

Федеральный исследовательский центр
«Карельский научный центр Российской академии наук»

ТРУДЫ

КАРЕЛЬСКОГО НАУЧНОГО ЦЕНТРА РОССИЙСКОЙ АКАДЕМИИ НАУК

№ 6, 2026

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ
И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Петрозаводск
2026

Научный журнал
**Труды Карельского научного центра
Российской академии наук**
№ 6, 2026
МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ
И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

ISSN 1997-3217 (печатная версия)
ISSN 2312-4504 (онлайн-версия)

Главный редактор
А. Ф. ТИТОВ, член-корр. РАН, д. б. н., проф.

Редакционный совет

О. Н. БАХМЕТ (зам. главного редактора), член-корр. РАН, д. б. н.; Н. А. БЕЛКИНА, д. г. н.; И. Н. БОЛОТОВ, член-корр. РАН, д. б. н.; А. В. ВОРОНИН, д. т. н., проф.; Х. ГАО, доктор математики, проф. (КНР); Н. В. ИЛЬМАСТ, д. б. н.; Ю. В. ЗАЙКА, д. ф.-м. н.; С. В. КРИВОВИЧЕВ, академик РАН, д. г.-м. н.; А. М. КРЫШЕНЬ, д. б. н.; О. Л. КУЗНЕЦОВ, д. б. н.; Ю. Н. ЛУКИНА, д. б. н.; В. В. МАЗАЛОВ, д. ф.-м. н., проф.; Н. Н. НЕМОВА, академик РАН, д. б. н., проф.; А. Н. РЕТТИЕВА, д. ф.-м. н.; С. А. СВЕТОВ, д. г.-м. н.; А. И. СЛАБУНОВ, д. г.-м. н.; Н. Ю. СУХОВИЛО, к. г. н. (Республика Беларусь); С. ЛИ, доктор геологии (КНР); Н. Н. ФИЛАТОВ, член-корр. РАН, д. г. н., проф.; Н. С. ЧЕРНЕЦОВ, член-корр. РАН, д. б. н.; В. В. ЩИПЦОВ, д. г.-м. н., проф.

Редакционная коллегия серии
«Математическое моделирование и информационные технологии»

В. А. ВАТУТИН, д. ф.-м. н., проф.; Ю. В. ЗАЙКА, д. ф.-м. н., проф.; А. Н. КИРИЛЛОВ, д. ф.-м. н., доцент;
О. В. ЛУКАШЕНКО (ответственный секретарь), к. ф.-м. н.; В. В. МАЗАЛОВ (ответственный редактор),
д. ф.-м. н., проф.; Ю. Л. ПАВЛОВ (зам. ответственного редактора), д. ф.-м. н., проф.; Л. А. ПЕТРОСЯН,
д. ф.-м. н., проф.; А. В. СОКОЛОВ, д. ф.-м. н., проф.

Издается с января 2009 г.

Адрес редакции: 185910, Петрозаводск, ул. Пушкинская, 11
Тел. (8142)762018; факс (8142)769600
E-mail: trudy@krc.karelia.ru
Электронная полнотекстовая версия: <http://transactions.krc.karelia.ru>; <http://journals.krc.karelia.ru>

© ФИЦ «Карельский научный центр РАН», 2026
© Институт прикладных математических
исследований КарНЦ РАН, 2026

Karelian Research Centre of the Russian Academy of Sciences

TRANSACTIONS

**of the KARELIAN RESEARCH CENTRE
of the RUSSIAN ACADEMY of SCIENCES**

No. 6, 2026

**MATHEMATICAL MODELING
AND INFORMATION TECHNOLOGIES**

Petrozavodsk
2026

Editor-in-Chief

A. F. TITOV, RAS Corr. Fellow, DSc (Biol.), Prof.

Editorial Committee

O. N. BAKHMET (Deputy Editor-in-Chief), RAS Corr. Fellow, DSc (Biol.); N. A. BELKINA, DSc (Geog.); I. N. BOLOTOV, RAS Corr. Fellow, DSc (Biol.); N. S. CHERNETSOV, RAS Corr. Fellow, DSc (Biol.); N. N. FILATOV, RAS Corr. Fellow, DSc (Geog.), Prof.; H. GAO, PhD (Phys.-Math.), Prof. (China); N. V. ILMAST, DSc (Biol.); Yu. V. ZAIKA, DSc (Phys.-Math.); S. V. KRIVOVICHEV, RAS Academician, DSc (Geol.-Miner.); A. M. KRYSHEN, DSc (Biol.); O. L. KUZNETSOV, DSc (Biol.); Yu. N. LUKINA, DSc (Biol.); V. V. MAZALOV, DSc (Phys.-Math.), Prof.; N. N. NEMOVA, RAS Academician, DSc (Biol.), Prof.; A. N. RETTIEVA, DSc (Phys.-Math.); V. V. SHCHIPTSOV, DSc (Geol.-Miner.), Prof.; S. A. SVETOV, DSc (Geol.-Miner.); A. I. SLABUNOV, DSc (Geol.-Miner.); N. Yu. SUKHOVILO, PhD (Geog.) (Republic of Belarus); X. LI, PhD (Geol.-Miner.) (China); A. V. VORONIN, DSc (Tech.), Prof.

Editorial Board of the

«Mathematical Modeling and Information Techn» Series

A. N. KIRILLOV, DSc (Phys.-Math.), Assistant Prof.; O. V. LUKASHENKO (Executive Secretary), PhD (Phys.-Math.); V. V. MAZALOV (Editor-in-Charge), DSc (Phys.-Math.), Prof.; Yu. L. PAVLOV (Deputy Editor-in-Charge), DSc (Phys.-Math.), Prof.; L. A. PETROSIAN, DSc (Phys.-Math.), Prof.; A. V. SOKOLOV, DSc (Phys.-Math.), Prof.; V. A. VATUTIN, DSc (Phys.-Math.), Prof.; Yu. V. ZAIKA, DSc (Phys.-Math.), Prof.

Published since January 2009

Monthly

Editorial Office address: 11 Pushkinskaya St., 185910 Petrozavodsk, Karelia, Russia
Tel. (8142)762018; fax (8142)769600
E-mail: trudy@krc.karelia.ru
Full-text electronic version: <http://transactions.krc.karelia.ru>; <http://journals.krc.karelia.ru>

© Karelian Research Centre of the Russian Academy of Sciences, 2026
© Institute of Applied Mathematical Research, Karelian Research Centre,
Russian Academy of Sciences, 2026

УДК 519.179.2

ГРАФ СОЦИАЛЬНЫХ СВЯЗЕЙ РУССКИХ ПИСАТЕЛЕЙ: СЕТЕВОЙ АНАЛИЗ ЛИТЕРАТУРНОГО СООБЩЕСТВА НА ОСНОВЕ КОДИФИКАТОРА ЕГЭ

А. В. Бородина

*Институт прикладных математических исследований КарНЦ РАН,
ФИЦ «Карельский научный центр РАН» (ул. Пушкинская, 11, Петрозаводск,
Республика Карелия, Россия, 185910)*

В статье предлагается подход к количественному анализу социальной структуры русских литературных сообществ на основе построения графа личных взаимодействий писателей. Актуальность исследования обусловлена дефицитом структурированных датасетов в данной предметной области и возрастающей ролью методов цифровых гуманитарных наук (Digital Humanities) в изучении социодинамики литературы. Эмпирической базой для построения математической модели сообщества послужил перечень авторов из официального «Кодификатора ЕГЭ по литературе» (2026 г.), что исключает преднамеренную связность авторов. С помощью нейросетевых моделей (DeepSeek, YandexGPT) и последующей верификации экспертом были собраны данные о социальных связях (личное общение, переписка, членство в союзах) для 102 русских писателей 1757–2003 гг. Проведенный анализ сетевой структуры социальных взаимосвязей между писателями выявил типичные для социальных сетей свойства графа. Для ранжирования писателей предложена комбинированная метрика, сочетающая локальную и глобальные центральности, показана ее корреляция с алгоритмом PageRank. Предложенный подход к построению графа социальных взаимодействий дает результаты, согласованные с экспертными источниками, и может быть использован для социодинамического анализа литературных сообществ. Методы анализа социальных сетей хорошо применимы для изучения литературных сообществ и позволяют предложить альтернативный подход к оценке статуса писателя через структуру личных связей.

Ключевые слова: социальные сети; случайные графы; нейросеть; сообщество писателей; коэффициент кластеризации; модулярность; ранжирование

Для цитирования: Бородина А. В. Граф социальных связей русских писателей: сетевой анализ литературного сообщества на основе кодификатора ЕГЭ // Труды Карельского научного центра РАН. 2026. № 6. С. 5–19. doi: 10.17076/mat2368

Финансирование. Финансовое обеспечение исследований осуществлялось из средств федерального бюджета на выполнение государственного задания КарНЦ РАН (Институт прикладных математических исследований КарНЦ РАН).

A. V. Borodina. A GRAPH OF SOCIAL CONNECTIONS OF RUSSIAN WRITERS: A NETWORK ANALYSIS OF THE LITERARY COMMUNITY BASED ON THE UNIFIED STATE EXAM CODIFIER

Institute of Applied Mathematical Research, Karelian Research Centre, Russian Academy of Sciences (11 Pushkinskaya St., 185910 Petrozavodsk, Karelia, Russia)

This paper proposes an approach to the quantitative analysis of the social structure of Russian literary communities by constructing a graph of personal connections among writers. The relevance of the study arises from the lack of structured datasets in this subject domain and the growing significance of digital humanities methodologies for investigating the sociodynamics of literature. The empirical basis for constructing a mathematical community model was the list of authors from the official “Unified State Exam Codifier in Literature” (2026). This source ensures representativeness while avoiding artificial connectedness among authors. Social connections data, including personal communication, correspondence, and membership in professional unions, were collected for 102 Russian writers active between 1757 and 2003. The data collection process employed neural network models (DeepSeek, YandexGPT) followed by expert verification. The results showed that the proposed method for constructing the social interactions graph yields results consistent with expert data and can be used for sociodynamic analysis of literary communities. Social network analysis methods are well suited for studying writer communities and offer an alternative approach to assessing a writer’s status through the structure of personal connections.

Key words: social networks; random graph; neural network; writers community; clustering coefficient; modularity; ranking

For citation: Borodina A. V. A graph of social connections of Russian writers: a network analysis of the literary community based on the Unified State Exam (USE) codifier. *Trudy Karelskogo nauchnogo tsentra RAN = Transactions of the Karelian Research Centre RAS*. 2026. No. 6. P. 5–19. doi: 10.17076/mat2368

Funding. The studies were funded from the federal budget through state assignment to the Karelian Research Centre RAS (Institute of Applied Mathematical Research, Karelian Research Centre RAS).

ВВЕДЕНИЕ

Литературоведы и историки при анализе литературных сообществ оперируют скорее качественными, а не количественными оценками. Исследование свойств сообщества, выделение в нем значимых групп (кластеров) требует выявления социальных связей между каждой парой авторов. Подобные задачи решаются достаточно успешно, если речь идет, например, о структуре социальной сети либо о научном сообществе, где связи – это ссылки цитирования или соавторство в статьях.

Как показали результаты поиска, для сообщества писателей достаточно сложно найти готовые датасеты. Задача сбора и описания социальных связей и динамики в писательских сообществах является актуальной на сегодняшний день. В качестве эмпирической базы для изучения русских писателей специалисты-социологи в основном используют фундаментальные издания: «Лексикон русской литера-

туры XX века» В. Казака [1], «Русские писатели. 1800–1917» [6], «Русские писатели XIX века» [7], «Русские писатели XX века» [8]. Основная группа русских авторов советского периода может быть определена по энциклопедическому справочнику «Мир русской культуры» (1997) [4], а профессиональное сообщество последних десятилетий советского периода – по «Справочнику Союза писателей СССР» [11].

На актуальность проблематики количественного описания русских литературных сообществ указывает возросший интерес к исследованиям в сфере цифровых гуманитарных наук (Digital Humanities) и поддержка проектов в этой области ведущими российскими научными фондами.

Большой социодинамический анализ сообщества русских писателей дореволюционного периода и литературной среды в советское время на основе авторской методики С. Я. Сущего можно найти в работах [12, 13].

Схожие задачи ставят специалисты Института русской литературы РАН (Пушкинский дом) г. Санкт-Петербурга и работают с тем же массивом данных (словарь «Русские писатели. 1800–1917» [6]) в рамках текущего научного проекта «Социальная история русской литературы XIX века. Разработка методологии описания и создание электронного ресурса» (период выполнения 2024–2026 гг.), финансируемого Российским научным фондом [2]. Основная задача проекта – создание электронной базы данных на основе биографического словаря с целью обобщить социальные характеристики авторов (сословие, образование, жанровый выбор, места публикаций и т. д.) и выявить значимые корреляции.

Кроме того, важной открытой задачей является разработка методики оценивания писательского труда и рейтинга писателя в литературном поле. Подробнее вопросы анализа статуса писателей и их борьбы за позиции в современном российском литературном поле можно найти в работах [9, 10]. Критерии оценки, по которым оценивается положение и успех писателей, выявляются эмпирически, на основе тематического анализа интервью с ведущими российскими литературными критиками и публицистами.

Интерес к социологическому портрету писателя и анализу литературного поля активно развивается и требует внедрения новых методов для исследования. Построение структуры социального взаимодействия между писателями в виде графа позволит исследовать свойства данного сообщества и применить известные алгоритмы ранжирования для выявления статуса или веса писателя, основываясь на его личных связях в профессиональном сообществе.

Поскольку задачу описания социально-исторических связей между писателями и построение математической модели всего сообщества за весь временной период развития русской литературы решить не представляется возможным, то далее речь пойдет о выделении некоторого подмножества писателей, для которых было бы возможно установить социальные связи и построить модель такого фрагмента общей (неизвестной) сети в виде неориентированного графа, в котором узлами будут персоналии (авторы, писатели), а ребро между узлами проводится в случае, когда два автора взаимодействовали лично: связаны персональным общением, перепиской, входили в один литературный союз и т. д.

В работе в качестве объекта для исследования и построения математической модели

предлагается выбрать множество авторов, регламентированное в официальном документе «Кодификатор Единого государственного экзамена по литературе» за 2026 г. [3]. Кодификатор – это официальный перечень тем, произведений и навыков, которые выносятся на итоговый экзамен по литературе в старшей школе. Список авторов в кодификаторе ЕГЭ по литературе не является фиксированным и регулярно обновляется, писатели включаются в перечень без учета социальных связей или принадлежности к какому-то одному литературному объединению, таким образом можно исключить связность «по построению». Перечень кодификатора содержит произведения из разных временных периодов. В кодификатор попадают произведения, которые изучаются в средней и старшей школе. Тем не менее следует учитывать, что список составляется из ключевых авторов, чье творчество является традиционной и неотъемлемой основой школьного курса литературы.

В работе проведено исследование структуры и свойств графа социального взаимодействия писательского сообщества на базе списка авторов, представленных в кодификаторе [3]. Предложен подход к проведению социодинамического анализа такого рода сообществ на основе выявления свойств графа. Для двух временных периодов – «XIX – начало XX в.» и «советский период» – проведен сравнительный анализ распределения писателей по жанровой специализации.

МЕТОДИКА СБОРА ДАННЫХ О СОЦИАЛЬНЫХ СВЯЗЯХ

Специалисты в области социологии и литературы отмечают, что на данный момент литературные персоналии изучены очень выборочно. Кроме того, недостаточно исследованы взаимосвязи между персоналиями и их группами, например между писателями и издателями, писателями и читателями, критикой и читателями, цензурой и издателями [12]. Результаты поиска в сети Интернет также обнаруживают отсутствие в открытом доступе готовых датасетов с такой информацией в русскоязычном сегменте.

Достаточно подробное исследование в области социодинамического анализа русского писательского сообщества можно найти в статьях С. Я. Сущего [12, 13], где отдельно рассмотрены два периода: «XIX – начало XX в.» (1781–1917 гг.) и «советский период» (1886–1985 гг.). В качестве основы для анализа использовались биографические словари русских писателей XIX – начала XX века, кото-

рые содержат сведения примерно о 2,5 тыс. персоналиях, а советский период представлен 760 персоналиями. Эти результаты будем считать эталонными для проведения сравнения с данными, вычисленными на подсети кодификатора, с этой целью из полного списка авторов кодификатора были выбраны только русские писатели (102 автора за 1757–2003 гг.), а границы периодов соответствуют установленным в [12, 13].

С помощью нейросети DeepSeek была построена матрица смежности, представляющая собой таблицу, где 1 указывает на наличие социальной связи: переписка, работа в наставничестве, личное знакомство и т. д.; 0 – если связи не выявлено. Далее для уже имеющихся данных проведен повторный поиск связей с помощью другой нейросети – YandexGPT. На основе поиска был составлен список с подробным описанием каждой связи и проведена верификация полученных данных экспертом в области литературы И. Н. Тихоновой¹. Благодаря такой верификации найдено еще около 20 связей, которые до этого не были учтены.

Для сравнительного анализа по жанровой принадлежности все писатели были упорядочены по периоду дебюта. Перечень жанров взят в соответствии с классификацией С. Я. Сущего и пронумерован: 1 – проза; 2 – поэзия; 3 – драма; 4 – перевод; 5 – публицистика; 6 – критика; 7 – литературоведение; 8 – сценарии; 9 – журналистика; 10 – мемуары.

Таблица 1. Пример расхождения мнений YandexGPT и DeepSeek в жанре проза по писателю Е. А. Евтушенко

Table 1. An example of the divergence of opinions between YandexGPT and DeepSeek in the prose for the writer Ye. A. Yevtushenko

DeepSeek	1	«Ягодные места», автобиографическая проза
YandexGPT	0	Поэт и публицист; проза не главная форма творчества

Для поиска жанровой принадлежности снова использовались две нейросетевые языковые модели YandexGPT и DeepSeek. Формат вывода результата включал таблицу, состоящую из трех полей: 1 – ФИО писателя; 2 – ответ о принадлежности писателя к определенному жанру (1 – да, 0 – нет); 3 – обоснование. Данные сохранялись в формате csv, и далее с помощью программы на языке python выявлялись несовпадения по вто-

рому полю. Результаты сравнения показали, что процент несовпадения ответов двух нейросетей в таблицах составляет 1,96–3,92 % в зависимости от жанра. В таблице 1 приведен пример несовпадения результатов поиска. В таком случае решение принималось на основе повторного поиска с указанием произведения, а также на основе мнения эксперта.

СВОЙСТВА ГРАФА СОЦИАЛЬНОГО ВЗАИМОДЕЙСТВИЯ ПИСАТЕЛЕЙ

Обозначим $G = (V, N)$ неориентированный граф социального взаимодействия писателей русского сегмента кодификатора ЕГЭ по литературе [3]. Вершины графа G – это писатели, их множество фиксировано $|V| = n$. Множество ребер графа $|E| = m$ отражает взаимодействия в сообществе писателей. Ребро между двумя вершинами проводится, если имеется личная связь между парой авторов, а именно: личное знакомство, переписка, соавторство, наставничество, принадлежность к одному литературному сообществу и т. д.

Многие сложные социальные сети реального мира обладают рядом общих базовых свойств. Проведем анализ типичных свойств для построенной структуры. Вычисления реализованы на языке python с использованием библиотеки NetworkX.

Разреженная структура графа

Известно, что граф является разреженным [5], если при наличии n вершин количество ребер $k \cdot n$ для константы $k > 1$ значительно меньше максимального числа возможных ребер в полном графе с тем же числом вершин. Данное свойство присуще многим социальным сетям и впервые было описано как ключевая особенность веб-графов в исследованиях Barabasi и Albert [15].

Для графа G число вершин $n = 102$, число ребер $m = 288$, что существенно меньше, чем число ребер в полном графе, которое равно $C_n^2 = n(n-1)/2 = 5151$. Отношение числа ребер графа G к максимально возможному числу ребер называют плотностью графа

$$D = \frac{2m}{n(n-1)} = 0,0198,$$

у разреженного графа плотность близка к 0.

Кроме того, средняя степень узлов

$$\bar{d} = \frac{1}{n} \sum_{i=1}^n d_i = 5,65,$$

где d_i – число ребер, инцидентных вершине с номером i , на порядки меньше числа узлов

¹И. Н. Тихонова – учитель русского языка и литературы МОУ «Лицей № 40», г. Петрозаводск.

графа G , что также указывает на его разреженность.

Свойство связности и закон «шести рукопожатий»

Граф G состоит из четырех компонент связности, наибольшая по размеру (гигантская) компонента связности $G_{giant} = (V_{giant}, E_{giant})$ состоит из 97 вершин. Три остальных компоненты содержат 1, 1 и 3 узла соответственно.

Закон «шести рукопожатий» («шести кликов», «мир тесен») – это свойство, присущее веб-графам и некоторым социальным графам, состоит в том, что среднее расстояние между любыми двумя узлами сети составляет около 6 звеньев. В классическом исследовании Jeffrey Travers и Stanley Milgram [25] средняя длина цепочки составляет около 5–6. Данные исследования по Facebook [17] дают среднее расстояние 4,57 для всех пользователей и 3,46 для жителей США.

Диаметр в подграфе G_{giant} равен 9, среднее расстояние 3,4, что говорит о тесноте связей и выполнении принципа «мир тесен».

Степенной закон распределения степеней вершин

Для структуры социальных сетей характерно, что существует небольшое количество узлов с большими степенями (влиятельные лица в социальной сети), а большинство узлов имеют очень малые степени. Распределение степеней вершин в таких сетях моделируется с помощью степенного закона распределения:

$$P(\deg(v) = d) \sim c \cdot d^{-\alpha},$$

где $\deg(\cdot)$ обозначает степень вершины $v \in V$, d – значение степени, а константа c определяется из условия нормировки (сумма вероятностей равна 1), $\alpha > 1$ [5, 15].

На рис. 1 показаны результаты вычисления оценки параметров степенного закона распределения степеней вершин графа G с помощью построения степенной регрессии. Уравнение регрессии статистически значимо ($p\text{-value} = 2,68e-07$ для теста Фишера), коэффициент детерминации $R^2 = 0,798$.

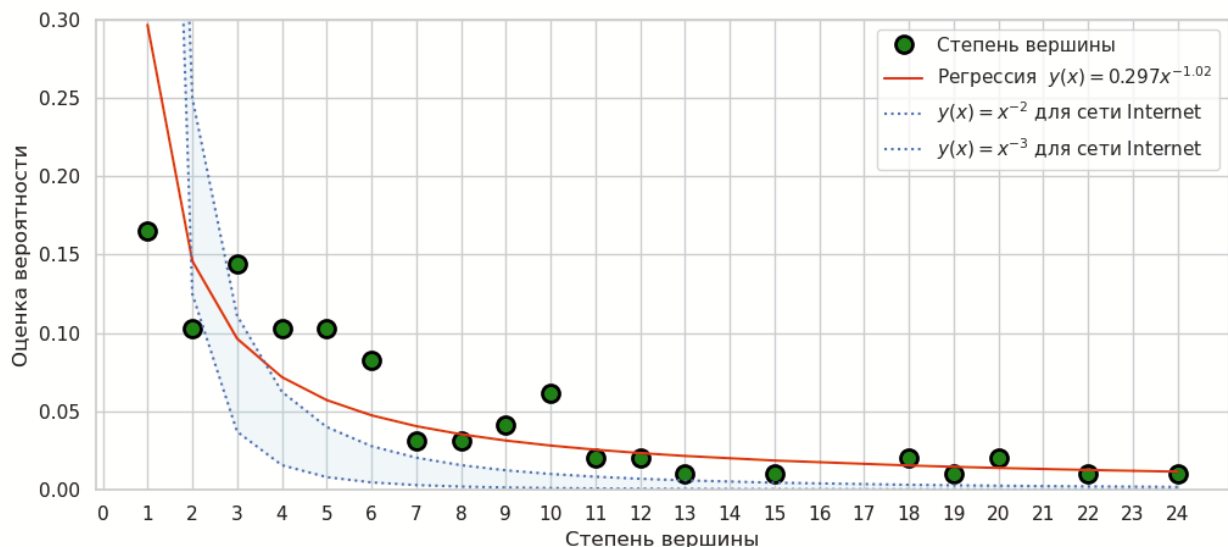


Рис. 1. Параметры степенной регрессионной модели для графа G

Fig. 1. The power-law regression model parameters for the graph G

В общем случае показатель степени α в распределении степеней вершин для графов социальных сетей чаще всего находится в интервале между 2 и 3. Это наиболее распространенное значение, которое возникает из-за механизма «предпочтительного присоединения» (когда новые участники с большей вероятностью знакомятся с уже популярными людьми). Именно для таких сетей характерно наличие «хабов» – узлов с большим количеством связей. Исследования реальных социальных

сетей показывают, что многие из них могут иметь значительно более низкие показатели масштабирования $\alpha \sim 1$ (например, [14, 23]).

Для рассматриваемой структуры сообщества писателей параметр $\alpha = 1,02$ не попал в классический диапазон [2; 3], это может говорить о сильном влиянии принципа взаимности, когда человек общается с теми, кто общается с ним, что характерно для социальных объединений, где связь – это личное общение.

Коэффициент кластеризации и выделение сообществ

Насколько хорошо граф поддается разбиению на отдельные группы-сообщества, можно оценить на основе модулярности. Для неориентированного графа G с матрицей смежности $A_{n \times n} = (A_{ij})_{i,j=1}^n$ и разбиения $c = \{c_1, \dots, c_k\}$ на k кластеров модулярность разбиения определяется формулой Ньюмана – Гирвана [21]:

$$Q(c) = \frac{1}{2m} \sum_{ij} (A_{ij} - \frac{d_i d_j}{2m}) I\{c_i = c_j\}, \quad (1)$$

где d_i – степень узла i , функция-индикатор $I\{c_i = c_j\}$ равна 1, если узлы i и j принадлежат одному сообществу, и 0, если иначе.

Если разбиение на сообщества известно, то модулярность целесообразно вычислять в виде:

$$Q(c) = \sum_{i=1}^k \left[\frac{L_i}{m} - \left(\frac{D_i}{2m} \right)^2 \right],$$

где k – общее количество сообществ, L_i – количество ребер внутри сообщества c_i , D_i – сумма степеней всех узлов, входящих в сообщество c_i .

1. Кластеризация методом Лувена.

Для разбиения на сообщества будем использовать алгоритм Лувена (Louvain algorithm), который имеет практическую эффективность при работе с разреженными графами и максимизирует модулярность $Q(c)$ [14, 16].

Алгоритм выявляет группы тесно связанных узлов в сети, где связи внутри группы плотнее, чем связи с остальным «миром». Его часто используют в социальных сетях (для поиска групп друзей), в биоинформатике (для анализа взаимодействия белков) и в финансовой сфере (для обнаружения мошеннических транзакций). Метод Лувена удобен тем, что не требует предварительного знания количества кластеров, но на поздних итерациях могут возникать плохо связанные кластеры. Кроме того, для него характерно повышенное число кластеров. В частности, на модельных данных, где разбиение на сообщества известно, алгоритм Лувена может разбивать истинные сообщества на более мелкие сообщества с большей модулярностью (см., например, [14]).

Стандартный алгоритм Лувена максимизирует модулярность, определенную формулой (1). В этом случае алгоритм может не обнаружить сообщества, размер которых меньше некоторого порогового значения (масштаб). Это обстоятельство может быть критично для больших сетей, где алгоритм будет объединять маленькие, но плотные (с точ-

ки зрения модулярности) сообщества в более крупные кластеры.

В работе для прироста модулярности в алгоритме Лувена вместо (1) использована обобщенная формула Райхардта – Борнхольдта [22] в виде:

$$Q(c) = \frac{1}{2m} \sum_{ij} (A_{ij} - r \frac{d_i d_j}{2m}) I\{c_i = c_j\}, \quad (2)$$

где параметр r (разрешение) управляет масштабом обнаруживаемых сообществ.

При $r = 1$ формула (2) является частным случаем классической модулярности. Выбор $r < 1$ способствует поиску крупных структур, на практике, как правило, выбирают $r \in [0, 5; 0, 8]$. Для обнаружения более мелких сообществ нужно выбирать $r > 1$, чем выше число, тем больше будет итоговых сообществ. Как правило, достаточно протестировать алгоритм для значений из диапазона $r \in [1, 2; 2, 0]$.

Для выбора оптимального значения параметра разрешения r для графа G используется эмпирический подход по методу «локтя» (Elbow method). На рис. 2 приведены результаты вычисления для $r \in [0, 5; 2, 1]$ с шагом 0,1. Оптимальное значение модулярности $Q = 0,6055$ достигается при $r \in [0, 9; 1, 2]$, число кластеров $k = 8$, что говорит о высокой модулярности и четко выраженной структуре сообществ.

Для верификации полученных результатов проведена кластеризация методом Лейдена.

2. Кластеризация методом Лейдена.

Поскольку алгоритм Лувена имеет недостатки и иногда может создавать несвязные сообщества, проведем кластеризацию графа G с использованием алгоритма Лейдена (Leiden algorithm), который как раз гарантирует, что все обнаруженные сообщества будут хорошо связаны внутри.

Метод был предложен в работе [24], он считается улучшенной версией метода Лувена и использует для определения модулярности обобщенную формулу (2). На рис. 3 приведены результаты вычислений для графа G . Оптимум модулярности $Q = 0,6256$ достигается при $r \in [0, 8; 1, 6]$, число кластеров $k = 8$.

Результаты экспериментов показывают, что для исследуемого графа сообщества писателей можно использовать для разбиения на кластеры разрешение $r = 1$ с максимальным значением модулярности. При этом число кластеров $k = 8$ и состав самих кластеров совпадает для обоих методов. На рис. 4 показана визуализация разбиения графа G на сообщества $c_i, i \in [1; 8]$.

