся людей в год t . Рассмотрим возрастающие относительно t функции:

$$\text{Int}B(t_0, t) = \sum_{\tau=t_0}^t B(\tau),$$

$$\text{Int}D(t_0, t) = \sum_{\tau=t_0}^t D(\tau),$$

где $\text{Int}B(t_0, t)$ – интегральный объем притока в систему за счет рождаемости на промежутке времени от t_0 до t , $\text{Int}D(t_0, t)$ – интегральный объем оттока из системы за счет смертности на промежутке времени от t_0 до t .

Тогда балансовое уравнение имеет следующий вид

$$N(t) = N(t_0) + \text{Int}B(t_0, t) - \text{Int}D(t_0, t), \quad t > t_0, \quad (1)$$

где $N(t)$ – значение численности населения на 1 января года t .

Цель настоящей работы – применение авторегрессионной модели скользящего среднего ARIMA для прогнозирования численности населения Земли, а также оценка качества и точности прогноза по полученной модели.

3. Результаты работы. Опишем основные результаты, полученные в рамках данной работы.

3.1. Исследование выбора оптимальных параметров. В работе проведено исследование по выбору оптимальных параметров модели ARIMA(p,d,q) для прогнозирования численности населения Земли. Для анализа использовались данные ООН о численности населения с 1950 г. по 2021 г. [7]. Первая задача исследования состояла в построении ретроспективного прогноза численности населения с

1980 г. по 2021 г. на данных с 1950 г. по 1979 г. Затем верхняя граница промежутка данных увеличивалась на один год, тем самым увеличивая объем данных и уменьшая горизонт прогнозирования. Для выбора оптимальных параметров было использовано два подхода:

1. Минимизация информационных критериев АИС и ВИС, где наилучшей оказалась модель ARIMA(0, 2, 0).
2. Визуальный анализ графиков автокорреляции и частичной автокорреляции, приведший к выбору модели ARIMA(0, 2, 1).

Для оценки точности прогноза y'_t по сравнению с фактическим значением y_t при использовании указанных подходов рассмотрим следующие характеристики:

- точность (%) = $(1 - \frac{|N(t) - \bar{N}(t)|}{N(t)})100\%$;

- средняя абсолютная процентная ошибка

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{\bar{y}_t - y_t}{y_t} \right| 100\%;$$

- средняя процентная ошибка $MPE = \frac{1}{n} \sum_{t=1}^n \frac{y_t - \bar{y}_t}{y_t} 100\%$.

Полученные характеристики точности прогноза численности населения планеты в период с 1980 г. по 2010 г. при минимизации информационных критериев АИС и ВИС и визуальном анализе графиков автокорреляции представлены в таблице 1.

Таблица 1. Характеристики точности ретроспективных прогнозов численности населения Земли в период с 1980 г. по 2010 г.

Год	Минимизация AIC/BIC			Визуальный анализ графиков		
	Точность	МАРЕ	МРЕ	Точность	МАРЕ	МРЕ
1980	97,89	3,40	3,40	97,89	3,40	3,40
1981	98,48	2,49	2,49	98,48	2,49	2,49
1982	98,89	1,84	1,84	98,89	1,84	1,84
1983	99,32	1,14	1,14	99,32	1,14	1,14
1984	99,84	0,23	0,17	99,84	0,23	0,17
1985	99,52	0,82	0,82	99,52	0,82	0,82
1986	99,83	0,25	0,18	99,83	0,25	0,18
1987	99,72	0,47	-0,30	99,72	0,47	-0,30
1988	99,72	1,23	-1,21	99,72	1,23	-1,21
1989	99,10	1,78	-1,78	99,10	1,78	-1,78
1990	99,24	1,52	-1,52	99,24	1,52	-1,52
1991	99,08	1,92	-1,92	99,08	1,92	-1,92
1992	99,00	2,16	-2,16	99,00	2,16	-2,16
1993	99,57	0,91	-0,91	99,57	0,91	-0,91
1994	99,74	0,54	-0,54	99,74	0,54	-0,54
1995	99,87	0,22	-0,19	99,87	0,22	-0,19
1996	99,88	0,20	0,08	99,88	0,20	0,08
1997	99,87	0,23	0,15	99,87	0,23	0,15
1998	99,35	1,63	1,63	99,35	1,63	1,63
1999	98,49	1,02	1,02	98,49	1,02	1,02
2000	99,47	1,42	1,42	99,47	1,42	1,42
2001	99,85	0,34	0,34	99,85	0,34	0,34
2002	99,85	0,35	0,35	99,85	0,35	0,35
2003	99,84	0,39	0,39	99,85	0,39	0,39
2004	99,83	0,45	0,45	99,84	0,45	0,45
2005	99,83	0,48	0,48	99,83	0,48	0,48
2006	99,86	0,40	0,40	99,86	0,40	0,40
2007	99,87	0,38	0,38	99,87	0,38	0,38
2008	99,90	0,26	0,23	99,90	0,26	0,24
2009	99,92	0,19	0,10	99,92	0,19	0,10
2010	99,92	0,17	-0,003	99,93	0,17	0,001

Из таблицы 1 видно, что точность построенных прогнозов достаточно высокая, что подтверждает эффективность рассматриваемого метода. В обоих подходах абсолютная процентная ошибка $МАРЕ < 1\%$ в более 60% случаев и $МАРЕ < 2\%$ в более 90% случаев. А разница ошибок между двумя используемыми подходами менее 0,01% говорит о равнозначности этих подходов в проведенном численном эксперименте.

3.2. Прогнозирование показателей смертности и рождаемости. Вторая задача исследования заключалась в построении прогноза годовых значений смертности и рождаемости на основе тех же данных с использованием модели ARIMA. Прогноз строился на основании данных статистики до 1989 г. включительно с интервалом в пять лет до 2015 г. В этом случае основным критерием выбора наилучшей модели ARIMA(1, 2, 1) использовалась средняя абсолютная процентная ошибка MAPE. В таблицах 2 и 3 приведены характеристики точности прогноза рождаемости и смертности населения планеты на основе полученной ARIMA-модели.

Таблица 2. Характеристики точности ретроспективных прогнозов рождаемости населения планеты

Год	Точность	MAPE	MPE
1990	91,18	16,30	-16,30
1995	97,89	2,15	0,58
2000	97,65	3,30	2,80
2005	98,03	2,76	-0,80
2010	97,72	5,24	-5,19
2015	98,06	4,68	-4,68

Таблица 3. Характеристики точности ретроспективных прогнозов смертности населения планеты

Год	Точность	MAPE	MPE
1990	94,93	8,36	8,36
1995	95,65	7,79	-7,28
2000	97,95	2,48	1,66
2005	97,70	4,00	3,95
2010	97,36	6,88	6,88
2015	97,80	6,54	6,54

3.3. Прогноз численности населения на основе уравнения демографического баланса. Третья задача исследования включала построение прогноза численности населения планеты с использованием уравнения демографического баланса (1), при ранее полученных прогнозах смертности и рождаемости. Точность этого прогноза оказалась значительно выше, чем точность отдельных прогнозов рождаемости и смертности. Результаты прогноза отражены в таблице 4.

Таблица 4. Характеристики точности результата ретроспективных прогнозов численности населения планеты

Год	Точность	МАРЕ	МРЕ
1990	95,72	4,84	-4,84
1995	99,16	0,96	0,96
2000	99,38	0,68	0,68
2005	99,90	0,15	0,003
2010	99,74	0,51	-0,51
2015	99,91	0,30	-0,30

Сравнивая полученный прогноз с прогнозом, представленным в таблице 1, в соответствующие года, видим, что существенное отличие есть только в 1990 г. В обоих прогнозах абсолютная процентная ошибка $МАРЕ < 1\%$ почти во все рассматриваемые годы, а ее значения различаются не более, чем на 0,8%, за исключением 1990 г.

4. Заключение. Результаты проведенных вычислительных экспериментов продемонстрировали возможность применения авторегрессионной модели скользящего среднего ARIMA для ретроспективного прогнозирования численности населения Земли непосредственно на основе статистических данных численности населения прошлых периодов.

В целом, результаты данного исследования подтверждают значимость использования моделей анализа временных рядов, в частности ARIMA, для прогнозирования ключевых демографических показателей. Это открывает путь для дальнейших исследований в данной области и может способствовать разработке более точных и надежных стратегий планирования на глобальном уровне.

Литература

1. Захаров В. В., Балыкина Ю. Е. Прогнозирование динамики эпидемии коронавируса (COVID-19) на основе метода прецедентов // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2020. Т. 16. № 3. С. 249–259.
2. Афанасьев В. Н. Анализ временных рядов и прогнозирование. Саратов: Ай Пи Эр Медиа, 2020. 286 с.

3. Захаров В. В., Балыкина Ю. Е. Балансовая модель эпидемии COVID-19 на основе процентного прироста // Информатика и автоматизация. 2021. № 5(20). С. 1034–1065.
4. Копнова Е. Д., Родионова Л. А. Статистические подходы к анализу и прогнозированию демографических данных // Известия Саратовского университета. Новая серия. Серия Экономика. Управление. Право. 2016. Т. 16. Вып. 3. С. 306–315.
5. Ульянова Д. С. Статистический анализ смертности от туберкулеза // Процессы управления и устойчивость. 2023. Т. 10. № 1. С. 477–481.
6. Сорокина П. Э. Исследование и прогнозирование продолжительности жизни // Процессы управления и устойчивость. 2022. Т. 9. № 1. С. 438–442.
7. Compact (most used: estimates and medium projections) // United Nations, Department of Economic and Social Affairs, Population Division (2022) [Электронный ресурс]: URL:<https://population.un.org/wpp/Download/Standard/MostUsed/> (дата обращения: 17.01.2024).

УДК 519.8

Бушуева А. С., Власова Т. В.

Применение эвристических алгоритмов для решения задачи маршрутизации с несколькими складами

1. Введение. Задача маршрутизации транспортных средств (ТС) является одной из самых изучаемых в последние десятилетия. Существует множество ее модификаций, связанных с ограничениями, например, на грузоподъемность [1], на количество транспортных средств [2], на возможность дозагрузки на промежуточных остановках и на количество депо [3], и многими другими. Еще одной модификацией является задача маршрутизации с вывозом и доставкой [4]. В 1969 г. впервые была сформулирована задача маршрутизации ТС с несколькими депо [5].

В данной работе рассматривается задача маршрутизации ТС с несколькими депо, в которой депо могут выступать в качестве промежуточных пунктов, необходимых для дозагрузки ТС. Такая задача была предложена в [6]. Примером может служить доставка из разных кафе или ресторанов, у которых нет собственного курьера.

2. Постановка задачи. Пусть имеется конечное множество поставщиков и потребителей. Каждый потребитель ждет заказ от определенного поставщика. Имеется одно ТС грузоподъемности Q . При этом ТС отправляется от одного и того же поставщика и возвращается обратно. Перед выездом известна информация обо всех потребительских запросах и координатах всех поставщиков и потребителей.

Необходимо построить полный замкнутый маршрут для ТС, при котором будут обслужены все клиенты. При этом минимизируется сумма затрат. Затраты в данной задаче определяются временем передвижения по пути между клиентами.

3. Математическая постановка задачи. Введем следующие

Бушуева Ангелина Сергеевна – студент, Санкт-Петербургский государственный университет; e-mail: angelinabush1702@gmail.com, тел.: +7(999)525-38-42

Власова Татьяна Владиславовна – доцент, Санкт-Петербургский государственный университет; e-mail: t.vlasova@spbu.ru, тел.: +7(812)428-41-19

обозначения. Задача задана на полном ориентированном графе $G = (V, A)$, где $V = C \cup P$ – множество вершин, $A = \{(i, j) : i, j \in V, i \neq j\}$ – множество дуг, $C = \{1, \dots, n\}$ – подмножество потребителей, $P = \{n+1, \dots, n+p\}$ – подмножество p поставщиков. Каждый потребитель ждет заказ от определенного поставщика, т. е. каждому элементу $i \in C$ ставится в соответствие элемент $n+l \in P$, где l – номер поставщика. Обозначим q_i – запрос потребителя i на объем груза; (i, j) – путь, проходящий через вершины i и j со стоимостью переезда c_{ij} .

Для решения задачи используется одно ТС с грузоподъемностью Q , которую нельзя превышать на маршруте. Будем считать, что выезд ТС происходит из вершины $n+1$ от поставщика с номером $l=1$, а после обслуживания всех клиентов заканчивается там же. При этом важен порядок обхода вершин: потребитель может быть обслужен только после посещения соответствующего поставщика и только один раз. Такой маршрут будем называть *полным*.

Пусть h – это маршрут, соединяющий поставщиков, при этом начальный и конечный поставщики могут совпадать. С каждым маршрутом h связан параметр продолжительности поездки

$$\pi_h = \sum_{(i,j) \in A} c_{ij} x_{ij}^h,$$

где $x_{ij}^h = 1$, если ТС проходит путь (i, j) , $x_{ij}^h = 0$ – в противном случае. Введем следующие целочисленные переменные:

- $e_{ih} = 1$, если клиент i находится на маршруте h , и $e_{ih} = 0$ – в противном случае ($i = \overline{1, n}$);
- $f_{hl} = 2$, если маршрут h начинается и заканчивается у поставщика l , $f_{hl} = 1$, если маршрут h либо только начинается, либо только заканчивается у поставщика l , $f_{hl} = 0$ – в противном случае ($l = \overline{1, p}$).

Пусть $T = \{h\}$ – множество всех маршрутов h , удовлетворяющих условиям

$$\sum_{i=1}^n e_{ih} q_i \leq Q, \quad \sum_{l=1}^p f_{hl} = 2.$$

Величина $y_l = 1$, если ТС покидает поставщика l , $y_l = 0$ – в противном случае; $I \subseteq T$ – множество маршрутов между разными поставщиками; параметр $z_l \in \mathbf{Z}$, $z_l \geq 0$, учитывает, сколько раз ТС посещает и покидает поставщика с номером l на множестве I ($l = \overline{1, p}$).

С учетом введенных обозначений запишем рассматриваемую задачу следующим образом:

$$\min \sum_{h=1}^{|T|} \pi_h, \quad (1)$$

$$\sum_{h=1}^{|T|} e_{ih} = 1, \quad i \in C, \quad (2)$$

$$\sum_{l=1}^p y_l \leq p, \quad (3)$$

$$\sum_{i \in CUP} x_{ij}^h - \sum_{i \in CUP} x_{ji}^h = 0, \quad j \in C, \quad i \neq j, \quad h \in T, \quad (4)$$

$$\sum_{h \in I} f_{hl} = 2z_l, \quad l = \overline{1, p}, \quad (5)$$

$$\sum_{i=1}^n e_{ih} q_i \leq Q, \quad h \in T. \quad (6)$$

Здесь (1) – минимизируемая функция затрат на полный маршрут. Ограничение (2) означает, что каждый клиент посещается не более одного раза. Ограничение (3) указывает на то, что ТС заедет к поставщику l только при необходимости дозагрузки, не считая случая возвращения в начальную точку после обслуживания всех клиентов. Ограничения (4) и (5) гарантируют, что ТС после приезда в промежуточную вершину покинет ее. Условие (6) – ограничение на объем заказов в ТС.

4. Результаты экспериментов. Для решения поставленной задачи были применены эвристические методы: жадный [7], генетический [8] и муравьиный алгоритмы [9]. Использовались тестовые данные [10], содержащие до 50 точек, из которых 47 принадлежат клиентам, и три – поставщикам.

Ниже представлена таблица с результатами численных экспериментов. В ней n – количество клиентов, p – количество складов, при каждом из значений которых генетический и муравьиный алгоритмы запускались по 20 раз. В третьем столбце таблицы представлены значения целевой функции при реализации жадного алгоритма (GRA); в четвертом и пятом – средние значения целевой функции при применении генетического (GA) и муравьиного (ACO) алгоритмов; в шестом и седьмом – лучшие значения целевой функции при

реализации генетического и муравьиного алгоритмов соответственно. В муравьином алгоритме использовалось 50 колоний по 50 муравьев; в генетическом – 50 популяций по 20 особей.

Таблица. Результаты работы методов

n	p	GRA	Среднее решение		Лучшее решение	
			GA	ACO	GA	ACO
1	2	3	4	5	6	7
10	2	101,55	87,15	93,4	71,49	89,21
20	2	251,86	274,14	233,61	227,78	224,23
30	2	386,44	420,89	312,96	395,31	299,77
40	2	528,92	598,14	433,96	526,49	414,49
47	2	747,14	911,3	637,713	870,2	622,09
40	3	543,42	614,9	445,25	585,97	428,4
47	3	808,22	925,9	640,48	893,74	613,57

Из таблицы видно, что есть случаи, когда лучшее значение целевой функции, полученное при применении генетического алгоритма, меньше, чем при реализации жадного алгоритма, но среднее решение было эффективнее только при $n = 10$. При этом жадный алгоритм дает худшие значения целевой функции, чем муравьиный алгоритм. Если сравнивать решения, полученные при применении муравьиного и генетического алгоритма, то муравьиный алгоритм дает худшее значение целевой функции только при $n = 10$.

По продолжительности работы жадный алгоритм затрачивает больше времени на поиск решения, чем генетический, а муравьиный алгоритм находит решение медленнее, чем вышеупомянутые методы. Таким образом, муравьиный алгоритм строит лучшие решения, но работает медленнее.

5. Заключение. В работе рассмотрена задача маршрутизации ТС с несколькими депо и промежуточными остановками. Для решения применены эвристические методы, в ходе сравнения которых выявлено, что для рассматриваемой задачи муравьиный алгоритм находит лучшие решения за приемлемое время.

Литература

1. Чабанец А. В., Власова Т. В. Сравнительный анализ муравьиного алгоритма и метода имитации отжига для задачи маршрутизации с ограничением на грузоподъемность // Процессы управления и устойчивость. 2020. Т. 7. № 1. С. 428–432.

2. Купинская А. И., Власова Т. В. Динамическая адаптация алгоритма поиска с запретами для задачи вывоза и доставки с несколькими транспортными средствами // Процессы управления и устойчивость. 2019. Т. 6. № 1. С. 435–439.
3. Markelova A., Allahverdyan A., Martemyanov A., Sokolova I., Petrosian O., Svirkin M. Applied routing problem for a fleet of delivery drones using a modified parallel genetic algorithm // Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes. 2022. Vol. 18. Iss. 1. P. 135–148.
4. Berbeglia G., Cordeau J. F., Gribkovskaia I., Laporte G. Static pickup and delivery problems: a classification scheme and survey // Springer. 2007. Vol. 15. P. 1–31.
5. Tillman F. A. The multiple terminal delivery problem with probabilistic demands // Transportation Science. 1969. Vol. 3. No 3. P. 192–204.
6. Crevier B., Cordeau J. F., Laporte G. The multi-depot vehicle routing problem with inter-depot routes // European Journal of Operational Research. 2007. Vol. 176. No 2. P. 756–773.
7. Wang Y. Review on greedy algorithm // Theoretical and Natural Science. 2023. Vol. 14. No 1. P. 233–239.
8. Holland J. H. Adaptation in natural and artificial systems. Ann Arbor: The university of Michigan press, 1975. 183 p.
9. Colomi A., Dorigo M., Maniezzo V. Distributed Optimization by Ant Colonies // Actes de la premiere conference europeenne sur la vie artificielle, Paris, France, Elsevier Publishing. 1991. P. 134–142.
10. Zunic E. Real-world MDVRP data with realistic constraints [Электронный ресурс]: URL:https://data.4tu.nl/articles/_/12713447/1 (дата обращения: 01.03.24).

УДК 519.218

Калинин П. Е.

Модель кратного тестирования в эксперименте с бинарным исходом

Рекомендовано к публикации доцентом Лежминой Е. А.

Введение. В классической модели принятия решений в зависимости от выбора стороны, принимающей решение, и в зависимости от реализовавшегося исхода «эксперимента» происходит получение выигрыша в широком смысле этого слова. И в некоторых ситуациях, встречающихся на практике, возможно проведение исследования, уточняющего вероятности результатов, но всё же не дающего однозначный ответ на вопрос, какое именно событие произойдёт после принятия решения. Использование дополнительной информации возможно за счёт применения байесовского подхода к моделированию вероятностного процесса [1].

Обзор литературы и описание модели. В работе [2] предлагается модель для уточнения вероятностей на основе байесовского подхода – с помощью вычисления условных вероятностей. Кроме того, байесовские методы уточнения вероятностей результата за счёт получения дополнительной информации широко распространены в экономических задачах [3, 4]. Суть и общая схема самого подхода подробно описаны в [1]. В литературе по принятию решений указывается, что на практике встречаются ситуации, в которых возможно проведения «простого» тестирования [2, 5] (один тест), а возможность кратного тестирования (последовательности тестов) не освещается. В данной статье предлагается построение абстрактной модели эксперимента с бинарным исходом и возможностью кратного тестирования (проведения последовательности тестов).

Модель бинарного исхода с простым тестированием. Пусть после некоторого эксперимента могут произойти два события: Q_1 с вероятностью p и Q_2 с вероятностью $(1 - p)$. Также возможно проведение тестирования, уточняющего вероятности исходов. Под точностью тестирования i будем понимать долю случаев, в которых тест

Калинин Павел Евгеньевич – студент, Санкт-Петербургский государственный университет; e-mail: st087365@student.spbu.ru, тел.: +7(981)958-75-13

даёт прогноз, который осуществится. На параметр i накладываются естественные ограничения: $i > \frac{1}{2}$.

Результаты тестирования будем обозначать I_1 и I_2 , и интерпретировать как события – информация от теста о прогнозировании событий Q_1 и Q_2 соответственно. Построим вспомогательную таблицу 1 вероятностей исходов теста и эксперимента.

Таблица 1. Вероятности исходов при простом тестировании

	Q_1	Q_2	Сумма
I_1	pi	$(1-p)(1-i)$	$pi + (1-p)(1-i)$
I_2	$p(1-i)$	$(1-p)i$	$p(1-i) + (1-p)i$
Сумма	p	$1-p$	1

В правом столбце находятся вероятности событий I_1 и I_2 . Вычислим условные вероятности появления события Q_1 :

$$P(Q_1|I_1) = \frac{P(Q_1 \cap I_1)}{P(I_1)} = \frac{pi}{pi + (1-p)(1-i)},$$

$$P(Q_1|I_2) = \frac{P(Q_1 \cap I_2)}{P(I_2)} = \frac{p(1-i)}{p(1-i) + (1-p)i}.$$

Таким образом, получаем общую расчётную схему для апостериорных вероятностей исхода эксперимента при условии появления одного из событий: I_1 или I_2 .

Модель бинарного исхода с кратным тестированием. Рассмотрим теперь возможность кратного тестирования, т. е. проведения ряда тестов, каждый из которых генерирует одно из событий: I_1 или I_2 . Для случая двух тестирований получим таблицу 2.

Таблица 2. Вероятности исходов при двойном тестировании

	Q_1	Q_2	Сумма
I_1, I_1	pi^2	$(1-p)(1-i)^2$	$pi^2 + (1-p)(1-i)^2$
I_1, I_2	$pi(1-i)$	$(1-p)i(1-i)$	$pi(1-i) + (1-p)i(1-i)$
I_2, I_1	$pi(1-i)$	$(1-p)i(1-i)$	$pi(1-i) + (1-p)i(1-i)$
I_2, I_2	$p(1-i)^2$	$(1-p)i^2$	$p(1-i)^2 + (1-p)i^2$
Сумма	p	$1-p$	1

Введём в рассмотрение функции

$$f(p, i) = \frac{pi}{pi + (1-p)(1-i)},$$

$$g(p, i) = \frac{p(1-i)}{p(1-i) + (1-p)i}.$$

Данные функции представляют собой апостериорную вероятность события Q_1 при осуществлении событий I_1 или I_2 соответственно.

Утверждение 1. *Последовательность событий I_1, I_2 в любом порядке не изменяет вероятность исхода эксперимента, т. е.*

$$f(g(p, i), i) = g(f(p, i), i) = p.$$

Доказательство. Проведём преобразования

$$\begin{aligned} f(g(p, i), i) &= \frac{\frac{p(1-i)}{p(1-i)+(1-p)i}i}{\frac{p(1-i)}{p(1-i)+(1-p)i}i + \left(1 - \frac{p(1-i)}{p(1-i)+(1-p)i}\right)(1-i)} = \\ &= \frac{\frac{p(1-i)i}{p(1-i)+(1-p)i}}{\frac{p(1-i)i}{p(1-i)+(1-p)i} + \frac{(1-p)i(1-i)}{p(1-i)+(1-p)i}} = \\ &= \frac{pi(1-i)}{pi(1-i) + (1-p)i(1-i)} = \frac{p}{p+1-p} = p. \end{aligned}$$

Случай $g(f(p, i), i) = p$ доказывается аналогично.

Утверждение 1 доказано.

Обозначим $f_n(p, i)$ и $g_m(p, i)$ – апостериорные вероятности события Q_1 после n событий I_1 или после m событий I_2 . Тогда

$$f_n(p, i) = \frac{p^n}{pi^n + (1-p)(1-i)^n}, \quad g_m(p, i) = \frac{p(1-i)^m}{p(1-i)^m + (1-p)i^m}.$$

Утверждение 2. *Если проведён $n-1$ тест и получено $n-1$ событие I_1 , то на шаге n получение события I_1 эквивалентно ситуации, в которой априорная вероятность события Q_1 равна $f_{n-1}(p, i)$ и проведён один тест с результатом I_1 , т. е.*

$$f_n(p, i) = f(f_{n-1}(p, i), i), \quad g_m(p, i) = g(g_{m-1}(p, i), i).$$

Доказательство. Проведём преобразования

$$\frac{\frac{pi^{n-1}}{pi^{n-1}+(1-p)(1-i)^{n-1}}i}{\frac{pi^{n-1}}{pi^{n-1}+(1-p)(1-i)^{n-1}}i + \left(1 - \frac{pi^{n-1}}{pi^{n-1}+(1-p)(1-i)^{n-1}}\right)(1-i)} =$$

$$\begin{aligned}
&= \frac{\frac{pi^n}{pi^{n-1} + (1-p)(1-i)^{n-1}}}{\frac{pi^n}{pi^{n-1} + (1-p)(1-i)^{n-1}} + \frac{pi^{n-1} + (1-p)(1-i)^{n-1} - pi^{n-1}}{pi^{n-1} + (1-p)(1-i)^{n-1}}(1-i)} = \\
&= \frac{pi^n}{pi^n + (1-p)(1-i)^n}.
\end{aligned}$$

Случай $g_m(p, i) = g(g_{m-1}(p, i), i)$ доказывается аналогично.

Утверждение 2 доказано.

Утверждение 3. Пусть проведена серия из $(n + m)$ тестов. По их итогам было получено n событий I_1 и m событий I_2 (в произвольной перестановке). Тогда апостериорная вероятность события Q_1 вычисляется следующим образом:

$$P(Q_1) = \begin{cases} f_{n-m}(p, i) & \text{при } n > m, \\ g_{m-n}(p, i) & \text{при } n < m, \\ p & \text{при } n = m. \end{cases}$$

Доказательство аналогично доказательству утверждения 2, по определению функций f и g .

Рассмотрим ситуацию, когда $N = n - m > 0$, и выясним, при каком значении N вероятность события Q_1 не меньше заданной вероятности ε .

Теорема. Апостериорная вероятность события Q_1 не меньше заданной вероятности ε тогда и только тогда, когда выполнено

$$N \geq \log_{\left(\frac{1-i}{i}\right)} \left(\frac{p(1-\varepsilon)}{\varepsilon(1-p)} \right).$$

Доказательство. Пусть $P(Q_1) \geq \varepsilon$. Тогда по утверждению 2, получаем

$$\begin{aligned}
\frac{pi^N}{pi^N + (1-p)(1-i)^N} &\geq \varepsilon, \\
pi^N &\geq \varepsilon pi^N + \varepsilon(1-p)(1-i)^N.
\end{aligned}$$

Следовательно,

$$\begin{aligned}
\left(\frac{1-i}{i}\right)^N &\leq \frac{p(1-\varepsilon)}{\varepsilon(1-p)}, \\
N &\geq \log_{\left(\frac{1-i}{i}\right)} \left(\frac{p(1-\varepsilon)}{\varepsilon(1-p)} \right).
\end{aligned}$$

Заключение. Построена модель кратного тестирования для уточнения вероятностей исхода эксперимента. В рамках построенной модели сформулировано несколько утверждений. Полученные результаты могут быть полезны с практической точки зрения для выведения взаимосвязи между оценками ожидаемого выигрыша и количеством проведённых предварительных тестирований.

Литература

1. Ахмадеев Д. Р., Хрунова Е. Е. Логическая схема байесовского подхода и вероятностного моделирования // XXII Международная конференция по мягким вычислениям и измерениям. 2019. С. 378–381.
2. Шаров В. Д. Применение байесовского подхода для уточнения вероятностей событий в автоматизированной системе прогнозирования и предотвращения авиационных происшествий // Управление большими системами. Вып. 43. 2013. С. 240–253.
3. Грошев Л. Н., Кознов А. Б., Рогожникова Н. Л. Байесовские методы принятия решений в экономике // Социально-экономические науки и гуманитарные исследования. 2015. Т. 6. С. 29–31.
4. Гусева М. Е., Силаев А. М. Использование байесовских методов для макроэкономического моделирования фаз бизнес-цикла // Вестник Санкт-Петербургского университета. Экономика. 2021. Т. 37. Вып. 2. С. 298–317.
5. Эддуос М., Стэнсфилд Р. Методы принятия решений / Пер. с англ. под ред. член-корр. РАН Елисейевой М.: Аудит, ЮНИТИ, 1997. 590 с.

УДК 519.863

Карпычев А. В.

Исследование возможности моделирования спроса со стороны потребителя электроэнергии в электрической сети

Рекомендовано к публикации профессором Крылатовым А. Ю.

1. Введение. Передача электроэнергии в электрической сети является одной из важнейших составляющих современной энергетической системы. С развитием технологий и ростом энергопотребления повышение эффективности передачи и распределения электроэнергии становится все более актуальной задачей. Математическое моделирование в данной области играет ключевую роль, позволяя исследовать сам процесс передачи электроэнергии, а также снизить потери и повысить энергоэффективность электрической сети в целом.

Электрической сетью в электроэнергетике является взаимосвязь между потребителем и производителем электроэнергии посредством кабельных сетей. Ключевой особенностью современного управления является способность регулирования и распределения объема передаваемой потребителю электроэнергии.

Математическое моделирование электрических сетей применяется на различных уровнях в зависимости от назначения самой модели. Это может быть как производственный уровень [1, 2], так и уровень регионов [3] и государства [4] в целом. Назначение модели устанавливается оптимизацией целевой функции, в рамках которой часто выступает максимизация прибыли, а также снижение нагрузки на производителя и уменьшение потерь в сети. Важно отметить, что для возможности составления математической модели электрической сети ключевым фактором является определение спроса со стороны потребителя. Проблемы моделирования спроса представлены в работах [5, 6]. Тем самым, при наличии созданной математической модели предоставляется только выбрать эффективное управление.

Карпычев Александр Витальевич – магистрант, Санкт-Петербургский государственный университет; e-mail: st121048@student.spbu.ru, тел.: +7(981)137-34-90

В данной работе разрабатывается математическая модель электрической сети и находится ее решение методом множителей Лагранжа. Определяется оптимальная цена, с помощью которой представляется возможным регулировать спрос на электроэнергию.

2. Математическое моделирование энергосистемы. Для анализа возможности моделирования спроса на электроэнергию со стороны потребителя формируется модель простой энергетической системы и находится оптимальное решение методом коэффициентов Лагранжа.

2.1. Определение математической модели. Функция прибыли F_i , определяемая спросом D_i со стороны потребителя i и стохастической составляющей ω_i (характеризующей экономическую ситуацию на уровне), формирует целевую функцию с учетом затрат на производство. Затраты на производство определяются стоимостью топлива ρ и мощностью генерируемой электроэнергии Y_j на энергоустановке j . Тогда целевая функция добавочной стоимости принимает вид

$$\sum_i F_i(D_i, \omega_i) - \sum_j \rho Y_j. \quad (1)$$

На значения целевой функции устанавливаются следующие ограничения:

1. Закон сохранения энергии:

$$\sum_j Y_j(t) - L(t) - \sum_i D_i(t) = 0,$$

где $L(t)$ – потери на транспортировку электроэнергии.

2. Ограничение по объему вырабатываемой мощности:

$$Y_j - K_j \tilde{a}_j \leq 0,$$

где K_j – максимальная мощность, генерируемая энергоустановкой j , \tilde{a}_j – коэффициент состояния электрогенератора.

3. Ограничение неотрицательности:

$$Y_j \geq 0.$$

4. Ограничения по напряжению на кабельной сети:

$$\sum_k \left(Z_k(t) - Z_{k,max} \right) \leq 0,$$

$$\sum_k \left(Z_k(t) - Z_{k,min} \right) \geq 0,$$

где Z – напряжение на кабеле k , $Z_{k,max}$ – максимальное и $Z_{k,min}$ – минимальное напряжение на кабеле k .

2.2. Формирование Лагранжиана. Из полученной модели (1) формируется функция Лагранжа

$$G(D) = f(D) + \lambda_l h_l(D) + \mu_n g_n(D). \quad (2)$$

Целевая функция

$$f(D) = \sum_i F_i(D_i, \omega_i) - \sum_j \rho Y_j.$$

Условия дополняющей жесткости:

$$\mu_1(Y_j - K_j \tilde{a}_j) = 0,$$

$$\mu_2(Y_j) = 0,$$

$$\mu_3 \left(\sum_k (Z_k(t) - Z_{k,max}) \right) = 0,$$

$$\mu_4 \left(\sum_k (Z_k(t) - Z_{k,min}) \right) = 0.$$

Дифференцируя Лагранжиан (2), получим решение первого порядка, имеющее следующий вид:

$$\lambda \left(1 + \frac{\partial L(t)}{\partial D_i(t)} \right) + \sum \frac{\partial Z_k(t)}{\partial D_i(t)} (\mu_3 - \mu_4) - \frac{\partial F_i(D)}{\partial D_i(t)} = 0.$$

При этом

$$\mu_{k,3} - \mu_{k,4} = \mu_{k,max} - \mu_{k,min} = \mu_k,$$

$$\frac{\partial F_i(D)}{\partial D_i(t)} = p_i^*,$$

где p_i^* – оптимальная цена для потребителя i .

Тогда окончательное решение

$$p_i^* = \lambda \left(1 + \frac{\partial L(t)}{\partial D_i(t)} \right) + \sum \frac{\partial Z_k(t)}{\partial D_i(t)} (\mu_k).$$

Тем самым, определяя функцию потерь $L(t)$ и функцию напряжения $Z_k(t)$ и устанавливая оптимальную цену p_i^* , можно регулировать спрос D_i со стороны потребителя.

2.3. Применение модели для управления электрической сетью России. Организация управления энергосетью Российской федерации осуществляется Системным оператором Единой энергетической сетью и его территориальными подразделениями. Централизованный контроль со стороны центра создает возможность для оптимального регулирования цен на электроэнергию и управления спросом потребителя. В связи с высоким износом [7] создается потребность в снижении нагрузки на некоторые единицы генерирующего оборудования, что удобно достигается за счет снижения спроса.