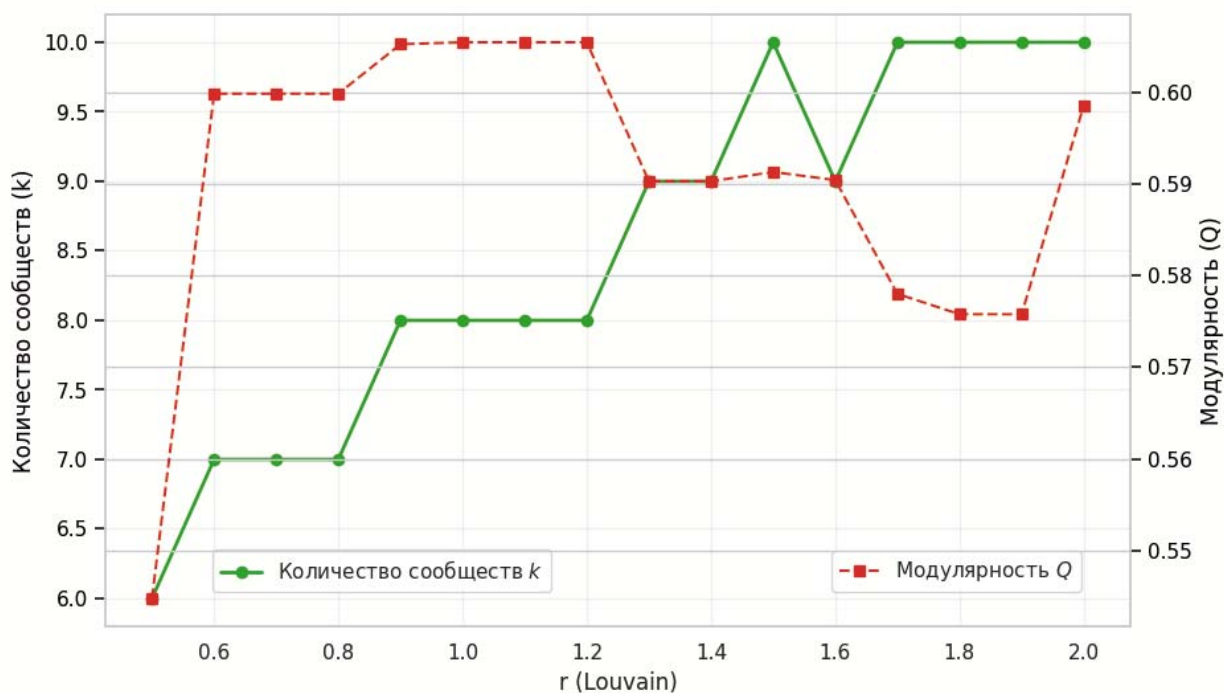


Рис. 2. Параметры разбиения графа G на кластеры методом Лувена
 Fig. 2. Parameters for clustering the graph G with the use of Louvain's method

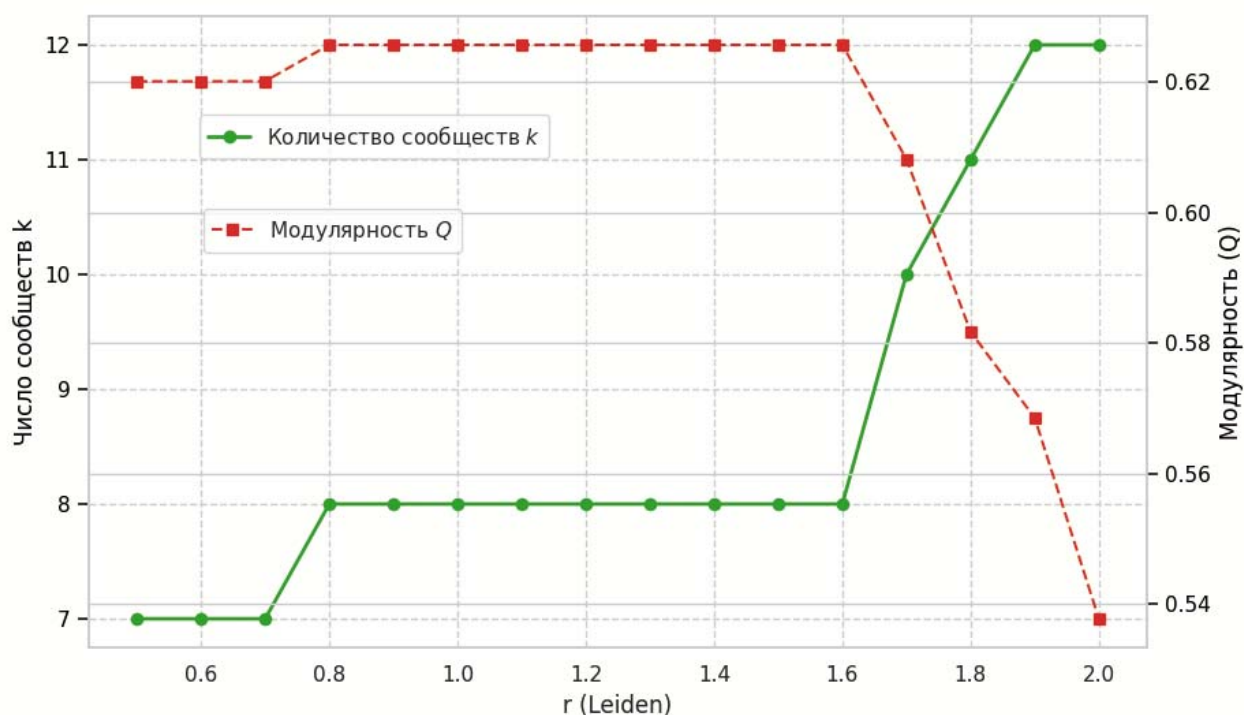


Рис. 3. Параметры разбиения графа G на кластеры методом Лейдена
 Fig. 3. Parameters for clustering the graph G with the use of Leiden's method

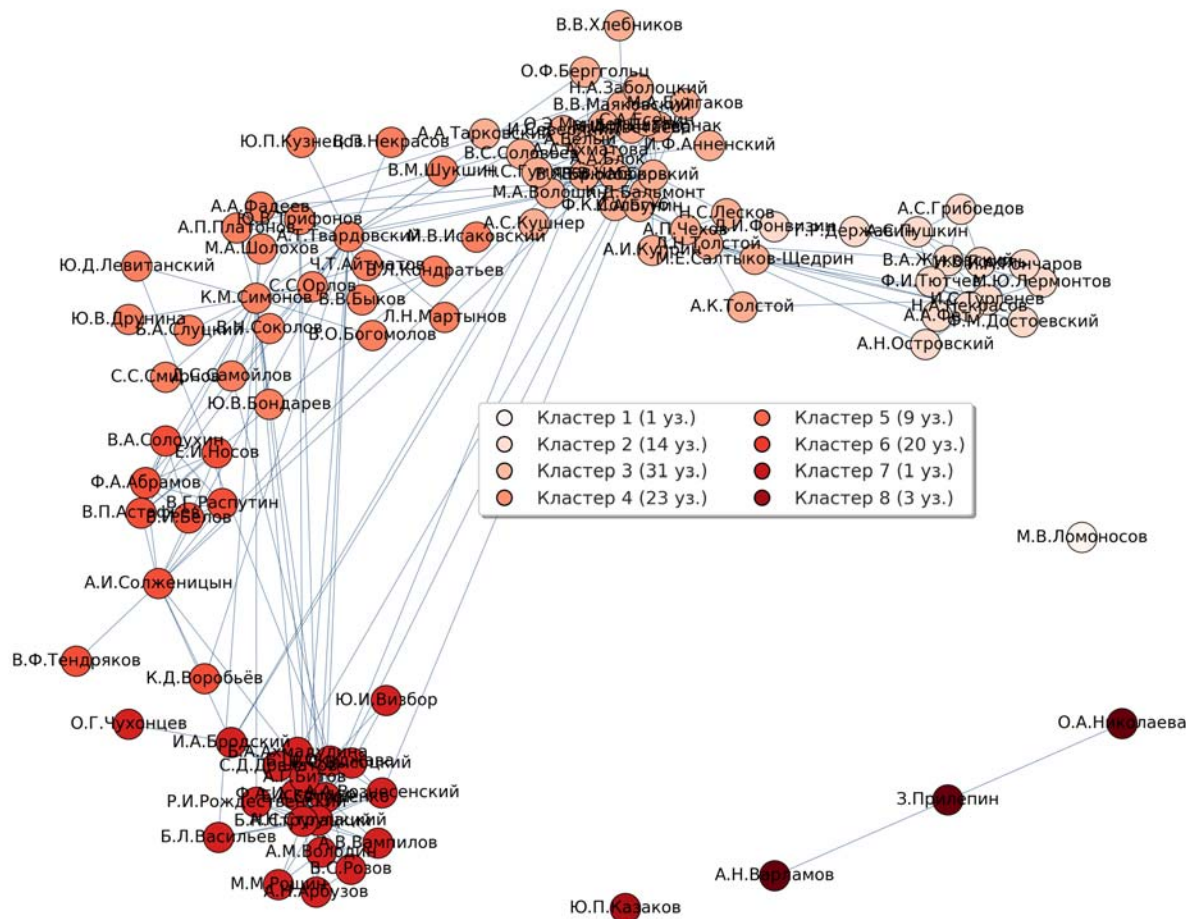


Рис. 4. Разбиение на кластеры $c_i, i \in [1; 8]$ графа G по методу Лувена
 Fig. 4. Clustering the graph G into $c_i, i \in [1; 8]$ with the use of Louvain's method

Ранжирование и выделение центральных узлов

Вопрос определения социального статуса писателей и репутации писательского труда является дискуссионным. В основном для определения рейтинга писателя учитывают мнение эксперта – литературного критика (см., например, в [10]). Анализ тесноты социального взаимодействия между писателями на основе личного общения может быть дополнительной альтернативной методикой для определения веса писателя, поскольку на профессиональную значимость личные связи оказывают значительное влияние. Ниже предлагается сравнить способы определения значимости узла на основе двух метрик: PageRank и комбинированной центральности по степеням.

1. Метрика PageRank. Алгоритм PageRank был разработан и успешно применяется с целью ранжирования веб-страниц в сетях большого объема с учетом важности (значимости), где важность страницы определяется важностью страниц, которые на нее ссылаются [14].

На текущий момент идея, лежащая в основе PageRank, применима во многих типах сообществ. Например, в академической среде импакт-фактор журнала зависит от того, сколько важных статей на него ссылаются. В социальных сетях авторитетный человек – это тот, на кого ссылаются (упоминают) другие авторитетные люди.

Данный подход может оказаться также полезным для определения социального статуса и репутации писателей, когда для рейтинга писателя анализируется не только содержание его произведений, а еще и структура социальных связей всей сети сообщества, куда входит данный писатель.

В работе был реализован метод вычисления метрики PageRank для построенного графа сообщества писателей G с учетом разбиения на 8 кластеров, в каждом кластере определены центральные узлы.

В табл. 2 приведены результаты вычисления метрики PR для центральных узлов кластеров. На рис. 5 показана структура сети с выделением узлов по размеру на основе метрики PageRank.

Таблица 2. Центральные узлы кластеров $c_i, i \in [1; 8]$ по метрике PageRank
 Table 2. Central nodes of clusters $c_i, i \in [1; 8]$ according to the PageRank metric

№	Центральные узлы Central nodes	PR
1	['М.В.Ломоносов']	1,000
2	['В.А.Жуковский']	0,129
3	['А.А.Ахматова']	0,092
3	['М.Горький']	0,091
4	['К.М.Симонов']	0,215
5	['А.И.Солженицын']	0,159
6	['А.Н.Стругацкий']	0,093
6	['Б.Н.Стругацкий']	0,093
7	['Ю.П.Казаков']	1,000
8	['З.Прилепин']	0,486

Метрика PageRank может быть использована для определения рейтинга писателя в социальной сети сообщества писателей, если под рейтингом понимать авторитет в профессиональном сообществе, измеренный через структуру связей. Метрика PageRank является глобальной, т. к. ее значение для конкретного узла зависит от структуры всего графа, а не только от его ближайшего окружения. Сравним результат ранжирования на основе PageRank с другим подходом, который учитывает как глобальные, так и локальные метрики.

2. Комбинированная метрика. В работе [20] предложена мера центральности Convex Mixed Centrality (COMC) на основе выпуклой комбинации локальных и глобальных характеристик узла сети. Аналогичный комбинированный подход с применением

Показаны узлы с PageRank > 0.059
(топ 30%)

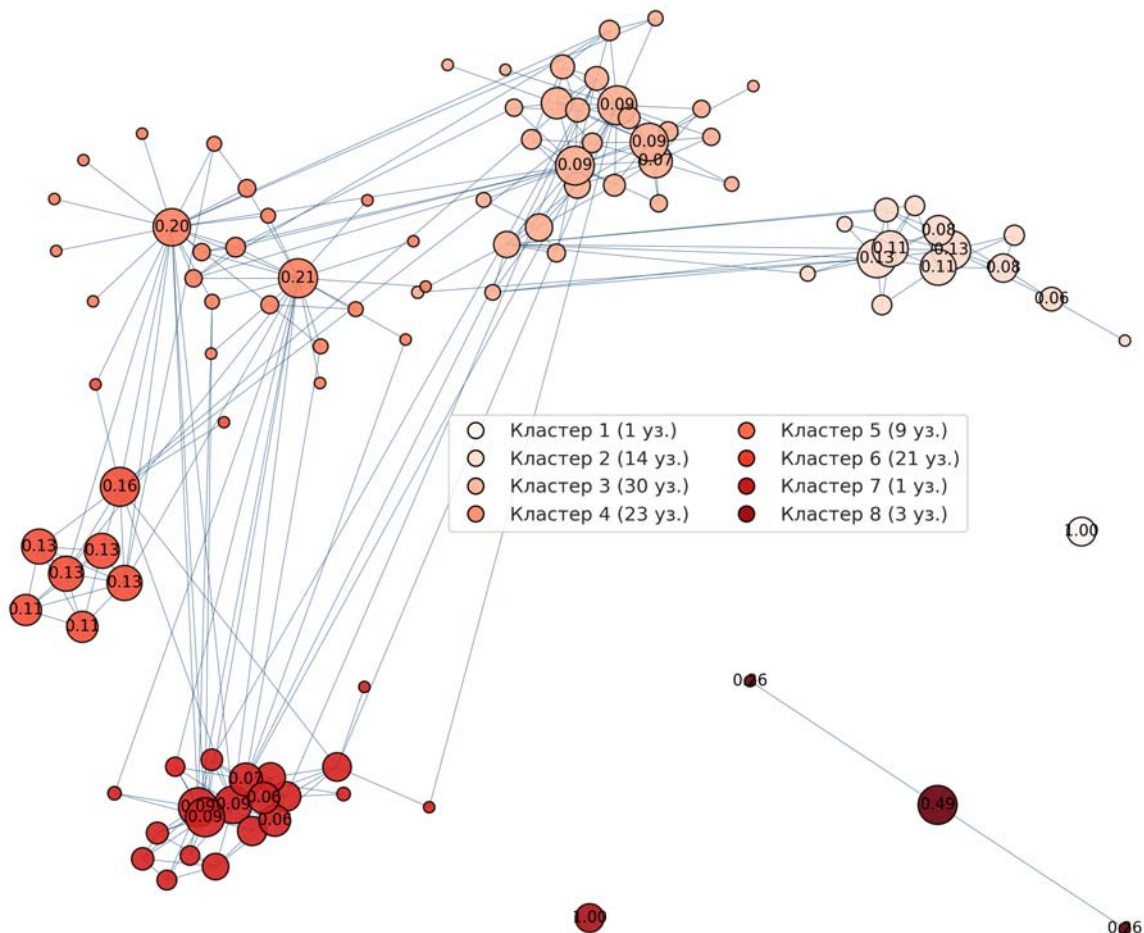


Рис. 5. Ранжирование узлов по метрике PageRank в кластерах $c_i, i \in [1; 8]$
 Fig. 5. Ranking in accordance with the PageRank metric in $c_i, i \in [1; 8]$

смешанных мер, вычисленных на локальной и глобальной структуре сети, описан и в других работах [18, 19]. Обобщенно метрику СОМС для узла i в неориентированном графе G можно представить в виде:

$$COMC(i) = \alpha \cdot C_{local}(i) + (1 - \alpha) \cdot C_{global}(i),$$

где $C_{local}(i)$ – коэффициент, отражающий локальную роль узла (например, его степень или локальную кластеризацию); C_{global} – коэффициент, отражающий глобальную роль узла (например, близость к другим узлам или посредничество); $\alpha \in [0; 1]$.

Метрика СОМС разработана для решения фундаментальной проблемы: традиционные меры центральности учитывают либо только локальные, либо только глобальные характеристики узлов, что приводит к снижению эффективности при работе с крупномасштабными сетями. Подход, основанный на комбинации СОМС, позволяет учесть узлы, находящиеся в промежуточных зонах сети, которые не являются ни явными центрами, ни изолированной периферией. Далее под термином СОМС будем понимать не одну фиксированную формулу, а семейство метрик, построенных по принципу выпуклой комбинации «локальности» и «глобальности». Такой подход к вычислению центральности узлов обеспечивает простой и мощный инструмент балансировки между микро- и макросвойствами сети, что особенно ценно при поиске структурно значимых объектов в социологических исследованиях, когда узел сети – это персоналия.

Учитывая специфику рассматриваемой сети сообщества писателей, комбинированный подход при выборе метрики, определяющей важность узла, достаточно логичен, поскольку для определения социального статуса писателя необходимо учитывать и его связи внутри социальной группы (литературные объединения, кружки, союзы писателей), и способность поддерживать социальные связи с ключевыми значимыми фигурами (редакторы, литературные агенты, издательства), способность влиять на общественное мнение (участие в фестивалях, лекциях, сотрудничество с медиа) и т. д.

В текущей работе для решения задачи ранжирования писателей и идентификации центральных узлов в кластерах предлагается использовать комбинированный показатель, основанный на взвешенной сумме метрик центральности (Degree, Betweenness, Closeness) узла i :

$$C_{DBC}(i) = \alpha_1 C_D(i) + \alpha_2 C_B(i) + \alpha_3 C_C(i), \quad (3)$$

где $\sum_{j=1}^3 \alpha_j = 1$. В локальную компоненту формулы (3) входит коэффициент $C_D(i)$, равный числу прямых связей (ребер) узла i с другими узлами сети (метрика Degree Centrality), который показывает центральность узла по степени и отражает его «локальную известность». Локальная компонента показывает, насколько вершина активна в своем ближайшем окружении. Глобальную компоненту формулы (3) формируют два коэффициента: $C_B(i)$ вычисляется как число кратчайших путей через узел i (метрика Betweenness Centrality); $C_C(i)$ вычисляется на основе среднего расстояния до всех узлов сети (метрика Closeness Centrality), которые в комбинации учитывают положение узла в общей топологии сети, его роль «моста» и скорость доступа к другим узлам. Глобальная компонента показывает, насколько вершина важна для всего графа в целом.

Выбор весов $\alpha_1 = 0,5$, $\alpha_2 = 0,3$, $\alpha_3 = 0,2$ обусловлен «природой» метрик, а именно: локальная компонента (Degree) играет доминирующую роль при анализе внутрикластерной структуры, тогда как глобальные компоненты (Betweenness и Closeness) дополняют оценку, учитывая межузловые взаимодействия за пределами непосредственного окружения узла, при этом обе компоненты вносят равный вклад в общую метрику C_{DBC} . В зависимости от целей ранжирования коэффициенты можно варьировать. В табл. 3 показаны центральные узлы по комбинированной метрике.

Таблица 3. Центральные узлы кластеров $c_i, i \in [1; 8]$ по комбинированной метрике
Table 3. Central nodes of clusters $c_i, i \in [1; 8]$ according to the convex metric

№	Центральные узлы Central nodes	C_{DBC}
1	[М.В.Ломоносов']	0,500
2	[В.А.Жуковский']	0,593
3	[М.Горький']	0,511
3	[А.А.Ахматова']	0,484
3	[А.А.Блок']	0,443
4	[К.М.Симонов']	0,671
4	[А.Т.Твардовский']	0,621
5	[А.И.Солженицын']	0,674
6	[А.Н.Стругацкий']	0,569
6	[Б.Н.Стругацкий']	0,569
7	[Ю.П.Казаков']	0,000
8	[З.Прилепин']	1,000

Результаты вычислений показали, что комбинированная метрика C_{DBC} для графа G дает сходный с PageRank рейтинг и может так-

же успешно быть использована для определения рейтинга писателей, тогда как распределение значимости узла по отдельным метрикам C_D , C_B , C_C может значительно отличаться от PageRank. В качестве примера такой ситуации можно рассмотреть результаты ранжирования в кластере c_6 (21 узел) из таблицы 4, где метрики PR и C_{COMC} дают схожий порядок между собой и отличный от метрики C_C .

Таблица 4. Ранжирование узлов кластера c_6 , сравнение метрик
Table 4. Ranking of cluster nodes c_6 , comparison of metrics

№ (PR)	№ (C_{DBC})	№ (C_C)
1 (0,093)	1 (0,569)	2 (0,741)
2 (0,093)	2 (0,569)	3 (0,741)
3 (0,087)	3 (0,556)	1 (0,769)
4 (0,069)	4 (0,423)	5 (0,645)
5 (0,063)	5 (0,400)	4 (0,667)
6 (0,063)	6 (0,389)	8 (0,625)
7 (0,057)	7 (0,364)	6 (0,645)
8 (0,057)	8 (0,364)	7 (0,645)
9 (0,055)	9 (0,349)	9 (0,606)
10 (0,053)	10 (0,313)	12 (0,526)
11 (0,047)	11 (0,291)	10 (0,540)
12 (0,035)	14 (0,223)	16 (0,487)
13 (0,034)	13 (0,231)	13 (0,526)
14 (0,033)	12 (0,233)	13 (0,526)
15 (0,029)	17 (0,195)	17 (0,476)
16 (0,028)	16 (0,203)	15 (0,512)
17 (0,028)	15 (0,205)	14 (0,526)
18 (0,018)	19 (0,131)	19 (0,408)
19 (0,017)	18 (0,138)	18 (0,444)
20 (0,014)	20 (0,095)	20 (0,351)
21 (0,014)	21 (0,095)	21 (0,350)

ТВОРЧЕСКО-ЖАНРОВАЯ СПЕЦИАЛИЗАЦИЯ СООБЩЕСТВА

Теперь сравним соотношение авторов по жанровой специализации в графе G с количественным анализом более обширной группы писателей, опубликованным в экспертных работах С. Я. Суцего [12, 13], которые считаем эталонными для выполнения сравнения. Для этого в рассматриваемом графе G выделим временные периоды по году дебюта писателей аналогично тем, которые рассматривает эксперт, а именно: 1 период – 1781–1917 гг., 2 период – 1886–1985 гг. Ряд указанных им закономерностей оказалось возможным обнаружить и на структуре графа G .

Распределение авторов по специализациям за два периода показано на рис. 6 и 7. Мож-

но отметить, что в программу по литературе в средней и старшей школе включены преимущественно авторы, дебютировавшие в прозе и поэзии, при этом процент дебютантов по кодификатору близок к данным, указанным в статьях С. Я. Суцего, а также сходной является динамика в преобладании жанров.

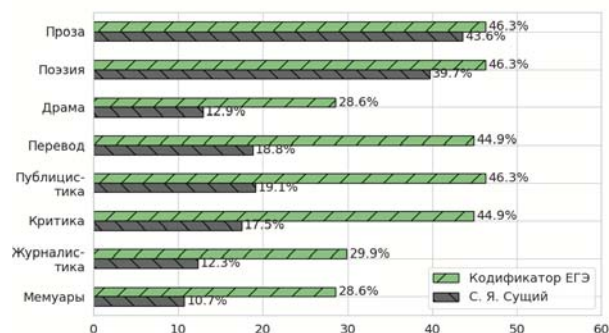


Рис. 6. Жанровая специализация литераторов XIX – начала XX в. (1781–1917 гг.)

Fig. 6. Genre specialization of writers of the 19th and early 20th centuries (1781–1917)

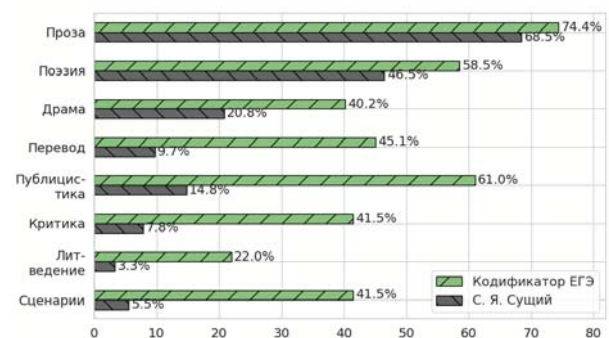


Рис. 7. Жанровая специализация литераторов советского периода (1886–1985 гг.)

Fig. 7. Genre specialization of writers of the Soviet period (1886–1985)

Так, согласно кодификатору, в русском писательском кругу XIX – начала XX в. попеременно доминировали поэзия и проза. В среднем за период и в прозе, и в поэзии дебютировали по 46,3 % (43,6 и 39,7 % – аналогичные показатели в [12] соответственно).

Остальные жанровые специализации также представлены в кодификаторе, но процент в целом существенно выше, чем описывает С. Я. Суций. Это можно объяснить спецификой поиска информации нейросетями DeepSeek, YandexGPT и разным принципом отнесения писателя к тому или иному жанру. В частности, в работах С. Я. Суцего в таблицах подсчитан именно процент дебютантов в том или ином жанре, тогда как нейросеть

относит писателя к определенному жанру тогда, когда у него имеются произведения в этом жанре, несмотря на то, что дебютировать он мог в другом жанре или совмещать специализации.

В частности, для публицистики по кодификатору оценка писателей, работавших в этом жанре, составила 46,3 % за 1 период против 19,1 % дебютантов в таблицах из работы [12]. Однако в тексте работы отмечается, что «... в 1850–1870-е гг. поэзия среди авторских дебютов оказалась даже не второй, но третьей специализацией, уступив вторую позицию публицистике. Почти каждый пятый литературный дебютант в середине века выбирал для профессиональной самореализации публицистику. Но и представители других творческих специализаций в 1850–1880-х гг. обращаются к ней куда чаще, чем их предшественники. Едва ли можно найти крупного русского писателя этого времени, который не отдал бы ей дань». Таким образом, полученный по кодификатору процент вполне обоснован.

В советский период, согласно кодификатору, в 1920-е гг. проза возвращает себе лидирующие позиции среди дебютантов – 66,2 % дебютируют в прозе, при этом 33,7 % совмещают прозу и поэзию (2/3 и 1/3 – аналогичные показатели в [13] соответственно). На всем

протяжении советского периода преобладала проза, процентные показатели по кодификатору близки к указанным диапазонам 50–90 % в прозе и 35–55 % в поэзии. Также данные, полученные по кодификатору, сопоставимы с данными по жанровой специализации всего профессионального сообщества русской литературы, которые представлены в «Справочнике Союза писателей СССР» [11], где основными жанровыми специализациями были проза и поэзия – более 70 %.

Отдельный интерес представляет динамика лидерства жанров проза и поэзия для периода XIX – начала XX в., описанная С. Я. Сущим. Вычисления на множестве писателей из кодификатора также показывают перелом в соотношении ведущих жанров в конце 1820-х – 1830-е гг., когда ведущая роль переходит от поэзии к прозе, при этом в жанровом распределении дебютантов в 1830-е гг. они делятся между прозой и поэзией практически в равной степени. Второй переломный момент приходится на Серебряный век русской литературы (1890–1920-е гг.), когда к 1910–1917 гг. доля дебютантов в прозе сократилась. На рис. 8 можно увидеть смену предпочтений между прозой и поэзией в одни и те же периоды по кодификатору и согласно эталонным экспертным данным из статьи [12].

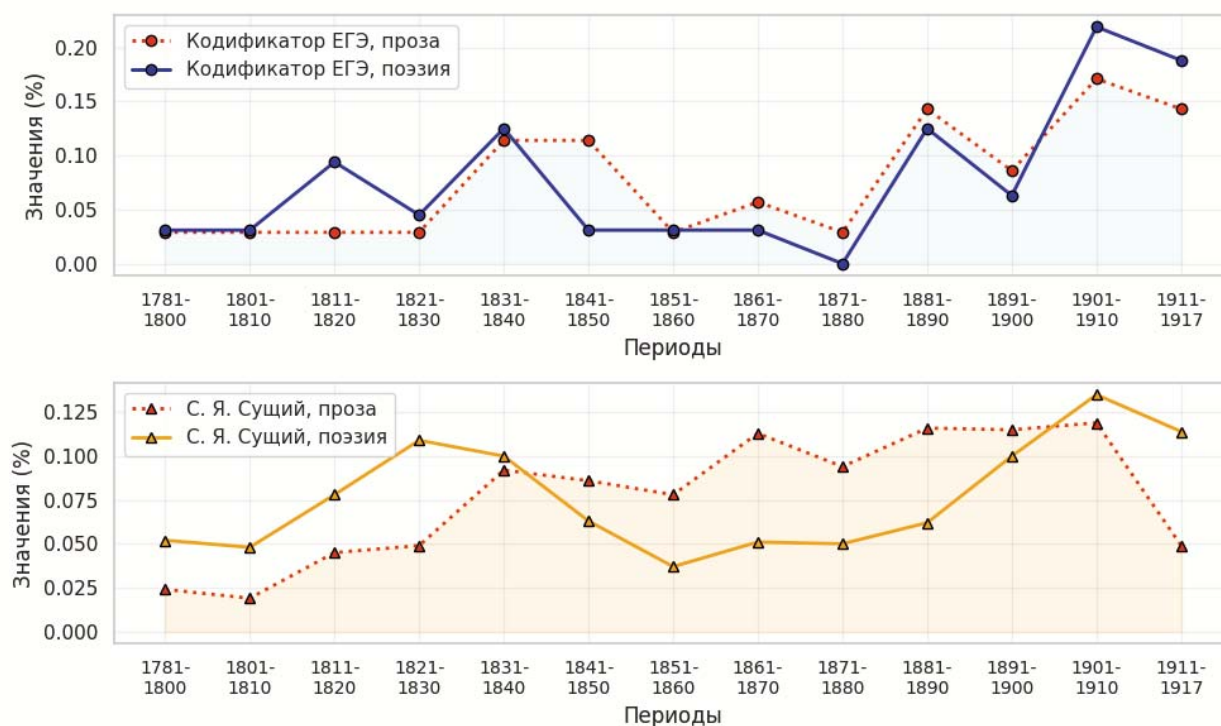


Рис. 8. Динамика доминирования жанров проза и поэзия за период XIX – начала XX в. (1781–1917 гг.)
Fig. 8. Leadership dynamics between prose and poetry from the 19th to the early 20th centuries (1781–1917)

ЗАКЛЮЧЕНИЕ

В работе предложен подход к построению модели профессионального сообщества литераторов на основе учета социальных связей между ними и метрики для ранжирования писателей на основе их социального веса.

Для построения математической модели в виде неориентированного графа в качестве множества вершин выбран набор писателей, чьи произведения учтены в официальном кодификаторе для проведения единого государственного экзамена по литературе. Наличие ребра между двумя вершинами графа определялось алгоритмом двухуровневого поиска с помощью нейросетей DeepSeek, YandexGPT с последующей верификацией описания связей экспертом.

Анализ структуры социального графа литераторов показал сходство свойств построенной сети с другими социальными сетями и веб-графами, а именно: разреженность, степенное распределение степеней вершин, свойство «малого мира», высокую кластеризацию и наличие плотных подграфов (сообществ). Граф состоит из четырех компонент связности, наибольшая из которых содержит 97 вершин из 102.

Для выделенного сообщества были исследованы показатели творческой специализации авторов в двух временных отрезках: XIX – начало XX в. и советский период в русской литературе.

Результаты сравнения с данными, приведенными в статьях С. Я. Сущего, показали идентичный характер соотношений между прозой и поэзией для выделенных временных периодов, а также сходную динамику смены лидирующих позиций между двумя стилями.

Таким образом, предложенный в работе способ построения набора данных для описания социальной сети литераторов показал хорошую согласованность с имеющимися в открытом доступе экспертными данными.

Кроме того, обнаруженные свойства позволяют рассматривать гипотезу, что построенный подграф по свойствам схож с графом, который мог быть получен в результате случайного блуждания.

Поскольку исходный граф всей сети литераторов неизвестен и не может быть построен, то достоверно утверждать, что подграф получен случайным блужданием, невозможно, однако можно отметить совместимость с этой гипотезой, т. к. наблюдаемая структура типична для случайного блуждания. Этот вопрос требует дополнительного изучения.

ЛИТЕРАТУРА

1. Казак В. Лексикон русской литературы XX века. М.: Культура, 1996. 512 с.
2. Карточка проекта фундаментальных и поисковых научных исследований, поддержанного Российским научным фондом // РНФ [Электронный ресурс]. URL: <https://rscf.ru/project/24-18-00761/> (дата обращения: 10.01.2026).
3. Кодификатор проверяемых требований к результатам освоения основной образовательной программы среднего общего образования и элементов содержания для проведения Единого государственного экзамена по литературе [Электронный ресурс]. URL: https://doc.fipi.ru/ege/demoversii-specifikacii-kodifikatory/2026/li_11_2026.zip (дата обращения: 15.02.2026).
4. Мир русской культуры. Энциклопедический справочник / Науч. ред. А. Н. Мячин. М.: Вече, 1997. 624 с.
5. Райгородский А. М. Математические модели интернета // Квант. 2012. №. 4. С. 12–16.
6. Русские писатели. 1800–1917: Биографический словарь: В 5 т./ Гл. ред. П. А. Николаев. М.: Сов. энцикл.; Большая рос. энцикл., 1989–2007.
7. Русские писатели, XIX век: Биобиблиографический словарь: в двух частях / Под ред. П. А. Николаева. М.: Просвещение, 1996.
8. Русские писатели XX века: Биографический словарь / Гл. ред. и сост. П. А. Николаев. М.: Большая рос. энцикл.; Рандеву-АМ, 2000. 806 с.
9. Рязанцев А. П. Российское литературное поле в XXI веке: оценка состояния в интервью с критиками // Социологические исследования. 2025. Т. 51, № 8. С. 132–141. doi: 10.7868/S3034601025080112
10. Рязанцев А. П. Социальный статус и репутация писателей в современной России: тематический анализ // Интеракция. Интервью. Интерпретация. 2023. Т. 15, № 1. С. 27–44. doi: 10.19181/inter.2023.15.1.2
11. Справочник Союза писателей СССР. М.: Советский писатель, 1981. 821 с.
12. Сущий С. Я. Русская литература XIX – начала XX в.: социодинамический анализ писательского сообщества // Социологические исследования. 2020. № 2. С. 128–143. doi: 10.31857/S013216250008500-3
13. Сущий С. Я. Русская литература советского периода – социодинамика писательского сообщества // Социологические исследования. 2021. № 2. С. 112–129. doi: 10.31857/S013216250009902-5
14. Avrachenkov K., Dreveton M. Statistical analysis of networks. Boston-Delft: Now Publ., 2020. 248 p. doi: 10.1561/9781638280514

15. Barabási A. L., Albert R. Emergence of scaling in random networks // *Science*. 1999. Vol. 286, no. 5439. P. 509–512. doi: 10.1126/science.286.5439.509

16. Blondel V. D., Guillaume J.-L., Lambiotte R., Lefebvre E. Fast unfolding of communities in large networks // *J. Stat. Mech.: Theory Exp.* 2008. October. P10008. doi: 10.1088/1742-5468/2008/10/P10008

17. Edunov S., Diuk C., Filiz I. O., Bhagat S., Burke M. Three and a half degrees of separation // Research at Facebook blog. 4 February 2016. URL: <https://research.facebook.com/blog/three-and-a-half-degrees-of-separation/> (дата обращения: 10.01.2026).