Помимо повышения надежности действующего оборудования, снижается необходимость использования пиковых генерирующих установок. При этом повышается эффективность работы станции и снижаются издержки, связанные с ремонтом и поддержкой работы таких установок.

Различные модели ценообразования можно применять на практике, они предусматривают изменение цены в разное время. Допустимо использовать почасовое, дневное или ежемесячное изменение цены для удобства потребителя и в зависимости от рассматриваемого региона. Возможность применения указанных методов достигается за счет высокого уровня цифровизации систем управления энергосетью [8].

3. Заключение. При использовании указанной математической модели генерирующая компания может устанавливать необходимый спрос на электроэнергию для различных станций или регионов. Тем самым снижается нагрузка на оборудование и повышается эффективность его использования [8]. Также экономическая выгода достигается за счет снижения необходимости в размещении резервных (пиковых) мощностей. Для Российской Федерации использование такого метода регулирования удобно в связи с наличием централизованного управления всей энергосетью. Также перераспределение нагрузки на отдельные станции позволяет повысить срок службы изношенного оборудования и эффективность работы станций.

Литература

1. Устюгов Н. В., Проталинский О. М. Математическая модель потребления электроэнергии в организационно-технической системе // Вестник АГТУ. 2020. № 3. С. 116–124.
2. Крестова Ю. И., Цирюльниченко Л. А. Математическое моделирование как эффективный инструмент прогнозирования и управления производственными процессами // Вестник Южно-Уральского государственного университета. Химические технологии. 2018. № 1. С. 5–13.
3. Арончик Г. И., Балтер В. П., Чертков Б. З. Математическое моделирование и оптимизация технико-экономических показателей региональной энергосистемы в условиях нечеткости исходной информации // Вестник Самарского государственного технического университета. Математика. 2020. № 16. С. 149–155.
4. Труфанов В. В., Ханаев В. В. Математическое моделирование нагрузок потребителей электроэнергии при оптимизации развития электроэнергетических систем // Электричество. 2008. № 9. С. 2–9.
5. Эйфельд М. А. Эконометрическое моделирование спроса на энергоресурсы в России // Экономика и бизнес. 2007. № 9. С. 204–206.
6. Bohn E. R., Caramanis C. M., Schweppe C. F. Optimal pricing in electrical networks over space and time // Wiley. 1984. No 3. P. 360–376.

7. Гаак В. К., Гаак А. В. Износ оборудования – возрастающая проблема теплоэнергетики // Национальные приоритеты России. 2017. № 1(23). С. 99–102.
8. Батенков К. А. Анализ и синтез структур сетей связи методом перебора состояний // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2022. Т. 18. № 3. С. 300–315.
9. Смирнов А. А., Карпычев А. В. и др. Исследования развития основных направлений технико-экономического совершенствования ГТУ // Актуальные проблемы морской энергетики. 2022. № 11. С. 164–168.

УДК 519.8

Король М. М.

Алгоритмы равновесного распределения потоков на улично-дорожной сети города

Рекомендовано к публикации профессором Крылатовым А. Ю.

1. Введение. В современном городском планировании и инженерии транспортных систем возникает необходимость в точном прогнозировании и оптимизации загруженности дорожной инфраструктуры. Одним из ключевых инструментов для достижения этой цели является задача поиска равновесного распределения потоков (Traffic Assignment Problem).

Эта задача сфокусирована на определении оптимальных маршрутов для каждой пары отправление – прибытие. Найденное равновесие представляет собой состояние, при котором ни на одной паре изменение распределения потоков не приведет к улучшению результата. Понимание этого позволяет эффективно управлять потоками, минимизируя заторы и оптимизируя использование инфраструктуры. Таким образом, поиск равновесного распределения потоков становится критически важным для современного городского планирования и управления транспортной инфраструктурой.

2. Задача поиска равновесного распределения потоков. Представим городскую сеть в виде ориентированного графа, где имеется множество вершин и множество ребер. Каждому ребру мы задаем монотонно возрастающую, неотрицательную функцию задержки, где поток представляет собой движение вдоль ребра. Также определим спрос с помощью набора пар отправление – прибытие. Для каждой пары обозначается величина потока, которая должна быть перемещена между отправлением и прибытием. Маршрут для каждой пары рассматривается как путь от отправления до прибытия, учитывающая последовательность вершин, через которые он проходит. Задача равновесного распределения потоков принимает форму нелинейной оптимизации с линейными ограничениями [1, 2].

Король Максим Максимович – магистрант, Санкт-Петербургский государственный университет; e-mail: baspresso@gmail.com, тел.: +7(963)802-68-76

Для более точного понимания процесса, представим городскую сеть как ориентированный граф $G = (V, E)$, где V – множество перекрестков, а $E \subseteq V \times V$ – множество дорог между соседними перекрестками. Также определим набор упорядоченных пар отправление–прибытие (origin–destination, OD) $W \subseteq V \times V$. Каждая OD-пара w характеризуется положительным спросом на передвижение F^w . Для каждой OD-пары w определим маршрут r^w как последовательность дуг, соединяющих отправление и прибытие. Для каждой OD-пары введем множество всех возможных маршрутов и величину потока для каждого маршрута. При решении задачи учитываем функции задержки $t_e(f_e)$ на дугах e и вычисляем задержку для каждого маршрута $t_r^w(f)$ как сумму задержек на дугах, через которые он проходит. В конечном итоге, формулируем задачу нелинейной оптимизации с целью минимизации суммарной задержки по всем маршрутам, при линейных ограничениях на спрос и потоки [3]:

$$\min_f \sum_e \int_0^{f_e} t_e(u) du$$

при ограничениях

$$\sum_{r \in R^w} f_r^w = F^w, \quad f_r^w \geq 0.$$

3. Методы перераспределения потоков. Существуют три основных подхода к перераспределению потоков: link-based [4], path-based [5], а также bush-based [6]. В данной статье рассмотрены и представлены результаты программных реализаций link-based и path-based методов.

Алгоритм 1. Link-based подход

```

1:  $f^0 \leftarrow$  feasible link-flow solution
2: while convergence criterion is not met do
3:    $d^i \leftarrow$  direction of descent
4:    $\tau^i \leftarrow$  step size
5:    $f^{i+1} \leftarrow f^i + \tau^i \cdot d^i$ 
6: end while
7: return  $f$ 

```

В 1-й строке алгоритма 1, реализующего link-based подход, выполняется поиск решения задачи путем использования метода all-or-nothing (AON) assignment, который учитывает загруженность ребер для определения кратчайших маршрутов и направлений потоков. После начальной итерации поток для каждой OD-пары направляется вдоль кратчайшего маршрута при нулевых потоках на ребрах. В результате получается распределение потоков f^0 .

В строках 2–6 выполняется цикл while, который завершается при различных условиях, таких как максимальная разница между максимальными и минимальными задержками для OD-пар или RGAP (relative gap), по формуле, учитывающей текущие значения потоков и минимальные задержки,

$$\text{RGAP} = 1 - \frac{\sum_{w \in W} F^w t_{\min}^w}{\sum_{e \in E} f_e t_e(f_e)},$$

где $t_{\min}^w = \min_{r \in R^w} t_r^w(f)$ – минимальная задержка среди всех доступных маршрутов r для OD-пары w задержка $t_r^w(f)$.

В строке 3 определяется направление спуска (direction of descent) d_i , которое может варьироваться в различных вариантах, таких как FW (Frank–Wolfe), CFW (Conjugate Frank–Wolfe) и т. д. В классическом алгоритме Frank–Wolfe d_i вычисляется как разница между потоками по методу AON и текущими значениями потоков и f^i .

В строке 4 выбирается оптимальный шаг спуска τ^i для минимизации целевой функции.

В строке 5 происходит перераспределение потоков с учетом ранее вычисленных направлений спуска d_i и шага спуска τ^i .

Далее рассмотрим path-based подход к решению задачи равновесного распределения потоков (алгоритм 2). Основной идеей данного метода является работа с используемыми маршрутами для каждой OD-пары w , поддержанием информации \hat{R}^w о них, а также находящихся на них потоков f_r^w .

В 1-й строке алгоритма, как и в link-based подходе, задается допустимое решение задачи (AON solution). Для каждой пары w проводится поиск наикратчайшего маршрута при нулевой загруженности сети. Далее, поток, равный величине спроса F^w , направляется вдоль данного маршрута. Маршруты для каждой пары будут добавляться в специальные множества \hat{R}^w , а распределение потоков вдоль них как f^w .

```

1:  $f, \hat{R} \leftarrow$  AON solution
2: while convergence criterion is not met do
3:   for each  $w \in W$  do
4:      $p \leftarrow$  shortest path between  $w$  on  $G$  congested by  $f$ 
5:      $\hat{R}^w = \hat{R}^w \cup p$ 
6:   end for
7:   while termination condition for each  $w$  is not met do
8:     for each  $w \in W$  do
9:        $f^w, \hat{R}^w \leftarrow \mathfrak{C}(f^w, \hat{R}^w)$ 
10:    end for
11:  end while
12: end while
13: return  $x$ 

```

Со 2-й по 12-ю строку проводится добавление новых маршрутов и перераспределение потоков вдоль них.

С 3-й по 6-ю строку для каждой пары w во множество \hat{R}^w добавляется кратчайший маршрут p при текущем распределении потоков f . При реализации данного алгоритма, с целью оптимизации процесса, поиск маршрутов проводился сразу для всех пар с одинаковым истоком посредством алгоритма Дейкстры.

Основная идея, на которой строятся дальнейшие действия состоит в следующем: если \exists пара w , у которой найдутся маршруты $r_1^w, r_2^w \in \hat{R}^w$, при этом $t_{r_1}^w > t_{r_2}^w$, а также $f_{r_1}^w > 0$, тогда перенос части потока с маршрута r_1^w на маршрут r_2^w приведет к уменьшению целевой функции [7].

С 7-й по 11-ю строку алгоритм пытается привести систему (f, \hat{R}) к такому состоянию, при котором для каждой пары w задержки на маршрутах $r \in \hat{R}^w$ станут равны, $t_{r_1}^w = t_{r_2}^w, r_1, r_2 \in \hat{R}^w \forall w \in W$.

Для этого в строках 8–10 поэтапно обходятся все пары w и на каждой из них в строке 9 потоки перераспределяются так, чтобы задержки на маршрутах для текущей пары стали равны между собой.

При реализации path-based для перераспределения потоков $\mathfrak{C}(f^w, \hat{R}^w)$ для каждой пары w использовались два различных метода, описанных в [8, 9].

4. Результаты программных реализаций. Реализация алгоритмов производилась на C++, код реализации представлен в репозитории [11]. Набором данных являлась сеть SiouxFalls [12]. Так как решение задачи равновесного распределения потоков может требовать высокой точности [10], при реализации использовался вещественный тип wide-decimal с установленной точностью 10^{-100} из репозитория [13].

На рис. 1 представлены результаты работы четырех алгоритмов FW (Frank–Wolfe), CFW (Conjugate Frank–Wolfe), Pb1 (Path-based метод [8]), Pb2 (Path-based метод [9]).

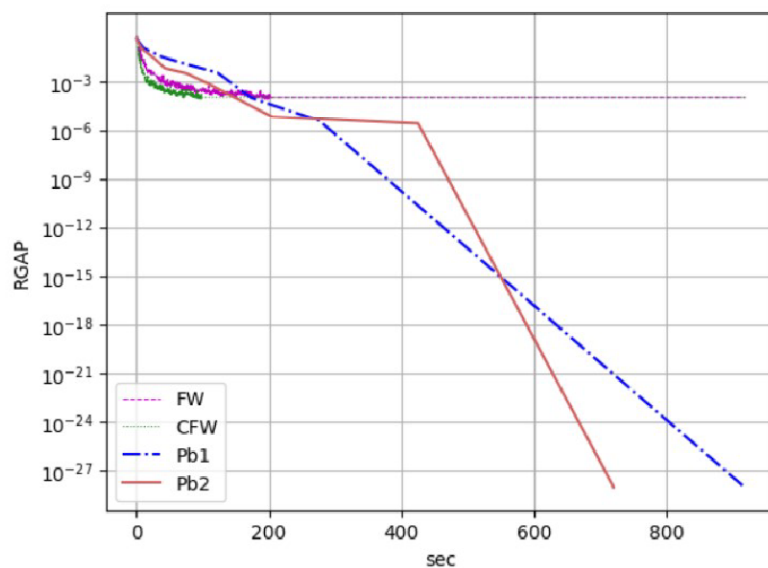


Рис. 1. Результаты программных реализаций

5. Заключение. Основным результатом данной работы является построение алгоритмов и демонстрация ими высокой точности. Дополнительно был проведен обзор существующих методов решения задачи равновесного распределения потоков.

6. Благодарность. Исследование выполнено при финансовой поддержке Российского научного фонда (проект РНФ No 22-71-10063 «Разработка интеллектуальных инструментов оптимизации мультимодальных систем распределения потоков в загруженных многопродуктовых сетях»).

Литература

1. Beckmann M., McGuire C., Winsten C. Studies in the Economics of Transportation. New Haven: Yale University Press, 1956. 226 p.
2. Krylatov A., Raevskaya A. Travel times equilibration procedure for route-flow traffic assignment problem // Learning and Intelligent Optimization. Springer. 2020. P. 68–79.
3. Крылатов А. Ю. Оптимальные стратегии управления транспортными потоками на сети из параллельных каналов // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2014. № 2. С. 120–129.
4. Mitradjieva M., Lindberg P. O. The stiff is moving – conjugate direction Frank-Wolfe methods with applications to traffic assignment // Transportation Science. 2013. Vol. 47. No 2. P. 280–293.
5. Krylatov A. Column generation for the equilibrium route-flow traffic assignment problem // Annals of Mathematics and Artificial Intelligence. 2022. Vol. 90. P. 697–713.
6. Perederieieva O., Ehrgott M., Raith A., Wang J. Y. T. A framework for and empirical study of algorithms for traffic assignment // Computers & Operations Research. 2015. Vol. 54. P. 90–107.
7. Devarajan S. A note on network equilibrium and noncooperative games // Transp. Res. Part B. 1981. Vol. 15. No 6. P. 421–426.

8. Крылатов А. Ю. Сведение задачи минимизации выпуклой сепарабельной функции с линейными ограничениями к задаче поиска неподвижной точки // Дискретный анализ и исследование операций. 2018. Т. 25. № 1. С. 75–97.
9. Krylatov A. Sequential path-equilibration algorithm for highly accurate traffic flow assignment in an urban road network // Optimization Methods and Software. 2023. Vol. 38. No 5. P. 1081–1104.
10. Bar-Gera H., Hellman F., Patriksson M. Computational precision of traffic equilibria sensitivities in automatic network design and road pricing // Procedia – Social and Behavioral Sciences. 2013. Vol. 80. P. 41–60.
11. Traffic assignment problem [Электронный ресурс]: URL:https://github.com/baspreso/Traffic_assignment_problem (дата обращения: 20.04.2024).
12. Transportation networks [Электронный ресурс]: URL:<https://github.com/bstabler/TransportationNetworks> (дата обращения: 20.04.2024).
13. Wide-decimal [Электронный ресурс]: URL:<https://github.com/skormanyos/wide-decimal> (дата обращения: 20.04.2024).

УДК 519.81

Новгородцев В. А.

О задаче управления портфелем в случае модели Хестона

Рекомендовано к публикации доцентом Кумачёвой С. Ш.

1. Введение. В данной работе применяется методология, разработанная в [1], для решения несогласованной во времени [2] задачи стохастического оптимального управления портфелем, содержащая метрику CVaR [3], в случае модели Хестона. В отличие от работы [1], динамика активов описывается многомерной моделью Хестона [4, 5] с постоянной матрицей корреляций. Функционал качества, кроме метрики CVaR, содержит еще и функционал из классической постановки.

В работе показано, что исходную задачу можно свести к задаче двухуровневой оптимизации. Однако для использования методов [1], необходимо сделать некоторые преобразования и перейти к возмущенной задаче двухуровневой оптимизации, внешнюю задачу которой можно решать градиентными методами. Также приводится сходимость решения возмущенной постановки к невозмущенной.

В качестве альтернативных способов решения задач стохастического оптимального управления отметим подходы, связанные с использованием обучения с подкреплением [6].

Данное исследование возможно применить в задачах управления портфелем на финансовых рынках.

2. Начальные построения. Рассмотрим $(\Omega, \mathcal{F}, \mathbb{P})$ – вероятностное пространство с фильтрацией $\{\mathcal{F}_t\}_{t \in [0, T]}$, которая непрерывна справа и такая, что \mathcal{F}_0 содержит множества меры нуль [7].

Новгородцев Виталий Александрович – аспирант, Санкт-Петербургский государственный университет; e-mail: vit.novgorodtcev@mail.ru, тел.: +7(921)771-68-41

Рассмотрим совместную динамику n активов в модели Хестона

$$\begin{cases} \frac{dS^i(t)}{S^i(t)} = \mu_i dt + \sqrt{v^i(t)} dW_i^S(t), \\ dv^i(t) = \kappa_i(\bar{v}_i - v^i(t))dt + \gamma_i \sqrt{v^i(t)} dW_i^v(t), \\ i = 1, \dots, n, \\ (S^1(0), \dots, S^n(0))^T = S_0, \\ (v^1(0), \dots, v^n(0))^T = v_0. \end{cases} \quad (1)$$

где процессы $S = (S^1, \dots, S^n)$ – стоимости активов; $v = (v^1, \dots, v^n)$ – волатильности активов; S_0 и v_0 – векторы начальных стоимостей и волатильностей активов; μ_i – коэффициент дрейфа актива i ; W_1^S, \dots, W_n^S – коррелированные броуновские движения, адаптированные к фильтрации $\{\mathcal{F}_t\}_{t \in [0, T]}$, соответствующие активам, имеющие корреляцию с постоянным коэффициентом $\rho_{i,j}$: $dW_i^S(t)dW_j^S(t) = \rho_{i,j}dt$ для всех $i, j = 1, \dots, n$; W_1^v, \dots, W_n^v – некоррелированные броуновские движения, адаптированные к фильтрации $\{\mathcal{F}_t\}_{t \in [0, T]}$ и соответствующие волатильностям активов. При этом для каждого актива i имеет место следующая корреляция между броуновскими движениями: $dW_i^S(t)dW_i^v(t) = \rho_{S,v}^i dt$. Величина $\kappa_i \geq 0$ – скорость возвращения к среднему, $\bar{v}_i \geq 0$ – долгосрочное среднее значение процесса волатильности, $\gamma_i > 0$ – волатильность волатильности.

Используя разложение Холецкого [5], исходные коррелированные броуновские движения $W_{Sv} = (W_1^v, \dots, W_n^v, W_1^S, \dots, W_n^S)^T$ можно выразить через некоррелированные броуновские движения $\bar{W} = (\bar{W}_1^v, \dots, \bar{W}_n^v, \bar{W}_1^S, \dots, \bar{W}_n^S)^T$, как в [8]:

$$W_{Sv} = L\bar{W} = \begin{pmatrix} E_n & 0_n \\ C_{S,v} & L_S^* \end{pmatrix} \bar{W},$$

где $C_{S,v} = \text{diag}(\rho_{S,v}^1, \dots, \rho_{S,v}^n)$, E_n – единичная матрица, L_S^* – нижняя треугольная $(n \times n)$ -матрица.

Далее введем построчное произведение $*$ вектора u размерности n на $(n \times m)$ -матрицу Q :

$$u * Q = \begin{pmatrix} u_1 q_{11} & \dots & u_1 q_{1m} \\ \dots & \dots & \dots \\ u_n q_{n1} & \dots & u_n q_{nm} \end{pmatrix}.$$

Тогда динамику S в (1) можно записать в виде

$$dS(t) * \frac{1}{S(t)} = \mu dt + \sqrt{v(t)} * (C_{S,v} \quad L_S^*) d\bar{W}(t),$$

где $dS(t) = (dS_1(t), \dots, dS_n(t))^T$, $\frac{1}{S(t)} = (\frac{1}{S_1(t)}, \dots, \frac{1}{S_n(t)})^T$, $\mu = (\mu_1, \dots, \mu_n)^T$, $\sqrt{v(t)} = (\sqrt{v^1(t)}, \dots, \sqrt{v^n(t)})^T$, а $(C_{S,v} \quad L_S^*)$ – $(n \times 2n)$ -матрица. Введем обозначение $\sigma(v(t)) = \sqrt{v(t)} * (C_{S,v} \quad L_S^*)$.

Введем множество \mathcal{A} всех допустимых управлений, т. е. таких, для которых $\mathbb{E} \int_0^T \|A_t\|^2 dt < \infty$, и \mathcal{F}_t – прогрессивно измеримых управляемых процессов [7] со значениями в компактном выпуклом множестве $\mathbb{A} = \{a \in \mathbb{R}^n \mid \sum_{i=1}^n a_i = 1, a_i \geq 0\}$.

Динамика стоимости портфеля X^A [1], составленного из n активов, под воздействием управления $A \in \mathcal{A}$ выглядит следующим образом:

$$\begin{cases} dv^1(t) = \kappa_1(\bar{v}_1 - v^1(t))dt + \gamma_1 \sqrt{v^1(t)} d\bar{W}_1^v(t), \\ \dots \\ dv^n(t) = \kappa_n(\bar{v}_n - v^n(t))dt + \gamma_n \sqrt{v^n(t)} d\bar{W}_n^v(t), \\ \frac{dX^A(t)}{X^A(t)} = A^T(t)\mu dt + A^T(t)\sigma(v(t))d\bar{W}(t), \\ X^A(0) = x_0, \quad v(0) = v_0. \end{cases}$$

Перейдем для удобства от процесса X^A к процессу $W = \ln X^A$, используя формулу Ито [5]:

$$\begin{cases} dv^1(t) = \kappa_1(\bar{v}_1 - v^1(t))dt + \gamma_1 \sqrt{v^1(t)} d\bar{W}_1^v(t), \\ \dots \\ dv^n(t) = \kappa_n(\bar{v}_n - v^n(t))dt + \gamma_n \sqrt{v^n(t)} d\bar{W}_n^v(t), \\ dW^A(t) = [A^T(t)\mu - \frac{1}{2}A^T(t)\sigma(v(t))\sigma(v(t))^T A(t)]dt + \\ \quad + A^T(t)\sigma(v(t))d\bar{W}(t), \\ v(0) = v_0, \\ W^A(0) = w_0 = \ln x_0. \end{cases} \quad (2)$$

Для любого $A \in \mathcal{A}$ существует единственное сильное решение W^A для уравнения (2) при любых начальных данных [7].

Пусть $\{W_s^{t,w_0,A} \mid t \leq s \leq T\}$ – решение уравнения (2) при фиксированных $(t, w_0) \in [0, T] \times \mathbb{R}$ и управлении $A \in \mathcal{A}$.

Далее проверим, что матрица $\sigma(v(t))\sigma^T(v(t))$ положительно определена. Заметим, что

$$\begin{aligned} & A^T(t)\sigma(v(t))\sigma^T(v(t))A(t) = \\ & = \left[A^T(t)\sqrt{v(t)} * (C_{S,v} \ L_S^*) \right] \cdot \left[A^T(t)\sqrt{v(t)} * (C_{S,v} \ L_S^*) \right]^T = \\ & = \left[A(t) * \sqrt{v(t)} \right]^T (C_{S,v} \ L_S^*) (C_{S,v} \ L_S^*)^T \left[A(t) * \sqrt{v(t)} \right]. \end{aligned}$$

Положительная определенность матрицы $\sigma(v(t))\sigma^T(v(t))$ сводится к положительной определенности матрицы $(C_{S,v} \ L_S^*)(C_{S,v} \ L_S^*)^T$. Путем элементарных преобразований получаем

$$(C_{S,v} \ L_S^*) (C_{S,v} \ L_S^*)^T = C_{S,v}^2 + L_S^*(L_S^*)^T. \quad (3)$$

Обе матрицы в правой части равенства (3) неотрицательно определены, поэтому сумма матриц неотрицательно определена. Тогда неотрицательно определена и $\sigma(v(t))\sigma^T(v(t))$. Заключаем, что $(-W^A(t))$ выпукло вниз по A [1].

3. Постановка. Рассмотрим задачу минимизации на множестве \mathcal{A} функционала [1]

$$\begin{aligned} V_{0,w_0} = \inf_{A \in \mathcal{A}} \left\{ \mathbb{E} \left[\int_0^T e^{-rt} h(-W_t^{0,w_0,A}, A(t)) dt \right] + \mathbb{E} \left[-W_T^{0,w_0,A} \right] + \right. \\ \left. + \lambda \cdot \text{CVaR}_\alpha \left(-W_T^{0,w_0,A} \right) \right\}, \end{aligned} \quad (4)$$

$$\text{CVaR}_\alpha(\xi) = \inf_{y \in \mathbb{R}} \mathbb{E} \left[y + \frac{1}{1-\alpha} (\xi - y)^+ \right],$$

где $\alpha \in (0, 1)$, $\lambda > 0$, $r > 0$ – дисконтирующий множитель; $A(t) \in \mathbb{A}$ – управление в момент времени t ; $h : \mathbb{R} \times \mathbb{A} \rightarrow \mathbb{R}$ – липшицева с не более чем квадратичным ростом, совместно выпукла вниз по первому и второму аргументам и неубывающая по первому аргументу функция для удержания процесса в определенном состоянии; значения $(0, w_0)$ у функции ценности V_{0,w_0} указывают на то, что необходимо знать значение функции ценности для момента времени $t = 0$ и начального состояния процесса $W^A(0) = w_0$.

Используя методологию работы [1], задачу (4) сведем к задаче двухуровневой оптимизации вида

$$\begin{aligned}
V_{0,w_0} &= \inf_{y \in \mathbb{R}} \bar{V}_{0,w_0}(y), \\
\bar{V}_{0,w_0}(y) &= \inf_{A \in \mathcal{A}} \left\{ \mathbb{E} \left[\int_0^T e^{-rt} h(-W_t^{0,w_0,A}, A(t)) dt - W_T^{0,w_0,A} + \right. \right. \\
&\quad \left. \left. + \lambda \left(y + \frac{1}{1-\alpha} (-W_T^{0,w_0,A} - y)^+ \right) \right] \right\}. \quad (5)
\end{aligned}$$

Таким образом, несогласованная во времени задача (4) свелась к задаче двухуровневой оптимизации. Чтобы иметь возможность решать эту задачу, используя подход из [1], необходимо добиться некоторых дополнительных свойств динамики (2) и функции $f(-W_T^{0,w_0,A}, y) = (y + \frac{1}{1-\alpha}(-W_T^{0,w_0,A} - y)^+)$: равномерной параболичности и равномерной полувогнутости [1]. Для получения этих свойств динамика (2) и функция f возмущаются, тогда задача (5) сводится к возмущенной задаче двухуровневой оптимизации с функцией ценности $\bar{V}_{0,w_0,\varepsilon}$, где $\varepsilon > 0$, в которой внутренняя задача является классической, поэтому ее можно решать через уравнения Гамильтона–Якоби–Беллмана (ГЯБ) [7], а задачу внешнего уровня можно решать через градиентные методы, поскольку можно провести аналогичные рассуждения, что и в [1], так как в данном случае $(-W^A(t))$ выпукло вниз по A , коэффициенты динамики (2) липшицевы и функция h обладает необходимыми свойствами. При этом справедлива оценка:

$$|\bar{V}_{0,w_0}(y_\varepsilon^*) - \bar{V}_{0,w_0}(y^*)| < C\varepsilon \quad (6)$$

для всех y_ε^* , которые доставляют минимум $\bar{V}_{0,w_0,\varepsilon}$, и y^* минимизирующее \bar{V}_{0,w_0} . При этом C зависит только от $h, f, \lambda, \alpha, \mathbb{A}, r$ и параметров в динамике (2).

Запишем возмущенную динамику, возмущенную f (через инфимум свертку [1]) и $\bar{V}_{0,w_0,\varepsilon}(y)$.

Возмущенная динамика

$$\begin{cases} dv^1(t) = \kappa_1(\bar{v}_1 - v^1(t))dt + \gamma_1\sqrt{v^1(t)}d\bar{W}_1^v(t), \\ \dots \\ dv^n(t) = \kappa_n(\bar{v}_n - v^n(t))dt + \gamma_n\sqrt{v^n(t)}d\bar{W}_n^v(t), \\ d\widehat{W}^{A,\varepsilon}(t) = [A^T(t)\mu - \frac{1}{2}A^T(t)\sigma(v(t))\sigma(v(t))^T A(t)]dt + \\ \quad + a^T(t)\sigma(v(t))d\bar{W}(t) + \varepsilon d\widehat{W}(t), \\ v(0) = v_0, \\ \widehat{W}^{A,\varepsilon}(0) = w_0, \end{cases} \quad (7)$$

где процесс \widehat{W} – адаптирован к фильтрации $\{\mathcal{F}_t\}_{t \in [0, T]}$ и некоррелирован с остальными броуновскими движениями в \bar{W} .

Инфимум-свертка $f_\varepsilon(g(\cdot), \cdot) : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ с $g(x) = -x$, $x \in \mathbb{R}$, такая же, как и в [1].

Возмущенная функция ценности

$$\begin{aligned} \bar{V}_{0, w_0, \varepsilon}(y) = \inf_{A \in \mathcal{A}} \left\{ \mathbb{E} \left[\int_0^T e^{-rt} h(-\widehat{W}_t^{0, w_0, A, \varepsilon}, A(t)) dt + \right. \right. \\ \left. \left. + f_\varepsilon(g(\widehat{W}_T^{0, w_0, A, \varepsilon}), y) \right] \right\}. \end{aligned} \quad (8)$$

4. Градиент. Используя подход из [1], можно показать, что функция $\bar{V}_{0, w_0, \varepsilon}$ полувогнута и выпукла вниз из-за вида динамики (7), вида функции (8), выпуклости вниз $-W^A(t)$ по A и свойств функции h и f_ε , тогда можно использовать градиентные методы, сходящиеся к глобальному минимуму. Также, следуя доказательствам в [1], можно получить, что градиент

$$\nabla_y \bar{V}_{0, w_0, \varepsilon}(y) = \mathbb{E} \nabla_y f_\varepsilon(g(\widehat{W}_T^{0, w_0, A^*, \varepsilon}), y), \quad \forall y \in \mathbb{R}, \quad (9)$$

где A^* – оптимальное управление задачи (8) при данном $y \in \mathbb{R}$, которое можно построить через решение уравнения ГЯБ в силу равномерной параболичности и выпуклости вниз уравнения ГЯБ [1].

Таким образом, фиксируя некоторый $y \in \mathbb{R}$ и найдя оптимальное управление A^* , можно найти градиент, используя методы Монте-Карло или линейные уравнения в частных производных [1].

5. Уравнение Гамильтона – Якоби – Беллмана. Значение $\bar{V}_{0,w_0,\varepsilon}(y)$ при любом фиксированом $y \in \mathbb{R}$ можно найти через решение уравнения ГЯБ. При этом решение существует при достаточно общих предположениях в вязкостном смысле [7, 9, 10].

Перейдем к записи уравнения ГЯБ для $\bar{V}_{0,w_0,\varepsilon}(y)$ в явном виде. В уравнении ГЯБ фигурирует выражение $\text{tr}(\sigma(x, t, a)\sigma^T(x, t, a)\nabla_x^2 v)$, где x – состояния в динамике (7), tr – след матрицы, $a \in \mathbb{A}$, ∇_x^2 – гессиан.

Введем обозначение $k = (k_1, \dots, k_{2n}) = a^T \sigma(v(t))$. В случае динамики (7) получаем

$$\begin{aligned} & \sigma(x, t, a)\sigma^T(x, t, a) = \\ & = \begin{pmatrix} \gamma_1^2 v^1 & 0 & \dots & 0 & k_1 \gamma_1 \sqrt{v^1} \\ 0 & \gamma_2^2 v^2 & \dots & 0 & k_2 \gamma_2 \sqrt{v^2} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \gamma_n^2 v^n & k_n \gamma_n \sqrt{v^n} \\ k_1 \gamma_1 \sqrt{v^1} & k_2 \gamma_2 \sqrt{v^2} & \dots & k_n \gamma_n \sqrt{v^n} & \sum_{i=1}^{2n} k_i^2 + \varepsilon^2 \end{pmatrix}. \end{aligned}$$

Введем обозначение $\partial_{ij}^2 v = \frac{\partial^2}{\partial v^i \partial v^j} v$. Тогда

$$\begin{aligned} \text{tr}(\sigma(x, t, a)\sigma^T(x, t, a)\nabla_x^2 v) &= \gamma_1^2 v^1 \partial_{11}^2 v + \dots + \gamma_n^2 v^n \partial_{nn}^2 v + \\ &+ k_1 \gamma_1 \sqrt{v^1} \partial_{W1}^2 v + \dots + k_n \gamma_n \sqrt{v^n} \partial_{Wn}^2 v + \\ &+ k_1 \gamma_1 \sqrt{v^1} \partial_{1W}^2 v + \dots + k_n \gamma_n \sqrt{v^n} \partial_{nW}^2 v + \left[\sum_{i=1}^{2n} k_i^2 + \varepsilon^2 \right] \partial_{WW}^2 v = \\ &= \sum_{i=1}^n \gamma_i^2 v^i \partial_{(v^i)^2}^2 v + \sum_{i=1}^n k_i \gamma_i \sqrt{v^i} \partial_{Wv^i}^2 v + \sum_{i=1}^n k_i \gamma_i \sqrt{v^i} \partial_{v^i W}^2 v + \\ &+ [a^T \sigma(v(t))\sigma(v(t))^T a + \varepsilon^2] \partial_{WW}^2 v. \end{aligned}$$

Уравнение ГЯБ принимает следующий вид:

$$\begin{aligned} \partial_t v - rv + \inf_{a \in \mathbb{A}} \left[\sum_{i=1}^n \kappa_i (\bar{v}_i - v^i(t)) \partial_{v^i} v + (a^T \mu - \frac{1}{2} a^T \sigma(v(t))\sigma(v(t))^T a) \partial_W v + \right. \\ \left. + \frac{1}{2} \left\{ \sum_{i=1}^n \gamma_i^2 v^i \partial_{(v^i)^2}^2 v + \sum_{i=1}^n k_i \gamma_i \sqrt{v^i} \partial_{Wv^i}^2 v + \sum_{i=1}^n k_i \gamma_i \sqrt{v^i} \partial_{v^i W}^2 v + \right. \right. \\ \left. \left. + [a^T \sigma(v(t))\sigma(v(t))^T a + \varepsilon^2] \partial_{WW}^2 v \right\} + h(-W, a) \right] = 0, \end{aligned}$$

$$\begin{aligned}
(t, v^1, \dots, v^n, W) &\in [0, T) \times \mathbb{R}_+^n \times \mathbb{R}, \\
k &= (k_1, \dots, k_{2n}) = a^T \sigma(v(t)), \\
v(T, v^1, \dots, v^n, W) &= f_\varepsilon(g(W), y).
\end{aligned}$$

Также известно [1, 7, 9], что

$$\bar{V}_{0, w_0, \varepsilon}(y) = v(0, v_0^1, \dots, v_0^n, w_0).$$

Таким образом, при каждом фиксированном $y \in \mathbb{R}$ можно решить внутреннюю задачу в (8), решив уравнение ГЯБ, и найдя соответствующее оптимальное управление и затем подставив его в динамику (7) для вычисления градиента методом Монте-Карло, используя выражение (9). Далее делается градиентный шаг. Градиентный метод сойдется к глобальному минимуму в силу выпуклости вниз функции $\bar{V}_{0, w_0, \varepsilon}$. При этом будет иметь место оценка сходимости (6) к решению исходной задачи V_{0, w_0} .