18. Hajarathaiah K., Enduri M. K., Anamalamudi S., Sangi A. R. Algorithms for finding influential people with mixed centrality in social networks // *Arab. J. Sci. Eng.* 2023. Vol. 48, no. 8. P. 10417–10428. doi: 10.1007/s13369-023-07619-w

19. Mohammad B. B., Dhuli S., Enduri M. K. Isolating centrality-based generalization of traditional centralities to discover vital nodes in complex networks // *Arab. J. Sci. Eng.* 2025. Vol. 50, no. 15. P. 12003–12025. doi: 10.1007/s13369-024-09628-9

20. Mohammad B. B., Dhuli V. S., Enduri M. K., Cenkeramaddi L. R. A novel convex combination-based mixed centrality measure for identification of influential nodes in complex networks // *IEEE Access*. 2024. Vol. 12. P. 123897–123920. doi: 10.1109/ACCESS.2024.3450296

21. Newman M. E. J., Girvan M. Finding and evaluating community structure in networks // *Phys. Rev. E*. 2004. Vol. 69, no. 2. Art. 026113. doi: 10.1103/PhysRevE.69.026113

22. Reichardt J., Bornholdt S. When are networks truly modular? // *Phys. D: Nonlinear Phenom.* 2006. Vol. 224, no. 1-2. P. 20–26. doi: 10.1016/j.physd.2006.09.009

23. Schnegg M. Reciprocity and the emergence of power laws in social networks // *Internat. J. Modern Phys. C*. 2006. Vol. 17, no. 07. P. 1067–1076. doi: 10.1142/S0129183106009473

24. Traag V. A., Waltman L., Van Eck N. J. From Louvain to Leiden: guaranteeing well-connected communities // *Sci. Rep.* 2019. Vol. 9, no. 1. Art. 5233. doi: 10.1038/s41598-019-41695-z

25. Travers J., Milgram S. An experimental study of the small world problem // *Sociometry*. 1969. Vol. 32, no. 4. P. 425–443.

REFERENCES

1. Kazak V. Lexicon of Russian literature of the 20th century. Moscow: Kul'tura; 1996. 512 p. (In Russ.)

2. Project card for fundamental and exploratory scientific research supported by the Russian Science Foundation. (In Russ.). *RNF = Russian Science*

Foundation. URL: <https://rscf.ru/project/24-18-00761/> (accessed: 10.01.2026).

3. Codifier of verifiable requirements for the results of mastering the basic educational program of secondary general education and content elements for the Unified State Examination in literature. (In Russ.). URL: https://doc.fipi.ru/ege/demoversii-specifikacii-kodifikatory/2026/li_11_2026.zip (accessed: 15.02.2026).

4. Myachin A. N. (ed.). The world of Russian culture. Encyclopedic reference book. Moscow: Veche; 1997. 624 p. (In Russ.)

5. Rajgorodskij A. M. Mathematical models of the Internet. *Kvant = Quantum*. 2012;4:12–16. (In Russ.)

6. Nikolaev P. A. (ed.). Russian writers. 1800–1917. Biographical dictionary: In 5 vol. Moscow: Sov. entsikl.; Bol'shaya ros. entsikl.; 1989–2007. (In Russ.)

7. Nikolaev P. A. (ed.). Russian writers of the 19th century: Biographical dictionary. Vol. 1–2. Moscow: Prosveshchenie; 1996. (In Russ.)

8. Nikolaev P. A. (ed.). Russian writers of the 20th century: Biographical dictionary. Moscow: Bol'shaya ros. entsikl.; Rendezvous-AM; 2000. 806 p. (In Russ.)

9. Ryazantsev A. P. The Russian literary field in the 19th century: an assessment of the state in interviews with critics. *Sotsiologicheskie issledovaniya = Sociological Research*. 2025;51(8):132–141. (In Russ.). doi: 10.7868/S3034601025080112

10. Ryazancev A. P. Social status and reputation of writers in modern Russia: thematic analysis. *Interaktsiya. Interv'yū. Interpretatsiya = Interaction. Interview. Interpretation*. 2023;15(1):27–44. (In Russ.). doi: 10.19181/inter.2023.15.1.2

11. Directory of the Union of writers of the USSR. Moscow: Sovetskii pisatel'; 1981. 821 p. (In Russ.)

12. Sushchij S. Ja. Russian literature of the 19th and early 20th centuries: asociodynamic analysis of the writers' community *Sotsiologicheskie issledovaniya = Sociological Research*. 2020;2:128–143. (In Russ.). doi: 10.31857/S013216250008500-3

13. Sushchiy S. Ya. Russian literature of the Soviet period: the sociodynamics of the writers' community. *Sotsiologicheskie issledovaniya = Sociological Research*. 2021;2:112–129. (In Russ.). doi: 10.31857/S013216250009902-5

14. Avrachenkov K., Dreveton M. Statistical analysis of networks. Boston-Delft: Now Publ.; 2020. 248 p. doi: 10.1561/9781638280514

15. Barabási A. L., Albert R. Emergence of scaling in random networks. *Science*. 1999;286(5439):509–512. doi: 10.1126/science.286.5439.509

16. Blondel V. D., Guillaume J.-L., Lambiotte R., Lefebvre E. Fast unfolding of communities in large networks. *J. Stat. Mech.: Theory Exp.* 2008;10:P10008. doi: 10.1088/1742-5468/2008/10/P10008

17. Edunov S., Diuk C., Filiz I. O., Bhagat S., Burke M. Three and a half degrees of separation. *Research at Facebook blog*. 4 February 2016. URL: <https://research.facebook.com/blog/three-and-a-half-degrees-of-separation/> (accessed: 10.01.2026).
18. Hajarathaiah K., Enduri M. K., Anamalamudi S., Sangi A. R. Algorithms for finding influential people with mixed centrality in social networks. *Arab. J. Sci. Eng.* 2023;48(8):10417–10428. doi: 10.1007/s13369-023-07619-w
19. Mohammad B. B., Dhuli S., Enduri M. K. Isolating centrality-based generalization of traditional centralities to discover vital nodes in complex networks. *Arab. J. Sci. Eng.* 2025;50(15):12003–12025. doi: 10.1007/s13369-024-09628-9
20. Mohammad B. B., Dhuli V. S., Enduri M. K., Cenkeramaddi L. R. A novel convex combination-based mixed centrality measure for identification of influential nodes in complex networks. *IEEE Access*. 2024;12:123897–123920. doi: 10.1109/ACCESS.2024.3450296
21. Newman M. E. J., Girvan M. Finding and evaluating community structure in networks. *Phys. Rev. E*. 2004;69(2):026113. doi: 10.1103/PhysRevE.69.026113
22. Reichardt J., Bornholdt S. When are networks truly modular? *Phys. D: Nonlinear Phenom.* 2006;224(1-2):20–26. doi: 10.1016/j.physd.2006.09.009
23. Schnegg M. Reciprocity and the emergence of power laws in social networks. *Internat. J. Modern Phys. C*. 2006;17(07):1067–1076. doi: 10.1142/S0129183106009473
24. Traag V. A., Waltman L., Van Eck N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci. Rep.* 2019;9(1):5233. doi: 10.1038/s41598-019-41695-z
25. Travers J., Milgram S. An experimental study of the small world problem. *Sociometry*. 1969;32(4):425–443.

Поступила в редакцию / received: 10.04.2026; принята к публикации / accepted: 22.05.2026.
 Автор заявляет об отсутствии конфликта интересов / The author declares no conflict of interest.

СВЕДЕНИЯ ОБ АВТОРЕ:

Бородина Александра Валентиновна
 канд. физ.-мат. наук, старший научный сотрудник
e-mail: borodina@krc.karelia.ru

CONTRIBUTOR:

Borodina, Alexandra
 Cand. Sci. (Phys.-Math.), Senior Researcher

УДК 532.516:532.546

НЕСТАЦИОНАРНОЕ ИЗОБАРИЧЕСКОЕ ТЕЧЕНИЕ ВЯЗКОЙ НЕСЖИМАЕМОЙ ЖИДКОСТИ В ТОНКОМ СЛОЕ С ПРОНИЦАЕМЫМИ ГРАНИЦАМИ

К. В. Губарева^{1*}, Е. Ю. Просвирыков^{2,3}, А. В. Еремин¹

¹ Самарский государственный технический университет (ул. Молодогвардейская, 244, Самара, Россия, 443100), *r.kristina2017@mail.ru

² Уральский федеральный университет имени первого Президента России Б. Н. Ельцина (ул. Мира, 19, Екатеринбург, Россия, 620062)

³ Институт машиноведения имени Э. С. Горкунова УрО РАН (ул. Комсомольская, 34, Екатеринбург, Россия, 620049)

Представлено точное решение нестационарных уравнений Навье – Стокса, описывающее трехмерное изобарическое течение вязкой несжимаемой жидкости в горизонтальном слое с проницаемыми границами. Скоростное поле ищется в классе Линя, что позволяет свести исходную систему к одномерным уравнениям для амплитудных функций, зависящих от вертикальной координаты и времени. Показано, что вертикальная компонента скорости постоянна и соответствует равномерному отсосу или вдуву. Получены аналитические выражения для эволюции полей скорости, включая нелинейную компоненту, описывающую вертикальную завихренность. Задача Штурма – Лиувилля для оператора конвекции-диффузии решается аналитически с учетом условия проскальзывания Навье на нижней границе. Для компоненты, возбуждаемой нелинейным источником, использован интеграл Дюамеля. Проведен параметрический анализ влияния вертикального числа Рейнольдса, модифицированного числа Тейлора, угла направления скорости на верхней границе и параметра проскальзывания на динамику течения. Установлены режимы с развитыми противотечениями и застойными точками. Показано, что при отрицательном вертикальном числе Рейнольдса (отсос) время релаксации убывает с ростом проскальзывания, что качественно отличается от классического случая прилипания. Полученные результаты могут быть использованы для верификации численных алгоритмов и при моделировании гидродинамических процессов в пористых средах, мембранных аппаратах и каталитических слоях, где важны нестационарные эффекты и наличие проницаемых границ.

Ключевые слова: точные решения; уравнения Навье – Стокса; нестационарное течение; проницаемые границы; условие проскальзывания; вертикальная завихренность; противотечения; время релаксации

Для цитирования: Губарева К. В., Просвирыков Е. Ю., Еремин А. В. Нестационарное изобарическое течение вязкой несжимаемой жидкости в тонком слое с проницаемыми границами // Труды Карельского научного центра РАН. 2026. № 6. С. 20–31. doi: 10.17076/mat2344

Финансирование. Исследование выполнено при финансовой поддержке Министерства науки и высшего образования Российской Федерации (тема № FSSE-2025-0006) в рамках государственного задания Самарского государственного технического университета (создание новых молодежных лабораторий).

**K. V. Gubareva^{1*}, E. Yu. Prosviryakov^{2,3}, A. V. Eremin¹.
UNSTEADY ISOBARIC FLOW OF A VISCOUS
INCOMPRESSIBLE FLUID IN A THIN LAYER WITH
PERMEABLE BOUNDARIES**

¹ Samara State Technical University (244 Molodogvardeyskaya St., 443100 Samara, Russia), *r.kristina2017@mail.ru

² Ural Federal University (19 Mira St., 620062 Yekaterinburg, Russia)

³ Gorkunov Institute of Engineering Science, Ural Branch of the Russian Academy of Sciences (34 Komsomolskaya St., 620049 Yekaterinburg, Russia)

An exact solution of unsteady Navier–Stokes equations is presented for a three-dimensional isobaric flow of a viscous incompressible fluid in a horizontal layer with permeable boundaries. The velocity field is sought in the Lin class, which reduces the original system to one-dimensional equations for amplitude functions depending on the vertical coordinate and time. It is shown that the vertical velocity component is constant, corresponding to uniform suction or injection. Analytical expressions for the evolution of the velocity fields are obtained, including the nonlinear component describing the vertical vorticity. The Sturm–Liouville problem for the convection–diffusion operator is solved analytically taking into account the Navier slip condition on the lower boundary. For the component excited by a nonlinear source, the Duhamel integral is used. Parametric analysis is performed for the impact of the vertical Reynolds number, the modified Taylor number, the angle of the velocity direction on the upper boundary, and the slip parameter on the flow dynamics. Regimes with developed counterflows and stagnation points are identified. It is shown that for a negative vertical Reynolds number (suction), the relaxation time decreases with increasing slip, which differs qualitatively from the classical no-slip case. The results obtained can be used for verifying numerical algorithms and in modelling hydrodynamic processes in porous media, membrane devices and catalytic layers, where unsteady effects and the presence of permeable boundaries are important.

Key words: exact solutions; Navier–Stokes equations; unsteady flow; permeable boundaries; slip condition; vertical vorticity; counterflows; relaxation time

For citation: Gubareva K. V., Prosviryakov E. Yu., Eremin A. V. Unsteady isobaric flow of a viscous incompressible fluid in a thin layer with permeable boundaries. *Trudy Karelskogo nauchnogo tsentra RAN = Transactions of the Karelian Research Centre RAS*. 2026. No. 6. P. 20–31. doi: 10.17076/mat2344

Funding. The study was supported by the Ministry of Science and Higher Education of the Russian Federation (theme No. FSSE-2025-0006) within state assignment to the Samara State Technical University (creation of new youth laboratories).

ВВЕДЕНИЕ

Гидродинамические процессы в слоистых средах с проницаемыми границами широко распространены в природе и технике. К ним относятся фильтрация жидкостей через пористые материалы, течение в мембранных аппаратах, работа каталитических реакторов, а также крупномасштабные движения в океане и атмосфере [1, 10, 17]. Во всех этих случаях возможность точного математического описания полей скорости позволяет прогнозировать поведение системы, оптимизировать конструкцию устройств и верифицировать численные модели [9, 11].

Уравнения Навье–Стокса, описывающие движение вязкой несжимаемой жидкости, до

сих пор остаются предметом интенсивных исследований. Несмотря на их нелинейность, существуют классы точных решений, которые дают ценную информацию о структуре течений и служат эталонами для проверки приближенных методов [4, 6, 18]. Одним из наиболее продуктивных подходов является использование решений в классе Линия [14], где компоненты скорости представляются в виде линейных форм относительно двух пространственных переменных с коэффициентами, зависящими от третьей координаты и времени [4, 18]. Такие решения позволяют описать трехмерные течения, включая эффекты вертикальной завихренности, и находят применение в геофизической гидродинамике [2, 16], хими-

ческой технологии [3] и микрофлюидике [15]. При этом следует подчеркнуть, что подобные точные решения служат в первую очередь эталонами для верификации численных алгоритмов и инструментом для качественного анализа влияния различных факторов, а не прямым описанием конкретных природных или технологических процессов во всей их сложности. В частности, они позволяют исследовать вклад таких эффектов, как проницаемость границ, проскальзывание и нестационарность, в упрощенной, но математически строгой постановке.

Особый интерес представляют нестационарные режимы, возникающие при изменении граничных условий или начальных возмущений. В таких задачах необходимо учитывать изменения скорости во времени, время релаксации к стационарному состоянию, а также эволюцию во времени областей обратных течений и застойных зон (которые могут существовать и в стационарных состояниях) [8, 12]. Эти явления существенно влияют на эффективность тепломассообмена и химических превращений в аппаратах с проницаемыми стенками, например, в каталитических слоях или мембранных реакторах. Разработка аналитических моделей, учитывающих нестационарность, позволяет установить количественные критерии перехода к стационарному режиму, что важно для проектирования технологических процессов.

Цель работы – построение и анализ точного нестационарного решения уравнений Навье – Стокса для изобарического течения в горизонтальном слое с проницаемыми границами. В отличие от известных стационарных решений [5] в работе учитывается изменение возмущений, вызванных начальными отклонениями от равновесия. Исследуется влияние вертикального числа Рейнольдса, модифицированного числа Тейлора, угла направления скорости на верхней границе и параметра проскальзывания Навье на динамику формирования противотечений и застойных точек. Особое внимание уделяется зависимости времени релаксации от этих параметров. Полученные результаты могут служить основой для верификации численных алгоритмов и для анализа гидродинамических процессов в пористых и мембранных системах, где важна точная оценка переходных режимов.

Постановка задачи

Рассматривается бесконечный горизонтальный слой вязкой несжимаемой жидкости толщины h . Система координат выбра-

на так, что ось Oz направлена вертикально вверх, а плоскости $z = 0$ и $z = h$ являются нижней и верхней границами. Границы считаются проницаемыми и недеформируемыми; через них может происходить равномерный отсос или вдув жидкости с постоянной вертикальной скоростью W . Жидкость предполагается изобарической – градиент давления отсутствует, что позволяет рассматривать чисто инерционно-вязкие течения. Такое приближение является стандартным для решений класса Линя [6, 14] и оправдано для тонких горизонтальных слоев ($\delta = h/l \ll 1$), где вертикальные вязкие напряжения доминируют над горизонтальными градиентами давления. Оценка порядков показывает, что условие $|\partial P / \partial x| \ll \mu |\partial^2 V_x / \partial z^2|$ выполняется при $\delta^2 \text{Re}_U \lesssim O(1)$. Для рассматриваемых далее параметров ($\delta = 0,01$) это условие удовлетворяется, что подтверждает применимость изобарической модели для описания течений в тонких слоях.

Исходная система уравнений Навье – Стокса и неразрывности имеет вид:

$$\frac{\partial \mathbf{V}}{\partial t} + (\mathbf{V} \cdot \nabla) \mathbf{V} = -\frac{1}{\rho} \nabla P + \nu \Delta \mathbf{V}, \quad \nabla \cdot \mathbf{V} = 0,$$

где $\mathbf{V} = (V_x, V_y, V_z)$ – вектор скорости, P – давление, ρ – плотность, ν – кинематическая вязкость. В изобарическом приближении $\nabla P = 0$, что существенно упрощает задачу.

Решение ищется в классе Линя, обобщенном на нестационарный случай [7]. Поле скоростей представляется в виде линейной комбинации по горизонтальным координатам:

$$\begin{aligned} V_x &= U(z, t) + y u(z, t), \\ V_y &= V(z, t), \\ V_z &= w(z, t), \\ P &= P_0(z, t). \end{aligned} \tag{1}$$

Все функции в (1) зависят только от вертикальной координаты z и времени t . Далее функции скорости записываются в виде $V_x(z, x, y, t)$, $V_y(z, x, y, t)$, $V_z(z, x, y, t)$, где первый аргумент – вертикальная координата z , далее следуют горизонтальные координаты x, y и время t . Такой порядок аргументов подчеркивает, что основное внимание в работе уделяется вертикальной структуре течения, а горизонтальные координаты играют роль параметров.

Представление (1) является частным случаем общего класса Линя [6, 14] и выбрано специально для описания течений с вертикальной завихренностью. Действительно, вертикальная компонента завихренности вычис-

ляется как $\omega_z = \partial V_y / \partial x - \partial V_x / \partial y$. Для принятого анзаца (V_y не зависит от x , а V_x линейно зависит от y) получаем $\omega_z = -u(z, t) \neq 0$. Таким образом, асимметрия в зависимости компонент скорости от горизонтальных координат не является ограничением общности, а отражает физическую природу рассматриваемого класса точных решений, позволяя моделировать вихревые структуры при сохранении пространственно-одномерной редукции уравнений Навье – Стокса.

Подстановка (1) в уравнения движения и неразрывности дает условие $\partial w / \partial z = 0$, откуда $w = w(t)$. Далее из уравнения для вертикальной компоненты при $\nabla P = 0$ следует $dw/dt = 0$, поэтому вертикальная скорость постоянна: $w(t) = W$.

Для оставшихся компонент получается система:

$$\frac{\partial u}{\partial t} + W \frac{\partial u}{\partial z} = \nu \frac{\partial^2 u}{\partial z^2}, \quad (2)$$

$$\frac{\partial V}{\partial t} + W \frac{\partial V}{\partial z} = \nu \frac{\partial^2 V}{\partial z^2}, \quad (3)$$

$$\frac{\partial U}{\partial t} + W \frac{\partial U}{\partial z} = \nu \frac{\partial^2 U}{\partial z^2} - Vu. \quad (4)$$

Уравнения (2) и (3) линейны и однородны, уравнение (4) линейно относительно U с неоднородностью, определяемой произведением уже найденных функций V и u .

Граничные и начальные условия

На нижней границе ($z = 0$) ставится условие проскальзывания Навье [13, 15]:

$$\alpha \frac{\partial V_x}{\partial z} \Big|_{z=0} = V_x \Big|_{z=0}, \quad \alpha \frac{\partial V_y}{\partial z} \Big|_{z=0} = V_y \Big|_{z=0},$$

где $\alpha = \beta/h$ – безразмерная длина скольжения, β – размерная длина скольжения. При $\alpha = 0$ получается классическое условие прилипания. Согласно экспериментальным данным [15], для гладких гидрофильных поверхностей β составляет единицы–десятки нанометров, что при толщине слоя $h \sim 1$ мм дает $\alpha \sim 10^{-6}$ – 10^{-4} . Для гидрофобных или микроструктурированных поверхностей β может достигать микрометров, и тогда $\alpha \sim 10^{-3}$ – 10^{-1} . На величину β влияют шероховатость, смачиваемость и адсорбция примесей. В работе рассматривается диапазон $\alpha \in [0, 5]$, охватывающий как практически важные малые значения, так и предельные случаи сильного проскальзывания. В функциях U, u, V эти условия

записываются как

$$\begin{aligned} \alpha \frac{\partial u}{\partial z}(0, t) &= u(0, t), \\ \alpha \frac{\partial V}{\partial z}(0, t) &= V(0, t), \\ \alpha \frac{\partial U}{\partial z}(0, t) &= U(0, t). \end{aligned} \quad (5)$$

На верхней границе ($z = 1$) задаются компоненты скорости как функции времени (в безразмерной форме). В общем случае:

$$\begin{aligned} u(1, t) &= u_1(t), \\ V(1, t) &= V_1(t), \\ U(1, t) &= U_1(t). \end{aligned} \quad (6)$$

Для постоянных граничных условий, как в [5], принимаются значения

$$u_1 = \frac{\text{Ta}}{2\text{Re}_U}, \quad V_1 = \sin \varphi, \quad U_1 = \cos \varphi,$$

где $\text{Ta} = 2\Omega l^2/\nu$ – модифицированное число Тейлора, Ω – вертикальная компонента завихренности, φ – угол, задающий направление скорости на верхней границе.

Подчеркнем, что выбор граничных условий на верхней границе не преследует цели точного воспроизведения какого-либо конкретного физического сценария (например, ветрового воздействия на океанический слой), а служит удобной параметризацией для исследования влияния сдвиговых и вращательных компонент скорости на динамику течения. Аналогичный подход характерен для классических задач Куэтта и Экмана, в которых при малых числах Рейнольдса допустимо пренебрежение нелинейными членами и анализ проводится на основе линеаризованных уравнений.

Начальные условия при $t = 0$:

$$\begin{aligned} u(z, 0) &= u_0(z), \\ V(z, 0) &= V_0(z), \\ U(z, 0) &= U_0(z). \end{aligned} \quad (7)$$

Для перехода к безразмерным переменным вводятся масштабы: толщина слоя h для вертикальной координаты, характерный горизонтальный масштаб l , масштаб скорости $[U]$. Тогда

$$\begin{aligned} z^* &= \frac{z}{h}, \quad t^* = \frac{t\nu}{h^2}, \quad x^* = \frac{x}{l}, \quad y^* = \frac{y}{l}, \\ u^* &= \frac{u}{[U]}, \quad V^* = \frac{V}{[U]}, \quad U^* = \frac{U}{[U]}, \quad W^* = \frac{Wh}{\nu} = \text{Re}_w. \end{aligned}$$

Вводятся числа Рейнольдса: $\text{Re}_w = Wh/\nu$ (вертикальное), $\text{Re}_U = [U]l/\nu$ (горизонтальное) и параметр геометрической анизотропии $\delta = h/l$.

Горизонтальный масштаб l определяет характерный размер области, в которой рассматриваемое течение описывается представлением (1). В силу линейной зависимости V_x от y данное решение является локальным и применимо лишь при $|x|, |y| \lesssim l$. Вне этой окрестности линейное приближение нарушается, и необходимо учитывать нелинейные эффекты либо градиенты давления. Никаких дополнительных граничных условий на горизонтальных границах не требуется, поскольку решение не претендует на описание конечного домена размером $l \times l$, а используется для анализа течения в локальной окрестности начала координат. Все последующие результаты относятся именно к этой ограниченной области, где поле скорости сохраняет физический смысл.

После подстановки и опускания звездочек система (2)–(4) принимает вид

$$\frac{\partial u}{\partial t} + \text{Re}_w \frac{\partial u}{\partial z} = \frac{\partial^2 u}{\partial z^2}, \quad (8)$$

$$\frac{\partial V}{\partial t} + \text{Re}_w \frac{\partial V}{\partial z} = \frac{\partial^2 V}{\partial z^2}, \quad (9)$$

$$\frac{\partial U}{\partial t} + \text{Re}_w \frac{\partial U}{\partial z} = \frac{\partial^2 U}{\partial z^2} - \delta^2 \text{Re}_U V u, \quad (10)$$

где $z \in [0, 1]$, $t \geq 0$. Граничные условия (5) и (6) и начальные условия (7) также записываются в безразмерной форме с сохранением тех же обозначений.

МЕТОД РЕШЕНИЯ

Стационарное решение

Для решения нестационарной задачи сначала находятся стационарные профили, отвечающие установившемуся режиму ($t \rightarrow \infty$). Уравнения (8)–(10) при $\partial/\partial t = 0$ дают (штрих обозначает производную по z)

$$u_{\text{ст}}'' - \text{Re}_w u_{\text{ст}}' = 0, \quad V_{\text{ст}}'' - \text{Re}_w V_{\text{ст}}' = 0,$$

$$U_{\text{ст}}'' - \text{Re}_w U_{\text{ст}}' = -\delta^2 \text{Re}_U V_{\text{ст}} u_{\text{ст}}.$$

Граничные условия для стационарной части совпадают с (5)–(6), но без зависимости от времени. Решения для $u_{\text{ст}}$ и $V_{\text{ст}}$ находятся явно:

$$u_{\text{ст}}(z) = u_1 \frac{e^{\text{Re}_w z} + \alpha \text{Re}_w - 1}{e^{\text{Re}_w} + \alpha \text{Re}_w - 1},$$

$$V_{\text{ст}}(z) = V_1 \frac{e^{\text{Re}_w z} + \alpha \text{Re}_w - 1}{e^{\text{Re}_w} + \alpha \text{Re}_w - 1}.$$

Для случая $\text{Re}_w = 0$ эти выражения упрощаются до линейных функций: $u_{\text{ст}} = u_1(z + \alpha)/(1 + \alpha)$, $V_{\text{ст}} = V_1(z + \alpha)/(1 + \alpha)$.

Стационарное поле $U_{\text{ст}}$ определяется из краевой задачи второго порядка с переменными коэффициентами, которая решается численно с помощью метода конечных разностей (функция `bvp4c` в MATLAB). Полученное решение $U_{\text{ст}}(z)$ используется в качестве базового состояния.

Преобразование уравнений и задача Штурма – Лиувилля

Для исследования нестационарной части введем замену, преобразующую конвективный член к виду без первой производной. Пусть

$$f(z, t) = e^{\frac{\text{Re}_w}{2} z} \psi(z, t),$$

где f обозначает любую из функций u , V или U . Подстановка в (8)–(9) приводит к уравнению

$$\frac{\partial \psi}{\partial t} = \frac{\partial^2 \psi}{\partial z^2} - \frac{\text{Re}_w^2}{4} \psi.$$

Граничные условия для ψ получаются из (5)–(6). На верхней границе:

$$\psi(1, t) = e^{-\frac{\text{Re}_w}{2}} f_1(t).$$

На нижней границе после преобразования условия (5) имеем

$$\alpha \frac{\partial \psi}{\partial z}(0, t) + \left(\frac{\alpha \text{Re}_w}{2} - 1 \right) \psi(0, t) = 0. \quad (11)$$

Для построения базиса собственных функций рассматривается однородная задача, в которой граничное условие на верхней границе заменяется нулевым: $\psi(1, t) = 0$. Это дает самосопряженную краевую задачу Штурма – Лиувилля для оператора $L\psi = -\psi'' + \frac{\text{Re}_w^2}{4}\psi$ с граничными условиями (11) и $\varphi_n(1) = 0$ для собственных функций $\varphi_n(z)$. Отметим, что нулевое условие используется только для определения спектра и базисных функций; неоднородность граничных условий (6) будет учтена в дальнейшем через стационарное решение и разложение отклонений по найденному базису. Собственные функции $\varphi_n(z)$ и собственные значения λ_n удовлетворяют

$$-\varphi_n'' + \frac{\text{Re}_w^2}{4} \varphi_n = \lambda_n \varphi_n, \quad n = 1, 2, \dots,$$

с теми же однородными граничными условиями. Общее решение имеет вид

$$\varphi_n(z) = A_n \sin(k_n z) + B_n \cos(k_n z),$$

$$k_n^2 = \lambda_n - \frac{\text{Re}_w^2}{4}.$$

Из граничных условий получается характеристическое уравнение для k_n :

$$\alpha k_n \cos k_n + \left(\frac{\alpha \text{Re}_w}{2} - 1 \right) \sin k_n = 0. \quad (12)$$

При $\alpha = 0$ уравнение упрощается до $\sin k_n = 0$, откуда $k_n = n\pi$, $\lambda_n = (n\pi)^2 + \text{Re}_w^2/4$. Для $\alpha > 0$ корни k_n находятся численно (методом Ньютона с выбором интервалов). После определения k_n строятся ненормированные функции

$$\varphi_n^{(0)}(z) = \sin(k_n z) + \frac{\alpha k_n}{1 - \alpha \text{Re}_w/2} \cos(k_n z),$$

которые затем нормируются условием $\int_0^1 \varphi_n^2(z) dz = 1$. Собственные функции ортогональны и образуют полную систему в $L_2[0, 1]$.

Решение для u и V

Уравнения (8)–(9) однородны, поэтому их решения можно представить в виде разложения по собственным функциям. С учетом замены $f = e^{\frac{\text{Re}_w}{2}z}\psi$ (где f – любая из функций u или V) имеем:

$$u(z, t) = u_{\text{ст}}(z) + e^{\frac{\text{Re}_w}{2}z} \sum_{n=1}^{\infty} C_n^u e^{-\lambda_n t} \varphi_n(z),$$

$$V(z, t) = V_{\text{ст}}(z) + e^{\frac{\text{Re}_w}{2}z} \sum_{n=1}^{\infty} C_n^V e^{-\lambda_n t} \varphi_n(z).$$

Коэффициенты C_n^u и C_n^V определяются из начальных условий (7) после разложения отклонений $u(z, 0) - u_{\text{ст}}(z)$ и $V(z, 0) - V_{\text{ст}}(z)$ на базис φ_n с весом $e^{-\frac{\text{Re}_w}{2}z}$:

$$C_n^u = \int_0^1 e^{-\frac{\text{Re}_w}{2}z} [u_0(z) - u_{\text{ст}}(z)] \varphi_n(z) dz,$$

$$C_n^V = \int_0^1 e^{-\frac{\text{Re}_w}{2}z} [V_0(z) - V_{\text{ст}}(z)] \varphi_n(z) dz.$$

Эти ряды сходятся экспоненциально быстро благодаря множителям $e^{-\lambda_n t}$.

Решение для U с интегралом Дюамеля

Уравнение (10) неоднородно: источник $Q(z, t) = -\delta^2 \text{Re}_U V(z, t) u(z, t)$ зависит от уже найденных полей. После замены $U = e^{\frac{\text{Re}_w}{2}z}\Psi$ и разложения Ψ по тем же собственным функциям

$$\Psi(z, t) = \sum_{n=1}^{\infty} S_n(t) \varphi_n(z),$$

для коэффициентов $S_n(t)$ получаются обыкновенные дифференциальные уравнения

$$\frac{dS_n}{dt} = -\lambda_n S_n + q_n(t),$$

где

$$q_n(t) = \int_0^1 e^{-\frac{\text{Re}_w}{2}z} [Q(z, t) - Q_{\text{ст}}(z)] \varphi_n(z) dz,$$

а $Q_{\text{ст}}(z) = -\delta^2 \text{Re}_U V_{\text{ст}}(z) u_{\text{ст}}(z)$ – стационарная часть источника. Начальные условия для

$S_n(t)$ получаются из начального возмущения $U_0(z) - U_{\text{ст}}(z)$:

$$S_n(0) = \int_0^1 e^{-\frac{\text{Re}_w}{2}z} [U_0(z) - U_{\text{ст}}(z)] \varphi_n(z) dz.$$

Решение этих уравнений записывается в форме интеграла Дюамеля:

$$S_n(t) = S_n(0) e^{-\lambda_n t} + \int_0^t e^{-\lambda_n(t-\tau)} q_n(\tau) d\tau.$$

Интеграл вычисляется численно методом трапеций на заданной сетке моментов времени t_k . После нахождения $S_n(t)$ поле $U(z, t)$ находится как

$$U(z, t) = U_{\text{ст}}(z) + e^{\frac{\text{Re}_w}{2}z} \sum_{n=1}^{\infty} S_n(t) \varphi_n(z).$$

Сходимость ряда обеспечивается экспоненциальным затуханием $e^{-\lambda_n t}$, поэтому для практических расчетов достаточно удерживать несколько десятков первых мод.