6. Заключение. Основными достижениями данной работы являются следующие. Разработанная в [1] методология успешно применена для несогласованной во времени задачи стохастического оптимального управления портфелем, в которой динамика активов описывается моделью Хестона, а функционал качества, кроме метрики CVaR, содержит функционал из классической постановки. Исходная задача свелась к задаче двухуровневой оптимизации, для решения которой были сделаны преобразования, что привело к возмущенной задаче двухуровневой оптимизации, задача внешнего уровня которой решается уже градиентными методами, а внутренняя – через уравнение Гамильтона–Якоби–Беллмана. Приводится сходимость решения возмущенной задачи к невозмущенной.

Исследование имеет широкие возможности применения в задачах управления портфелем на финансовых рынках.

Литература

1. Miller C. W., Yang I. Optimal control of conditional value-at-risk in continuous time // SIAM J, Control Optim. 2017. Vol. 55. No 2. P. 856–884.

2. Bjork T., Murgoci A. A general theory of markovian time inconsistent stochastic control problems [Электронный ресурс]: URL:https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1694759 (дата обращения: 20.03.2024).
3. Uryasev S., Rockafellar R. T. Conditional value-at-risk: optimization approach // Stochastic Optimization: Algorithms and Applications. Applied Optimization. 2001. Vol. 54. P. 411–435.
4. Heston S.L. A closed-form solution for options with stochastic volatility with applications to bond and currency options // The Review of Financial Studies. 1993. Vol. 6. Iss. 2. P. 327–343.
5. Grzelak L. A. Mathematical Modeling and Computation in Finance. London: World Scientific, 2020. 556 p.
6. Новгородцев В. А., Кудин П. С., Кумачева С. Ш. Об оптимизации решений с помощью прогнозного модуля в обучении с подкреплением // Процессы управления и устойчивость. 2023. Т. 10. № 1. С. 454–461.
7. Fleming W. H., Soner H. M. Controlled Markov Processes and Viscosity Solutions. 2 ed. New York: Springer, 2006. 429 p.
8. Wadman W.S. An advanced Monte Carlo method for the multi-asset Heston model [Электронный ресурс]: URL:<https://www.researchgate.net/publication/325952691> (дата обращения: 20.03.2024).
9. Touzi N. Optimal Stochastic Control, Stochastic Target Problems, and Backward SDE. 1 ed. New York: Springer, 2012. 214 p.
10. Nikitin F. F. Viscosity solutions and programmed iteration method for Isaacs equation // Vestnik of Saint Petersburg University. Applied Mathematics. Computer Science. Control Processes. 2014. No 2. P. 84–92.

УДК 519.83

Русакова К. В.

Сравнение стратегий с короткой памятью в игре «дилемма заключенного»

Рекомендовано к публикации профессором Парилиной Е. М.

1. Введение. В бесконечно повторяющихся играх игроки могут использовать знание о своих действиях и действиях противника на предыдущих шагах для выбора стратегии на текущем шаге игры. И люди, и компьютеры не имеют возможности запоминать бесконечно много информации, и возникает проблема хранения памяти о предыдущих шагах. В статье рассмотрена бесконечно повторяющаяся игра «дилемма заключенного» с игроками, имеющими возможность запоминать информацию о не более, чем одном шаге.

2. Описание стратегий. Рассмотрим одновременную игру «дилемма заключенного» [1, 2] с параметрами $c < d < a < b$ и матрицей выплат, представленной в таблице 1.

Таблица 1. Матрица выплат

	C	D
C	(a, a)	(c, b)
D	(b, c)	(d, d)

Для бесконечно повторяющейся игры «дилемма заключенного» рассмотрим шесть стратегий поведения [3], которые обозначим через $S_1, S_2, S_3, S_4, S_5, S_6$. Стратегия поведения представляет собой функцию η , зависящую от номера шага t и истории $h(t)$ [4]. Стратегия поведения ниже описана для первого игрока, для второго описывается аналогично, поскольку игроки симметричны в игре. Изучаемые стратегии поведения имеют вид:

1. Стратегия S_1 – *All C* предписывает всегда сотрудничать, т. е.

$$\eta_1(t, h(t)) \equiv C.$$

Русакова Ксения Владимировна – студент, Санкт-Петербургский государственный университет; e-mail: st097257@student.spbu.ru, тел.: +7(911)167-39-27

2. Стратегия S_2 – *All D* предписывает всегда предавать, т. е.

$$\eta_2(t, h(t)) \equiv D.$$

3. Стратегия S_3 – *TFT* (*Tit-for-Tat*, «око-за-око») предписывает игроку на первом шаге сотрудничать, на остальных – повторять действие противника на предыдущем шаге, т. е.

$$\eta_3(t, h(t)) = \begin{cases} C, & t = 1 \text{ и } (t \geq 2 : h(t) = \{(x_i, y_i)_{i=1}^{t-1} | y_{t-1} = C\}), \\ D, & \text{в противном случае.} \end{cases}$$

Стратегия предполагает немедленное наказание за отказ противника от сотрудничества, при этом за одно предательство игрок наказывается только один раз и, если противник возвращается к сотрудничеству, то стратегия *TFT* предписывает сотрудничать.

4. Стратегия S_4 – *Suspicious Tit-for-Tat* (*STFT*) [5] предписывает на первом шаге предавать, на последующих шагах – повторять действие противника, т. е.

$$\eta_4(t, h(t)) = \begin{cases} D, & t = 1 \text{ и } (t \geq 2 : h(t) = \{(x_i, y_i)_{i=1}^{t-1} | y_{t-1} = D\}), \\ C, & \text{в противном случае.} \end{cases}$$

Стратегия похожа на *TFT*, за исключением первого шага, на котором данная стратегия предписывает предавать, как бы предупреждая соперника о последствиях предательства.

5. Стратегия S_5 – «Скользкий» на первом шаге сотрудничает, на последующих выполняет противоположное действие противника, т. е.

$$\eta_5(t, h(t)) = \begin{cases} C, & t = 1 \text{ и } (t \geq 2 : h(t) = \{(x_i, y_i)_{i=1}^{t-1} | y_{t-1} = D\}), \\ D, & \text{в противном случае.} \end{cases}$$

Когда игрок, использующий эту стратегию, видит, что на предыдущем шаге его противник сотрудничает, то, предполагая, что противник на данном шаге поступит также, он хочет воспользоваться шансом и предать его. При этом стратегия предписывает сотрудничество, если противник предал на предыдущем шаге.

6. Стратегия S_6 на первом шаге предписывает предавать, на последующих предписывает выбирать действие, противоположное противнику, т. е.

$$\eta_6(t, h(t)) = \begin{cases} D, & t = 1 \text{ и } (t \geq 2 : h(t) = \{(x_i, y_i)_{i=1}^{t-1} | y_{t-1} = C\}), \\ C, & \text{в противном случае.} \end{cases}$$

Выигрыши игрока в бесконечно повторяющейся игре считаются с учетом дисконтирующего фактора δ , одинакового для обоих игроков:

$$E_i(x, y) = \sum_{n=1}^{\infty} \delta^{n-1} K_i(x_n, y_n). \quad (1)$$

В таблице 2 приведена память, т. е. количество предыдущих шагов, информацию о которых использует каждая из стратегий.

Таблица 2. Память k , требуемая для реализации стратегии

	S_1	S_2	S_3	S_4	S_5	S_6
k	0	0	1	1	1	1

3. Результаты игры. Смоделируем ситуации, в которых встречаются игроки, играющие в соответствии со стратегиями $S_1, S_2, S_3, S_4, S_5, S_6$.

Теорема. *Выигрыши игроков для каждой пары стратегий представлены в таблице 3. Ситуация (S_2, S_2) является абсолютным равновесием по Нэшу.*

Доказательство. Рассмотрим результат игры при встрече двух игроков, играющих стратегии S_1 . На каждом шаге обе стратегии предписывают выбрать C , поэтому на каждом шаге оба игрока получают выигрыш $K_1(C, C) = K_2(C, C) = a$, тогда выигрыши игроков в бесконечно повторяющейся игре по формуле (1):

$$E_1(S_1, S_1) = E_2(S_1, S_1) = \sum_{n=1}^{\infty} \delta^{n-1} a = a \sum_{n=1}^{\infty} \delta^{n-1} = \frac{a}{1-\delta}.$$

Для остальных пар стратегий выигрыш считается аналогичным образом.

Второе утверждение теоремы следует из Народной теоремы [6], так как на каждом шаге оба автомата играют равновесие по Нэшу в базовой игре.

Обозначим

$$E(S_5, S_3) = \left(\frac{a + b\delta + d\delta^2 + c\delta^3}{1 - \delta^4}, \frac{a + c\delta + d\delta^2 + b\delta^3}{1 - \delta^4} \right), \quad (2)$$

$$E(S_5, S_4) = \left(\frac{c + a\delta + b\delta^2 + d\delta^3}{1 - \delta^4}, \frac{b + a\delta + c\delta^2 + d\delta^3}{1 - \delta^4} \right), \quad (3)$$

$$E(S_6, S_3) = \left(\frac{b + d\delta + c\delta^2 + a\delta^3}{1 - \delta^4}, \frac{c + d\delta + b\delta^2 + a\delta^3}{1 - \delta^4} \right), \quad (4)$$

$$E(S_6, S_4) = \left(\frac{d + c\delta + a\delta^2 + b\delta^3}{1 - \delta^4}, \frac{d + b\delta + a\delta^2 + c\delta^3}{1 - \delta^4} \right). \quad (5)$$

С учетом обозначений (2)–(5) в таблице 3 представлены выигрыши игроков в бесконечно повторяющейся игре. Знак «'» у номера формулы обозначает смену выплат местами в этой формуле.

Таблица 3. Выигрыши игроков, где $(x, y)' = (y, x)$

	S_1	S_2	S_3	S_4	S_5	S_6
S_1	$\frac{(a,a)}{1-\delta}$	$\frac{(c,b)}{1-\delta}$	$\frac{(a,a)}{1-\delta}$	$(c + \frac{a\delta}{1-\delta}, b + \frac{a\delta}{1-\delta})$	$(a + \frac{c\delta}{1-\delta}, a + \frac{b\delta}{1-\delta})$	$\frac{(c,b)}{1-\delta}$
S_2	$\frac{(b,c)}{1-\delta}$	$\frac{(d,d)}{1-\delta}$	$(b + \frac{d\delta}{1-\delta}, c + \frac{d\delta}{1-\delta})$	$\frac{(d,d)}{1-\delta}$	$\frac{(b,c)}{1-\delta}$	$(d + \frac{b\delta}{1-\delta}, d + \frac{c\delta}{1-\delta})$
S_3	$\frac{(a,a)}{1-\delta}$	$(c + \frac{d\delta}{1-\delta}, b + \frac{d\delta}{1-\delta})$	$\frac{(a,a)}{1-\delta}$	$(\frac{c+b\delta}{1-\delta^2}, \frac{b+c\delta}{1-\delta^2})$	(2)'	(4)'
S_4	$(b + \frac{a\delta}{1-\delta}, c + \frac{a\delta}{1-\delta})$	$\frac{(d,d)}{1-\delta}$	$(\frac{b+c\delta}{1-\delta^2}, \frac{c+b\delta}{1-\delta^2})$	$\frac{(d,d)}{1-\delta}$	(3)'	(5)'
S_5	$(a + \frac{b\delta}{1-\delta}, a + \frac{c\delta}{1-\delta})$	$\frac{(c,b)}{1-\delta}$	(2)	(3)	$(\frac{a+d\delta}{1-\delta^2}, \frac{a+d\delta}{1-\delta^2})$	$\frac{(c,b)}{1-\delta}$
S_6	$\frac{(b,c)}{1-\delta}$	$(d + \frac{c\delta}{1-\delta}, d + \frac{b\delta}{1-\delta})$	(4)	(5)	$\frac{(b,c)}{1-\delta}$	$(\frac{d+a\delta}{1-\delta^2}, \frac{d+a\delta}{1-\delta^2})$

В таблице 4 для каждой пары стратегий представлено количество раундов в долях от всех раундов, в которых реализуется кооперативная пара стратегий (C, C) . Будем называть это число *индексом кооперации* [7].

Таблица 4. Индексы кооперации для различных наборов стратегий

	S_1	S_2	S_3	S_4	S_5	S_6
S_1	1	0	1	1	0	0
S_2	0	0	0	0	0	0
S_3	1	0	1	0	1/4	1/4
S_4	1	0	0	0	1/4	1/4
S_5	0	0	1/4	1/4	1/2	0
S_6	0	0	1/4	1/4	0	1/2

4. Матрица победителей. Исследуем условия, как часто стратегия является выигрышной, т. е. если противник использует какую-либо стратегию $S_1, S_2, S_3, S_4, S_5, S_6$ с равной вероятностью, то какова вероятность получения большего выигрыша по сравнению с противником. Пусть $b = 1, c = 0$, т. е. $0 < d < a < 1$.

В таблице 5 представлен анализ каждой пары стратегий. Знаком «+» обозначена ситуация, в которой выигрыш первого игрока больше выигрыша второго, «=» – равен, «-» – меньше. Последние два столбца представляют собой вероятности первого игрока встретить такого игрока, что выигрыш у первого игрока будет больше $P(+)$ или не меньше $P(+, =)$, чем выигрыш противника.

Таблица 5. Сравнение выигрышей игроков при $b = 1, c = 0$

	S_1	S_2	S_3	S_4	S_5	S_6	$P(+)$	$P(+, =)$
S_1	=	-	=	-	-	-	0	1/3
S_2	+	=	+	=	+	+	2/3	1
S_3	=	-	=	-	-	-	0	1/3
S_4	+	=	+	=	+	+	2/3	1
S_5	+	-	+	-	=	-	1/3	1/2
S_6	+	-	+	-	+	=	1/2	2/3

Из таблицы 5 можно сделать вывод, что использование стратегии S_2 и S_4 приносит больший выигрыш по сравнению со стратегиями S_1, S_3, S_5 и S_6 при встрече со случайным противником. Менее хороший, но лучший чем S_1, S_3, S_5 результат показывает стратегия S_6 , которая позволяет получить не меньший, чем противник, выигрыш с вероятностью $2/3$.

5. Заключение. В данной работе рассмотрены различные стратегии поведения с короткой памятью в бесконечно повторяющейся игре «дилемма заключенного», рассмотрены результаты их взаимодействия. Изучены ситуации возникновения кооперации между двумя стратегиями как на каждом шаге, так и с какой-либо периодичностью. В дальнейшем предполагается рассмотреть взаимодействие

стратегий, которые являются автоматами, при условии наличия затрат на хранение памяти.

Литература

1. Писарева А. М. Построение стратегии наказания в повторяющихся играх «дилемма заключенного» // Процессы управления и устойчивость. 2023. Т. 10. № 1. С. 472–476.
2. Петросян Л. А., Зенкевич Н. А., Шевкопляс Е. В. Теория игр. СПб.: БХВ-Петербург, 2012. 432 с.
3. Pei H. Reputation Effects under Short Memories // Proceedings of the 24th ACM Conference on Economics and Computation. 2023. P. 1046–1046.
4. D’Arcangelo C., Andreozzi L., Faillo M. Human players manage to extort more than the mutual cooperation payoff in repeated social dilemmas // Scientific Reports. 2021. Vol. 11. No. 1. Art. no 16820.
5. García J., Van Veelen M. No strategy can win in the repeated prisoner’s dilemma: linking game theory and computer simulations // Frontiers in Robotics and AI. 2018. Vol. 5. Art. no 102.
6. Печерский С. Л., Беляева А. А. Теория игр для экономистов. Вводный курс. СПб.: Изд-во Европейского университета, 2001. 342 с.
7. Suna P., Parilina E. M., Gaob H. Two-stage network games modeling the belt and road initiative // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. 2022. Т. 18. № 1. С. 87–98.

UDC 519.83

Bykova M. S.

Payment schemes and their properties for cooperative stochastic games

Recommended for publication by professor Parilina E. M.

1. Introduction. There are various approaches to sustain cooperation over time, particularly, in [1, 2] searching for payment schemes satisfying some desirable properties is considered for the class of finite-horizon dynamic games and games played over event trees, respectively. The objective of this paper is to define payment scheme for sustaining cooperation in cooperative infinite-horizon stochastic games and generalise some properties, presented in the articles mentioned above, to this class of games. Stochastic infinite-horizon games were introduced by Shapley L. S. [3].

2. Model of a cooperative infinite-horizon stochastic game. In this section the model of a cooperative infinite-horizon stochastic game is briefly described [4].

The model is presented by following elements:

1. A finite set of players $N = \{1, \dots, n\}$;
2. A finite non-empty set of states $\Omega = \{\omega_1, \dots, \bar{\omega}\}$;
3. For each player $i \in N$, a finite non-empty set of strategies A_i^ω in state $\omega \in \Omega$ with element $a_i^\omega \in A_i^\omega$ which is called an action of player $i \in N$ in this state. The vector of players' strategies $a^\omega = (a_i^\omega : i \in N)$ is called an action profile in state $\omega \in \Omega$. $A^\omega = A_1^\omega \times \dots \times A_n^\omega$ is a set of action profiles in state $\omega \in \Omega$.
4. A finite payoff function $K_i^\omega : \prod_{k \in N} A_k^\omega \rightarrow \mathbb{R}$ of player $i \in N$ in state $\omega \in \Omega$.
5. A transition function $p(\cdot | \omega, a^\omega) : \Omega \times A^\omega \rightarrow \Delta(\Omega)$ from state $\omega \in \Omega$ and action profile $a^\omega \in \prod_{i \in N} A_i^\omega$, where $\Delta(\Omega)$ is a probability distribution on set Ω .