На основе полученных полей скорости вычисляются интегральные характеристики. Процент области с противотечениями определяется как доля вертикальной координаты z , в которой $U(z, t) < 0$, усредненная по пространству в каждый момент времени. Застойные точки – положения z , где $U(z, t) = 0$ (для фиксированных x, y), находятся интерполяцией между узлами сетки. Время релаксации $T_{\text{rel}} = 1/\lambda_1$, где λ_1 – наименьшее собственное значение, характеризует скорость затухания возмущений.

Параметрические исследования проводятся для различных жидкостей (вода, масло SAE 30, глицерин при 20 °C) и варьирования управляющих параметров Re_w , Ta , φ , α . Все расчеты выполняются в MATLAB 2023b с использованием разработанного алгоритма.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Прежде чем перейти к анализу профилей скорости, оценим выполнимость изобарического приближения для выбранных жидкостей. Как показано в разделе «Постановка задачи», условие малости горизонтальных градиентов давления имеет вид $\delta^2 \text{Re}_U \lesssim O(1)$. Для воды ($\text{Re}_U \approx 10^5$) получаем $\delta^2 \text{Re}_U = 10$; для масла SAE 30 ($\text{Re}_U \approx 10^3$) – 0,1; для глицерина ($\text{Re}_U \approx 10^2$) – 0,01. Таким образом, для более вязких жидкостей (масло, глицерин) изобарическое приближение выполняется с большим запасом, тогда как для воды его применимость ограничена, но все же допустима в рамках оценочного анализа. Это качественно согласуется с наблюдаемым ниже поведением решений: для воды релаксация происходит быстрее, а профили скорости менее чувствительны к вариациям параметров.

На рис. 1 представлена эволюция профиля $U(z, t)$ для трех жидкостей при $Ta = 10^6$, $\varphi = -\pi/4$, $\alpha = 0,01$. Для воды ($Re_w = -10$) релаксация происходит быстро: уже к моменту $t = 0,1$ профиль близок к стационарному. Для масла и глицерина ($|Re_w| \ll 1$) процесс идет значительно медленнее, что отражается в большем времени релаксации (табл.). На всех профилях видна зона обратного течения ($U < 0$) вблизи верхней границы, которая со временем уменьшается, но не исчезает полностью.

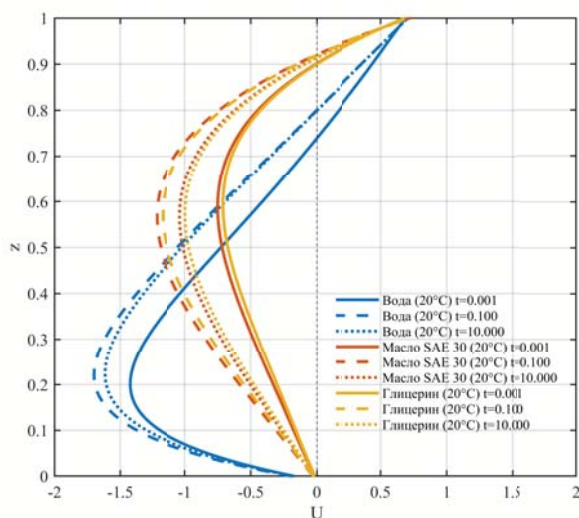


Рис. 1. Эволюция профиля $U(z, t)$ для трех жидкостей. Сплошные линии – $t = 0,001$, штриховые – $t = 0,1$, пунктирные – $t = 10$. Параметры: $Ta = 10^6$, $\varphi = -\pi/4$, $\alpha = 0,01$

Fig. 1. Evolution of the $U(z, t)$ profile for three fluids. Solid lines – $t = 0,001$, dashed – $t = 0,1$, dotted – $t = 10$. Parameters: $Ta = 10^6$, $\varphi = -\pi/4$, $\alpha = 0,01$

Характеристики жидкостей и параметры течения при $\alpha = 0,01$, $Ta = 10^6$, $\varphi = -\pi/4$

Fluid properties and flow parameters for $\alpha = 0,01$, $Ta = 10^6$, $\varphi = -\pi/4$

Жидкость, 20 °C Fluid	Re_w	Доля, ($U < 0$), % Fraction	T_{rel}
Вода Water	-10,00	78,7	0,0285
Масло SAE 30 Oil	-0,10	91,0	0,0993
Глицерин Glycerin	-0,01	91,0	0,0993

На рис. 2 показаны распределения $U(z, t)$ в координатах (z, t) : цветом изображено значение скорости, а красной линией отмечена гра-

ница $U = 0$. По этим графикам хорошо видно, как область обратного течения ($U < 0$) изменяется во времени. Для воды граница $U = 0$ быстро достигает стационарного положения, тогда как для вязких жидкостей процесс растянут во времени. Стационарная граница для воды находится на $z \approx 0,8$, для масла и глицерина – на $z \approx 0,91$. Процент области с $U < 0$ в среднем составляет 78,7 % для воды и около 91 % для масла и глицерина.

Параметрический анализ для воды представлен на рис. 3. Зависимость доли противотечений и времени релаксации от Re_w (рис. 3, а) показывает, что максимальная доля обратного течения (около 91 %) достигается при $Re_w \approx -2$. При увеличении Re_w в положительную сторону доля резко падает до 41 % при $Re_w = 4$. Время релаксации минимально при больших $|Re_w|$ ($T_{rel} = 0,0285$ при $Re_w = -10$) и возрастает до 0,0993 при $Re_w = 0$. Такое поведение объясняется конкуренцией вязкой диффузии и вертикального конвективного переноса. При $Re_w = 0$ конвекция отсутствует, и возмущения затухают только за счет диффузии, что дает наибольшее время релаксации. При $Re_w \neq 0$ (отсос или вдув) конвективный унос возмущений ускоряет их вынос из слоя или подавление в пограничной области, сокращая T_{rel} . С ростом $|Re_w|$ вклад конвекции усиливается и время релаксации монотонно убывает. Математически это отражается в структуре собственных значений задачи Штурма – Лиувилля: $\lambda_n = k_n^2 + Re_w^2/4$. Слагаемое $Re_w^2/4$ всегда положительно и увеличивает скорость затухания возмущений.

График зависимости доли противотечений от φ (рис. 3, б) периодичен с периодом 2π . Противотечения отсутствуют (0 %) при $\varphi \in [0^\circ; 90^\circ]$, когда горизонтальная компонента скорости на верхней границе $U_1 = \cos \varphi \geq 0$. Максимальная доля противотечений (100 %) достигается при $\varphi \in [180^\circ; 270^\circ]$, когда $U_1 = \cos \varphi < 0$ и течение направлено вдоль отрицательной оси x . В переходных диапазонах $\varphi \in [90^\circ; 180^\circ]$ и $[270^\circ; 360^\circ]$ доля противотечений монотонно изменяется от 0 до 100 % и обратно.

Зависимость времени релаксации от параметра проскальзывания α (рис. 3, с) при $Re_w = -10$ показывает убывание T_{rel} с ростом α : от 0,0287 при $\alpha = 0$ до 0,0255 при $\alpha = 5$. Это качественно отличается от случая $Re_w = 0$, где T_{rel} возрастает с α . Причина заключается в несимметричном влиянии знака Re_w на спектр задачи Штурма – Лиувилля при наличии проскальзывания. Хотя формально $\lambda_n = k_n^2 + Re_w^2/4$, корни k_n определяются из харак-

теристического уравнения (12), которое содержит Re_w линейно. При $\alpha = 0$ (прилипание) уравнение не зависит от знака Re_w и $k_n = n\pi$. При $\alpha > 0$ знак Re_w влияет на k_n и для отрицательных Re_w (отсос) увеличение α приводит к росту k_1 и, следовательно, λ_1 , ускоряя

релаксацию. Для положительных Re_w (вдув) эффект обратный. Таким образом, наблюдаемое поведение является следствием совместного влияния конвекции и граничного условия проскальзывания.

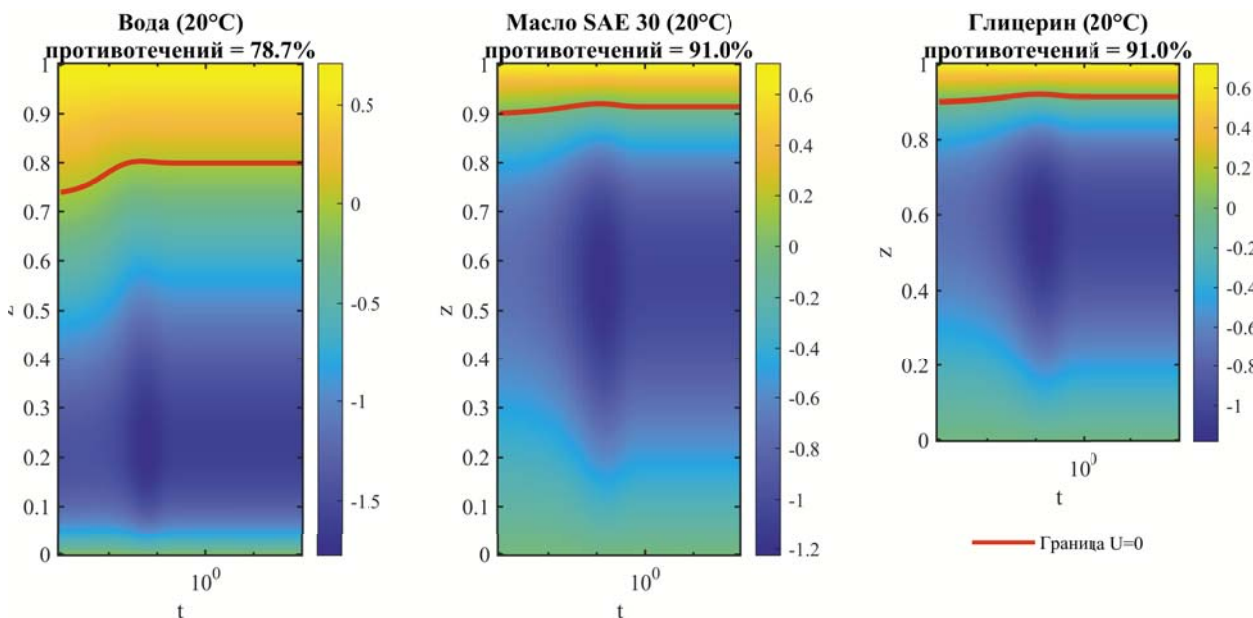


Рис. 2. Области противотечений на плоскости (z, t) . Красная линия – граница $U = 0$. Параметры: $Ta = 10^6$, $\varphi = -\pi/4$, $\alpha = 0,01$

Fig. 2. Counterflow regions in the (z, t) plane. Red line – boundary $U = 0$. Parameters: $Ta = 10^6$, $\varphi = -\pi/4$, $\alpha = 0.01$

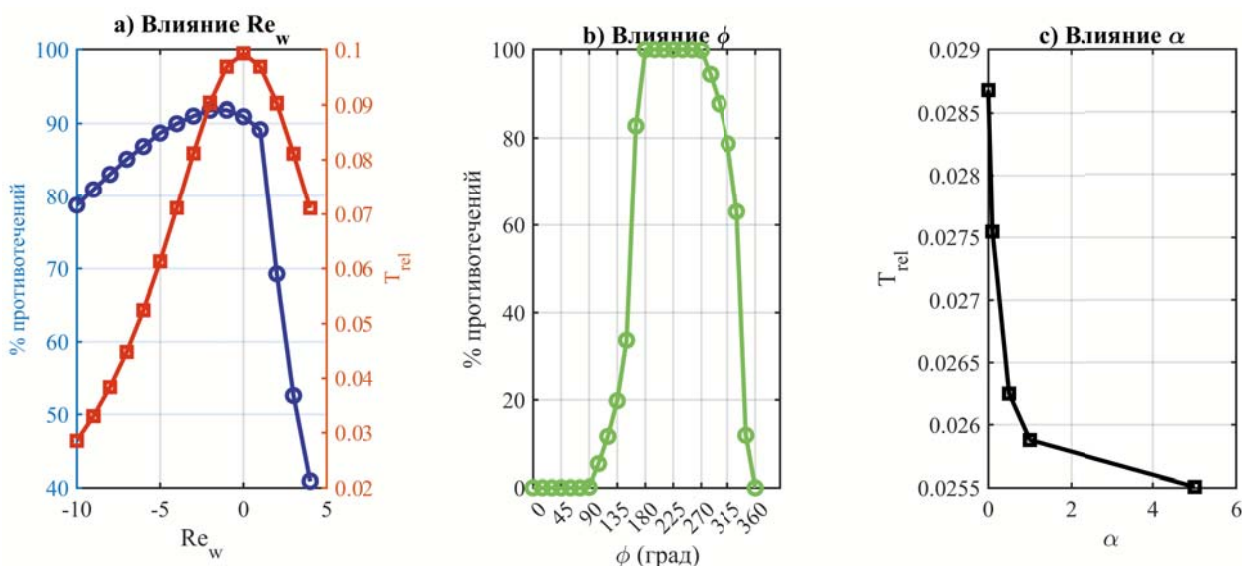


Рис. 3. Параметрические зависимости для воды: а) влияние Re_w ($Ta = 10^6$, $\varphi = -\pi/4$, $\alpha = 0,01$); б) влияние φ ($Re_w = -10$, $Ta = 10^6$, $\alpha = 0,01$); в) влияние α ($Re_w = -10$, $Ta = 10^6$, $\varphi = -\pi/4$)

Fig. 3. Parametric dependencies for water: а) influence of Re_w ($Ta = 10^6$, $\varphi = -\pi/4$, $\alpha = 0.01$); б) influence of φ ($Re_w = -10$, $Ta = 10^6$, $\alpha = 0.01$); в) influence of α ($Re_w = -10$, $Ta = 10^6$, $\varphi = -\pi/4$)

На рис. 4 представлены линии тока в плоскости (x, z) при $y = 0$ для всех трех жидкостей в начальный ($t = 0,001$) и конечный ($t = 100$) моменты времени. Линии тока имеют наклонную структуру, обусловленную совместным действием горизонтального сдвига и вертикального отсоса ($Re_w < 0$). Для воды ($|Re_w| = 10$) наклон линий тока более выражен, тогда как для масла и глицерина

($|Re_w| \ll 1$) линии тока практически горизонтальны. Цветовая карта показывает, что область возвратного течения ($V_x < 0$, синий цвет) формируется в нижней части слоя и занимает около 20 % толщины для воды и более 90 % для масла и глицерина, что согласуется с данными таблицы и рис. 2. К моменту $t = 100$ структура течения стабилизируется, что подтверждает выход на стационарный режим.

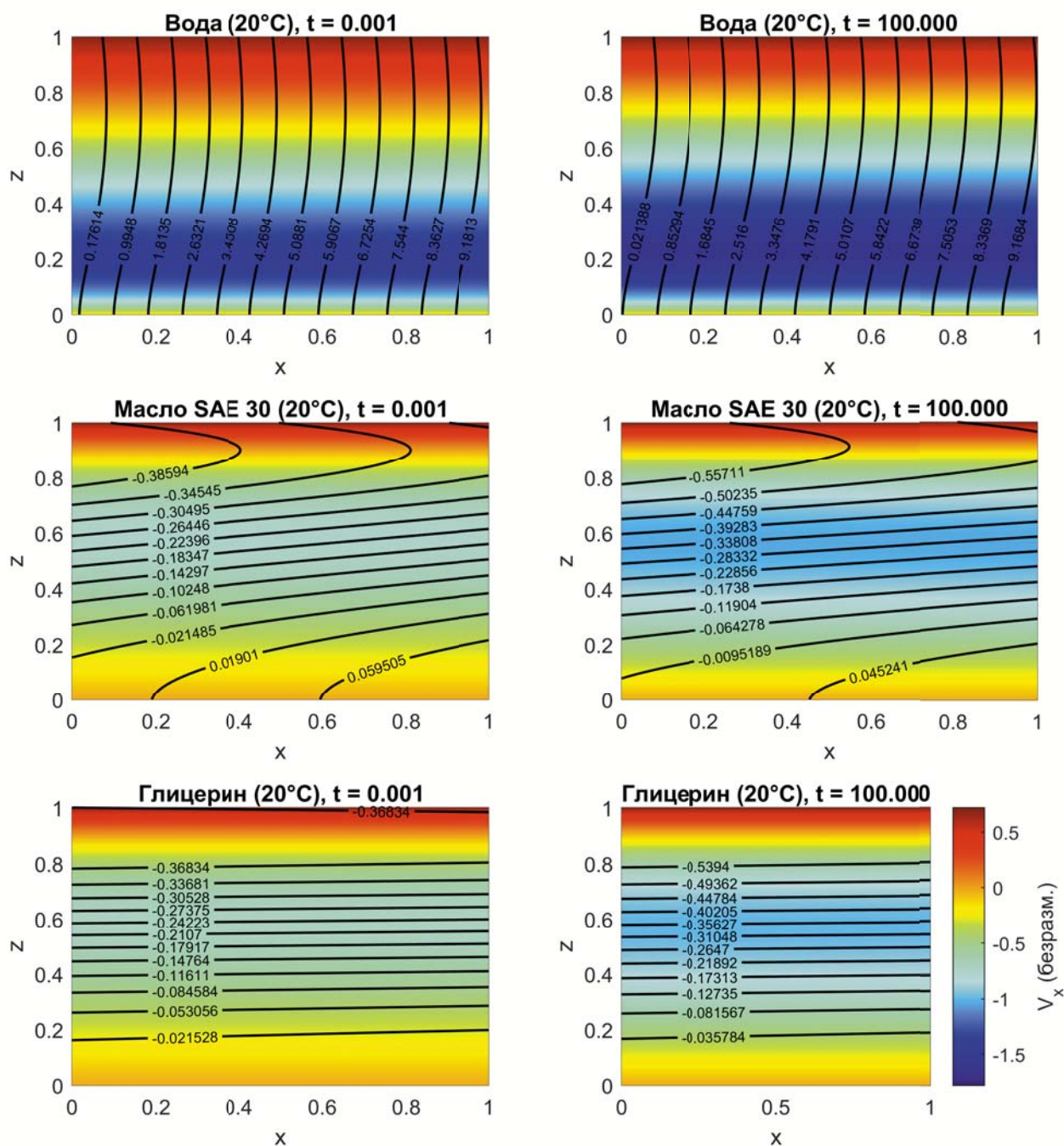


Рис. 4. Линии тока в плоскости (x, z) при $y = 0$ для трех жидкостей
Fig. 4. Streamlines in the (x, z) -plane at $y = 0$ for three fluids

ЗАКЛЮЧЕНИЕ

В работе построено точное решение нестационарных уравнений Навье–Стокса для трехмерного изобарического течения вязкой несжимаемой жидкости в горизонтальном слое с проницаемыми границами. Решение принадлежит классу Линя и обобщает известное стационарное решение на случай зависимости от времени.

Выведена замкнутая система уравнений для амплитудных функций $u(z, t)$, $V(z, t)$, $U(z, t)$. Показано, что вертикальная скорость постоянна и определяется режимом отсоса/вдува. Задача сведена к уравнению теплопроводности с потенциалом; решение представлено в виде рядов по собственным функциям задачи Штурма–Лиувилля с условиями Навье на нижней границе. Для компоненты U , возбуждаемой неоднородным источником, применен интеграл Дюамеля, что позволило точно учесть влияние нестационарного источника.

Проведен анализ для трех жидкостей (вода, масло SAE 30, глицерин) при варьировании Re_w , Ta , φ и α . Установлены режимы с развитыми противотечениями и застойными точками. Показано, что при отрицательном вертикальном числе Рейнольдса ($Re_w = -10$) время релаксации убывает с ростом параметра проскальзывания α , что качественно отличается от классического случая $Re_w = 0$.

Полученные результаты могут быть использованы в нескольких направлениях. Во-первых, найденное точное решение служит эталоном для верификации численных методов, предназначенных для моделирования нестационарных течений с проницаемыми границами и условиями частичного проскальзывания. Во-вторых, выявленная зависимость времени релаксации от α и знака Re_w (ускорение релаксации с ростом α при отсосе и замедление при вдуве) дает качественный критерий для управления переходными режимами в мембранных аппаратах и микрофлюидных устройствах. В-третьих, построенные карты противотечений и застойных точек позволяют оценить области возможного накопления частиц в слое, что важно для проектирования каталитических реакторов и фильтрационных систем с проницаемыми стенками.

Отметим также, что прямое экспериментальное воспроизведение рассмотренного течения сопряжено с известными трудностями (необходимость создания однородного вертикального потока через проницаемые границы и контролируемого неоднородного поля скорости на верхней поверхности слоя). Полученные аналитические решения могут служить

качественными ориентирами при планировании будущих экспериментальных исследований в этой области.

ЛИТЕРАТУРА

1. Аристов С. Н., Шварц К. Г. Вихревые течения адвективной природы во вращающемся слое жидкости. Пермь: Изд-во Перм. гос. ун-та, 2006. 155 с.
2. Аристов С. Н., Шварц К. Г. Эволюция ветровой циркуляции в неизотермическом океане // Океанология. 1990. Т. 30, № 4. С. 562–566.
3. Губарева К. В., Просвирыков Е. Ю., Еремин А. В. Аналитическое решение уравнений Навье–Стокса для описания неоднородного течения Куэтта с квадратичным профилем в слое с проницаемыми границами // Advanced Engineering Research (Rostov-on-Don). 2026. Т. 26, № 1. С. 22–42. doi: 10.23947/2687-1653-2026-26-1-2242
4. Полянин А. Д., Жуков А. И. Методы разделения переменных и точные решения нелинейных уравнений математической физики. М.: ИПМех РАН, 2020. 384 с.
5. Привалова В. В., Просвирыков Е. Ю. Нелинейное изобарическое течение вязкой несжимаемой жидкости в тонком слое с проницаемыми границами // Вычислительная механика сплошных сред. 2019. Т. 12, № 2. С. 230–242. doi: 10.7242/1999-6691/2019.12.2.20
6. Aristov S. N., Knyazev D. V., Polyinin A. D. Exact solutions of the Navier–Stokes equations with the linear dependence of velocity components on two space variables // Theoretical Foundations of Chemical Engineering. 2009. Vol. 43, no. 5. P. 642–662. doi: 10.1134/S0040579509050066
7. Aristov S. N., Prosviryakov E. Y. Unsteady layered vortical fluid flows // Fluid Dyn. 2016. Vol. 51, no. 2. P. 148–154. doi: 10.1134/S0015462816020034
8. Bogoyavlenskij O. The new effect of oscillations of the total angular momentum vector of viscous fluid // Phys. Fluids. 2022. Vol. 34. Art. 083108. doi: 10.1063/5.0101870
9. Burmasheva N. V., Prosviryakov E. Yu. Exact solutions to Navier–Stokes equations describing a gradient nonuniform unidirectional vertical vortex fluid flow // Dynamics. 2022. Vol. 2. P. 175–186. doi: 10.3390/dynamics2020009
10. Drazin P. G., Riley N. The Navier–Stokes equations: a classification of flows and exact solutions. Cambridge: Cambridge Univ. Press, 2006. 196 p.
11. Ershkov S., Burmasheva N., Leshchenko D. D., Prosviryakov E. Yu. Exact solutions of the Oberbeck–Boussinesq equations for the description of shear thermal diffusion of Newtonian fluid flows // Symmetry. 2023. Vol. 15, no. 9. Art. 1730. doi: 10.3390/sym15091730

12. Feng X., Lu X., He Y. Difference finite element method for the 3D steady Navier–Stokes equations // *SIAM J. Numer. Anal.* 2023. Vol. 61, no. 1. P. 167–193. doi: 10.1137/21M1450872
13. Itoh S., Tanaka N., Tani A. The initial value problem for the Navier–Stokes equations with general slip boundary conditions in Holder spaces // *J. Math. Fluid Mech.* 2003. Vol. 5. P. 275–301. doi: 10.1007/s00021-003-0074-6
14. Lin C. C. Note on a class of exact solutions in magneto-hydrodynamics // *Arch. Rational Mech. Anal.* 1957. Vol. 1. P. 391–395. doi: 10.1007/BF00298016
15. Neto C., Evans D. R., Bonaccorso E., Butt H.-J., Craig V. S. J. Boundary slip in Newtonian liquids: a review of experimental studies // *Rep. Progr. Phys.* 2005. Vol. 68. P. 2859–2897. doi: 10.1088/0034-4885/68/12/R05
16. Shtern V. Counterflows. Paradoxical fluid mechanics phenomena. Cambridge: Cambridge Univ. Press, 2012. 470 p. doi: 10.1017/CBO9781139226516
17. Temam R. Navier–Stokes equations: theory and numerical analysis. Amsterdam: North-Holland, 2001. 500 p.
18. Wang C. Y. Exact solutions of the steady-state Navier–Stokes equations // *Annu. Rev. Fluid Mech.* 1991. Vol. 23. P. 159–177. doi: 10.1146/annurev.fl.23.010191.001111
19. Wang C. Y. Exact solutions of the unsteady Navier–Stokes equations // *Appl. Mech. Rev.* 1989. Vol. 42, no. 11S. P. 269–282. doi: 10.1115/1.3152400

REFERENCES

1. Aristov S. N., Shvarts K. G. Vortical flows of advective nature in a rotating fluid layer. Perm: Perm State University; 2006. 155 p. (In Russ.)
2. Aristov S. N., Shvarts K. G. Evolution of wind circulation in a non-isothermal ocean. *Okeanologiya = Oceanology*. 1990;30(4):562–566. (In Russ.)
3. Gubareva K. V., Prosviryakov E. Yu., Eremin A. V. Analytical solution of the Navier–Stokes equations for describing an inhomogeneous Couette flow with a quadratic velocity profile in a layer with permeable boundaries. *Advanced Engineering Research (Rostov-on-Don)*. 2026;26(1):22–42. (In Russ.). doi: 10.23947/2687-1653-2026-26-1-2242
4. Polyanin A. D., Zhurov A. I. Separation of variables and exact solutions of nonlinear equations of mathematical physics. Moscow: IPMekh RAN; 2020. 384 p. (In Russ.)
5. Privalova V. V., Prosviryakov E. Yu. Nonlinear isobaric flow of a viscous incompressible fluid in a thin layer with permeable boundaries. *Vychislitel'naya mekhanika sploshnykh sred = Computational Continuum Mechanics*. 2019;12(2):230–242. (In Russ.). doi: 10.7242/1999-6691/2019.12.2.20

6. Aristov S. N., Knyazev D. V., Polyanin A. D. Exact solutions of the Navier–Stokes equations with the linear dependence of velocity components on two space variables. *Theoretical Foundations of Chemical Engineering*. 2009;43(5):642–662. doi: 10.1134/S0040579509050066
7. Aristov S. N., Prosviryakov E. Y. Unsteady layered vortical fluid flows. *Fluid Dyn.* 2016;51(2):148–154. doi: 10.1134/S0015462816020034
8. Bogoyavlenskij O. The new effect of oscillations of the total angular momentum vector of viscous fluid. *Phys. Fluids*. 2022;34:083108. doi: 10.1063/5.0101870
9. Burmasheva N. V., Prosviryakov E. Yu. Exact solutions to Navier–Stokes equations describing a gradient nonuniform unidirectional vertical vortex fluid flow. *Dynamics*. 2022;2:175–186. doi: 10.3390/dynamics2020009
10. Drazin P. G., Riley N. The Navier–Stokes equations: a classification of flows and exact solutions. Cambridge: Cambridge Univ. Press; 2006. 196 p.
11. Ershkov S., Burmasheva N., Leshchenko D. D., Prosviryakov E. Yu. Exact solutions of the Oberbeck–Boussinesq equations for the description of shear thermal diffusion of Newtonian fluid flows. *Symmetry*. 2023;15(9):1730. doi: 10.3390/sym15091730
12. Feng X., Lu X., He Y. Difference finite element method for the 3D steady Navier–Stokes equations. *SIAM J. Numer. Anal.* 2023;61(1):167–193. doi: 10.1137/21M1450872
13. Itoh S., Tanaka N., Tani A. The initial value problem for the Navier–Stokes equations with general slip boundary conditions in Holder spaces. *J. Math. Fluid Mech.* 2003;5:275–301. doi: 10.1007/s00021-003-0074-6
14. Lin C. C. Note on a class of exact solutions in magneto-hydrodynamics. *Arch. Rational Mech. Anal.* 1957;1:391–395. doi: 10.1007/BF00298016
15. Neto C., Evans D. R., Bonaccorso E., Butt H.-J., Craig V. S. J. Boundary slip in Newtonian liquids: a review of experimental studies. *Rep. Progr. Phys.* 2005;68:2859–2897. doi: 10.1088/0034-4885/68/12/R05
16. Shtern V. Counterflows. Paradoxical fluid mechanics phenomena. Cambridge: Cambridge University Press; 2012. 470 p. doi: 10.1017/CBO9781139226516
17. Temam R. Navier–Stokes equations: theory and numerical analysis. Amsterdam: North-Holland; 2001. 500 p.
18. Wang C. Y. Exact solutions of the steady-state Navier–Stokes equations. *Annu. Rev. Fluid Mech.* 1991;23:159–177. doi: 10.1146/annurev.fl.23.010191.001111
19. Wang C. Y. Exact solutions of the unsteady Navier–Stokes equations. *Appl. Mech. Rev.* 1989;42(11S):269–282. doi: 10.1115/1.3152400

Поступила в редакцию / received: 25.03.2026; принята к публикации / accepted: 29.04.2026.
Авторы заявляют об отсутствии конфликта интересов / The authors declare no conflict of interest.

СВЕДЕНИЯ ОБ АВТОРАХ:

Губарева Кристина Владимировна

канд. техн. наук, доцент кафедры «Промышленная теплоэнергетика»

e-mail: r.kristina2017@mail.ru

Просвиряков Евгений Юрьевич

д-р физ.-мат. наук, доцент, профессор кафедры «Информационные технологии и системы управления» УрФУ; заведующий сектором нелинейной вихревой гидродинамики ИМАШ

e-mail: evgen_pros@mail.ru

Еремин Антон Владимирович

д-р техн. наук, доцент, проректор по научной работе, заведующий кафедрой «Промышленная теплоэнергетика»

e-mail: a.v.eremin@list.ru

CONTRIBUTORS:

Gubareva, Kristina

Cand. Sci. (Eng.), Associate Professor

Prosviryakov, Evgenii

Dr. Sci. (Phys.-Math.), Associate Professor, Professor, UrFU; Sector Head, IES UB RAS

Eremin, Anton

Dr. Sci. (Eng.), Associate Professor, Vice-Rector for Science, Head of Department

УДК 519.6:539.2

ПАРАМЕТРИЧЕСКАЯ ИДЕНТИФИКАЦИЯ МОДЕЛИ ТЕРМОДЕСОРБЦИИ ВОДОРОДА МЕТОДОМ СОПРЯЖЕННЫХ УРАВНЕНИЙ

Ю. В. Заика, Е. К. Костикова*

*Институт прикладных математических исследований КарНЦ РАН,
ФИЦ «Карельский научный центр РАН» (ул. Пушкинская, 11, Петрозаводск,
Республика Карелия, Россия, 185910), *kostikova@krc.karelia.ru*

Представлен метод параметрической идентификации модели термодесорбции водорода из конструкционных материалов на основе помехоустойчивых интегральных операторов обработки измерений и теории сопряженных задач математической физики. Особенность модели заключается в том, что производная по времени входит не только в уравнение диффузии, но и в нелинейные граничные условия. В простейшем случае при преобразовании к эквивалентному интегро-дифференциальному уравнению это приводит к появлению производных дробного порядка. В общей терминологии теории динамических систем это содержательный пример системы с памятью. В физических терминах: поверхностная и объемная концентрации «помнят» свою предысторию.

Ключевые слова: термодесорбция водорода; нелинейные краевые задачи; динамические граничные условия; сопряженные уравнения

Для цитирования: Заика Ю. В., Костикова Е. К. Параметрическая идентификация модели термодесорбции водорода методом сопряженных уравнений // Труды Карельского научного центра РАН. 2026. № 6. С. 32–46. doi: 10.17076/mat2381

Финансирование. Финансовое обеспечение исследований осуществлялось из средств федерального бюджета на выполнение государственного задания КарНЦ РАН (Институт прикладных математических исследований КарНЦ РАН).

Yu. V. Zaika, E. K. Kostikova*. PARAMETRIC IDENTIFICATION OF A HYDROGEN THERMAL DESORPTION MODEL BY THE CONJUGATE EQUATIONS METHOD

*Institute of Applied Mathematical Research, Karelian Research Centre, Russian Academy of Sciences (11 Pushkinskaya St., 185910 Petrozavodsk, Karelia, Russia), *kostikova@krc.karelia.ru*

The paper presents a method for parametric identification of a model of hydrogen thermal desorption from structural materials based on noise-resistant integral measurement processing operators and the theory of conjugate problems of mathematical physics. The special feature of the model is that the time derivative is included not only in the diffusion equation, but also in the nonlinear

boundary conditions. In the simplest case, when transformed into an equivalent integro-differential equation, this leads to the emergence of fractional derivatives. In the general terminology of the dynamical systems theory, this is a meaningful example of a memory system. In physical terms: concentrations on the surface and in the bulk ‘remember’ their previous history.

Key words: hydrogen thermal desorption; nonlinear boundary value problems; dynamic boundary conditions; conjugate equations

For citation: Zaika Yu. V., Kostikova E. K. Parametric identification of hydrogen thermal desorption model by the conjugate equations method. *Trudy Karelskogo nauchnogo tsentra RAN = Transactions of the Karelian Research Centre RAS*. 2026. No. 6. P. 32–46. doi: 10.17076/mat2381

Funding. The studies were funded from the federal budget through state assignment to the Karelian Research Centre RAS (Institute of Applied Mathematical Research, Karelian Research Centre RAS).

ВВЕДЕНИЕ

Вычислительное материаловедение играет всевозрастающую роль в разработке технологий водородной энергетики, а в отдаленной перспективе и термоядерной (в частности, реализуется проект ITER). Одним из эффективных методов исследования взаимодействия конструкционных материалов с изотопами водорода является термодесорбционная спектроскопия (ТДС). Образец (рассматриваем тонкую пластину из материала с металлическими свойствами), предварительно насыщенный растворенным водородом, медленно нагревается в вакуумной камере. С помощью масс-спектрометра регистрируется поток дегазации. Анализ подлечит ТДС-спектр – зависимость плотности десорбционного потока от текущей температуры. Но эксперименты по водородопроницаемости проводятся на тонких мембранах, а выводы необходимо обобщать на «стенки защитных конструкций». Помимо численных методов решения соответствующих краевых задач, необходимы помехоустойчивые алгоритмы решения обратных задач параметрической идентификации по текущим экспериментальным данным. Методики оценки диффузионных параметров хорошо разработаны. Гораздо труднее оценить поверхностные параметры, которые зависят и от конкретных условий эксплуатации материала.

В основе данной работы – экспериментальные результаты и теоретические исследования, изложенные в монографиях [1, 2, 5–8].

МОДЕЛЬ ТДС-ДЕГАЗАЦИИ И СОПРЯЖЕННЫЕ УРАВНЕНИЯ

Рассмотрим два варианта модели термодесорбции из образца конструкционного матери-

ала (с металлическими свойствами) в форме пластины, которая предварительно насыщена растворенным атомарным водородом. Модель с *поверхностной десорбцией* (ПД):

$$\frac{\partial c}{\partial t} = D(t) \frac{\partial^2 c}{\partial x^2} - a_1(t)c + a_2(t)z,$$

$$(t, x) \in Q_* = (0, t_*) \times (0, \ell),$$

$$\frac{\partial z}{\partial t} = a_1(t)c - a_2(t)z, \quad a_1(0)\bar{c} = a_2(0)\bar{z},$$

$$c(0, x) = \bar{c}, \quad z(0, x) = \bar{z},$$

$$c(t, x) = c(t, \ell - x), \quad c_{0,\ell}(t) = g(t)q(t),$$

$$\frac{dq}{dt} = \mu s(t)p(t) - J(t) + D(t) \frac{\partial c}{\partial x} \Big|_{x=0},$$

$$J(t) = b(t)q^2(t), \quad t \in [0, t_*].$$

Здесь $c(t, x)$ – концентрация растворенного атомарного водорода, $c_{0,\ell} = c|_{x=0,\ell}$; z – концентрация захваченных в дефекты (ловушки) атомов H; a_1 и a_2 – коэффициенты поглощения и выделения водорода ловушками; J – моделируемая плотность десорбционного потока водорода (в молекулярной форме) с поверхности. Для упрощения обозначений в зависимости от контекста записываем, например, $D(T)$ или $D(t)$, подразумевая $T = T(t)$. Для определенности считаем зависимость от температуры аррениусовской, включая формально коэффициент быстрого растворения g (но без ограничения знака «энергии активации» E_g). При необходимости можно учесть слабую зависимость газокинетического коэффициента $\mu(T) \propto 1/\sqrt{T}$, «включая» ее в $s(T)$. Если молекулу H_2 («налетающую» на поверхность частицу) рассматривать как пару атомов H, то для единообразия подсчета в атомах лучше написать $2s$ вместо s . Тогда s будет безразмерным параметром (доля адсорби-

ровавшихся атомов). Но для задачи оценивания значений (s или $2s$) это непринципиально.

Измеряемое давление $p(t)$ связано с плотностью десорбционного потока $J(t)$ интегральным уравнением первого рода (с параметрами конкретной вакуумной установки). Полагаем, что вспомогательная задача $p \mapsto J$ решена предварительно и входными данными для обратной задачи являются функции $p(t)$, $J(t)$.

Отличие модели с *объемной десорбцией* (ОД) в граничном условии:

$$\mu s(t)p(t) - J(t) = -D(t)c_x(t, 0), \quad J(t) = b(t)c_0^2(t).$$

Таким образом, в выкладках для ПД-задачи следует формально полагать $g = 1$, $\dot{q} = 0$ и заменять концентрацию $q(t)$ на $c_0(t) = c_\ell(t)$. Коэффициенты десорбции (эффективной комбинации) в обоих случаях обозначаем b .

До начала ТДС-эксперимента пластину насыщают водородом при постоянной температуре $\bar{T}$ и постоянном достаточно высоком давлении $\bar{p}$. После окончания переходных процессов получаем соотношения

$$\mu s \bar{p} = b \bar{q}^2, \quad g \bar{q} = \bar{c}, \quad a_1 \bar{c} = a_2 \bar{z} \quad (T = \bar{T}),$$

где $\bar{c}$, $\bar{z}$ – установившиеся концентрации диффундирующего (атомарного) и захваченного ловушками водорода. Константы $\bar{c}$, $\bar{z}$ априори неизвестны. После быстрого охлаждения образца (отключения тока нагрева) и откачки газа из камеры происходит повторный нагрев пластины по монотонному (обычно линейному) закону $T(t) \in [T^-, T^+]$. Время окончания эксперимента t_* определяется практически полной дегазацией: $J(t) \approx 0$, $t \geq t_*$.

Одним из эффективных подходов к решению обратных задач математической физики является метод сопряженных уравнений. Адаптируем общую схему метода к рассматриваемому классу задач параметрической идентификации. Пусть заданы уравнение $c_t = \mathcal{A}c$ в области $Q_* = (0, t_*) \times (0, \ell)$ и краевые условия, где $\mathcal{A}$ – интегро-дифференциальный оператор. Считаем $\mathcal{A}$ линейным, а граничные условия в общем случае нелинейны. Используя интегрирование по частям, для произвольной гладкой функции $\psi(t, x)$ получим

$$\begin{aligned} 0 &= \int_0^{t_*} \int_0^\ell \psi(t, x) [c_t - \mathcal{A}c] dx dt \\ &= \int_0^{t_*} \int_0^\ell c(t, x) [\psi_t - \mathcal{A}^* \psi] dx dt + f. \end{aligned}$$

Слагаемое f содержит величины, связанные с краевыми условиями. Если фиксировать решение $\psi(t, x)$ сопряженного уравнения $\psi_t =$

$\mathcal{A}^* \psi$, то получим соотношение $f = 0$, содержащее искомые параметры и экспериментальные данные. Для сопряженного уравнения не фиксируются краевые условия. Выбирая различные «простые» решения $\psi = \psi_i(t, x)$ из бесконечного их числа, получаем систему уравнений $f_i = 0$ для оценивания $D, \dots, s$.

Уравнения для оценки параметров

Вначале приведем выкладки для общего случая. Когда дефектами можно пренебречь ($a_i = 0$) и ставится задача оценки основных параметров $\{D, b, g, s\}$, выражения упрощаются. Проинтегрируем явно линейное уравнение для $z(t, x)$ и подставим полученное выражение в уравнение диффузии. Используя интегрирование по частям, для достаточно гладкой функции $\psi(t, x)$ получим

$$\begin{aligned} 0 &= \int_0^{t_*} \int_0^\ell \psi(t, x) \\ &\times \left[\frac{\partial c}{\partial t} - D(t) \frac{\partial^2 c}{\partial x^2} + a_1(t)c(t, x) - \dots \right] dx dt = \dots \\ &= -\bar{z} \int_0^{t_*} \left[\int_0^\ell \psi(t, x) dx \right] a_2(t) \exp \left\{ -\int_0^t a_2 ds \right\} dt \\ &- \int_0^{t_*} \int_0^\ell c(t, x) \left[\frac{\partial \psi}{\partial t} + D(t) \frac{\partial^2 \psi}{\partial x^2} - a_1(t)\psi(t, x) \right. \\ &+ a_1(t) \int_t^{t_*} a_2(\tau) \exp \left\{ \int_\tau^t a_2(s) ds \right\} \psi(\tau, x) d\tau \left. \right] dx dt \\ &+ \int_0^{t_*} [\psi(t, 0) + \psi(t, \ell)] D \frac{\partial c}{\partial x} \Big|_{x=0} dt \\ &+ \int_0^{t_*} c_0(t) D \frac{\partial \psi}{\partial x} \Big|_0^\ell dt - \bar{c} \int_0^\ell \psi(0, x) dx. \quad (1) \end{aligned}$$

С целью исключения двойного интеграла подчиним функцию $\psi(t, x)$ сопряженному линейному однородному уравнению

$$\begin{aligned} \frac{\partial \psi}{\partial t} &= -D \frac{\partial^2 \psi}{\partial x^2} + a_1 \psi \\ &- a_1 \int_t^{t_*} a_2(\tau) \exp \left\{ \int_\tau^t a_2 ds \right\} \psi(\tau, x) d\tau. \quad (2) \end{aligned}$$

Подставим в соотношение (1) с учетом уравнения (2) выражение концентрации $c_0(t)$ через плотность десорбции $J = b[c_0/g]^2$ и выражение плотности диффузионного потока $Dc_x(t, 0)$ из граничного условия:

$$\begin{aligned} &- \bar{z} \int_0^{t_*} \langle \dots \rangle dt - \bar{c} \int_0^\ell \psi(0, x) dx \\ &+ \int_0^{t_*} g \sqrt{b^{-1} J D} \frac{\partial \psi}{\partial x} \Big|_0^\ell dt + \int_0^{t_*} \{ \dot{q} - \mu s p + b q^2 \} \\ &\times [\psi(t, \ell) + \psi(t, 0)] dt = 0 \Rightarrow -\bar{z} \dots \end{aligned}$$

$$\begin{aligned}
& -\bar{c} \dots + \int_0^{t_*} \dots dt + q(t)[\psi(t, \ell) + \psi(t, 0)]|_0^{t_*} \\
& + \int_0^{t_*} \{J - \mu sp\}[\psi(t, \ell) + \psi(t, 0)] dt \\
& - \int_0^{t_*} q(t)[\dot{\psi}(t, \ell) + \dot{\psi}(t, 0)] dt = 0.
\end{aligned}$$

Поверхностная концентрация $q(t)$ исключается в силу

$$bq^2 = J, \quad g(T_0)q(0) = \bar{c}, \quad T_0 \equiv T(0), \quad q(t_*) = 0.$$

Кроме того, следует подставить $\bar{c}$, $\bar{z}$ с учетом начального насыщения:

$$b\bar{c}^2 g^{-2} = \mu s \bar{p} \mapsto \bar{c}, \quad a_2 \bar{z} = a_1 \bar{c} \mapsto \bar{z} \quad (T = \bar{T}).$$

Несмотря на внешнюю громоздкость, имеем существенное преимущество перед исходной системой дифференциальных связей. После выбора «пробных» $\psi(t, x)$ и подстановки

$$D(t) = D_0 u^{E_D}, \dots, \quad s(t) = s_0 u^{E_s},$$

$u(t) \equiv \exp\{-1/[RT(t)]\}$, получаем уравнения вида

$$f(D_0, E_D, \dots, s_0, E_s) = 0.$$

Давление $p(t)$ и плотность десорбции $J(t)$, определяемые по результатам измерений, входят под знаком интеграла, что важно с вычислительной точки зрения. Варьируя условия насыщения $\bar{p}$, $\bar{T}$, «управление» $u(t)$ (т. е. температуру $T(t)$ на отрезке времени $[0, t_*]$) и выбирая различные решения $\psi(t, x)$ линейного сопряженного уравнения (2) без краевых условий, получаем необходимое количество уравнений для оценивания параметров $D_0, \dots, E_s$.

Выбор сопряженных функций

1. Положим $\psi(t, x) = \beta(t)x$. Подставляя в сопряженное уравнение (2), получим

$$\begin{aligned}
\dot{\beta}(t) &= a_1(t)\beta(t) \\
& - a_1(t) \int_t^{t_*} a_2(\tau) \exp\left\{\int_\tau^t a_2(s) ds\right\} \beta(\tau) d\tau.
\end{aligned}$$

Обозначая $\gamma(t) = -\int_t^{t_*} \langle \dots \rangle d\tau$, перейдем к линейной системе

$$\begin{aligned}
\dot{\beta}(t) &= a_1(t)[\beta(t) + \gamma(t)], \\
\dot{\gamma}(t) &= a_2(t)[\beta(t) + \gamma(t)], \quad \gamma(t_*) = 0.
\end{aligned}$$

Отсюда определяем функцию $\beta(t)$:

$$\begin{aligned}
\frac{d}{dt}(\beta + \gamma) &= [a_1 + a_2](\beta + \gamma), \quad \dot{\beta} = a_1[\beta + \gamma], \\
\beta(t) + \gamma(t) &= \beta(t_*) \exp\left\{-\int_t^{t_*} (a_1 + a_2) ds\right\}, \\
\beta(t) &= \beta(t_*) \left[1 - \int_t^{t_*} a_1(\tau) \right. \\
& \quad \times \exp\left\{-\int_\tau^{t_*} [a_1(s) + a_2(s)] ds\right\} d\tau \Big].
\end{aligned}$$

С учетом однородности по $\psi(t, x)$ примем нормировку $\beta(t_*) = 1$. Далее нетрудно конкретизировать уравнение $f(D_0, \dots, E_s) = 0$ (запись $f(x)|_a^b$ означает разность $f(b) - f(a)$).

Фактически в уравнении отсутствуют диффузионные параметры D_0, E_D при таком выборе функции ψ . Коэффициент диффузии $D(T)$ может быть известен из независимых экспериментов, когда условия таковы, что диффузия фактически является единственным лимитирующим фактором. При $T(t) = T_e = \text{const}$ (если после вакуумирования осуществить скачкообразный повторный нагрев до температуры экстракции T_e) соотношения упрощаются. Например, для $T_e = \bar{T}$ получаем

$$\begin{aligned}
\beta(t) &= [a_1 + a_2]^{-1} \left[a_2 + a_1 \right. \\
& \quad \times \exp\left\{(a_1 + a_2)(t - t_*)\right\} \Big], \quad a_i = a_i(\bar{T}), \\
f &\equiv \sqrt{\mu s \bar{p}} g b^{-1/2} \left\{ [\ell g + 2][2g]^{-1} \beta(0) \right. \\
& \quad + \ell a_1 [2(a_1 + a_2)]^{-1} [1 - \exp\{-(a_1 + a_2)t_*\}] \Big\} \\
& \quad + \int_0^{t_*} \left\{ (\mu s p - J)\beta + \sqrt{b^{-1}J} \dot{\beta} \right\} dt = 0.
\end{aligned}$$

Скачкообразный нагрев чреват возникновением существенного градиента температуры в образце и необходимостью коррекции диффузионного уравнения. Целесообразно модифицировать эксперимент: после насыщения образца водородом следует, не изменяя температуру $\bar{T}$, сбросить давление с помощью дополнительной вакуумируемой емкости. При необходимости температуры насыщения и дегазации могут быть различными, изменения в выкладках не принципиальны. Нестационарный эксперимент с медленным нагревом более информативен.

2. Рассмотрим вариант $\psi(t, x) = \beta(t) \sin \sigma x (\beta(t) \cos \sigma x)$:

$$\dot{\beta}(t) = [D(t)\sigma^2 + a_1(t)]\beta(t) - a_1(t) \int_t^{t_*} a_2(\tau) \exp \left\{ \int_\tau^t a_2(s) ds \right\} \beta(\tau) d\tau.$$

Определив $\gamma(t) = -\int_t^{t_*} \langle \dots \rangle d\tau$, преобразуем уравнение в систему

$$\dot{\beta} = [D\sigma^2 + a_1]\beta + a_1\gamma, \quad \dot{\gamma} = a_2[\beta + \gamma], \quad \gamma(t_*) = 0.$$

Если дегазацию проводить при постоянной температуре $T(t) = \bar{T}$, то можно выписать явные аналитические выражения. Характеристический полином системы при $D = D(\bar{T})$, $a_i = a_i(\bar{T})$ равен $\lambda^2 - (a_1 + a_2 + D\sigma^2)\lambda + D\sigma^2 a_2$. Дискриминант $(a_1 + a_2 + D\sigma^2)^2 - 4D\sigma^2 a_2 = 4a_1 a_2 + (a_1 - a_2 + D\sigma^2)^2$ положителен, имеется два различных положительных корня:

$$2\lambda_{1,2} = a_1 + a_2 + D\sigma^2 \mp [4a_1 a_2 + (a_1 - a_2 + D\sigma^2)^2]^{1/2}.$$

Следовательно, с точностью до множителя

$$\beta(t) = (\lambda_2 - \lambda_1)^{-1} \left[(a_2 - \lambda_1) \exp\{\lambda_1(t - t_*)\} - (a_2 - \lambda_2) \exp\{\lambda_2(t - t_*)\} \right].$$

Параметр σ можно брать равным $m\pi/(2\ell)$.

3. Аналогично для $\psi = \beta(t) \exp \sigma x$ и варианта $T(t) = \bar{T}$. Функция $\beta(t)$ определяется теми же формулами с заменой D на $-D$.

Итак, способами 1–3 можно получить любое количество уравнений $f = 0$, варьируя σ . Это количество можно «удвоить», рассматривая отрезок $[0, \ell_0]$, $\ell_0 = \ell/2$, и граничное условие $c_x(t, \ell_0) = 0$ в силу симметрии концентрации в пластине. Типовые выкладки не приводим. Отметим лишь, что появится произведение $\psi_x(t, \ell_0)c(t, \ell_0)$, в котором о концентрации $c(t, \ell_0)$ информации нет. Это влечет дополнительное ограничение $\psi_x(t, \ell_0) = 0$.

Интеграл от функции $\psi(t_*, x)c(t_*, x)$ полагался равным нулю в силу $c(t, x) \rightarrow 0$ с ростом времени t . Тем самым неявно речь шла только о тех решениях сопряженного уравнения, которые, если и растут с увеличением времени, то относительно медленно. Иначе следует аппроксимировать распределение $c(t_*, x)$.

ИДЕНТИФИКАЦИЯ БЕЗ УЧЕТА ДЕФЕКТОВ

Будем считать обратимый захват и сорбцию второстепенными факторами ($a_i \approx 0$, $\mu sr \approx 0$). Параметры a_i можно оценить после

того, как найдено удовлетворительное согласие с экспериментом выбором D , g , b , s .

В ОД-модели граничное условие $Dc_x(t, 0) = bc_0^2(t)$ в пределе ($t \rightarrow +0$) формально противоречит равномерному начальному распределению $c(0, x) = \varphi(x) = \bar{c}$, $x \in [0, \ell]$. На самом деле в процессе вакуумирования перед ТДС-дегазацией функция $\varphi(x)$ успевает немного «прогнуться». В алгоритмах идентификации $\varphi(x)$ будет использоваться только интегрально (в соответствии с обобщенным решением), искажением можно пренебречь.

Объемная десорбция

Граничные условия $Dc_x|_{0,\ell} = \pm bc_{0,\ell}^2$ нелинейны. Но по постановке обратной задачи функция $J(t) = bc_{0,\ell}^2$ известна. Поэтому формально можем рассмотреть линейную краевую задачу II рода, заменив $bc_{0,\ell}^2$ на плотность десорбции $J(t)$. Сделаем замену времени $t' = \int_0^t D(s) ds$ и после преобразований оставим прежнее обозначение t :

$$\frac{\partial c}{\partial t} = \frac{\partial^2 c}{\partial x^2}, \quad \frac{\partial c}{\partial x} \Big|_{x=0} = \tilde{J}(t), \quad \frac{\partial c}{\partial x} \Big|_{\ell} = -\tilde{J}(t),$$

$$\tilde{J}(t) = D^{-1}(t)J(t), \quad J(t) = b(t)c_0^2(t) = b(t)c_\ell^2(t).$$

Можно было бы сразу перейти к безразмерному показателю времени диффузии t'/ℓ^2 и нормировкам x/ℓ , $c/\bar{c}$, но в дальнейших выкладках это не является принципиальным.

По постановке ТДС-эксперимента в режиме «быстрое охлаждение – медленный нагрев» имеем $J(0) \approx 0$ ($J(0) \ll J_{\max}$). Решение краевой задачи представимо в форме

$$c(t, x) = \bar{c} - \int_0^t \tilde{J}(\tau) \{G(x, t; 0, \tau) + G(x, t; \ell, \tau)\} d\tau,$$

$$G(x, t; y, \tau) = \frac{1}{\ell}$$

$$+ \frac{2}{\ell} \sum_{n=1}^{\infty} \exp \left\{ -\frac{n^2 \pi^2}{\ell^2} (\tau - t) \right\} \cos \frac{n\pi x}{\ell} \cos \frac{n\pi y}{\ell},$$

где G – функция Грина. Представление $c_0(t)$:

$$c_0(t) = \bar{c} - \int_0^t \tilde{J}(\tau) K(t - \tau) d\tau,$$

$$K(s) = \frac{2}{\ell} \left[1 + 2 \sum_{n=2m}^{\infty} \exp \left\{ -\frac{n^2 \pi^2}{\ell^2} s \right\} \right].$$

В исходном времени аргумент $t - \tau$ заменится интегралом $\int_\tau^t D(s) ds$, а вместо $\tilde{J}(\tau)$ будет $J(\tau)$. Выражение $c_0(t)$ следует подставить в равенство $J(t) = b(t)c_0^2(t)$ ($\sqrt{J} = c_0\sqrt{b}$). Функция $J(t)$ известна, поэтому получаем семейство уравнений для оценки параметров:

$$\Phi(t; D_0, E_D, b_0, E_b) = 0, \quad t \in [t_1, t_2] \subset (0, t_*).$$

Та же зависимость с учетом $c_x(t, \ell_0) = 0$ получается для $[0, \ell_0]$ ($\ell_0 = \ell/2$ – центр симметрии). Имеется в виду «эквивалентная» задача

$$c_t = Dc_{xx}, \quad x \in (0, \ell_0), \quad c(0, x) = \varphi(x) = \bar{c}, \\ Dc_x(t, 0) = bc_0^2, \quad c_x(t, \ell_0) = 0.$$

Формально можно считать, что обратная задача параметрической идентификации существенно упрощена: осталось подобрать константы в явной формуле. Но D входит в выражение под символами интеграла и ряда, что делает этот этап нетривиальным. Ограничиваясь частичной суммой ряда, можно вычислять интегралы типа свертки

$$J_n(t) = \int_0^t J(\tau) \exp \left\{ -\frac{n^2 \pi^2}{\ell^2} \int_\tau^t D(s) ds \right\} d\tau = O(n^{-2})$$

($n = 2m, m \in \mathbb{N}$) как решения линейных задач

$$\dot{J}_n(t) = -\frac{n^2 \pi^2}{\ell^2} D(t) J_n(t) + J(t), \quad J_n(0) = 0.$$

При численной реализации уравнения следует нормировать на $J_{\max}$. Чтобы не оперировать частичной суммой с большим числом слагаемых, для малых s целесообразно использовать другое представление ряда:

$$1 + 2 \sum_{n=1}^{\infty} \exp \{ -n^2 \pi^2 \theta \} = \frac{1}{\sqrt{\pi \theta}} \sum_{-\infty}^{\infty} \exp \left\{ -\frac{n^2}{\theta} \right\}, \\ \theta = \frac{4}{\ell^2} \int_\tau^t D(\xi) d\xi.$$

Если равновесная концентрация $\bar{c}$ известна, а поток регистрируется в относительных единицах ($J(t)/J_{\max}$), то плотность термодесорбции $J_{\max}$ находится из материального баланса:

$$\bar{c}\ell = 2 \int_0^{t_*} J(\tau) d\tau \Rightarrow 2 \int_0^{t_*} J(\tau) J_{\max}^{-1} d\tau = \bar{c}\ell J_{\max}^{-1}.$$

Определение коэффициента десорбции

Рассмотренное выше параметрическое тождество $\Phi(t; D_0, E_D, b_0, E_b) \equiv 0$ позволяет составлять любое количество уравнений $f(D_0, E_D, b_0, E_b) = 0$. Оценка значений D_0, E_D, b_0, E_b по одной экспериментальной кривой возможна, но требует решения системы уравнений численными методами. Поставим цель: ценою дополнительной экспериментальной информации свести вычисления к минимуму. Это повысит и точность оценивания.

Будем исходить из того, что техника определения равновесных концентраций в диапазоне $\bar{c} = \bar{c}(\bar{p}, T) \propto \sqrt{\bar{p}}$ (закон Сиверта) отработана. Коэффициент диффузии при фиксированной температуре достаточно надежно определяется методом проницаемости, когда за счет большого перепада давлений на входной и выходной сторонах мембраны при относительно высоких температурах и не слишком малых ℓ достигается диффузионный режим проницаемости. По времени запаздывания t_0 (точка пересечения с осью t асимптоты графика количества N , проникшего сквозь мембрану) находят $D = \ell^2/(6t_0)$. При этом не обязательно достижение входной концентрации $c_0 \approx \bar{c}$ – главное, чтобы переходные процессы на входе относительно быстро приводили к установлению $c_0(t) \approx c_0 = \text{const}, t > \varepsilon$ ($\varepsilon \ll t_*$). Неявно предполагается, что этот начальный всплеск $c_0(t)$ существенно не изменит начальное распределение $\varphi(x) = 0$ (хотя в модели входной поток $-Dc_x(t, 0)$ формально не ограничен). Варьирование температуры образца позволяет оценить D_0, E_D .

Гораздо сложнее определять параметры поверхностных процессов, поскольку SLR (surface limited regime) характерен для невысоких температур и давлений. При этом градиент концентрации мал, модель упрощается (в классе ОДУ), но падает точность измерений. Самое интересное (окрестность пика потока) происходит на этапе активного «соизмеримого» взаимодействия диффузии и десорбции. Чтобы повысить точность оценивания и упростить математическое обеспечение, считаем, что помимо ТДС-кривой $J(t)$ известны равновесная концентрация $\bar{c}$ (при заданных условиях насыщения $\bar{p}, \bar{T}$) и температурная зависимость $D = D(T)$ (в аррениусовском случае D_0, E_D). Необязательно, что это потребует дополнительных экспериментов. Параметры поверхностных процессов существенно зависят от внешних условий: окислы, примеси, шероховатость... В этом одна из причин разброса оценок. Объемные же параметры $\bar{c}, D$ «старого» материала известны и отражены в справочной литературе. Изменение свойств поверхности может быть и целенаправленным: например, напыление защитного (микро)слоя.

По заданному закону нагрева $T = T(t)$ определим функции

$$\alpha(t) = \exp \left\{ \frac{\pi^2}{\ell^2} \int_0^t D(T(s)) ds \right\},$$

$$J_n(t) = \alpha^{-n^2}(t) \int_0^t J(\tau) \alpha^{n^2}(\tau) d\tau, \quad n = 2m.$$

При достаточном количестве J_n

$$c_0(t) \approx \bar{c} - 2\ell^{-1}S(t) - 4\ell^{-1}\{J_2(t) + \dots + J_{2k}(t)\},$$

$$S(t) \equiv \int_0^t J(\tau) d\tau.$$

Алгоритм идентификации модели сводится к подбору b_0, E_b из условия $J(t) = b_0 \exp\{-E_b/[RT]\}c_0^2(t)$, $t \in [t_1, t_2]$. Ориентируясь на значения E_b в несколько десятков кДж, представим экспоненту в форме $u^\nu(t)$, $u(t) = \exp\{-10^4/[RT(t)]\}$ ($E_b = 10\nu$ кДж). После логарифмирования приходим к линейному по $\ln b_0$ и ν соотношению:

$$\Lambda(t) \equiv \ln \{J(t)c_0^{-2}(t)\}$$

$$= \ln b_0 + \nu \ln u(t) = \ln b_0 - \nu 10^4[RT(t)]^{-1}.$$

На плоскости $\{T^{-1}, \Lambda\}$ имеем отрезок прямой, по пересечениям которой с осями координат находим значения $\ln b_0, \nu$ (b_0, E_b).

Сопряженные уравнения

Подробнее остановимся на эксперименте, когда можно считать $\varphi(x) \approx \bar{c}$, $J(0) \approx 0$ ($J(0) \ll J_{\max}$): насыщение, быстрое охлаждение, вакуумирование, медленный нагрев. Обычно $T(t) = T_0 + vt$ ($T(t_m) = T_{\max} \Rightarrow v = 0, t \geq t_m$). В интегральном измерении несущественно рассогласование при $t \rightarrow +0$ равномерного насыщения с граничным условием $Dc_x(t, 0) = b(t)c_0^2(t)$ ($J = bc_0^2$).

Соотношение вида (1), получаемое преобразованием по частям двойного интеграла от $\psi[c_t - Dc_{xx}]$, после подстановки решения уравнения $\psi_t = -D(t)\psi_{xx}$ упрощается:

$$0 = -\bar{c} \int_0^\ell \psi(0, x) dx + \int_0^{t_*} c_0 D \frac{\partial \psi}{\partial x} \Big|_0 dt$$

$$+ \int_0^{t_*} [\psi(t, 0) + \psi(t, \ell)] D \frac{\partial c}{\partial x} \Big|_0 dt.$$

Подразумеваем $(\psi c)|_{t_*} \approx 0$, т. е. «пробная» функция ψ если и растет, то значительно медленнее, чем убывает концентрация на этапе остаточной дегазации. Варианты $\psi = 1$, $\psi = x$ дают материальный баланс. Для $\psi(t, x) = \beta(t) \exp \sigma x$ с учетом $J = bc_0^2$ получаем

$$X[1 + \exp \sigma \ell] + Y\sigma[\exp \sigma \ell - 1] = \frac{\bar{c}}{\sigma}[\exp \sigma \ell - 1],$$

$$\beta(t) = \exp \left\{ -\sigma^2 \int_0^t D d\tau \right\}, \quad \bar{c}\ell = 2 \int_0^{t_*} J(\tau) d\tau,$$

$$X \equiv \int_0^{t_*} \beta J dt, \quad Y \equiv \int_0^{t_*} \beta D \sqrt{b^{-1}} J dt.$$

Параметр σ варьируем по порядку $\sigma \ell \sim 1$. Это касается как функции $\exp\{\sigma x\}$, так и $\beta(t)$. В случае $\sigma = 1/\ell$ в аргументе экспоненты $\beta(t)$ – безразмерное характеристическое «время диффузии». Замена σ на $-\sigma$ не меняет функцию $\beta(t)$, значения X, Y и уравнение с точностью до множителя. При предельном переходе $\sigma \rightarrow 0$ получаем $\psi = 1$.

Аналогично получаем линейные по X, Y уравнения для $\psi = \beta(t) \sin \sigma x$ ($\cos$),

$$\beta(t) = \exp \left\{ \sigma^2 \int_0^t D(\tau) d\tau \right\}.$$

Здесь уже функция $\beta(t)$ растет, поэтому следует проанализировать условие равенства нулю интеграла от $\psi(t_*, x)c(t_*, x)$ по $x \in [0, \ell]$. В силу симметричности $c(t, x)$ относительно $x = \ell_0 \equiv \ell/2$ этого можно добиться, например, выбором $\sigma \ell = 2\pi$ (для $\sin$) и $\sigma \ell = \pi$ (для $\cos$). Но при этом уравнения вырождаются. По сходным мотивам не рассматриваем краевую задачу на отрезке $[0, \ell_0]$ с условием симметрии $c_x(t, \ell_0) = 0$ и ограничение $\psi_x(t, \ell_0) = 0$. Следует воспользоваться приближением распределения $c(t_*, x)$, например, параболическим по x выражением $\tilde{c}(t_*, x) = \tilde{c}(t_*, x; D, b)$ (см. далее).