Bykova Maria Sergeevna – graduate student, Saint Petersburg State University; e-mail: st076609@student.spbu.ru, phone: +7(902)700-92-94

The work was performed with financial support of the Russian Science Foundation grant No 22-11-00051, <https://rscf.ru/project/22-11-00051/>

6. An initial state of the game $\pi_0 = (\pi_0^{\omega_1}, \dots, \pi_0^{\omega_n})$ defined by probability distribution, where π_0^ω is a probability that the state ω is realized in the first time period, $\sum_{\omega \in \Omega} \pi_0^\omega = 1$.

The time discrete and the number of periods of time is infinite. The game is playing as follows:

1. Before the game starts, the initial state is chosen from Ω set according to probability distribution π_0 , that is, π_0^ω is the probability of stochastic game begins in ω state.
2. At the first time period the ω state is realized. Players choose their actions simultaneously. Player $i \in N$ choose the strategy $a_i^\omega \in A_i^\omega$. Therefore, the action profile $a^\omega = (a_i^\omega : i \in N) \in A_1^\omega \times \dots \times A_n^\omega$ is realized. In which case the player i obtain the payoff $K_i^\omega(a^\omega)$. As soon as the players get to know that the action profile a^ω occurred the game comes over next state $\omega' \in \Omega$ with probability $p(\omega'|\omega, a^\omega)$.
3. In the second time period the state $\omega' \in \Omega$ is realized with probability $p(\omega'|\omega, a^\omega)$. Player $i \in N$ choose the action $a_i^{\omega'} \in A_i^{\omega'}$. Therefore, the action profile $a^{\omega'} = (a_i^{\omega'} : i \in N) \in A_1^{\omega'} \times \dots \times A_n^{\omega'}$ is realized. In which case the player i obtain the payoff $K_i^{\omega'}(a^{\omega'})$.
4. Further, the game continues according to the scheme described above.

Strategy of player $i \in N$ is a function $\phi_i = \{\phi_i(k)\}_{k=1}^\infty$ and $\phi_i(k) : h(k) \times \Omega \rightarrow \Delta(A_i^\omega)$, where $h(k)$ is a history of stage k , which is a set of pairs of states and action profiles realized before the stage k : $((\omega(1), a(1)), (\omega(2), a(2)), \dots, (\omega(k-1), a(k-1)))$. We denote by Φ_i the set of strategies of player i , and by $\phi = (\phi_i : i \in N)$, $\phi_i \in \Phi_i$, the strategy profile in stochastic game.

The discounted payoff of player i in stochastic infinite-horizon game is defined as mean value of payoff according to initial state of the game $\omega(1)$ and strategy profile η :

$$E_i(\phi) = E^{\omega(1), \phi} \sum_{k=1}^{\infty} \delta^{k-1} K_i^{\omega(k)}(a(k)).$$

Players agree to cooperate and jointly choose their strategies in any state $\omega \in \Omega$ to maximise their total discounted payoffs over the entire horizon. The existence of maximum of total discounted cooperative

payoff is proved in [3]. Besides, the situation for which this maximum is achieved is a stationary strategies profile.

Stationary strategy $\eta_i = \{\eta_i(k)\}_{k=1}^{\infty}$, $\eta_i(k) : \Omega \rightarrow \Delta(A_i^{\omega})$, does not depend on the discrete-time parameter and forces player to choose the same strategy in the same state regardless of the period's history. We denote the stationary strategies profile in stochastic game by $\eta = (\eta_i : i \in N)$, and the set of stationary strategies of player i by H_i .

In that case, the transition function can be represented by matrix of size $\bar{\omega} \times \bar{\omega}$:

$$\Pi(\eta) = \begin{pmatrix} p(\omega_1|\omega_1, a^{\omega_1}) & \cdots & p(\bar{\omega}|\omega_1, a^{\omega_1}) \\ p(\omega_1|\omega_2, a^{\omega_2}) & \cdots & p(\bar{\omega}|\omega_2, a^{\omega_2}) \\ \cdots & \cdots & \cdots \\ p(\omega_1|\bar{\omega}, a^{\bar{\omega}}) & \cdots & p(\bar{\omega}|\bar{\omega}_1, a^{\bar{\omega}}) \end{pmatrix}.$$

Let $\eta^* = (\eta_1^*, \dots, \eta_n^*)$ be the vector of cooperative stationary strategies maximising total discounted payoffs, that is:

$$\max_{\eta \in \prod_{i \in N} H_i} \sum_{i \in N} E_i(\eta) = \sum_{i \in N} E_i(\eta^*),$$

where $E_i(\eta) = (E_i^{\omega_1}(\eta), \dots, E_i^{\bar{\omega}}(\eta))$ is the vector of payoffs of player i in any subgame. Then, $\eta_i^*(\omega) = a_i^{*\omega}$ and $a^{*\omega} = (a_i^{*\omega} : i \in N)$ is an action profile realized when players choose their optimal cooperative strategies η_i^* , $i \in N$.

In case of stationary players' strategies, cooperative payoff satisfies following recurrent equation:

$$V(N) = \max_{\eta \in \prod_{i \in N} H_i} \sum_{i \in N} E_i(\eta) = \sum_{i \in N} K_i(a^*) + \delta \Pi(\eta^*)V(N), \quad (1)$$

where $K_i(a^*) = (K_i^{\omega_1}(a^{*\omega_1}), \dots, K_i^{\bar{\omega}}(a^{*\bar{\omega}}))$.

We can obtain the matrix form of $V(N)$ from (1):

$$V(N) = (I - \delta \Pi(\eta^*))^{-1} \sum_{i \in N} K_i(a^*), \quad (2)$$

where I is an identity matrix.

Taking into account the choice of initial state in accordance with probability distribution π_0 the payoff of maximal coalition in the described stochastic game is defined as follows:

$$\bar{V}(N) = \pi_0 V(N).$$

3. Properties of payment schemes in stochastic games. In this section, the payment schemes for cooperative infinite-horizon stochastic games are defined, that is, the set of payments for any state and any time period of cooperative game. These payments can be built to satisfy some desirable properties. The formulation of them is provided below as well.

New stochastic game may be defined in accordance with some function

$$\beta_i^\omega(k, a^\omega) = \begin{cases} \beta_i^\omega(k, a^{*\omega}), & \text{if } a^\omega = a^{*\omega}, \\ K_i^\omega(k, a^\omega), & \text{if } a^\omega \neq a^{*\omega}, \end{cases}$$

which is called *payment function* of player i in state ω . This function determines the payment that player $i \in N$ receives in state ω at time k . Let $\beta_i(k, a) = (\beta_i^{\omega_1}(k, a^{\omega_1}), \dots, \beta_i^{\omega_n}(k, a^{\omega_n}))$ be the vector of payments of player i .

Definition 1. The set $\mathcal{B} = (\beta_i(k, a^*) : i \in N, k = 1, 2, \dots)$ is called a *payment scheme*.

Now and further we assume payments to be *stationary*, that is, $\beta_i(k, a^\omega) = \beta_i(a^\omega), \forall i \in N, k = 1, 2, \dots$

Payments of player i in subgame G^ω satisfy recurrent equation:

$$B_i^\omega(\eta^*) = \beta_i^\omega(a^{*\omega}) + \delta \sum_{\omega' \in \Omega} p(\omega' | \omega, a^{*\omega}) B_i^{\omega'}(\eta^*).$$

Let $B_i(\eta^*) = (B_i^{\omega_1}(\eta^*), \dots, B_i^{\omega_n}(\eta^*))$ be the payments of player i in a vector form. Then, the total discounted players' payoff received according to payment scheme \mathcal{B} in a new stochastic game satisfies the following recurrent equation:

$$\sum_{i \in N} B_i(\eta^*) = \sum_{i \in N} \beta_i(a^{*\omega}) + \delta \Pi(\eta^*) \sum_{i \in N} B_i(\eta^*). \quad (3)$$

Definition 2. A sequence of pairs of time periods and states realized at corresponding time period with accordance to cooperative strategy profile η^* , that is,

$$\Psi(k, \omega(k)) = (\psi_\tau = (\tau, \omega(\tau)) : \tau = 1, \dots, (k-1))$$

is called a *cooperative trajectory* of the game.

Denote by $S(k)$ the capitalized *savings* under cooperative trajectory $\Psi(k, \omega(k))$ given by

$$S(k) = \sum_{\tau=1}^{k-1} \frac{1}{\delta^{k-\tau}} \sum_{i \in N} \left(K_i^{\omega(\tau)} \left(a^{*\omega(\tau)} \right) - \beta_i^{\omega(\tau)} \left(a^{*\omega(\tau)} \right) \right).$$

P1: Feasibility. \mathbf{B} is *feasible* if

$$\sum_{i \in N} \beta_i^{\omega(k)} \left(a^{*\omega(k)} \right) \leq \sum_{i \in N} K_i^{\omega(k)} \left(a^{*\omega(k)} \right) + S(k) \quad (4)$$

for any time period k of the game. Condition (4) means that in each stage of the game the total payments at the stage k made according to payment scheme \mathbf{B} should not exceed the total cooperative payoff received in this stage plus savings done in this stage. At the first stage we have

$$\sum_{i \in N} \beta_i^{\omega(1)} \leq \sum_{i \in N} K_i^{\omega} \left(a^{*\omega(0)} \right)$$

because no savings are available, i.e., $S(1) = 0$.

P2: Stage budget balance (SBB). \mathbf{B} is *stage budget balanced* if

$$\sum_{i \in N} \beta_i^{\omega} \left(a^{*\omega(k)} \right) = \sum_{i \in N} K_i^{\omega} \left(a^{*\omega(k)} \right), \quad \omega \in \Omega. \quad (5)$$

Under condition (5), total payments in any state ω equal to optimal joint payoff in this state, i.e., this payoff is fully allocated between the players.

P3: Efficiency. \mathbf{B} is *efficient* if

$$\sum_{i \in N} B_i(\eta^*) = V(N).$$

This property ensures expected total discounted payments to the players in all stages over the entire horizon equals to total cooperative discounted payoff.

Approval 1. *If payment scheme \mathbf{B} is stage budget balanced, then it satisfies the Feasibility and Efficiency properties.*

Proof. When payment scheme is stage budget balanced, feasibility property holds with equality, therefore, there is no need to make savings,

that is $S(k) = 0, \forall k$. In case of stationary payment functions, we can obtain the matrix form of total payments $\sum_{i \in N} B_i(\eta^*)$ from (3):

$$\sum_{i \in N} B_i(\eta^*) = (I - \delta \Pi(\eta^*))^{-1} \sum_{i \in N} \beta_i(a^*),$$

then, efficiency property is obviously satisfied as payment scheme is stage budget balanced.

Note 1. Due to the lack of terminal period of time, triviality and positive terminal savings properties from paper [2] make no sense for the model of stochastic infinite-horizon game.

To proceed, we make following assumptions:

A1: If a player deviates from cooperation trajectory and cooperation breaks down in some state $\omega(k)$, i.e. the strategy profile shifts from η^* to $(\hat{\eta}_i, \eta_{N \setminus i}^*)$, the players switch to their Nash equilibrium strategies in the remaining subgame starting from the next state $\omega(k+1)$.

A2: There exists a unique Nash Equilibrium in each subgame starting from any state or a mechanism for selecting a unique equilibrium if there are many.

Note 2. The existence of Nash equilibrium strategy profile is proved in [5].

Let η^{ne} be the Nash equilibrium strategy profile, then, $\eta^{ne}(\omega) = a^{ne, \omega}$ and $a^{ne} = (a^{ne, \omega_1}, \dots, a^{ne, \omega})$. Payoff of player i in subgame G^ω in the Nash equilibrium is

$$E_i^\omega(\eta^{ne}) = K_i^\omega(a^{ne}) + \delta \sum_{\omega' \in \Omega} p(\omega' | \omega, a^{ne}) E_i^{\omega'}(\eta^{ne}).$$

For each state ω we denote by \hat{a}_i^ω the solution of

$$\arg \max_{\hat{a}_i^\omega \in A_i^\omega} \left\{ K_i^\omega(\hat{a}_i^\omega, a_{N \setminus i}^\omega) + \delta \sum_{\omega' \in \Omega} p(\omega' | \omega, a^{\omega'}) K_i^{\omega'}(\hat{a}_i^\omega, a_{N \setminus i}^\omega) \right\},$$

that is, the best reply action of player i to $a_{N \setminus i}^*$. Then, $\hat{a}_i = (\hat{a}_i^{\omega_1}, \dots, \hat{a}_i^{\omega})$ is the vector of optimal actions of player i in each state of the game.

P4: Dynamic individual rationality (DIR). \mathcal{B} is *dynamically individually rational* if for any state $\omega \in \Omega$ and for all $i \in N$

$$(I - \delta \Pi(\eta^*))^{-1} \beta_i(a^*) \geq (I - \delta \Pi(\eta^{ne}))^{-1} K_i(a^{ne}). \quad (6)$$

If the inequality (6) is satisfied then expected player's payoff in the subgame starting from any state when the payment scheme \mathcal{B} is accepted is greater or equal to the Nash equilibrium outcome of player i in subgame starting from the correspondent state, that is, no player finds optimal to deviate from cooperative trajectory in any subgame.

P5: Stability against individual deviation (SAID). \mathcal{B} is stable against individual deviation if for all $i \in N$

$$(I - \delta\Pi(\eta^*))^{-1}\beta_i(a^*) \geq \max_{\hat{a}_i \in A_i} \left\{ K_i(\hat{a}_i, a_{N \setminus i}^*) + \delta\Pi(\hat{\eta}_i, \eta_{N \setminus i}^*) E_i(\eta^{ne}) \right\} \\ := BR_i, \quad (7)$$

where matrix $\Pi(\hat{\eta}_i, \eta_{N \setminus i}^*)$ is defined as follows:

$$\Pi(\hat{\eta}_i, \eta_{N \setminus i}^*) = \begin{pmatrix} p(\omega_1|\omega_1, (\hat{a}_i^{\omega_1}, a_{N \setminus i}^{*\omega_1})) & \cdots & p(\bar{\omega}|\omega_1, (\hat{a}_i^{\omega_1}, a_{N \setminus i}^{*\omega_1})) \\ p(\omega_1|\omega_2, (\hat{a}_i^{\omega_2}, a_{N \setminus i}^{*\omega_2})) & \cdots & p(\bar{\omega}|\omega_2, (\hat{a}_i^{\omega_2}, a_{N \setminus i}^{*\omega_2})) \\ \cdots & \cdots & \cdots \\ p(\omega_1|\bar{\omega}, (\hat{a}_i^{\bar{\omega}}, a_{N \setminus i}^{*\bar{\omega}})) & \cdots & p(\bar{\omega}|\bar{\omega}_1, (\hat{a}_i^{\bar{\omega}}, a_{N \setminus i}^{*\bar{\omega}})) \end{pmatrix},$$

and $BR_i = (BR_i^{\omega_1}, \dots, BR_i^{\bar{\omega}})$.

In what follows, we will assume that the maximum operator applied to a matrix or vector computes the maximum row by row. By the first assumption mentioned above, after the deviation is identified, the game is played noncooperatively from the next stage onward. Expression in the LHS of the inequality (7) is a cooperative payoff of player i in matrix form, that is, each component of the resulting vector in LHS equals to cooperative payoff of player i in the subgame starting from correspondent state. The RHS of the (7) is a value player can get by unilateral deviation from the cooperative trajectory also presented in matrix form.

We can combine DIR and SAID properties and obtain the following condition:

$$(I - \delta\Pi(\eta^*))^{-1}\beta_i(a^*) \geq \max \{ (I - \delta\Pi(\eta^{ne}))^{-1}K_i(a^{ne}), BR_i \}.$$

4. Minimal payment scheme. In this section, we introduce a non-trivial payment scheme that satisfies Feasibility, DIR and SAID properties.

Definition 3. A feasible payment scheme \mathcal{B} that satisfies

$$(I - \delta\Pi(\eta^*))^{-1}\beta_i(a^*) = \max \{ (I - \delta\Pi(\eta^{ne}))^{-1}K_i(a^{ne}), BR_i \},$$

is called a *minimal* payment scheme and denoted by \mathcal{B}_{min} .

Therefore, players' payments are calculated by formula:

$$\beta_i(a^*) = (I - \delta\Pi(\eta^*)) \max\{(I - \delta\Pi(\eta^{ne}))^{-1}K_i(a^{ne}), BR_i\} \quad \forall i \in N. \quad (8)$$

If \mathcal{B}_{min} exists, it satisfies the feasibility property by definition, and DIR holds with equality which implies the SAID property if $(I - \delta\Pi(\eta^{ne}))^{-1}K_i(a^{ne}) \geq BR_i(k)$, or SAID holds with equality, implying the DIR property if $(I - \delta\Pi(\eta^{ne}))^{-1}K_i(a^{ne}) < BR_i(k)$. A \mathcal{B}_{min} scheme ensures that no player has an incentive to deviate along the cooperative state trajectory.