Подытожим вышеизложенное схемой оценки коэффициентов модели $D(T), b(T), s(T)$ ($D_0, E_D, b_0, E_b, s_0, E_s$). Как минимум нужно два эксперимента с различными температурами насыщения $\bar{T}$, чтобы при известном коэффициенте десорбции $b(T)$ из соотношений

$$\bar{c}\ell = 2 \int_0^{t_*} J(\tau) d\tau, \quad \mu s \bar{p} = b \bar{c}^2 \quad (T = \bar{T})$$

определялись параметры s_0 и E_s . Желательно варьировать давление $\bar{p}$, чтобы значение $\bar{c}$ существенно изменилось, и скорость нагрева v . Для оценки величин D_0, E_D, b_0, E_b берем по два уравнения ($\sigma = \sigma_1, \sigma_2$) вида

$$f \equiv \frac{\sigma \varkappa}{\bar{c}} X + \frac{\sigma^2}{\bar{c}} Y - 1 = 0,$$

$$\varkappa = \frac{\exp \sigma \ell + 1}{\exp \sigma \ell - 1} \quad (\psi = \beta(t) \exp \sigma x).$$

Заметим, что функция $\beta(t)$ зависит только от D_0, E_D, σ . Параметрами эксперимента являются $\bar{p}, \bar{T}, v, T_0, T_{\max}$, а значения σ задаются численно. Во избежание плохой обусловленности системы можно добавить семейство уравнений $f'_\sigma = 0$. Производные выражений X, Y по переменной σ вычисляются под знаком интеграла. Из-за большого разброса по-

рядков величин перейдем к комплексам переменных и параметров, например:

$$\begin{aligned}\hat{f} &\equiv \sigma \ell \varkappa \hat{X} + \hat{Y} - 1 = 0, \\ \hat{X} &= \frac{1}{t_*} \int_0^{t_*} \hat{J}(t) \beta(t) dt, \quad \hat{Y} = \frac{1}{t_*} \int_0^{t_*} \beta \hat{D} \sqrt{\hat{b}^{-1}} \hat{J} dt, \\ \hat{J} &= \frac{J t_*}{\bar{c} \ell}, \quad \hat{b} = \frac{b \bar{c} t_*}{\ell}, \quad \hat{D} = D \sigma^2 t_*, \quad \sigma \ell \sim 1.\end{aligned}$$

Далее можно следовать МНК:

$$F(D_0, E_D, b_0, E_b) = \sum \alpha_i \hat{f}_i^2 \rightarrow \min, \quad \alpha_i > 0.$$

РЕЗУЛЬТАТЫ ВЫЧИСЛИТЕЛЬНЫХ ЭКСПЕРИМЕНТОВ

Тестирование «ступенчатой» схемы идентификации проводилось в следующей последовательности. Сначала численно моделировалась плотность десорбции [4] и коэффициенты модели «забывались». Вычисления контролировались материальным балансом:

$$\begin{aligned}&\int_0^\ell [\varphi(x) + \psi(x)] dx \\ &= \int_0^\ell [c(t, x) + z(t, x)] dx + 2 \int_0^t J(\tau) d\tau, \quad J = b c_{0,\ell}^2.\end{aligned}$$

Оценка поверхностных параметров (десорбции и растворения)

Рассмотрим модель ТДС-дегазации пластины с поверхностной десорбцией:

$$\begin{aligned}\frac{\partial c}{\partial t} &= D(T) \frac{\partial^2 c}{\partial x^2}, \quad t \in (0, t_*), \quad x \in (0, \ell), \\ c(0, x) &= \bar{c}(x), \quad x \in [0, \ell], \\ c_{0,\ell}(t) &= g(T) q_{0,\ell}(t), \quad c_x(t, \ell) = -c_x(t, 0), \\ \dot{q}_0(t) &= -b(T) q_0^2(t) + D(T) c_x(t, 0), \\ \dot{q}_\ell(t) &= -b(T) q_\ell^2(t) - D(T) c_x(t, \ell), \\ D(T) &= D_0 \exp\{-E_D/[RT]\}, \\ b(T) &= b_0 \exp\{-E_b/[RT]\}, \quad g = \text{const}, \\ T(t) &= T_0 + vt, \quad c_0(t) = g q_{0,\ell}(t) = c_\ell(t), \\ J(t) &= b(T) q_0^2(t) = b(T) q_\ell^2(t), \quad J(t_*) \approx 0.\end{aligned}$$

Здесь остановимся на решении обратной задачи идентификации трех поверхностных параметров по одной кривой ТДС-дегазации.

Напомним, что используя интегрирование по частям, для гладкой функции $\psi(t, x)$

$$\begin{aligned}0 &= \int_0^{t_*} \int_0^\ell \psi(t, x) [c_t - \mathcal{A}c] dx dt \\ &= \int_0^{t_*} \int_0^\ell c(t, x) [\psi_t - \mathcal{A}^* \psi] dx dt + f.\end{aligned}$$

Слагаемое f содержит величины, связанные с краевыми условиями. Если фиксировать решение $\psi(t, x)$ сопряженного уравнения $\psi_t = \mathcal{A}^* \psi$, то получим соотношение $f = 0$, содержащее искомые параметры и экспериментальные данные. Выбирая различные «простые» решения $\psi = \psi_i(t, x)$ из бесконечного их числа, получаем систему для оценивания параметров. На решениях уравнения $\psi_t = -D\psi_{xx}$

$$\begin{aligned}0 &= \int_0^\ell \psi(t_*, x) c(t_*, x) dx \\ &- \bar{c} \int_0^\ell \psi(0, x) dx + \int_0^{t_*} c_0 D \frac{\partial \psi}{\partial x} \Big|_{x=0}^\ell dt \\ &+ \int_0^{t_*} (\dot{q} + J) [\psi(t, 0) + \psi(t, \ell)] dt. \quad (3)\end{aligned}$$

Выберем $\psi(t, x) = \beta(t) \exp \sigma x$, $\beta(t) = \exp\{-\sigma^2 \int_0^t D(\tau) d\tau\}$. Тогда первым слагаемым можно пренебречь в силу $\psi \rightarrow 0$, $c \rightarrow 0$ с ростом t . Слагаемое, содержащее $\dot{q}$, следует проинтегрировать по частям с учетом $\dot{\beta} = -D\sigma^2 \beta$. Начальная концентрация $\bar{c} = g\bar{q}$ ($q(0) = \bar{q}$) находится с точностью до неизвестного пока параметра $z = g\ell_0$ из баланса

$$\bar{c}\ell + 2\bar{q} = \frac{\bar{c}\ell}{z} [z+1] = 2 \int_0^{t_*} J(\tau) d\tau \equiv 2S_*, \quad 2\ell_0 = \ell.$$

После технических преобразований соотношение (3) примет вид

$$\begin{aligned}f &\equiv \sigma \ell_0 \varkappa \int_0^{t_*} \beta J dt - S_* \frac{\sigma \ell_0 \varkappa + z}{1+z} \\ &+ \frac{\sigma \ell_0 \varkappa + z}{\sqrt{b_0}} \sigma^2 \int_0^{t_*} \beta D \sqrt{J} \left(\exp \left\{ \frac{1}{2RT} \right\} \right)^{E_b} dt = 0,\end{aligned}$$

$\varkappa \equiv \coth \sigma \ell_0$. Уравнение $f = 0$ нормируем величиной $J_{\max} t_*$:

$$\frac{\bar{X}(\sigma)}{1 + \chi(\sigma \ell_0) z} + \frac{Y(E_b, \sigma)}{\sqrt{b_0}} - \frac{\bar{S}_*}{z+1} = 0,$$

$$\bar{X}(\sigma) \equiv \frac{1}{t_*} \int_0^{t_*} \frac{J(t)}{J_{\max}} \beta(t; \sigma, D) dt,$$

$$Y(E_b, \sigma) \equiv \frac{\sigma^2}{t_*} \int_0^{t_*} D(t) \beta(t; \sigma, D) \frac{\sqrt{J(t)}}{\sqrt{J_{\max}}} u^{E_b} dt,$$

$$\sqrt{b_0} = \sqrt{b_0} \sqrt{J_{\max} t_*}, \quad \chi(\sigma \ell_0) \equiv \frac{\tanh \sigma \ell_0}{\sigma \ell_0},$$

$$\bar{S}_* \equiv \frac{1}{t_*} \int_0^{t_*} \frac{J(t)}{J_{\max}} dt, \quad u = \exp \{1/(2RT)\}.$$

Введенные обозначения безразмерные. Уравнение линейное по переменной $1/\sqrt{b_0}$:

$$\frac{1}{\sqrt{b_0}} = \left(\frac{\bar{S}_*}{z+1} - \frac{\bar{X}(\sigma)}{\chi(\sigma\ell_0)z+1} \right) Y^{-1}(E_b, \sigma).$$

Выбирая значения σ_1 и σ_2 , выражаем $1/\sqrt{b_0}$ из одного уравнения и подставляем во второе:

$$\frac{1 - X(\sigma_1) \left[1 + \frac{1 - \chi_1}{z^{-1} + \chi_1} \right]}{Y(E_b, \sigma_1)} - \frac{1 - X(\sigma_2) \left[1 + \frac{1 - \chi_2}{z^{-1} + \chi_2} \right]}{Y(E_b, \sigma_2)} = F(\sigma_1, z, E_b) - F(\sigma_2, z, E_b),$$

$X(\sigma) = \bar{X}(\sigma)\bar{S}_*^{-1}$, $\chi_i \equiv \chi(\sigma_i\ell_0)$. Информация $J(t)$ (плотность десорбции) входит только под знаками интегралов, что обеспечивает помехоустойчивость. С учетом погрешностей измерений задача решается численно в заданном диапазоне $z \in [z^-, z^+]$, $E_b \in [E_b^-, E_b^+]$ (априорные оценки по порядкам величин):

$$\mathcal{G}(\sigma_1, \sigma_2; z, E_b) = (F(\sigma_1, z, E_b) - F(\sigma_2, z, E_b))^2 \rightarrow \min.$$

При известном E_b параметр z находится как больший положительный корень уравнения:

$$Az^2 + Bz + C = 0, \quad (4)$$

$$A = Y(\sigma_2)\chi_2 B_1 - Y(\sigma_1)\chi_1 B_2,$$

$$B = Y(\sigma_2)[\chi_2 A_1 + B_1] - Y(\sigma_1)[\chi_1 A_2 + B_2],$$

$$C = Y(\sigma_2)A_1 - Y(\sigma_1)A_2,$$

$$A_1 = 1 - X(\sigma_1), \quad A_2 = 1 - X(\sigma_2),$$

$$B_1 = \chi_1 - X(\sigma_1), \quad B_2 = \chi_2 - X(\sigma_2).$$

Далее по одному из уравнений (или как среднее арифметическое по двум) вычисляется b_0 :

$$b_0^{1/2} = \frac{1}{\sqrt{J_{\max}}} \cdot \frac{Y(E_b, \sigma) [\bar{S}_* t_*]^{-1}}{\frac{1}{z+1} - \frac{X(\sigma)}{\chi(\sigma\ell_0)z+1}}. \quad (5)$$

Используя нормированные уравнения, восстанавливаем E_b и z . Для вычисления предэкспоненциального множителя десорбции требуется знать $J_{\max}$ в физических единицах.

Для проверки работоспособности алгоритма идентификация параметров проводилась с обработкой численных спектров. Задача этого тестирования – дать рекомендации по выбору параметров (σ) сопряженных функций.

Рис. 1 иллюстрирует чувствительность к изменению безразмерных параметров $\sigma\ell_0$, z соотношений, входящих в уравнение для оценки поверхностных параметров материала. На рис. 1 (a) показана зависимость $\chi(\sigma\ell_0)$, по горизонтальной оси варьируется $\sigma\ell_0$, масштаб линейный, по вертикальной оси отмечены значения χ , масштаб логарифмический. На врезке в рис. 1 (a) показан тот же график, диапазоны по осям неизменны, но по горизонтальной оси масштаб логарифмический, а по вертикальной – линейный. На рис. 1 (b) показано, как влияют параметры $\sigma\ell_0$, z на значение выражения в скобках [...]. В последующих иллюстрациях эта скобка будет фигурировать как множитель, поэтому, выбирая σ_1 и σ_2 , ориентируемся на диапазон, в котором [...] > 1, но не слишком велико. Необходимо указать диапазон изменений параметров сопряженных функций: как именно следует варьировать σ (или безразмерные величины $\sigma\ell_0$). Отметим, что графики на рис. 1 не имеют привязки к конкретному ТДС-спектру/материалу. Представлена качественная зависимость для широкого диапазона значений $\sigma\ell_0$, z .

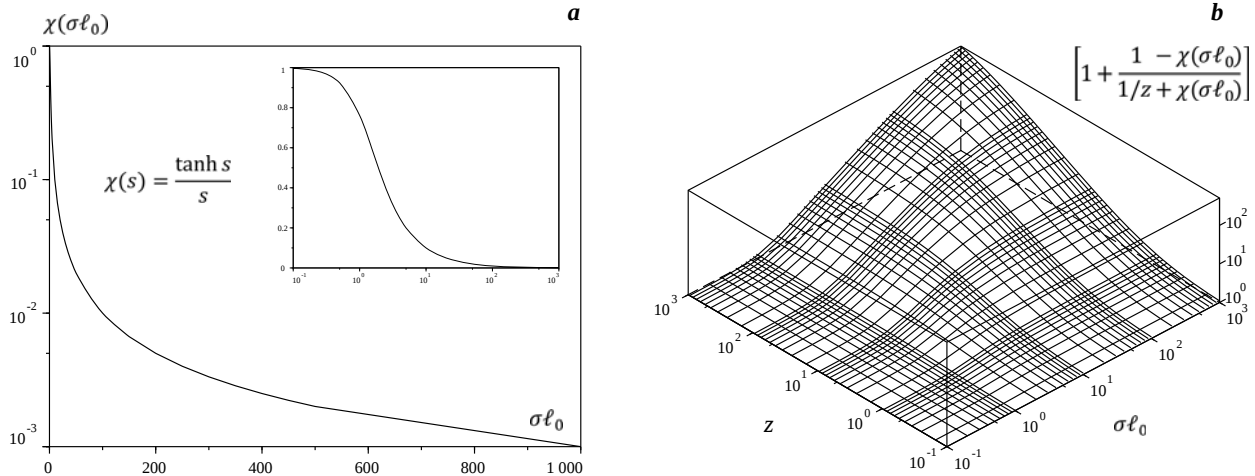


Рис. 1. Выбор параметров сопряженных функций, влияние $(\sigma\ell_0)$, z
Fig. 1. Selection of conjugate functions parameters, effects of $(\sigma\ell_0)$, z

Для тестирования методов численного моделирования ТДС-спектра и последующей параметрической идентификации поверхностных параметров опирались на данные по никелю [5]. В расчетах использовались следующие значения параметров: $\ell = 0,1$ см; $[E] = \text{kJ/mol}$; $\dot{T} = 0,1$ К/с; $b_0 = 1,53 \cdot 10^{-14}$ см²/с; $E_b = 43,2$; $D_0 = 7,5 \cdot 10^{-3}$ см²/с; $E_D = 40$; $g_0 = 100$ см⁻¹; $E_g = 0$; $T_0 = 300$ К; $\bar{c} = 10^{18}$ ат.Н/см³. Фиксируем шаг по времени 1 с; по пространственной переменной используем сетку из 500 точек с равным шагом по толщине пластины; время эксперимента – 8000 с. На рис. 2 (b) справа представлен полученный численный (модель-

ный) спектр, кривая обозначена J_{calc} и отмечена жирной линией на графике. Использование разностной схемы при моделировании десорбции на указанной сетке привело к рассогласованию материального баланса на уровне 0,152 %. Кроме того, выполнялась идентификация поверхностных параметров по зашумленному спектру, см. рис. 2 (a) слева (кривая J_{noisy}) – на спектр добавлен шум, не превосходящий 20 % от текущей плотности десорбционного потока; сглаженному зашумленному спектру, см. рис. 2 (b) (кривая J_{smooth} , отмечена тонкой линией на графике) – это спектр J_{noisy} после сглаживания помех. Сглаживание кривой выполнялось по 9 точкам [3].

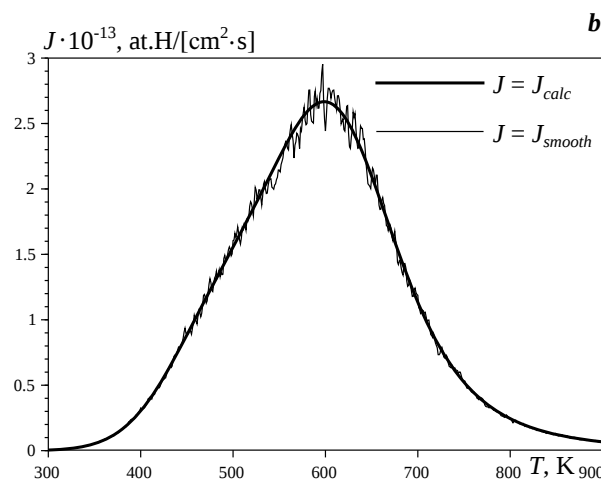
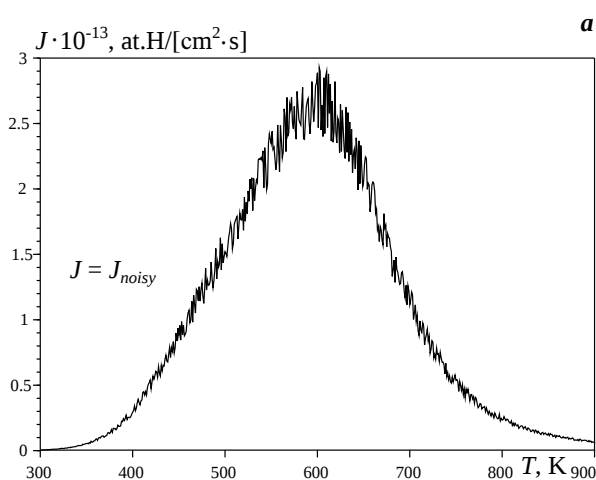


Рис. 2. Численные ТДС-спектры
Fig. 2. Numerical TDS spectra

На рис. 3 (b–d) показаны графики сопряженных функций вида $\psi = \beta(\sigma, t) \exp \sigma x$ ($\beta(\sigma, t) = \exp\{-\sigma^2 \int_0^t D(\tau) d\tau\}$) при фиксированных σ . Пространственная координата представлена на графиках в нормированном виде x/ℓ . Исходная краевая диффузионная задача ТДС-дегазации пластины по постановке симметрична относительно $x = \ell_0 = \ell/2$. В сопряженных функциях симметрии нет. По времени также представлена зависимость в безразмерных единицах – долях от общего времени эксперимента t_* . На рис. 3 (b–d) видно, что с ростом значения σ область значений сопряженных функций становится шире. На рис. 3 (a) показаны графики соответствующих поверхностей $\beta(\sigma, t)$, представленный диапазон значений σ достаточно широкий и содержит значения, указанные на рис. 3 (b–d).

На рис. 4 показаны линии уровня по трем значениям (1 %, 10 %, 20 %) для относительных погрешностей оценок E_b и z . Варьируются σ_1 и σ_2 . По осям отмечены σ и безразмерные величины $(\sigma \ell_0)$. На координатной плоскости параметры σ_1 и σ_2 взяты с одинаковым шагом, но со сдвигом, т. е. $\sigma_1 \neq \sigma_2$. На графиках уровни зафиксированы, ориентируясь на примерные уровни экспериментальных погрешностей. В предлагаемом вычислительном алгоритме потребуется использовать пару фиксированных значений σ_1, σ_2 (или $\sigma_1 \ell_0, \sigma_2 \ell_0$). На рисунках видно, что выбирать их можно из широкого диапазона в 2–3 порядка, что позволяет оценить поверхностные параметры с относительными погрешностями, которые не превосходят типичных экспериментальных погрешностей измерений.

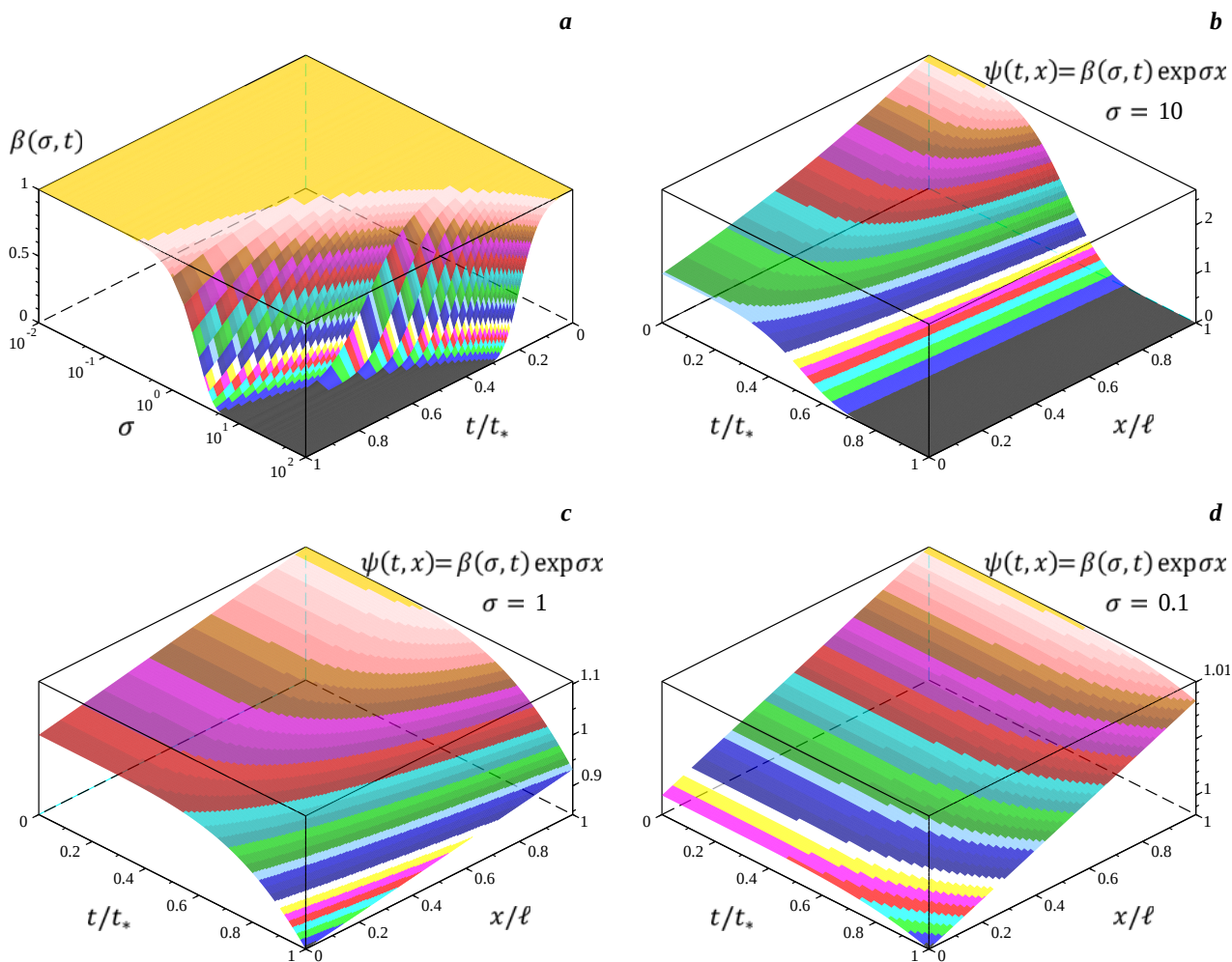


Рис. 3. Сопряженные функции, влияние параметров
Fig. 3. Conjugate functions, parameter effects

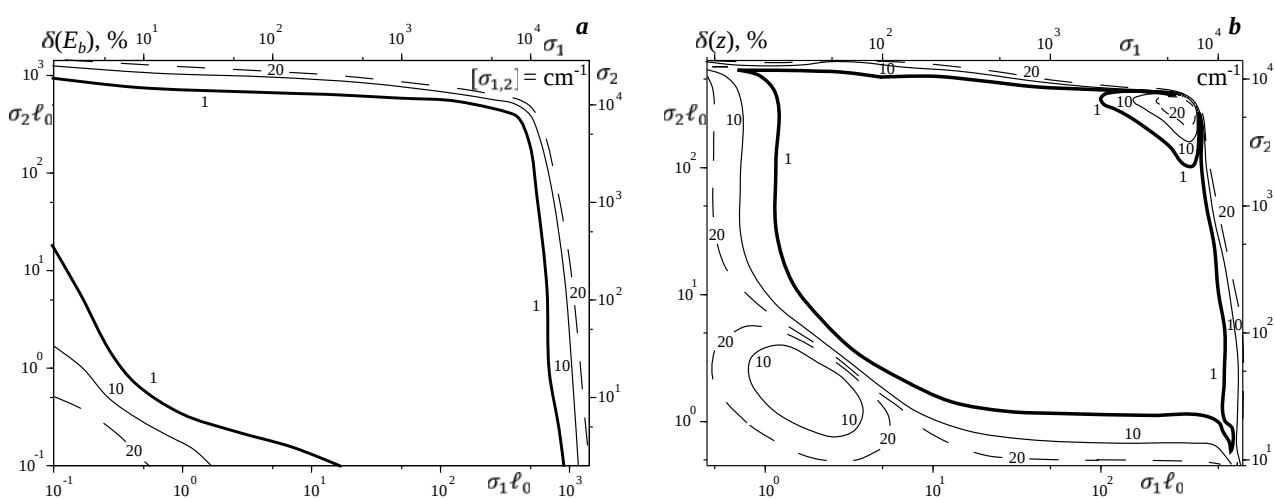
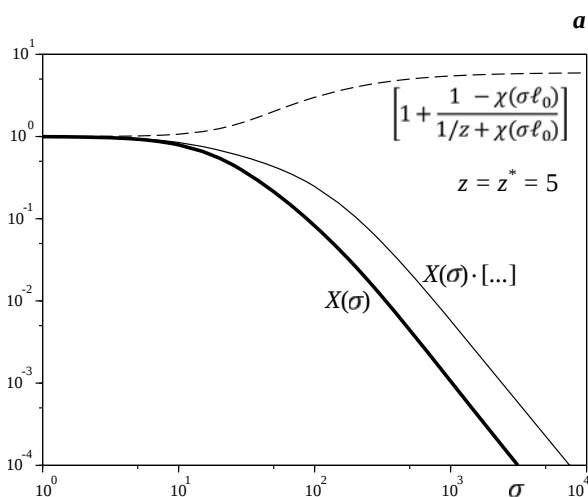


Рис. 4. Линии уровня. Погрешности оценок E_b, z [%]. Влияние σ_1, σ_2 ($\sigma_1 \ell_0, \sigma_2 \ell_0$)
Fig. 4. Level lines. Estimation errors of E_b, z [%]. Effects of σ_1, σ_2 ($\sigma_1 \ell_0, \sigma_2 \ell_0$)

Рис. 5 иллюстрирует влияние параметров на значения безразмерных комплексов X (рис. 5, *a*) и Y (рис. 5, *b*). На рисунке слева видно, что при малых значениях $\sigma : 10^0 - 10^1$ значения функции X и множителя [...] при X изменяются несущественно и сравнимы с 1.



На рисунке справа представлен график поверхности Y . Разброс значений по варьируемым параметрам весьма широкий и охватывает характерный диапазон значений E_b энергий активации поверхностной десорбции водорода для различных конструкционных материалов.

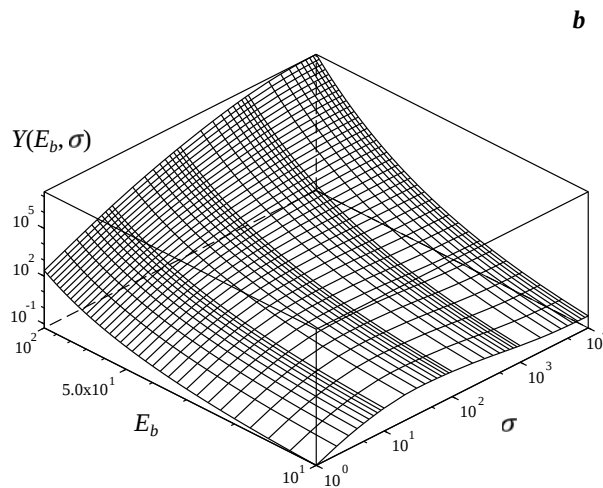


Рис. 5. Безразмерные комплексы X , Y . Влияние параметров
Fig. 5. Dimensionless complexes X , Y . Parameter effects

На рис. 6 (*a*, *b*) представлены графики задействованных в оценочных уравнениях безразмерных комплексов X и Y в симметричной форме, варьируются параметры сопряженных функций σ_1, σ_2 . Величины X, Y не имеют прямого физического смысла, но использовались из вычислительных соображений, чтобы соответствующие сопряженные функции «не давили» отдельные участки измерений. Должна быть соизмеримость сомножителей, в этом косвенный физический смысл. На рис. 6 (*c*, *d*) показаны сечения функции F , когда один из оцениваемых параметров принимает истинное значение. Слева на рис. 6 (*c*) изображена поверхность F при фиксированном истинном z^* ; E_b варьируется в пределах одного порядка, но сдвиг E_b меняет значения F на порядки (разброс значений на рисунке слева $10^{-6} - 1$). Влияние E_b на кинетику дегазации велико. Напротив, на рис. 6 (*d*) показаны поверхности F при фиксированном истинном E_b^* . Варьирование σ и z в пределах трех порядков приводит к небольшому немонотонным изменениям F , разброс значений $0,0034 - 0,0058$.

В алгоритме выполняется поиск минимума функции $\mathcal{G}$, с варьированием E_b и z из широкого физически обоснованного диапазона, при фиксированной паре значений σ_1, σ_2 . Поиск выполняли итерационно (достаточно

3–4 итераций): сначала решали задачу в широком диапазоне значений E_b и z . Затем, получив первое приближение оценок, решение задачи повторяли несколько раз в окрестности приближений, сужая область поиска и уменьшая шаг по варьируемым E_b, z . Каждое следующее решение задачи поиска параметров E_b, z , доставляющих минимум функции $\mathcal{G}$, уменьшало значение $\min \mathcal{G}$ на порядки.

На рис. 7 показаны сечения графика $\mathcal{G}$. На графиках слева и справа одинаковые диапазоны по координатным осям. На врезках к рисункам показаны те же кривые, но по вертикальным осям приведен тот же диапазон в линейной шкале. Кривые на рисунках пронумерованы по порядку используемых значений изменяемых параметров (слева – σ , справа – z). На рис. 7 (*a*) слева фиксировано истинное значение z^* , показаны три кривые для различных значений первого аргумента функции $\mathcal{G}$. Соответственно: кривая 1 построена при $\sigma_1 = 100$, кривая 2 – $\sigma_2 = 200$, кривая 3 – $\sigma_3 = 1000$. Справа на рис. 7 (*b*) изменяются значения третьего параметра функции $\mathcal{G}$ – z , кривые на рисунке и врезке аналогично пронумерованы. Лучшим значением варьируемого параметра является то, при котором минимум функции $\mathcal{G}$ меньше.

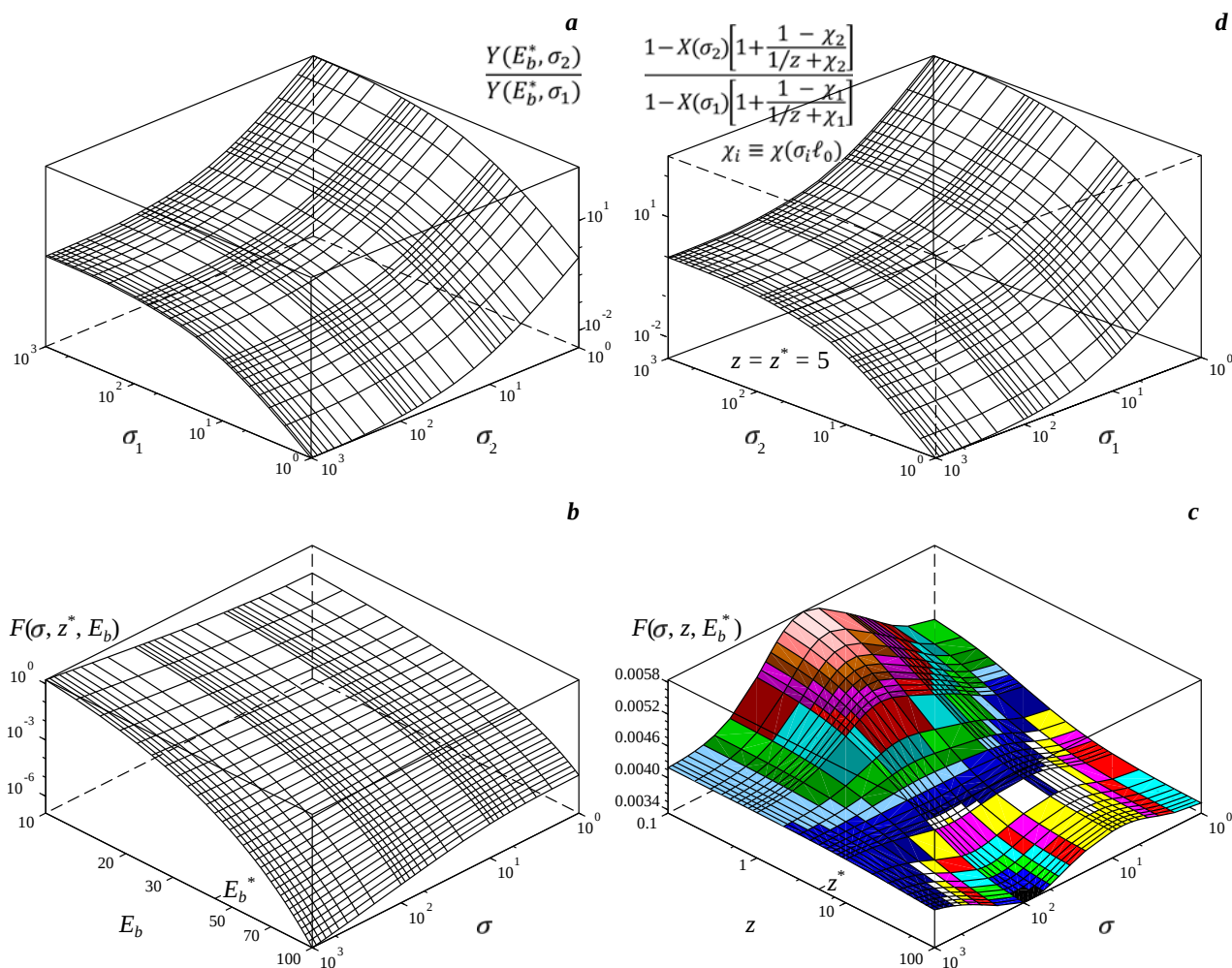


Рис. 6. Влияние параметров. Варьируем параметры $\sigma_{1,2}$ сопряженных функций и оцениваемые E_b, z
 Fig. 6. Parameter effects. We vary the parameters $\sigma_{1,2}$ of the conjugate function and E_b, z to be estimated

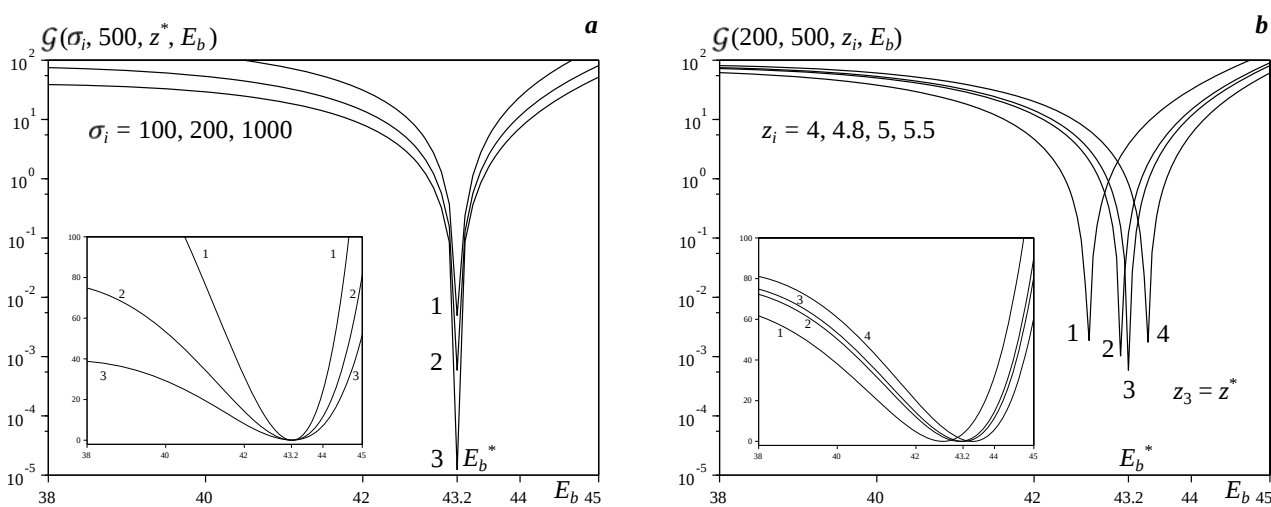


Рис. 7. Поиск энергии активации E_b , сечения. Влияние параметров
 Fig. 7. Search for activation energy E_b , cross-section. Parameter effects

В таблице приведены полученные оценки трех поверхностных параметров по одному ТДС-графику. Рассматривались три варианта ТДС-кривой: модельный численный спектр, зашумленный и затем сглаженный (ТДС-спектры приведены на рис. 2). Отметим, что благодаря использованию интегральных операторов не требуется специально сглаживать

зашумленную экспериментальную кривую. Алгоритм эффективно работает и на исходной ТДС-кривой с помехами, разница в относительных погрешностях незначительная. Решение квадратного уравнения (4) после оценки E_b несущественно смещает оценку z . Можно сразу решать задачу оценки пары значений (E_b, z) , а затем по формуле (5) вычислять параметр b_0 .

Оценки поверхностных параметров
Estimates of surface parameters

Параметр Parameter	Исходные данные Input data	Полученные оценки (их относительные погрешности) Parameter estimates (their relative errors)		
		J_{calc}	J_{noisy}	J_{smooth}
E_b , kJ/mol	43, 2	43, 127 (0, 169 %)	43, 246 (0, 106 %)	43, 255 (0, 127 %)
$g(= z/\ell_0)$, s ⁻¹	100	100, 168 (0, 168 %)	97, 03 (2, 97 %)	97, 244 (2, 756 %)
b_0 , cm ² /s	$1, 53 \cdot 10^{-14}$	$1, 395 \cdot 10^{-14}$ (8, 824 %)	$1, 453 \cdot 10^{-14}$ (5, 033 %)	$1, 49 \cdot 10^{-14}$ (2, 614 %)
$\mathcal{G}(10^2, 10^3, z, E_b)$		$1, 517 \cdot 10^{-14}$	$4, 868 \cdot 10^{-12}$	$9, 443 \cdot 10^{-14}$

ЗАКЛЮЧЕНИЕ

Предложен и протестирован вычислительный алгоритм идентификации по ТДС-спектру трех поверхностных параметров (соответствующих десорбции и растворению) термодесорбции изотопов водорода из конструкционных материалов методом сопряженных уравнений. Численные эксперименты показали, что алгоритм помехоустойчив, идентификация по зашумленному спектру заметно не ухудшает оценки поверхностных параметров. Программная реализация алгоритма и процедур, иллюстрирующих этапы алгоритма, выполнена в среде Scilab. Проверена работоспособность алгоритма решения обратной задачи на численных ТДС-спектрах, полученных при решении соответствующей нелинейной краевой задачи с динамическими граничными условиями. Поверхностные параметры восстанавливаются по численному ТДС-спектру с погрешностями, не превосходящими погрешности измерений в физических экспериментах. Изложенный метод позволяет восстанавливать энергию активации поверхностной десорбции и коэффициент растворения по ТДС-спектру (измеренному даже в относительных единицах), поскольку оценочные уравнения используют безразмерные величины. При необходимости оценить предэкспо-

ненту коэффициента поверхностной десорбции требуется знать значения десорбционного потока в физических величинах.

ЛИТЕРАТУРА

1. *Взаимодействие водорода с металлами* / Ред. А. П. Захаров. М.: Наука, 1987. С. 177–206.
2. *Водород в металлах: в 2 т.* / Под ред. Г. Алефельда, В. Фелькля; пер. с англ. М.: Мир, 1981.
3. *Волков Е. А. Численные методы.* М.: Наука, 1987. 248 с.
4. *Заика Ю. В., Костинова Е. К.* Разностная схема для краевой задачи ТДС-дегазации с динамическими граничными условиями // Ученые записки Петрозаводского государственного университета. Сер. Естественные и технические науки. 2009. № 7(101). С. 65–70.
5. *Изотопы водорода. Фундаментальные и прикладные исследования* / Ред. А. А. Юхимчук. Саратов: РФЯЦ-ВНИИЭФ, 2009. 697 с.
6. *Кунин Л. Л., Головин А. И., Суровой Ю. И., Хохрин В. М.* Проблемы дегазации металлов. М.: Наука, 1972. 324 с.
7. *Марчук Г. И.* Сопряженные уравнения и анализ сложных систем. М.: Наука, 1992. 336 с.
8. *Писарев А. А., Цветков И. В., Маренков Е. Д., Ярмо С. С.* Проницаемость водорода через металлы. М.: МИФИ, 2008. 144 с.

REFERENCES

1. *Zakharov A. P. (ed.)*. Interaction of hydrogen with metals. Moscow: Nauka; 1987. P. 177–206. (In Russ.)
2. *Alefeld G., Völkl J. (eds.)*. Hydrogen in metals. Berlin-Heidelberg-New York: Springer-Verlag; 1978.
3. *Volkov E. A.* Numerical methods. Moscow: Nauka; 1987. 248 p. (In Russ.)
4. *Zaika Yu. V., Kostikova E. K.* A difference scheme for the TDS-degassing boundary value problem with dynamic boundary conditions. *Uchenye zapiski Petrozavodskogo gosudarstvennogo universiteta. Ser. Estestvennye i tekhnicheskie nauki = Proceedings of Petrozavodsk State University. Ser. Natural and Technical Sciences*. 2009;7(101):65–70. (In Russ.)
5. *Yukhimchuk A. A. (ed.)*. Hydrogen isotopes. Fundamental and applied research. Sarov: RFYATs-VNIEHF; 2009. 697 p. (In Russ.)
6. *Kunin L. L., Golovin A. I., Surovoy Yu. I., Khokhrin V. M.* Problems of metal degassing. Moscow: Nauka; 1972. 324 p. (In Russ.)
7. *Marchuk G. I.* Conjugate equations and analysis of complex systems. Moscow: Nauka; 1992. 336 p. (In Russ.)
8. *Pisarev A. A., Tsvetkov I. V., Marenov E. D., Yarko S. S.* Hydrogen permeability through metals. Moscow: MIFI; 2008. 144 p. (In Russ.)

Поступила в редакцию / received: 17.04.2026; принята к публикации / accepted: 22.05.2026.
Авторы заявляют об отсутствии конфликта интересов / The authors declare no conflict of interest.

СВЕДЕНИЯ ОБ АВТОРАХ:

Заика Юрий Васильевич

д-р физ.-мат. наук, профессор, руководитель лаборатории моделирования природно-технических систем

e-mail: zaika@krc.karelia.ru

Костикова Екатерина Константиновна

канд. физ.-мат. наук, старший научный сотрудник

e-mail: kostikova@krc.karelia.ru

CONTRIBUTORS:

Zaika, Yury

Dr. Sci. (Phys.-Math.), Professor, Head of Laboratory of Natural-Tech. System Modelling

Kostikova, Ekaterina

Cand. Sci. (Phys.-Math.), Senior Researcher

УДК 515.12

О ТОПОЛОГИЧЕСКОЙ РАЗМЕРНОСТИ МАКСИМАЛЬНЫХ СЦЕПЛЕННЫХ СИСТЕМ

А. В. Иванов

*Институт прикладных математических исследований КарНЦ РАН,
ФИЦ «Карельский научный центр РАН» (ул. Пушкинская, 11, Петрозаводск,
Республика Карелия, Россия, 185910)*

В работе дано определение топологической размерности $\dim \xi$ максимальных сцепленных систем (м.с.с.) замкнутых подмножеств метризуемого компакта X . По определению $\dim \xi$, есть точная нижняя грань нижних размерностей квантования м.с.с. ξ по всем метрикам, совместимым с топологией X . Предложенное определение $\dim \xi$ мотивировано теоремой Понтрягина – Шнирельмана, характеризующей топологическую размерность метризуемого компакта как инфимум нижних емкостных размерностей этого компакта по всем совместимым метрикам. Для любого метризуемого компакта X и любой м.с.с. ξ выполняется неравенство $\dim \xi \leq \dim X$. При этом для конечномерного метризуемого компакта X справедлива следующая теорема о промежуточных значениях размерности $\dim \xi$: для любого целого числа k , удовлетворяющего неравенствам $0 \leq k \leq \dim X$, существует м.с.с. ξ_k такая, что $\dim \xi_k = k$. Доказано, что для всех рассмотренных ранее м.с.с. с известными значениями размерностей квантования размерность $\dim \xi$ является целым числом. В частности, построенные по технологии Е. В. Кашубы м.с.с. $\xi(A, B)$ всегда нульмерны в топологическом смысле.

Ключевые слова: метрический компакт; суперрасширение; размерность квантования; емкостная размерность; теорема Понтрягина – Шнирельмана

Для цитирования: Иванов А. В. О топологической размерности максимальных сцепленных систем // Труды Карельского научного центра РАН. 2026. № 6. С. 47–53. doi: 10.17076/mat2282

Финансирование. Финансовое обеспечение исследований осуществлялось из средств федерального бюджета на выполнение государственного задания КарНЦ РАН (Институт прикладных математических исследований КарНЦ РАН).

A. V. Ivanov. ON THE TOPOLOGICAL DIMENSION OF MAXIMAL LINKED SYSTEMS

*Institute of Applied Mathematical Research, Karelian Research Centre, Russian
Academy of Sciences (11 Pushkinskaya St., 185910 Petrozavodsk, Karelia, Russia)*

We define the topological dimension $\dim \xi$ of maximal linked systems (m.l.s.) of closed subsets of a metrizable compact space X . By definition, $\dim \xi$ is the greatest lower bound of the lower quantization dimensions of the m.l.s. ξ over all metrics

compatible with the topology of X . The proposed definition of $\dim \xi$ is motivated by the Pontryagin–Shnirelman theorem, which characterizes the topological dimension of a metrizable compact space as the infimum of the lower box dimensions of this compactum over all compatible metrics. For any metrizable compact space X and any m.l.s. ξ , the inequality $\dim \xi \leq \dim X$ holds. Moreover, for a finite-dimensional metrizable compactum X , the following theorem on intermediate values of the dimension $\dim \xi$ holds: for any integer k satisfying the inequalities $0 \leq k \leq \dim X$, there exists a m.l.s. ξ_k such that $\dim \xi_k = k$. It is proved that for all previously considered maximal linked systems with known values of the quantization dimensions, the dimension $\dim \xi$ is an integer. In particular, m.l.s. $\xi(A, B)$ constructed using E. V. Kashuba’s technology are always zero-dimensional in the topological sense.

Keywords: metric compactum; superextension; quantization dimension; box dimension; Pontryagin–Schnirelmann theorem

For citation: Ivanov A. V. On the topological dimension of maximal linked systems. *Trudy Karelskogo nauchnogo tsentra RAN = Transactions of the Karelian Research Centre RAS*. 2026. No. 6. P. 47–53. doi: 10.17076/mat2282

Funding. The studies were funded from the federal budget through state assignment to the Karelian Research Centre RAS (Institute of Applied Mathematical Research, Karelian Research Centre RAS).

В настоящей работе дано определение топологической размерности максимальных сцепленных систем (м.с.с.) – элементов суперрасширения λX метризуемого компакта X . В основе определения лежит понятие размерности квантования $\dim_\lambda(\xi, \rho)$ м.с.с. ξ метрического компакта (X, ρ) , которое впервые было рассмотрено (под названием «порядок метрической аппроксимации») в совместной статье автора и О. В. Фомкиной [5]. Размерность $\dim_\lambda(\xi, \rho)$ является частным случаем общего понятия размерности квантования точек пространств вида $\mathcal{F}(X)$, где $\mathcal{F}$ – полунормальный метризуемый функтор, а X – метрический компакт [3]. В ту же общую функториальную схему укладывается и классическое понятие емкостной размерности $\dim_B F$ замкнутого подмножества F метрического компакта (X, ρ) . В такой ситуации естественно предполагать наличие общих свойств у размерностей квантования м.с.с. и емкостных размерностей.

В 1932 году Л. С. Понтрягин и Л. Г. Шнирельман доказали следующую фундаментальную теорему:

Теорема 1 [8]. *Топологическая размерность $\dim X$ метризуемого компакта X совпадает с точной нижней гранью нижних емкостных размерностей этого компакта по всем метрикам, совместимым с топологией X .*

Этот результат мотивирует исследование точной нижней грани (нижних) размерностей квантования точек $\xi \in \lambda X$ по всем совместимым метрикам на X , которую (по аналогии) можно назвать топологической размерностью

$\dim \xi$ м.с.с. ξ . При этом оказывается, что всегда $\dim \xi \leq \dim X$. В работе доказано (теорема 3), что для любого метризуемого компакта X конечной размерности $\dim X = n$ и любого целого числа k , удовлетворяющего неравенствам $0 \leq k \leq n$, существует м.с.с. $\xi_k \in \lambda X$, для которой $\dim \xi_k = k$.

В статье А. А. Иванова [4] доказана следующая теорема о промежуточных значениях нижней размерности квантования максимальных сцепленных систем:

Теорема 2 [4]. *Для любого метрического компакта (X, ρ) и любого неотрицательного числа b , не превосходящего нижней емкостной размерности X , существует м.с.с. $\xi_b \in \lambda X$, нижняя размерность квантования которой равна b .*

Этот результат был получен путем построения конкретных систем ξ_b с использованием конструкции $\xi(A, B)$, предложенной Е. В. Кашубой [2]. В настоящей работе показано (теорема 4), что топологическая размерность всех м.с.с. вида $\xi(A, B)$ равна нулю. Таким образом, во всех известных ранее примерах и конструкциях, дающих конкретные числовые значения размерностей квантования максимальных сцепленных систем, топологическая размерность м.с.с. оказывается целым числом. Но остается открытым следующий вопрос: верно ли, что $\dim \xi$ есть целое число (или бесконечность) для любой м.с.с. $\xi \in \lambda X$?

В дальнейшем рассматриваются только метризуемые компакты, замыкание множества обозначается квадратными скобками. Си-

система ξ замкнутых подмножеств компакта X называется сцепленной, если любые два ее элемента имеют непустое пересечение. Множество максимальных (по включению) сцепленных систем («суперрасширение X ») обозначается через λX и наделяется топологией, открытую предбазу которой образуют множества вида $O(U) = \{\xi : \text{существует } F \in \xi \text{ такое, что } F \subset U\}$, где U открыто в X . В каждом элементе $F \in \xi$ содержится минимальный (по включению) элемент ξ . Замыкание объединения всех минимальных элементов ξ называется носителем данной системы и обозначается через $\text{supp}(\xi)$. М.с.с. с одноточечным носителем – это в точности системы вида $\xi_x = \{F = [F] : x \in F\}$, где $x \in X$. Если отождествить каждую точку $x \in X$ с м.с.с. ξ_x , мы получим топологическое вложение X в λX . Множество максимальных сцепленных систем, носитель которых содержит не более n точек, обозначается через $\lambda_n X$ ($n \in \mathbb{N}$). Объединение $\bigcup_{n \in \mathbb{N}} \lambda_n X$ всегда всюду плотно в λX . Если на X задана совместимая метрика ρ , то продолжение ρ_λ этой метрики на λX (выше отмечено, что X можно считать подмножеством λX) определяется по формуле:

$$\rho_\lambda(\xi, \eta) = \inf\{\varepsilon : B(F, \varepsilon) \in \eta$$

для любого $F \in \xi\}$,

где $B(F, \varepsilon) = \{x : \rho(x, F) \leq \varepsilon\}$ – замкнутая ε -окрестность F . Доказательство всех указанных выше фактов можно найти в [7, гл. 7].

Пусть (X, ρ) – метрический компакт, $\xi \in \lambda X$ и $\varepsilon > 0$. Определим число $N(\xi, \varepsilon, \rho)$ как наименьшую мощность носителя ε -приближения ξ по метрике ρ_λ :

$$N(\xi, \varepsilon, \rho) = \min\{n : \rho_\lambda(\xi, \lambda_n X) \leq \varepsilon\}.$$

Если м.с.с. ξ имеет бесконечный носитель, то $N(\xi, \varepsilon, \rho)$ неограниченно возрастает при $\varepsilon \rightarrow 0$. Скорость этого возрастания характеризуют верхняя и нижняя размерности квантования ξ по метрике ρ :

$$\overline{\dim}_\lambda(\xi, \rho) = \overline{\lim}_{\varepsilon \rightarrow 0} \frac{\log N(\xi, \varepsilon, \rho)}{-\log \varepsilon},$$

$$\underline{\dim}_\lambda(\xi, \rho) = \underline{\lim}_{\varepsilon \rightarrow 0} \frac{\log N(\xi, \varepsilon, \rho)}{-\log \varepsilon}.$$

Очевидно, что всегда $\underline{\dim}_\lambda(\xi, \rho) \leq \overline{\dim}_\lambda(\xi, \rho)$. В случае равенства используют обозначение $\dim_\lambda(\xi, \rho)$ и говорят о размерности квантования ξ .

Нам понадобится также понятие емкостной размерности замкнутого подмножества метрического компакта (X, ρ) . Пусть множество

F замкнуто в X и $\varepsilon > 0$. Обозначим через $N(F, \varepsilon, \rho)$ наименьшее число точек в ε -сети для F по метрике ρ . Верхняя и нижняя емкостные размерности множества F определяются по формулам:

$$\overline{\dim}_B(F, \rho) = \overline{\lim}_{\varepsilon \rightarrow 0} \frac{\log N(F, \varepsilon, \rho)}{-\log \varepsilon},$$

$$\underline{\dim}_B(F, \rho) = \underline{\lim}_{\varepsilon \rightarrow 0} \frac{\log N(F, \varepsilon, \rho)}{-\log \varepsilon}.$$

Легко проверить, что $N(F, \varepsilon, \rho)$ – это в точности наименьшее число точек в ε -приближении F по метрике Хаусдорфа ρ_H . Таким образом, с точки зрения теории метризуемых полунормальных функторов [3] приведенные выше определения размерностей ξ и F совпадают.

В [5] доказано, что для любой м.с.с. $\xi \in \lambda X$ выполняются неравенства:

$$\begin{aligned} \overline{\dim}_\lambda(\xi, \rho) &\leq \overline{\dim}_B(\text{supp}(\xi), \rho), \\ \underline{\dim}_\lambda(\xi, \rho) &\leq \underline{\dim}_B(\text{supp}(\xi), \rho). \end{aligned} \quad (1)$$

Как уже было отмечено, согласно теореме Понтрягина–Шнирельмана

$$\dim X = \inf\{\underline{\dim}_B(X, \rho) :$$

ρ – совместимая метрика на $X\}$,

где $\dim X$ – топологическая размерность компакта X .

Определение. Пусть $\xi \in \lambda X$. Топологической размерностью $\dim \xi$ м.с.с. ξ будем называть следующую величину:

$$\dim \xi = \inf\{\underline{\dim}_\lambda(\xi, \rho) :$$

ρ – совместимая метрика на $X\}$.

В силу неравенства (1), монотонности емкостных размерностей по замкнутым подмножествам [6, гл. 2] и теоремы Понтрягина–Шнирельмана для любого X и любой м.с.с. $\xi \in \lambda X$ имеет место неравенство

$$\dim \xi \leq \dim X.$$

Следующая теорема о промежуточных значениях $\dim \xi$ вытекает из результатов работы [5].

Теорема 3. Пусть X – конечномерный компакт размерности $\dim X = n$. Тогда для любого целого числа k , удовлетворяющего неравенствам $0 \leq k \leq n$, существует м.с.с. $\xi_k \in \lambda X$, для которой $\dim \xi_k = k$.

Доказательство. Зафиксируем в X собственное замкнутое подмножество F размерности k . Такое подмножество в конечномерном компакте всегда существует [1, гл. 5]. Выберем точку

$x \in X \setminus F$ и рассмотрим следующую сцепленную систему замкнутых подмножеств X :

$$\eta = \{F\} \cup \{\{x, y\} : y \in F\}.$$

Как известно [5], система η содержится в единственной м.с.с., состоящей из надмножеств элементов η , которую (следуя [5]) обозначим через $\xi(x, F)$. При этом в силу предложения 4 из [5] имеет место равенство

$$\dim_\lambda(\xi(x, F), \rho) = \dim_B(F, \rho) \quad (2)$$

для любой совместимой метрики ρ на X . Инфимум левой части равенства (2) по определению равен $\dim \xi(x, F)$. Известно, что емкостная размерность F , вычисленная относительно X по метрике ρ , заданной на X , совпадает с емкостной размерностью F относительно метрики $\rho|_F$. В то же время по теореме Хаусдорфа [9] любая совместимая метрика на F продолжается до совместимой метрики на X . Таким образом, согласно теореме Понтрягина–Шнирельмана инфимум правой части (2) равен $\dim F$. Таким образом, $\xi_k = \xi(x, F)$ – искома система. $\square$

Замечание 1. Если $\dim X = \infty$, мы можем утверждать только, что существует $\xi \in \lambda X$ размерности $\dim \xi = \infty$. Для построения такой системы достаточно взять собственное замкнутое подмножество $F \subset X$ бесконечной размерности, точку $x \in X \setminus F$ и положить $\xi = \xi(x, F)$. Для бесконечномерного X для конечных значений k вопрос о справедливости утверждения теоремы 3 остается открытым. Дело в том, что существуют бесконечномерные метризуемые компакты, в которых любое замкнутое подмножество имеет размерность либо 0, либо бесконечность [1, гл. 9].

В дальнейшем нам понадобится следующая конструкция, предложенная Е. В. Кашубой [2]. Пусть $A = \{a_i : i \in \mathbb{N}\}$ и $B = \{b_i : i \in \mathbb{N}\}$ – непересекающиеся последовательности в X , состоящие из попарно различных точек, и $[A] \cap [B] \neq \emptyset$. Для $i \in \mathbb{N}$ положим

$$A_i = \{a_1, \dots, a_i, b_i\}, \quad B_i = \{b_1, \dots, b_i, a_{i+1}\}.$$

Очевидно, что система

$$\eta(A, B) = \{A_i, B_i : i \in \mathbb{N}\} \cup \{[A], [B]\}$$

является сцепленной. В [2] показано, что $\eta(A, B)$ содержится в единственной м.с.с. $\xi(A, B)$, для которой множества A_i, B_i являются минимальными (по включению) элементами. Если при этом $[A] \cap B = [B] \cap A = \emptyset$, то минимальными элементами м.с.с. $\xi(A, B)$ являются также $[A]$ и $[B]$. Следующая лемма является обобщением леммы 2 из [4].

Лемма 1. Если для системы $\xi(A, B)$ при некотором $i > 1$ выполняется неравенство $\rho(b_i, \{a_2, \dots, a_i\}) \leq \varepsilon$, то

$$N(\xi(A, B), \varepsilon, \rho) \leq 2i. \quad (3)$$

Доказательство. Рассмотрим систему $\{A_k : k \leq i\} \cup \{B_k : k < i\} \cup \{B_i \setminus \{a_{i+1}\}\}$. Легко проверить, что эта система является сцепленной и содержится в единственной м.с.с. ν , элементы которой являются надмножествами исходной системы. Таким образом, лишь один минимальный элемент м.с.с. ν – множество $B_i \setminus \{a_{i+1}\}$ – не содержится в системе $\xi(A, B)$. Рассмотрим множество $H = B(B_i \setminus \{a_{i+1}\}, \varepsilon)$. В силу неравенства $\rho(b_i, \{a_2, \dots, a_i\}) \leq \varepsilon$ найдется точка a_k ($2 \leq k \leq i$), которая содержится в H . И тогда $B_{k-1} \subset H$, откуда следует, что $H \in \xi(A, B)$.

Итак, $B(F, \varepsilon) \in \xi(A, B)$ для любого $F \in \nu$, следовательно, $\rho_\lambda(\xi(A, B), \nu) \leq \varepsilon$. Неравенство (3) тем самым доказано, поскольку $|\text{supp}(\nu)| = 2i$. $\square$

Лемма 2. Пусть X – метризуемый компакт, в котором даны две непересекающиеся последовательности попарно различных точек $A = \{a_i : i \in \mathbb{N}\}$, $B = \{b_i : i \in \mathbb{N}\}$, имеющие общий предел p . Тогда для любой последовательности положительных чисел $\{\varepsilon_i : i \in \mathbb{N}\}$, монотонно сходящейся к нулю, на X существует совместимая с топологией метрика ρ такая, что $\rho(a_i, b_i) \leq \varepsilon_i$ для любого i .

Доказательство. Если разность $X \setminus (A \cup B)$ конечна, то компакт X гомеоморфен сходящейся последовательности, и тогда искомую метрику ρ можно определить следующим образом:

$$\begin{aligned} \rho(x, y) &= \varepsilon_1 \text{ при } x \in X \setminus (A \cup B \cup \{p\}) \text{ и } y \neq x; \\ \rho(a_k, b_k) &= \varepsilon_k; \\ \rho(a_k, y) &= \rho(b_k, y) = \varepsilon_k \text{ при } y \in \{a_i : i > k\} \cup \{b_i : i > k\} \cup \{p\}. \end{aligned}$$

Проверка аксиом метрики для ρ и совместимости этой метрики с топологией X не представляет труда.

Пусть теперь $|X \setminus (A \cup B)| \geq \omega_0$. Зафиксируем счетную базу $\mathcal{B}$ компакта X . Упорядоченную пару (U, V) элементов $\mathcal{B}$ назовем допустимой, если

- 1) $[U] \subset V$;
- 2) $p \in U \cup (X \setminus [V])$.

Для каждой допустимой пары (U, V) выберем непрерывную функцию $f_{(U, V)} : X \rightarrow [0, 1]$, удовлетворяющую следующим условиям:

$$\begin{aligned} f_{(U, V)}(x) &= 0 \text{ при } x \in [U]; \\ f_{(U, V)}(x) &= 1 \text{ при } x \in X \setminus V. \end{aligned}$$

Обозначим множество таких функций через $\mathcal{D}$. Легко проверить, что для любых двух различных точек $x, y \in X$ существует допустимая

пара (U, V) , для которой $x \in U$, $y \in X \setminus V$, и значит, $f_{(U,V)}(x) = 0$, $f_{(U,V)}(y) = 1$. (Другими словами, семейство функций $\mathcal{D}$ разделяет точки X .)

Для каждого $k \in \mathbb{N}$ положим

$$\mathcal{D}_k = \{f \in \mathcal{D} : f(a_i) = f(b_i), \quad i \geq k\}.$$

Если в допустимой паре (U, V) множество V не пересекается с $A \cup B \cup \{p\}$, то $f_{(U,V)} \in \mathcal{D}_1$. Поскольку разность $X \setminus (A \cup B)$ бесконечна, отсюда следует, что множество $\mathcal{D}_1$ также бесконечно. В силу условия 2) из определения допустимой пары каждая функция $f \in \mathcal{D}$ принимает значение 0 или 1 в некоторой окрестности точки p . Следовательно, любая функция $f \in \mathcal{D}$ попадает в множества $\mathcal{D}_k$ при достаточно больших значениях k .

Итак,

$$|\mathcal{D}_1| = \omega_0; \quad \mathcal{D}_1 \subset \mathcal{D}_2 \subset \dots \subset \mathcal{D}_k \subset \dots; \\ \bigcup_{k \in \mathbb{N}} \mathcal{D}_k = \mathcal{D}.$$

Зафиксируем теперь возрастающую последовательность натуральных чисел n_k такую, что

$$\frac{1}{2^{n_k}} \leq \varepsilon_k.$$

Занумеруем элементы $\mathcal{D}$ натуральными числами

$$\mathcal{D} = \{f_i : i \in \mathbb{N}\}$$

так, что $f_i \in \mathcal{D}_k$ для любого $i \leq n_k$ и $k \in \mathbb{N}$. При такой нумерации $f_i(a_k) = f_i(b_k)$ при $i \leq n_k$.

Рассмотрим теперь диагональное произведение Φ отображений f_i :

$$\Phi = \Delta_{i \in \mathbb{N}} f_i : X \rightarrow \prod_{i \in \mathbb{N}} [0, 1]_i,$$

действующее по формуле $\Phi(x) = (f_i(x) : i \in \mathbb{N})$. Поскольку семейство функций $\mathcal{D}$ разделяет точки X , отображение Φ является топологическим вложением компакта X в гильбертов куб $Q = \prod_{i \in \mathbb{N}} [0, 1]_i$. Зададим на Q стандартную метрику ρ , определяемую по формуле

$$\rho(x, y) = \sum_{i \in \mathbb{N}} \frac{|x_i - y_i|}{2^i},$$

где x_i , y_i — i -е координаты точек $x, y \in Q$.

Покажем, что ρ является искомой метрикой на X , если отождествить точки X с их образами в Q при отображении Φ . В самом деле, как

¹Здесь m_i и k_i — индексы точек множеств A и B , соответствующие исходной нумерации, относительно которой построена м.с.с. $\xi(A, B)$.