5. Example. We illustrate the construction of payment scheme with an example of two-player infinite-horizon stochastic game with two possible states $\Omega = \{\omega_1, \omega_2\}$, where in each state the game in normal form is defined:

$$\begin{aligned} \Gamma^{\omega_1} &= \langle N, A_1^{\omega_1} = \{a_1, a_2\}, A_2^{\omega_1} = \{b_1, b_2\}, K_1^{\omega_1}, K_2^{\omega_1} \rangle, \\ \Gamma^{\omega_2} &= \langle N, A_1^{\omega_2} = \{c_1, c_2\}, A_2^{\omega_2} = \{d_1, d_2\}, K_1^{\omega_2}, K_2^{\omega_2} \rangle, \end{aligned}$$

where $K_i^{\omega_1}$ and $K_i^{\omega_2}$, $i \in N$, are defined by matrices

$$\omega_1: \begin{array}{cc} & \begin{array}{cc} b_1 & b_2 \end{array} \\ \begin{array}{c} a_1 \\ a_2 \end{array} & \begin{pmatrix} (10, 10) & (0, 12) \\ (12, 0) & (2, 5) \end{pmatrix}, \end{array} \quad \omega_2: \begin{array}{cc} & \begin{array}{cc} d_1 & d_2 \end{array} \\ \begin{array}{c} c_1 \\ c_2 \end{array} & \begin{pmatrix} (15, 15) & (0, 20) \\ (20, 0) & (7, 1) \end{pmatrix}. \end{array}$$

Transition probabilities from states ω_1 and ω_2 are described by transition probability matrices

$$\begin{aligned} \omega_1: & \begin{array}{cc} & \begin{array}{cc} b_1 & b_2 \end{array} \\ \begin{array}{c} a_1 \\ a_2 \end{array} & \begin{pmatrix} (0.3, 0.7) & (1, 0) \\ (0, 1) & (0.5, 0.5) \end{pmatrix}, \\ \\ \omega_2: & \begin{array}{cc} & \begin{array}{cc} d_1 & d_2 \end{array} \\ \begin{array}{c} c_1 \\ c_2 \end{array} & \begin{pmatrix} (0.7, 0.3) & (0, 1) \\ (1, 0) & (0.5, 0.5) \end{pmatrix}, \end{aligned}$$

where the element (k, l) of the matrix consists of transition probabilities from state ω_j , $j = \{1, 2\}$, to states ω_1, ω_2 under condition that first player chooses strategy $a_k^{\omega_j}$ and second player chooses strategy $a_l^{\omega_j}$. Let δ be the common discount factor, that is, $\delta = 0.9$.

We calculate payments using expression (8). First, we can obtain optimal cooperative strategies η^* using expression (1) and calculate the cooperative payoff of the game using formula (2):

$$V(N) = \begin{pmatrix} 246.32 \\ 253.68 \end{pmatrix}.$$

To calculate payments we also need to determine Nash equilibrium strategy profile $\eta^{ne} = \{\eta_1^{ne}, \eta_2^{ne}\}$, where $\eta_i^{ne} = \{\eta_i^{ne, \omega} : \omega \in \Omega\}$, $i = 1, 2$, from given payoff matrices:

$$\eta^{ne} = \{(a_2, c_2), (b_2, d_2)\}.$$

The players' payoffs at the Nash equilibrium are:

$$E_1(\eta^{ne}) = (I - \delta\Pi(\eta^{ne}))^{-1}K_1(a^{ne}) = \begin{pmatrix} 42.5 \\ 47.5 \end{pmatrix},$$

$$E_2(\eta^{ne}) = (I - \delta\Pi(\eta^{ne}))^{-1}K_2(a^{ne}) = \begin{pmatrix} 32 \\ 28 \end{pmatrix}.$$

Further, we compute BR_i , $i = 1, 2$, which is the maximal payoff that player i can achieve by one-stage deviation. Taking into account assumption A1, we obtain:

$$BR_1 = \max_{\hat{a}_1 \in A_1} \{K_1(\hat{a}_1, a_2^*) + \delta\Pi(\hat{\eta}_1, \eta_2^*)E_1(\eta^{ne})\} = \begin{pmatrix} 54.75 \\ 58.25 \end{pmatrix},$$

$$BR_2 = \max_{\hat{a}_2 \in A_2} \{K_2(a_1^*, \hat{a}_2) + \delta\Pi(\eta_1^*, \hat{\eta}_2)E_2(\eta^{ne})\} = \begin{pmatrix} 40.8 \\ 45.2 \end{pmatrix}.$$

Now, we can calculate the payments by (8):

$$\beta_1(a^*) = \begin{pmatrix} 3.27 \\ 8.03 \end{pmatrix}, \quad \beta_2(a^*) = \begin{pmatrix} 1.308 \\ 7.292 \end{pmatrix}.$$

Finally, we check if our payment scheme satisfies Feasibility, Efficiency, SBB, DIR and SAID properties and conclude that the constructed feasible payment scheme satisfies DIR and SAID properties but doesn't satisfy SBB and Efficiency properties.

6. Conclusion. In the paper, we considered the model of discrete-time cooperative infinite-horizon stochastic game and presented the set

of payment scheme properties for this class of games. A payment scheme satisfying Feasibility, DIR, and SAID properties by construction was proposed and the way this scheme can be constructed was provided. In conclusion, we illustrated our results with an example of a simple infinite-horizon stochastic game.

References

1. Parilina E. M., Zaccour G. Payment schemes for sustaining cooperation in dynamic games // *Journal of Economic Dynamics and Control*. 2022. Vol. 139. No 139. P. 1200–1216.
2. Parilina E. M., Zaccour G. Payment schemes for sustaining cooperation in dynamic games played over event trees // *European Journal of Operational Research*. 2023. Vol. 313. No 313. P. 1200–1216.
3. Shapley L. S. Stochastic Games // *Proceedings of National Academy of Sciences of the USA*. 1953. Vol. 39. No 39. P. 1095–1100.
4. Parilina E. M. Solutions of cooperative stochastic games with transferable payoffs // PhD diss. Saint Petersburg State University. 2018. P. 329.
5. Fink A. M. Equilibrium in a stochastic n -person game // *Journal of Science of the Hiroshima University*. 1964. A-I 28. P. 89–93.

UDC 519.832.3

Lorits E. M.

Evolutionary model with environmental feedback and agent opinions

Recommended for publication by associate professor Gubar E. A.

1. Introduction. Evolutionary game theory is widely used in many scientific fields. For example, evolutionary games are used to simulate the interaction of large numbers of agents in a network [2,3]. In medicine, evolutionary games can be used, for example, to find methods to fight cancer [9] or to solve the problem of vaccinating the population [6].

Papers [4,5] consider the population and the changes that occur in that population, taking into account the environment and the state of that environment. The paper explores the idea of how the state of the environment and agents' opinions about it affect the state of the system. The population, the environment and agents' opinions form a hierarchical structure, where a change in one parameter of the system responsible for the state of the environment, the population or agents' opinions, causes a change in the remaining elements of the system.

2. An evolutionary game. Consider a population of size N existing in a finite space. It is assumed that the state of the population changes as a result of random pairwise interactions between its agents. It is also assumed that the number of agents is large and that each individual agent has no significant effect on the population [7,8]. Another assumption is that the population has two types of behaviour that agents can follow. An agent's choice of the i th type of behaviour is similar to the choice of the i th pure strategy in a non-cooperative game. It leads to a partition of the population into two subgroups. The agents of each subgroup are programmed to use the same pure strategy. The state of the population is defined as a vector $x_N(t) = (x_1(t), x_2(t))$, where each component $x_i(t)$ is the fraction of the population using the pure strategy i . This vector can be thought of as a mixed population strategy [7]. Let $x(t) = x_1(t)$, then $x_2(t) = 1 - x(t)$. The payoffs of the agents refer to the number of offspring (in biological systems) or the number of successors

Lorits Ekaterina Mikhailovna – graduate student, Saint Petersburg state University; e-mail: st077304@student.spbu.ru, phone: +7(952)237-91-92

(in economic and social systems) that follow the net strategy i . Over time, random pairwise meetings between agents occur in the population.

The outcomes of these encounters can be described by a bimatrix game [1]. Traditionally, in evolutionary games, the process is usually considered from the point of view of the first player, so the framework below is formulated in terms of the first player. Let e^i be the vector corresponding to the i -th net strategy of the player. The i th element is one and all others are zero. We introduce the function $u(e^i, x_N) = e^i A x_N$, $i = 1, \dots, n$ as the expected payoff of the agent with pure strategy i when facing a random opponent. This payoff depends on the population state vector x_N . Based on the payoff of the randomly chosen agent, the corresponding average payoff of the population is determined

$$u(x_N, x_N) = \sum_{i \in K} x_i u(e^i, x_N), i = 1, \dots, n. \quad (1)$$

The change in population composition corresponds to a change in the proportion of agents adhering to net strategy i . These changes are described by the replicator dynamics $\dot{x} = x(1-x)(u(e^1, x_N) - u(e^2, x_N))$, depending on the fractional distribution of players in the population x_N and the first player's payoff matrix A . In this paper, the population is assumed to be dependent on environmental feedbacks that affect the expected payoffs of agents. The resources available to the agents are considered as the environment. The state of the environment is described by the parameter $n(t)$, $n \in [0, 1]$, where $n = 0$ ($n = 1$) when the environment is completely consumed (replenished). The change in the state of the environment is determined by the dynamics $\dot{n} = n(1-n)(\theta x + \psi(1-x))$ proposed in [5]. It depends on the change in the state of the population. Agents following the first pure strategy replenish resources n at a rate $\theta > 0$, while agents following the second pure strategy cause resources to be depleted at a rate $\psi = -1$.

The player payoff matrix $A_n = nA_1 + (1-n)A_0$ establishes the relationship between the population and the state of the environment [4].

When $n = 1$ ($n = 0$), the game is defined by the payoff matrix A_1 (A_0). The matrices A_1 and A_0 are set so that the non-cooperative game, which takes place between agents at random encounters in the population, has different Nash equilibrium positions in replenished and depleted environments.

An article [5] examines changes in population state using replicator dynamics, which takes into account feedback from the environment and depends on public awareness of the state of the environment. In contrast to this study, the current work assumes that each agent has its own personal opinion about the state of the environment, but does not have reliable information about it.

Consider the case where an agent can hold one of two opinions, m_1 or m_2 , regardless of the strategy it chooses. We define the distribution of opinions in the population as a vector $y_N(t) = (y_1(t), y_2(t))$, where each component $y_i(t)$ is the proportion of agents in the population holding opinion m_i . Without loss of generality, we denote $y_1 = y(t)$, $y_2(t) = 1 - y(t)$. The process of opinion distribution in a population can be described by mean dynamics, which allows changes in the population to be described by a proportional imitation rule [8].

Let's introduce a matrix B , whose elements b_{ij} represent the degree of confidence of the agent with opinion m_i in the agent pursuing strategy j . Based on a pairwise imitation protocol [3], an imitation protocol (2) has been designed to describe changes of opinion popularity in the population. The imitation protocol is:

$$p_{ij} = \left[y_j \sum_{l=1}^2 x_l u(e^l, x_N, A_n) b_{jl} - y_i \sum_{l=1}^2 x_l u(e^l, x_N, A_n) b_{il} \right]_0^1. \quad (2)$$

Thus, we get the dynamics, which describes the popularity of the opinions in the population based on the mean dynamics (3) and pairwise imitation protocol:

$$\dot{y} = (1 - y)p_{21} - yp_{12}. \quad (3)$$

It is assumed that the state of the population depends on the popularity of opinions. Under this assumption, the first player's payoff matrix can be rewritten in the form $A_y = yA_1 + (1 - y)A_0$, which is obtained from the matrix A_n if the parameter n representing the state of the environment is replaced by the fraction of players y with opinion m_1 .

Thus, an evolutionary game with environment-opinion feedback can be represented as

$$\begin{cases} \dot{x} = x(1-x)(u(e^1, x_N, A_y) - u(e^2, x_N, A_y)), \\ \dot{n} = n(1-n)(\theta x + \psi(1-x)), \\ \dot{y} = (1-y)p_{21} - yp_{12}. \end{cases} \quad (4)$$

3. Example. The Prisoner's Dilemma. In the current experiment, the Prisoner's Dilemma in its economic interpretation is chosen as the base game describing the interaction of agents. In this game, the first strategy corresponds to the player's choice to cooperate and the second to the choice to defend.

Since the change in the state of the population as a function of the agents' opinions is considered in the model according to payoff matrix A_n , it is necessary to introduce two payoff matrices

$$A_0 = \begin{pmatrix} 3.5 & 0.75 \\ 2 & 0.75 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 4 & 1.25 \\ 4.5 & 1.25 \end{pmatrix}, \quad (5)$$

for which the following relations are true:

$$a_{11}^0 > a_{21}^0, \quad a_{12}^0 > a_{22}^0, \quad a_{11}^1 < a_{21}^1, \quad a_{12}^1 < a_{22}^1. \quad (6)$$

The Prisoner's Dilemma game is characterized by the existence of a single Nash equilibrium state, and given the relations (6), for the game given by the matrix A_1 (A_0), this state is (e^2, e^2) ((e^1, e^1)).

In the current numerical experiment, the parameters of the system (4) take values:

Table 1. System parameters.

Param.	Value	Description
θ	2	The rate of replenishment of resources by cooperates
ψ	-1	The rate at which defenders consume resources
x_0	0.5	Proportion of cooperators in the population at the initial time moment
y_0	0.6	Proportion of agents with an opinion m_1 at initial time moment
n_0	0.3	Environmental condition at initial time moment
B	$\begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}$	Agents with opinion m_1 (m_2) are assumed to trust only co-operators (defenders)

As it can be seen from the graph in Figure 1a), at the initial time moment, the proportion of cooperating agents decreases to zero as the environment begins to enrich.

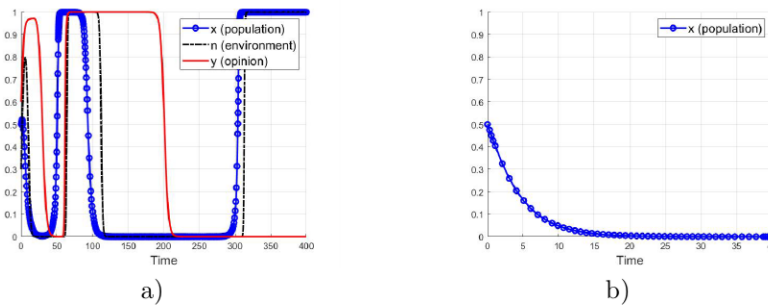


Figure 1. Population a) with, b) without the influence of the environment and the opinions of agents

However, as the proportion of defending agents in the population increases, the environmental resources decrease to zero. After several oscillations, the system reaches an equilibrium state (e^1, e^1) , all players have the opinion m_2 , the environment is replenished, i.e. $n = 1$.

Through a series of numerical experiments, it was found that the environment and the opinions of the agents have a significant impact on the stationary position of the population. In most cases, a change in the initial value of the population parameter, the environment, or the popularity of the opinions causes a change in the stationary position that the system reaches. The choice of the confidence matrix also plays an important role in the results of the simulation of changes in the population depending on the environment and the opinions of the agents.

References

1. Petrosyan L. A., Zenkevich N. A., Shevkoplyas E. V. Game Theory. St. Petersburg: BHV-Petersburg, 2012. 432 p. (in Russian)

2. Kurnosykh Z. A., Gubar E. A. Modeling of the evolutionary game taking into account the network structure // Control processes and stability. 2017. Vol. 4. No 1. P. 631–635. (in Russian)
3. Riehl J. R., Cao M. Control of stochastic evolutionary games on networks // IFAC. 2015. Vol. 48. Iss. 22. P. 76–81.
4. Weitz J. S., Eksin C., Paarporn K. et al. An oscillating tragedy of the commons in replicator dynamics with game-environment feedback // PNAS. 2016. Vol. 113. No 47. P. E7518–E7525.
5. Paarporn K., Eksin C. et al. Optimal control policies for evolutionary dynamics with environmental feedback // IEEE Conference on Decision and Control (CDC). 2018. P. 1905–1910.
6. Zhu Q., Gubar E., Altman E. Preface to special issue on dynamic games for modeling and control of epidemics // Dynamic Games and Applications. 2022. Vol. 12. Iss. 1. P. 1–6.
7. Weibull J. W. Evolutionary Game Theory. Cambridge: MIT Press, 1995. 265 p.
8. Sandholm W.H. Population Games and Evolutionary Dynamics. Cambridge: MIT Press, 2010. 616 p.
9. Bayer P., Gatenby R. et al. Coordination games in cancer // PLoS ONE. 2022. Vol. 17. Iss. 1. Art. no e0261578.

UDC 338.314.053.4

Wang D., Guo Z.

Single-factor multi-stage production control model based on linear optimal control

Recommended for publication by professor Smirnov N. V.

1. Introduction. This study proposes a new linear optimal control method for the single-factor multi-stage production issue in the textile industry. It aims to maximize revenue while considering resource constraints. The approach simplifies the complex control problem into a linear programming problem based on works by [1, 2]. Case studies demonstrate the effectiveness of adjusting key factors during production to optimize production and maximize profits within the budget. The article combines market dynamics and ensures both mathematical and commercial feasibility of production strategies. It provides practical guidance for production optimization in the textile industry and other sectors, supported by actual data.

2. Theoretical foundation. In controlled production, altering strategy affects revenue. With x_i as the i -th product's output and p_i its price, we aim for greater revenue via multi-stage control, focusing on one controllable factor, unlike single-stage control in [3]. The production rate is defined as:

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x} + \mathbf{B}(t)u + \mathbf{F}(t), \quad (1)$$

where, \dot{x}_i represents the production rate of product i . Control u is referred to as the influencing factor, which could be the working time, the number of equipment, the hiring or firing of employees, and so on. $F(n)$ represents the basic production rate of the n -th state variable. $\mathbf{x} \in R^n$, $u \in R^1$. The dimensions of A , B , and F are respectively $(n \times n)$, $(n \times 1)$ and $(n \times 1)$.

We want to maximize terminal revenue within the interval $[t_*, t^*]$.

$$\mathbf{J}(u) = \mathbf{p}^T \mathbf{x}(t^*), \quad (2)$$

Wang Dan – PhD student, Saint Petersburg State University; e-mail: st102920@student.spbu.ru, phone: +7(981)806-33-95

Guo Zhao – PhD student, Saint Petersburg state University; e-mail: 13709267195@163.com, phone: +7(981)683-74-98

where, $p = p_1, \dots, p_n$.

In multi-stage discrete control, we assume the control variable u is constant during each production stage, such as additives or labor, changing only at each new batch. $u(t)$ is treated as discrete constants over time.

$$T_u = \{t_*, t_* + h, \dots, t_* + kh, \dots, t_* + (N - 1)h\}.$$

We refer to these as switching points, T_u - the set of switching points.

$$\begin{cases} u(t) = u_k, & t \in [t_* + kh, t_* + (k + 1)h], \\ u_k = \text{constant}, & k = 0, \dots, N - 1, \quad h = \frac{t^* - t_*}{N} \end{cases} \quad (3)$$

where N is a natural number. h is the step size of the quantization period. It represents the time interval used to produce each batch of goods. According to equation (3), we can rewrite the Cauchy form of $x(t^*)$ as

$$\sum_{k=0}^{N-1} \int_{t_* + kh}^{t_* + (k+1)h} \mathbf{p}^T \mathbf{Y}(t^*) \mathbf{Y}^{-1}(\tau) (\mathbf{B}(\tau)u(\tau) + \mathbf{F}(\tau)) d\tau \rightarrow \max_u. \quad (4)$$

Let:

$$\mathbf{z}_p^T(t) = \mathbf{p}^T \mathbf{Y}(t^*) \mathbf{Y}^{-1}(t),$$

$$c_h(t_* + kh) = \int_{t_* + kh}^{t_* + (k+1)h} \mathbf{z}_p^T(\tau) \mathbf{B}(\tau) d\tau.$$

Then, $c_h(t)$ can be combined as

$$C = (c_h(t_*), \dots, c_h(t_* + kh), \dots, c_h(t_* + (N - 1)h))^T.$$

In summary, we can rewrite functional (2) as

$$C^T U \rightarrow \max. \quad (5)$$

For final state, we have boundary condition [1]

$$\mathbf{H}\mathbf{x}(t^*) = \mathbf{g}, \quad (6)$$

where $g \in R^m$, rank $H = m < n$, and H is $(m \times n)$ matrix. Here row h_i of matrix H represents the cost of producing the i th product. This means that equation (6) represents that the total cost of all products

produced by the factory is equal to the total budget cost g . For this equation,

$$\int_{t_*}^{t^*} \mathbf{HY}(t^*) \mathbf{Y}^{-1}(\tau) (\mathbf{B}(\tau)\mathbf{u}(\tau) + \mathbf{F}(\tau)) d\tau = \mathbf{g} - \mathbf{HY}(t^*) \mathbf{Y}^{-1}(t_*) \mathbf{x}_0. \quad (7)$$

Let:

$$\begin{aligned} \mathbf{HY}(t^*) \mathbf{Y}^{-1}(t) &= \mathbf{G}(t), \\ \mathbf{d}_h(t) &= \int_{t_*+kh}^{t_*+(k+1)h} \mathbf{G}(\tau)\mathbf{B}(\tau)d\tau. \end{aligned} \quad (8)$$

Then, $d_h(t)$ will have $N - 1$ solutions, which are:

$$\mathbf{D} = (\mathbf{d}_h(t_*), \dots, \mathbf{d}_h(t_* + kh), \dots, \mathbf{d}_h(t_* + (N - 1)h)).$$

In conclusion, we can rewrite the equation as:

$$\mathbf{D}\mathbf{U} = \mathbf{g}_0.$$

We can only control the production rate to a limited extent, that is,

$$|u_{kh}| \leq c, \quad k = 0, \dots, N - 1,$$

where, c is a constant. It represents the maximum control we can take.

3. Experimental results. The production situation in the textile industry was statistically analyzed by [4]. The enterprise's production line manufactures two types of products: cotton cloth and jute cloth. It is assumed that each worker operates for 8 hours per day, producing 120 meters of cotton cloth and 157 meters of jute cloth per hour, respectively. Due to limited factory inventory, the production rate decreases by 0.021 meters per hour for every meter of cotton cloth produced and by 0.017 meters per hour for every meter of jute cloth produced. With an initial inventory of zero for both types of cloth, the daily production is set to 960 meters of cotton cloth and 1256 meters of jute cloth. The production rates for both types of cloth can be controlled by adjusting the number of working hours for employees each day. These can be expressed,

$$\begin{aligned} \dot{x}_1 &= -0.021x_1 + 120u + 960, \\ \dot{x}_2 &= -0.017x_2 + 157u + 1256, \end{aligned} \quad (9)$$

where the selling price per meter of cotton cloth is $p_1 = 15$ yuan, and the cost price is $h_1 = 8$ yuan. The selling price per meter of jute cloth is $p_2 = 7$ yuan, and the cost price is $h_2 = 4.5$ yuan. Our production budget cost is $g = 400000$ yuan.

We can increase or decrease the working hours of each employee by up to two hours per day, i.e. $|u| \leq 2$. We can maximize the total production revenue within a month by adjusting the daily production targets, i.e. we divide U into $N = 30$ segments. $h = \frac{t^* - t_*}{N} = \frac{30 - 0}{30} = 30$.

In summary, our original linear optimal control problem can be written as:

$$\begin{aligned} J(u) &= 15x_1(t^*) + 7x_2(t^*) \rightarrow \max_u, \\ \dot{x} &= \begin{bmatrix} -0.0021 & 0 \\ 0 & -0.0017 \end{bmatrix} x + \begin{bmatrix} 120 \\ 157 \end{bmatrix} u + \begin{bmatrix} 960 \\ 1256 \end{bmatrix}, \\ x(t_*) &= \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \begin{bmatrix} 8 \\ 4.5 \end{bmatrix} x(t^*) = 400000, \quad |u| \leq 2, \\ \mathbf{x} \in \mathbb{R}^n, \quad u \in \mathbb{R}^1, \quad \text{rank } \mathbf{H} &= m < n, \quad t \in [0, 30]. \end{aligned}$$

This problem was reduced to the following one:

$$\begin{aligned} \mathbf{C}^T \mathbf{U} &\rightarrow \max, \\ \mathbf{D} \mathbf{U} &= \mathbf{g}_0, \\ |u_{kh}| &\leq 2, \quad k = 0, \dots, 29, \end{aligned} \tag{10}$$

where x_1 represents cotton cloth, and x_2 represents jute cloth. We have calculated the daily increase or decrease in working hours u_k through Python, where k is the number of days.

As shown in Figure 1 (a), the black part represents the daily working hours that should be increased. The gray part represents the daily working hours that should be decreased. We found that there are 7 days where the working hours are reduced by 2 hours each day, 22 days where overtime is 2 hours each day, and one day where overtime is 0.45 hours. This can maximize our final profit within a month. In addition, we have drawn a three-dimensional graph of the inventory of two types of fabrics versus the number of days in Figure 1 (b). Among them, the inventory of cotton cloth is on the x axis, the inventory of jute cloth is on the y axis, and the number of days is on the z axis.

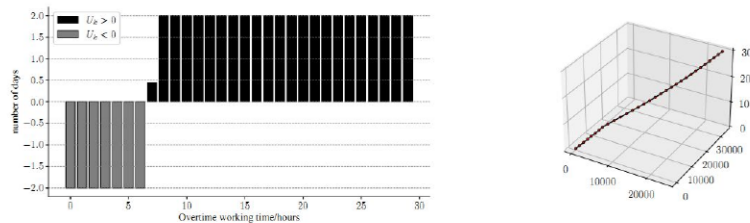


Figure 1. (a) Number of overtime hours, (b) Maximum daily earnings

Finally, at $t^* = 30$ days, our maximum total profit $J(u) = 877684.2997$ yuan.

4. Conclusion. This study presents a single-factor, multi-stage production control model suitable for the textile industry, which transforms production factors such as production rate, cost, and selling price into a linear programming problem. Through the adjustment of working hours, namely the control variable u , profit can be maximized within the budget. Experimental results show that appropriate adjustment of working hours can effectively manage production rates and maximize profits. This model provides a new tool for optimizing the production process, and future research will consider more variables and external factors to enhance the applicability of the model.

References

1. Gabasov R., Kirillova F. M., Balashevich N. V. On the synthesis problem for optimal control systems // *SIAM Journal on Control and Optimization*. 2000. Vol. 39. No 4. P. 1008–1042.
2. Gabasov R., Kirillova F. M., Balashevich N. V. Synthesis of adaptive optimal controls for linear dynamic systems // *Computational Optimal Control*. 1994. P. 83–88.
3. Hishamuddin H., Sarker R. A., Essam. D. A disruption recovery model for a single stage production-inventory system // *European Journal of Operational Research*. 2012. Vol. 222. No 3. P. 464–473.
4. Campo E. A., Cano J. A., Gmez-Montoya R. A. Linear programming for aggregate production planning in a Textile Company // *Fibres and Textiles in Eastern Europe*. 2018. Vol. 26. No 5. P. 13–19.

Contents

Plenary Session	7
Invited speakers	7
<i>Kostromin S. A.</i> Control processes in fundamental science.....	7
<i>Krylatov A. Yu.</i> Optimization models and methods for equilibrium flow assignment in transportation networks	9
1. Mathematical theory of control processes	35
<i>Andriyanova N. R.</i> Fixed-time stability of a switched multiagent system	35
<i>Babanov I. V.</i> Optimal control of a three-mass oscillatory system	42
<i>Gaponov D. N.</i> Stability of linear time delay systems. Constructive algebra approach	48
<i>Kamintseva V. S.</i> Implementation of the optimal control synthesis approach for dynamic objects	56
<i>Makoveeva P. E., Nesterov S. D.</i> Lyapunov – Krasovskii functionals for a linear system with distributed parameters	63
2. Mathematical methods in mechanics and physics	71
<i>Belyaeva A. A., Knyazev N. A.</i> Influence of modifications of computational models of permanent dipole magnets of synchrotron radiation source on the optical characteristics of a charged particle beam	71
<i>Vekovtsev V. V., Jordan A. A., Nikiforov K. A.</i> Development of a digital model of the field emission system	81
<i>Vialtseva V. V., Pronina Yu. G.</i> On estimating 3D density of parallel cracks from their traces on the surface.....	88
<i>Grosheva A. D., Znamenshchikova N. S.</i> Assessment of the possibility of developing three-dimensional computational models for precision magnetic systems comprising solenoids and ferromagnetic screens using integral and differential approaches.....	93
<i>Kazakova A.</i> Minimum variance for exponential distribution	99

<i>Kikenov A. A.</i> Mathematical model of tiltrotor	107
<i>Kondratyeva A. D.</i> Self-weight buckling of a nanocantilever on an elastic substrate	112
<i>Nabiev T. K., Shymanchuk D. V.</i> Construction of the trajectories of a group of UAVs in formation	121
<i>Belenok A. I.</i> Ritz method for building parametric representation of cost optimal trajectory	129
<i>Mullayanov T. O.</i> Investigating elastic response of nanostructured solid surface using boundary perturbation method	134
<i>Xing J.</i> MS-stability comparison of Runge–Kutta–Chebyshev methods application to SDEs	139
3. Mathematical models of medical and biological systems	145
<i>Davydenko B. N.</i> Predicting the result of hypermetropia correction using CHAID analysis	145
<i>Legkova I. A.</i> Detection of abnormalities and heart rhythm disturbances in patient ECG data	150
<i>Khaltobina M. M., Elizarova M. V.</i> Analysis of cell survival using the RMR model	155
<i>Khomenko I. E.</i> Converting quantitative data into binary data using adaptive two-way evaluation to build a patient survival model	163
<i>Valieva N. L., Goncharova A. B.</i> Breast cancer classification in mammograms	168
4. Information and computer technologies	173
<i>Antropova E. G.</i> Solving optimization problems using neural networks	173
<i>Babkina A. V.</i> Medical ultrasound images speckle noise removing ..	179
<i>Gorbunov V. I., Salimov T. A.</i> Problems of practical application of data science results and approaches to their solution	187
<i>Guseva M. Yu.</i> Calculation of an analytical solution for a three-dimensional system of ordinary differential equations with hysteresis and perturbation	195

<i>Kovalev A. Yu., Kalinin P. E.</i> An alternative algorithm for generating Perlin noise	200
<i>Kolesov V. V.</i> Algorithms for movie genres classification	205
<i>Kuklo R. A.</i> The application of a clustering method based on time and geographic data to reduce the amount of transmitted data	212
<i>Mosina M. A.</i> Investigation of workload distribution methods using the ant colony optimization	218
<i>Nabiev T. R.</i> Application of computer vision methods for detecting and tracking drones in video	226
<i>Neronov R. M.</i> Solving the problem of exploratory text analysis through clustering and approximation of the clusters number ...	231
<i>Nikitin A. V.</i> Application of neural network methods for segmentation of diseases of cultivated plants by color images in precision agriculture	236
<i>Pavlov A. V.</i> Methods of fine-tuning: RAG vs LoRA	241
<i>Pavlov M. V.</i> Gameplay testing methods in game development based on neural network reinforcement learning	247
<i>Ponomarev Yu. Yu.</i> Issues of accurate estimates of tree characteristics in metadata modeling in general-purpose information systems	253
<i>Poteshkin E. P.</i> Parameters selection in Monte Carlo SSA	262
<i>Rudovich K. V.</i> Library for creation and implementation genetic algorithms in C++	270
<i>Khromov N. A., Golyandina N. E.</i> High-order MSSA for signal extraction	278
<i>Chaikin G. A.</i> Tools development to track experiments for knowledge tracing in continuous learning	287
<i>Huang L.</i> Deep residual neural network based on feature extraction for ultrasound image classification of focal liver lesions	293
5. Control problems in social and economical systems ...	299
<i>Anikushin V. A., Redko A. Yu.</i> Application of the ARIMA model to forecast the world population	299

<i>Bushueva A. S., Vlasova T. V.</i> The application of heuristic algorithms for solving the problem of routing with multiple warehouses.....	307
<i>Kalinin P. E.</i> A multiple test model in a binary outcome experiment.....	312
<i>Karpychev A. V.</i> Investigation of the potential for modeling consumer demand for electricity in the electrical grid.....	317
<i>Korol M. M.</i> Equilibration algorithms for traffic flow assignment in an urban road network.....	323
<i>Novgorodtcev V. A.</i> On the portfolio management problem in the case of the Heston model.....	330
<i>Rusakova K. V.</i> Comparison of strategies with short memory in the game “Prisoner’s dilemma”.....	339
<i>Bykova M. S.</i> Payment schemes and their properties for cooperative stochastic games.....	345
<i>Lorits E. M.</i> Evolutionary model with environmental feedback and agent opinions.....	355
<i>Wang D., Guo Z.</i> Single-factor multi-stage production control model based on linear optimal control.....	361
Contents	366
Author index.....	370

Именной указатель

Андриянова Н. Р.	35	Калинин П. Е.	200, 312
Аникушин В. А.	299	Каминцева В. С.	56
Антропова Е. Г.	173	Карпычев А. В.	317
Бабанов И. В.	42	Кикенов А. А.	107
Бабкина А. В.	179	Князев Н. А.	71
Беляева А. А.	71	Ковалев А. Ю.	200
Бушуева А. С.	307	Колесов В. В.	205
Вековцев В. В.	81	Кондратьева А. Д.	112
Власова Т. В.	307	Король М. М.	323
Вяльцева В. В.	88	Костромин С. А.	7
Гапонов Д. Н.	48	Крылатов А. Ю.	9, 11
Голяндина Н. Э.	278	Кукло Р. А.	212
Горбунов В. И.	187	Легкова И. А.	150
Грошева А. Д.	93	Мосина М. А.	218
Гусева М. Ю.	195	Набиев Т. К.	121
Давыденко Б. Н.	145	Набиев Т. Р.	226
Елизарова М. В.	155	Неронов Р. М.	231
Знаменщикова Н. С.	93	Никитин А. В.	236
Иордан А. А.	81	Никифоров К. А.	81
Казакова А.	99	Новгородцев В. А.	330

Павлов А. В.	241	Шиманчук Д. В.	121
Павлов М. В.	247	Belenok A. I.	129
Пономарев Ю. Ю.	253	Bykova M. S.	345
Потешкин Е. П.	262	Goncharova A. B.	168
Пронина Ю. Г.	88	Guo Z.	361
Редько А. Ю.	299	Huang L.	293
Рудович К. В.	270	Lorits E. M.	355
Русакова К. В.	339	Makoveeva P. E.	63
Салимов Т. А.	187	Mullayanov T. O.	134
Халтобина М. М.	155	Nesterov S. D.	63
Хоменко И. Е.	163	Valieva N. L.	168
Хромов Н. А.	278	Wang D.	361
Чайкин Г. А.	287	Xing J.	139

Ежегодное научное издание

ПРОЦЕССЫ УПРАВЛЕНИЯ И УСТОЙЧИВОСТЬ

Том 11(27)

Труды 55-й международной научной конференции
аспирантов и студентов

Редакторы текста на русском языке: *А. С. Ерёмин, Т. Е. Смирнова*

Редакторы текста на английском языке: *А. С. Ерёмин, М. А. Смирнова*

Компьютерная верстка: *М. Н. Смирнов, М. А. Смирнова*

Подписано в печать 24.06.2024
Усл. печ. л. 21,59. Тираж 150 экз.

198504, Россия, Санкт-Петербург, Петергоф, Университетский просп., 35