уже было отмечено, $f_i(a_k) = f_i(b_k)$ при $i \leq n_k$. Следовательно,

$$\rho(\Phi(a_k), \Phi(b_k)) = \sum_{i \in \mathbb{N}} \frac{|f_i(a_k) - f_i(b_k)|}{2^i} \\ \leq \sum_{i > n_k} \frac{1}{2^i} = \frac{1}{2^{n_k}} \leq \varepsilon_k,$$

что и требовалось доказать. $\square$

Теорема 4. Топологическая размерность любой м.с.с. вида $\xi(A, B)$ равна нулю.

Доказательство. Пусть $A = \{a_n : n \in \mathbb{N}\}$, $B = \{b_n : n \in \mathbb{N}\}$ и p — общая предельная точка множеств A и B . Выберем последовательности $A' = \{a_{m_i} : i \in \mathbb{N}\}$ и $B' = \{b_{k_i} : i \in \mathbb{N}\}$, сходящиеся к p ¹, так, что выполняются неравенства

$$m_1 > 1 \text{ и } m_i < k_i < m_{i+1} \text{ при } i \in \mathbb{N}. \quad (4)$$

Положим далее

$$\varepsilon_i = \frac{1}{(2k_i)^i}. \quad (5)$$

Согласно лемме 2, примененной к последовательностям точек A' , B' и числовой последовательности ε_i , зададим на X метрику ρ , для которой

$$\rho(a_{m_i}, b_{k_i}) \leq \varepsilon_i. \quad (6)$$

В силу неравенств (4) и (6) для любого i

$$\rho(b_{k_i}, \{a_2, \dots, a_{k_i}\}) \leq \varepsilon_i.$$

Следовательно, по лемме 1

$$N(\xi(A, B), \varepsilon_i, \rho) \leq 2k_i. \quad (7)$$

Из (5) и (7) следует оценка нижней размерности квантования системы $\xi(A, B)$ по метрике ρ :

$$\underline{\dim}_\lambda(\xi(A, B), \rho) \leq \underline{\lim}_{i \rightarrow \infty} \frac{\log N(\xi(A, B), \varepsilon_i, \rho)}{-\log \varepsilon_i} \\ \leq \underline{\lim}_{i \rightarrow \infty} \frac{\log 2k_i}{\log(2k_i)^i} = 0.$$

Таким образом, $\dim \xi(A, B) = 0$.

Замечание 2. При доказательстве теоремы 3 мы использовали м.с.с. $\xi(x, F)$, для которой $\dim \xi(x, F) = \dim F = k$. При этом множество F было единственным минимальным по

включению элементом $\xi(x, F)$ ненулевой размерности. Отмеченный факт позволяет предположить, что в подобной ситуации для любой м.с.с. ξ будет иметь место равенство $\dim \xi = \dim F$, если F – единственный минимальный элемент ξ , размерность которого отлична от нуля. Однако с помощью теоремы 4 нетрудно показать, что это не так. Соответствующий контрпример можно построить в любом n -мерном компакте X для любого $k = \dim F \leq n$.

В самом деле, пусть $\dim X = n$ и $k \leq n$. В компакте X можно выбрать связное замкнутое подмножество G размерности k [1, гл. 5]. Возьмем две различные точки $x, y \in G$ и их непересекающиеся окрестности U и V . В силу теоремы суммы для размерности $\dim$ хотя бы одно из множеств $G \setminus U$ или $G \setminus V$ имеет размерность k . Будем считать, что k -мерно первое множество, и положим

$$F = G \setminus U.$$

Поскольку G связно, граница U содержит точки F , и мы можем выделить последовательность $A = \{a_n : n \in \mathbb{N}\}$ точек U , сходящуюся к некоторой точке $p \in F$. Пусть $B = \{b_n : n \in \mathbb{N}\}$ – счетное всюду плотное подмножество F , не содержащее точку p . Рассмотрим систему $\xi(A, B)$. Очевидно, что $[A] \cap B = [B] \cap A = \emptyset$, следовательно, множество минимальных элементов системы $\xi(A, B)$ совпадает с системой $\eta(A, B)$, а значит, содержит единственный элемент положительной размерности – множество F . При этом в силу теоремы 4 $\dim \xi(A, B) = 0$. Таким образом, система $\xi(A, B)$ – искомый контрпример. $\square$

Напомним, что теорема 3 выше названа теоремой о промежуточных значениях размерности максимальных сцепленных систем в конечномерном компакте. Однако это наименование не совсем корректно. Действительно, для любой м.с.с. $\xi \in \Lambda X$ имеет место неравенство $\dim \xi \leq \dim X$ и, как установлено в теореме 3, для каждого значения k из промежутка $[0, \dim X]$ существует м.с.с. размерности k . Но в теореме 3 идет речь только о целых значениях. В рассмотренных выше примерах значения $\dim \xi$ также были исключительно целыми, что справедливо (как легко проверить) и для всех м.с.с., рассмотренных в работе [5]. При этом остается открытым следующий важный вопрос:

Верно ли, что размерность $\dim \xi$ равна целому числу (или бесконечности) для любой м.с.с. ξ ?

ЛИТЕРАТУРА

1. Александров П. С., Пасынков Б. А. Введение в теорию размерности. М.: Наука, 1975. 576 с.
2. Вакулова (Кашуба) Е. В. О носителях максимальных сцепленных систем // Труды ПГУ. Математика. 2004. Вып. 11. С. 3–8.
3. Иванов А. В. О промежуточных значениях размерностей квантования идемпотентных мер // Труды института математики и механики УрО РАН. 2024. Т. 30, № 3. С. 139–148. doi: 10.21538/0134-4889-2024-30-3-139-148
4. Иванов А. А. О размерности квантования максимальных сцепленных систем // Сиб. матем. журн. 2024. Т. 65, № 3. С. 517–523. doi: 10.33048/smzh.2024.65.306
5. Иванов А. В., Фомкина О. В. О порядке метрической аппроксимации максимальных сцепленных систем и емкостных размерностях // Труды Карельского научного центра РАН. 2019. № 7. С. 5–14. doi: 10.17076/mat1034
6. Песин Я. Б. Теория размерности и динамические системы: современный взгляд и приложения. М.-Ижевск: Ин-т компьютер. исслед., 2002. 404 с.
7. Федорчук В. В., Филиппов В. В. Общая топология. Основные конструкции. М.: Изд-во Моск. ун-та, 1988. 252 с.
8. Pontryagin L., Shnirelman L. On one metric property of dimension // Ann. Math. 1932. Vol. 33. P. 156–162.
9. Toruńczyk H. A short proof of Hausdorff's theorem on extending metrics // Fundam. Math. 1973. Vol. 77. P. 191–193.

REFERENCES

1. Aleksandrov P. S., Pasynkov B. A. Introduction to dimension theory. Moscow: Nauka; 1975. 576 p. (In Russ.)
2. Vakulova (Kashuba) Ye. V. On the supports of maximal linked systems. *Trudy Petrozavodskogo gosudarstvennogo universiteta. Matematika = Proceedings of Petrozavodsk State University. Ser. Mathematics*. 2004;11:3–8. (In Russ.)
3. Ivanov A. V. On intermediate values of quantization dimensions of idempotent measures. *Trudy instituta matematiki i mekhaniki UrO RAN = Proceedings of the Institute of Mathematics and Mechanics, Ural Branch RAS*. 2024;30(3):139–148. doi: 10.21538/0134-4889-2024-30-3-139-148
4. Ivanov A. A. On the quantization dimension of maximal linked systems. *Siberian Mathematical Journal*. 2024;65(3):575–581. doi: 10.1134/S0037446624030066
5. Ivanov A. V., Fomkina O. V. On the order of metric approximation of maximal linked systems and box dimensions. *Trudy Karelskogo nauchnogo tsentra*

RAN = Transactions of the Karelian Research Centre RAS. 2019;7:5–14. (In Russ.). doi: 10.17076/mat1034

6. *Pesin Y. B.* Dimension theory in dynamical systems. Contemporary views and applications. Chicago: University of Chicago Press; 1997. 304 p.

7. *Fedorchuk V. V., Filippov V. V.* General topology. Basic constructions. Moscow: Izd-vo Mosk. un-ta; 1988. 252 p.

8. *Pontryagin L., Shnirelman L.* On one metric property of dimension. *Ann. Math.* 1932;33:156–162.

9. *Torunczyk H.* A short proof of Hausdorff's theorem on extending metrics. *Fundam. Math.* 1973;77:191–193.

Поступила в редакцию / received: 13.01.2026; принята к публикации / accepted: 17.03.2026.
Автор заявляет об отсутствии конфликта интересов / The author declares no conflict of interest.

СВЕДЕНИЯ ОБ АВТОРЕ:

Иванов Александр Владимирович

д-р физ.-мат. наук, профессор, ведущий научный сотрудник

e-mail: alvlivanov@krc.karelia.ru

CONTRIBUTOR:

Ivanov, Aleksander

Dr. Sci. (Phys.-Math.), Professor, Leading Researcher

УДК 517.91

УПРАВЛЯЕМАЯ ДИНАМИКА ПРОЦЕССА БИООЧИСТКИ СТОЧНЫХ ВОД ПРИ ПЕРЕМЕННОМ ВХОДНОМ ПОТОКЕ ЗАГРЯЗНЕНИЙ

А. Н. Кириллов, А. М. Сазонов*

*Институт прикладных математических исследований КарНЦ РАН,
ФИЦ «Карельский научный центр РАН» (ул. Пушкинская, 11,
Петрозаводск, Республика Карелия, Россия, 185910), *sazon-tb@mail.ru*

В работе представлена модель системы биологической очистки осадков сточных вод. Поставлена задача стабилизации процесса биоочистки при кусочно-постоянном входном потоке загрязнений для общего вида трофических функций, описывающих процесс окисления и входные потоки. Модель представляет собой систему нелинейных обыкновенных дифференциальных уравнений. Посредством замены переменных построена более простая для анализа вспомогательная система линейных дифференциальных уравнений. Для нее показано существование периодического решения. На основе данного результата решена задача стабилизации исходной системы биоочистки. Получен вид управлений, обеспечивающих локализацию периодического решения. Показано, что для любых начальных данных происходит синхронизация решений системы.

Ключевые слова: биологическая очистка; динамические системы; стабилизация; периодические решения; синхронизация решений

Для цитирования: Кириллов А. Н., Сазонов А. М. Управляемая динамика процесса биоочистки сточных вод при переменном входном потоке загрязнений // Труды Карельского научного центра РАН. 2026. № 6. С. 54–60. doi: 10.17076/mat2326

Финансирование. Финансовое обеспечение исследований осуществлялось из средств федерального бюджета на выполнение государственного задания КарНЦ РАН (Институт прикладных математических исследований КарНЦ РАН).

A. N. Kirillov, A. M. Sazonov*. CONTROLLED DYNAMICS OF THE BIOLOGICAL WASTEWATER TREATMENT PROCESS UNDER VARIABLE POLLUTION INPUT FLOW

*Institute of Applied Mathematical Research, Karelian Research Centre, Russian Academy of Sciences (11 Pushkinskaya St., 185910 Petrozavodsk, Karelia, Russia), *sazon-tb@mail.ru*

The paper presents a model of the biological wastewater treatment system. The problem of the wastewater treatment process stabilization for the piecewise-constant pollutant inflow for general form trophic functions describing the oxidation process

and the input flow is studied. The model has the form of a system of nonlinear ordinary differential equations (ODEs). By changing variables, we construct an intermediary system of linear ODEs more useful for the analysis. The existence of a periodic solution to the intermediary system is proved. Based on this result, the problem of the stabilization of the original wastewater treatment system is solved. The form of the controls providing localization of the periodic solution is obtained. It is demonstrated that for any initial data, the timing of the system's solutions is synchronized.

Key words: biological treatment; dynamical systems; stabilization; periodic solution; synchronization of solutions

For citation: Kirillov A. N., Sazonov A. M. Controlled dynamics of the biological wastewater treatment process under variable pollution input flow. *Trudy Karelskogo nauchnogo tsentra RAN = Transactions of the Karelian Research Centre RAS*. 2026. No. 6. P. 54–60. doi: 10.17076/mat2326

Funding. The studies were funded from the federal budget through state assignment to the Karelian Research Centre RAS (Institute of Applied Mathematical Research, Karelian Research Centre RAS).

ВВЕДЕНИЕ

Математическое моделирование динамики процесса биологической очистки сточных вод на основе применения активного ила является актуальной задачей в области охраны окружающей среды [1–9]. Сложность процесса очистки сточных вод связана с различными аспектами биологического, физико-химического, технологического характера. Попытки учесть в предлагаемых моделях как можно большее число параметров и связей между ними приводят к невозможности их аналитического и качественного исследования.

В данной работе предложена упрощенная качественная модель системы биологической очистки сточных вод. Предложенная модель описывает систему биоочистки, состоящую из аэротенка, отстойника и звена рециркуляции. На вход системы подается субстрат-загрязнитель и окисляющий его активный ил (биомасса). Аэротенк состоит из n последовательно соединенных компартментов, имеющих номера i , $i = 1, \dots, n$, биомасса и субстрат переходят из i -го компартмента в $i + 1$ -й. При этом происходит процесс окисления субстрата биомассой, в результате чего концентрация субстрата уменьшается, а биомассы – растет с увеличением i .

В статье представлено развитие исследований [2] для переменного входного потока субстрата-загрязнителя. Изменение входного потока субстрата можно интерпретировать, например, как различную нагрузку на очистную систему в зависимости от времени суток, поскольку ночью нагрузка меньше, или как увеличение нагрузки вследствие некоторой

нестандартной критической ситуации. Естественно суточные изменения считать периодическими, а критические – однократными.

Доказана теорема о существовании периодического решения для общего вида трофических функций в модели биоочистки, описанной в [2], определен его конкретный вид. На основе доказанной теоремы решена задача стабилизации процесса биоочистки. Поставлена и решена задача управления с целью локализации периодического решения. Показано, что для любых начальных данных решения системы приближаются к периодическому.

ДИНАМИКА ПРОЦЕССА БИООЧИСТКИ

Постановка задачи

Рассмотрим систему, описывающую процесс биологической очистки сточных вод, являющуюся многомерным обобщением модели [4]:

$$\begin{cases} \dot{x}_1 = au + f(x_1, s_1) - (u + b)x_1, \\ \dot{s}_1 = R(t) - \frac{1}{Y}f(x_1, s_1) - (u + b)s_1, \\ \dot{x}_i = f(x_i, s_i) + (u + b)x_{i-1} - (u + b)x_i, \\ \dot{s}_i = -\frac{1}{Y}f(x_i, s_i) + (u + b)s_{i-1} - (u + b)s_i, \end{cases} \quad (1)$$

где $s_i(t), x_i(t)$ – концентрации субстрата-загрязнителя и микроорганизмов активного ила в i -м компартменте соответственно, $f(x_i, s_i)$ – трофическая функция, $i = 1, \dots, n$, Y – коэффициент утилизации субстрата-загрязнителя в биомассу микроорганизмов, a – концентрация микроорганизмов на входе, b, u – скорости субстрата и ила на входе.

При этом входной поток субстратов $R(t)$ полагается кусочно-постоянным:

$$R(t) = \begin{cases} R_1, t \in [2kT, (2k+1)T], \\ R_2, t \in [(2k+1)T, 2(k+1)T], \\ k = 0, 1, \dots \end{cases}$$

где $0 < R_2 < R_1$ – постоянные, $2T > 0$ – период входного потока. Отметим, что такой вид $R(t)$ отражает суточные изменения входного потока загрязнений: больший поток R_1 наблюдается в дневное время, а меньший R_2 – в ночное.

Управлением является скорость u поступления ила на входе. В силу технических ограничений, присущих процессу биоочистки, полагаем, что $u = u(t)$ кусочно-постоянная функция такая, что $u(t) \in [0, M]$, где $M > 0$ – постоянная. При этом временные промежутки постоянства $u(t)$ совпадают с промежутками постоянства $R(t)$:

$$u(t) = \begin{cases} u_1, t \in [2kT, (2k+1)T], \\ u_2, t \in [(2k+1)T, 2(k+1)T], \\ k = 0, 1, \dots \end{cases}$$

Поставим задачу стабилизации процесса биоочистки, то есть нахождения управления $u(t) \in [0, M]$ такого, что

$$s_n(t, u) \leq c, \quad x_n(t, u) \leq r, \quad t \geq 0, \quad (2)$$

где $s_n(t, u), x_n(t, u)$ – концентрации субстрата и микроорганизмов в n -м (последнем) компартменте, соответствующие управлению u , c, r – положительные постоянные. Условия (2) отражают требования к соблюдению допустимых ограничений на потоки субстрата и активного ила на выходе из системы биоочистки, которые определяют постоянные c и r соответственно. Отметим, что по сути это задача об устойчивости по Лагранжу.

Периодическое решение

Обозначим $D = au + YR$ входной поток субстрата-загрязнителя и активного ила. Проведя замену переменных в системе (1) $z_i = x_i + Ys_i$, получим линейную систему вида

$$\begin{cases} \dot{z}_1 = D - (u + b)z_1 = g_1(z), \\ \dot{z}_i = (u + b)z_{i-1} - (u + b)z_i = g_i(z), \\ i = 2, \dots, n. \end{cases} \quad (3)$$

В данном параграфе будет показано существование периодического решения у системы с переключениями (3). Наличие периодического решения у системы позволяет говорить о некотором цикле процесса биоочистки, то есть стационарном режиме этого процесса.

Теорема 1. У системы с переключениями (3) существует периодическое решение.

Доказательство. Рассмотрим систему (3) при $n = 2$.

На промежутке $t \in [0, T]$ она имеет следующее решение:

$$\begin{cases} z_1(t) = E_1 + (z_1^0 - E_1)e^{-\gamma_1 t}, \\ z_2(t) = E_1 + ((z_1^0 - E_1)\gamma_1 t + (z_2^0 - E_1))e^{-\gamma_1 t}, \end{cases}$$

а на промежутке $t \in [T, 2T]$:

$$\begin{cases} z_1(t) = E_2 + (z_1^1 - E_2)e^{-\gamma_2(t-T)}, \\ z_2(t) = E_2 + ((z_1^1 - E_2)\gamma_2(t-T) + (z_2^1 - E_2))e^{-\gamma_2(t-T)}, \end{cases}$$

где

$$E_j = \frac{D_j}{u_j + b}, \quad D_j = au_j + YR_j, \quad \gamma_j = u_j + b, \\ z_j^0 = z_j(0), \quad z_j^1 = z_j(T), \quad j = 1, 2.$$

Решение системы с переключениями (3) с начальной точкой $z^0 = (z_1^0, z_2^0)$ будет периодическим с периодом $2T$, если будут выполнены следующие равенства:

$$z_1(0) = z_1(2T), \quad z_2(0) = z_2(2T).$$

Тогда для определения начальной точки получим следующую систему:

$$\begin{cases} z_1^1 = E_1 + (z_1^0 - E_1)e^{-\gamma_1 T}, \\ z_2^1 = E_1 + ((z_1^0 - E_1)\gamma_1 T + (z_2^0 - E_1))e^{-\gamma_1 T}, \\ z_1^0 = E_2 + (z_1^1 - E_2)e^{-\gamma_2 T}, \\ z_2^0 = E_2 + ((z_1^1 - E_2)\gamma_2 T + (z_2^1 - E_2))e^{-\gamma_2 T}. \end{cases} \quad (4)$$

Решая систему (4), получаем

$$z_1^0 = \frac{E_2(1 - e^{-\gamma_2 T}) + E_1 e^{-\gamma_2 T}(1 - e^{-\gamma_1 T})}{1 - e^{-(\gamma_1 + \gamma_2)T}},$$

$$\begin{aligned} z_2^0 &= \frac{1}{1 - e^{-(\gamma_1 + \gamma_2)T}} (z_1^0(\gamma_1 + \gamma_2)T e^{-(\gamma_1 + \gamma_2)T} \\ &+ E_1 e^{-\gamma_2 T}(1 - e^{-\gamma_1 T} + \gamma_2 T - (\gamma_1 + \gamma_2)T e^{-\gamma_1 T}) \\ &+ E_2(1 - (\gamma_2 T + 1)e^{-\gamma_2 T})). \end{aligned} \quad (5)$$

При $T > 0$ имеем $z_1^0 > 0$. Покажем, что $z_2^0 > 0$. Из (5) имеем

$$\begin{aligned} z_2^0 &= \frac{1}{1 - e^{-(\gamma_1 + \gamma_2)T}} (z_1^0(\gamma_1 + \gamma_2)T e^{-(\gamma_1 + \gamma_2)T} \\ &+ E_1 e^{-\gamma_2 T}(1 - (\gamma_1 T + 1)e^{-\gamma_1 T} + \gamma_2 T(1 - e^{-\gamma_1 T})) \\ &+ E_2(1 - (\gamma_2 T + 1)e^{-\gamma_2 T})). \end{aligned} \quad (6)$$

Для функции $f(x) = (x+1)e^{-x}$ имеем

$$f(0) = 1, \quad f'(x) = -xe^{-x} < 0 \quad \forall x > 0,$$

поэтому $f(x) < 1, \forall x > 0$. Таким образом, из (6) имеем $z_2^0 > 0$ при $T > 0$.

Для доказательства теоремы в n -мерном случае требуется аналогично повторить данную процедуру, последовательно решая системы относительно $z_i, i = 3, \dots, n$. $\square$

Согласно теореме 1, если подобрать u_1, u_2 так, что $\max z_n^j = \max(x_n^j + Ys_n^j) \leq L$ на $[0, 2T]$, где $L = \min(c, r)$, если $Y \geq 1$, или $\frac{L}{Y} = \min(c, r)$, если $Y < 1$, то для таких u_1, u_2 будет выполнено (2), то есть решена задача стабилизации.

ЛОКАЛИЗАЦИЯ ПЕРИОДИЧЕСКОГО РЕШЕНИЯ

Рассмотрим задачу локализации периодического решения $z_i^{per}(z^0, t)$, состоящую в определении управлений $u_1 \geq 0, u_2 \geq 0$, обеспечивающих выполнение неравенств:

$$c_1 \leq z_i^{per}(z^0, t) \leq c_2,$$

где c_1, c_2 – положительные постоянные.

Теорема 2. Пусть $E_1 < E_2$. Приведенные ниже управления u_1, u_2 решают задачу локализации периодического решения, то есть $c_1 \leq z_i^{per}(z^0, t) \leq c_2$.

Если $c_1 < c_2 < a, R_2 < R_1 \leq \frac{c_1 b}{Y}$, то

$$\begin{cases} \frac{Y R_1 - c_1 b}{c_1 - a} \leq u_1 < \frac{c_2 b - Y R_1}{a - c_2}, \\ \frac{(Y R_2 - ab)u_1 + Y b(R_2 - R_1)}{Y R_1 - ab} < u_2 \leq \frac{c_2 b - Y R_2}{a - c_2}. \end{cases}$$

Если $c_1 < c_2 < a, \frac{c_1 b}{Y} < R_1 < \frac{c_2 b}{Y}, R_2 \leq \frac{c_1 b}{Y}$, то

$$\begin{cases} 0 \leq u_1 < \frac{c_2 b - Y R_1}{a - c_2}, \\ \frac{(Y R_2 - ab)u_1 + Y b(R_2 - R_1)}{Y R_1 - ab} < u_2 \leq \frac{c_2 b - Y R_2}{a - c_2}. \end{cases}$$

Если $c_1 < a < c_2, R_2 < R_1 \leq \frac{c_1 b}{Y}$, то

$$\begin{cases} u_1 \geq \frac{Y R_1 - c_1 b}{c_1 - a}, \\ u_2 > \frac{(Y R_2 - ab)u_1 + Y b(R_2 - R_1)}{Y R_1 - ab}. \end{cases}$$

Если $c_1 < a < c_2, \frac{c_1 b}{Y} < R_1 < \frac{ab}{Y}, R_2 \leq \frac{ab}{Y}$, то

$$\begin{cases} u_1 \geq 0, \\ u_2 > \frac{(Y R_2 - ab)u_1 + Y b(R_2 - R_1)}{Y R_1 - ab}. \end{cases}$$

Если $c_1 < a < c_2, R_1 > \frac{ab}{Y}, \frac{ab}{Y} < R_2 \leq \frac{c_2 b}{Y}$, то

$$\begin{cases} u_1 > \frac{Y b(R_1 - R_2)}{Y R_2 - ab}, \\ 0 \leq u_2 < \frac{(Y R_2 - ab)u_1 + Y b(R_2 - R_1)}{Y R_1 - ab}. \end{cases}$$

Если $c_1 < a < c_2, R_1 > R_2 > \frac{c_2 b}{Y}$, то

$$\begin{cases} u_1 > \frac{c_2 b - Y R_1}{a - c_2}, \\ \frac{c_2 b - Y R_2}{a - c_2} \leq u_2 < \frac{(Y R_2 - ab)u_1 + Y b(R_2 - R_1)}{Y R_1 - ab}. \end{cases}$$

Если $a < c_1 < c_2, R_1 > \frac{c_1 b}{Y}, \frac{ab}{Y} < R_2 \leq \frac{c_2 b}{Y}$, то

$$\begin{cases} \frac{Y b(R_1 - R_2)}{Y R_2 - ab} < u_1 \leq \frac{Y R_1 - c_1 b}{c_1 - a}, \\ 0 \leq u_2 < \frac{(Y R_2 - ab)u_1 + Y b(R_2 - R_1)}{Y R_1 - ab}. \end{cases}$$

Если $a < c_1 < c_2, R_1 > \frac{c_1 b}{Y}, R_2 > \frac{c_2 b}{Y}$, то

$$\begin{cases} \frac{c_2 b - Y R_1}{a - c_2} < u_1 \leq \frac{Y R_1 - c_1 b}{c_1 - a}, \\ \frac{c_2 b - Y R_2}{a - c_2} \leq u_2 < \frac{(Y R_2 - ab)u_1 + Y b(R_2 - R_1)}{Y R_1 - ab}. \end{cases}$$

В остальных случаях решение не локализуемо.

Доказательство. Так как по условию $E_1 < E_2$, то имеем

$$(Y R_1 - ab)u_2 < Y R_2(u_1 + b) - b(au_1 + Y R_1). \quad (7)$$

Из (3) очевидно, что множество $\{E_1 \leq z_i \leq E_2\}$ инвариантно для системы (3). Тогда из теоремы 1 имеем, что $E_1 \leq z_i^{per}(z^0, t) \leq E_2$, поскольку периодическая траектория не может покинуть инвариантное множество. Таким образом, чтобы решить задачу локализации периодического решения, требуются u_1, u_2 такие, что $c_1 \leq E_1, E_2 \leq c_2$. Итак, получим

$$\begin{cases} c_1 \leq \frac{au_1 + Y R_1}{u_1 + b}, \\ \frac{au_2 + Y R_2}{u_2 + b} \leq c_2. \end{cases} \quad (8)$$

Перепишем (8) в виде

$$\begin{cases} (c_1 - a)u_1 \leq Y R_1 - c_1 b, \\ (a - c_2)u_2 \leq c_2 b - Y R_2. \end{cases} \quad (9)$$

Рассмотрим случай $c_1 < c_2 < a$. Из (9)

$$\begin{cases} u_1 \geq \frac{Y R_1 - c_1 b}{c_1 - a}, \\ u_2 \leq \frac{c_2 b - Y R_2}{a - c_2}. \end{cases} \quad (10)$$

Пусть $R_2 < R_1 \leq \frac{c_1 b}{Y}$. Тогда из (7)

$$u_2 > \frac{u_1(Y R_2 - ab) + Y b(R_2 - R_1)}{Y R_1 - ab} > 0. \quad (11)$$

Поскольку для выполнения второго неравенства из (10) и неравенства (11) необходимо

$$\frac{u_1(Y R_2 - ab) + Y b(R_2 - R_1)}{Y R_1 - ab} < \frac{c_2 b - Y R_2}{a - c_2},$$

получим ограничение на u_1 :

$$u_1 < \frac{c_2 b - Y R_1}{a - c_2}.$$

Несложно показать, что при указанных условиях в п. 1 теоремы 2

$$0 < \frac{Y R_1 - c_1 b}{c_1 - a} < \frac{c_2 b - Y R_1}{a - c_2}.$$

Таким образом, доказаны ограничения на управления в первом случае.

Пусть $R_1 > \frac{c_1 b}{Y}$. Поскольку $\frac{Y R_1 - c_1 b}{c_1 - a} < 0$, из (10)

$$\begin{cases} u_1 \geq 0, \\ u_2 \leq \frac{c_2 b - Y R_2}{a - c_2}. \end{cases}$$

Учитывая (7),

$$\begin{aligned} \frac{u_1(Y R_2 - ab) + Y b(R_2 - R_1)}{Y R_1 - ab} &< u_2 \\ &\leq \frac{c_2 b - Y R_2}{a - c_2}, \end{aligned}$$

следовательно,

$$u_1 < \frac{c_2 b - Y R_1}{a - c_2}.$$

Можно показать, что при условиях п. 2 теоремы 2 имеют место следующие неравенства:

$$\begin{aligned} \frac{c_2 b - Y R_i}{a - c_2} &> 0, \quad i = 1, 2, \\ \frac{u_1(Y R_2 - ab) + Y b(R_2 - R_1)}{Y R_1 - ab} &> 0. \end{aligned}$$

Следовательно, доказаны соответствующие условия на управления.

В случае $R_1 > \frac{c_1 b}{Y}$, $R_2 > \frac{c_2 b}{Y}$ из (10) имеем $u_2 < 0$, что противоречит условию неотрицательности управления. Случай $R_1 < \frac{c_1 b}{Y}$, $R_2 > \frac{c_2 b}{Y}$ невозможен, поскольку $R_2 < R_1$.

Анализ других случаев $a < c_1 < c_2$, $c_1 < a < c_2$ аналогичен. $\square$

Замечание. Отметим, что в теореме 1 рассматривается случай $E_1 < E_2$. Анализ случая $E_1 \geq E_2$ полностью аналогичен и с целью избежать громоздких выкладок, ухудшающих восприятие статьи, не приводится в работе.

СИНХРОНИЗАЦИЯ РЕШЕНИЙ

Теорема 3. Для любых начальных точек $z^0, \tilde{z}^0$ таких, что $\tilde{z}_i^0 = z_i^0 + \delta$, $i = 1, 2$, $\delta > 0$: $|z_i(\tilde{z}^0, T) - z_i(z^0, T)| < \delta$, $i = 1, \dots, n$, где $R(t) = \text{const}$, $t \in [kT, (k+1)T]$, $k = 0, 1, \dots$

Доказательство. Используя выражение для решения, имеем

$$\begin{aligned} |z_1(\tilde{z}^0, T) - z_1(z^0, T)| &= |E_1 + (z_1^0 + \delta - E_1)e^{-(u+b)T} - E_1 - (z_1^0 - E_1)e^{-(u+b)T}| \\ &= \delta e^{-(u+b)T} < \delta. \end{aligned}$$

$$\begin{aligned} |z_2(\tilde{z}^0, T) - z_2(z^0, T)| &= |E_1 + ((z_1^0 + \delta - E_1)(u+b)T + (z_2^0 + \delta - E_1))e^{-(u+b)T} - E_1 - ((z_1^0 - E_1)(u+b)T + (z_2^0 - E_1))e^{-(u+b)T}| \\ &= \delta((u+b)T + 1)e^{-(u+b)T}. \end{aligned} \tag{12}$$

Как было показано в доказательстве теоремы 1, справедливо $(x+1)e^{-x} < 1$, $\forall x > 0$. Таким образом, из (12), $|z_2(\tilde{z}^0, T) - z_2(z^0, T)| < \delta$. Для $i = 3, \dots, n$ доказательство аналогично вышеизложенному. $\square$

Следствие. Если периодическое решение является устойчивым, то происходит стабилизация решения. Таким образом, периодическое решение является притягивающим.

Далее приведен результат численного эксперимента для случая $n = 2$, отражающий утверждение теоремы 3 относительно синхронизации решений. Здесь $a = 1$, $b = 1$, $Y = 1$, $R_1 = 2$, $R_2 = 1$, $u_1 = 1$, $u_2 = 2$, $z^0 = (2, 1)$, $\tilde{z}^0 = (1.4, 1.6)$. Отметим, что решения для данных начальных точек $z^0, \tilde{z}^0$ не периодические.

Как видно из приведенных численных результатов (рис. 1, 2), решения системы приближаются друг к другу.

ЗАКЛЮЧЕНИЕ

В работе представлена модель управления системой биологической очистки осадков сточных вод с трофическими функциями общего вида. Поставлена задача стабилизации процесса биоочистки при кусочно-постоянном входном потоке загрязнений. Показано существование периодического решения для вспомогательной системы, полученной при помощи замены переменных, определен его конкретный вид. Получен вид управлений, позволяющий локализовать периодическое решение. Данный результат свидетельствует о цикличности процесса биоочистки и позволяет решить задачу стабилизации процесса биоочистки для исходной системы. Показано, что для любых начальных данных происходит синхронизация решений системы. Проведен численный эксперимент, подтверждающий этот результат.

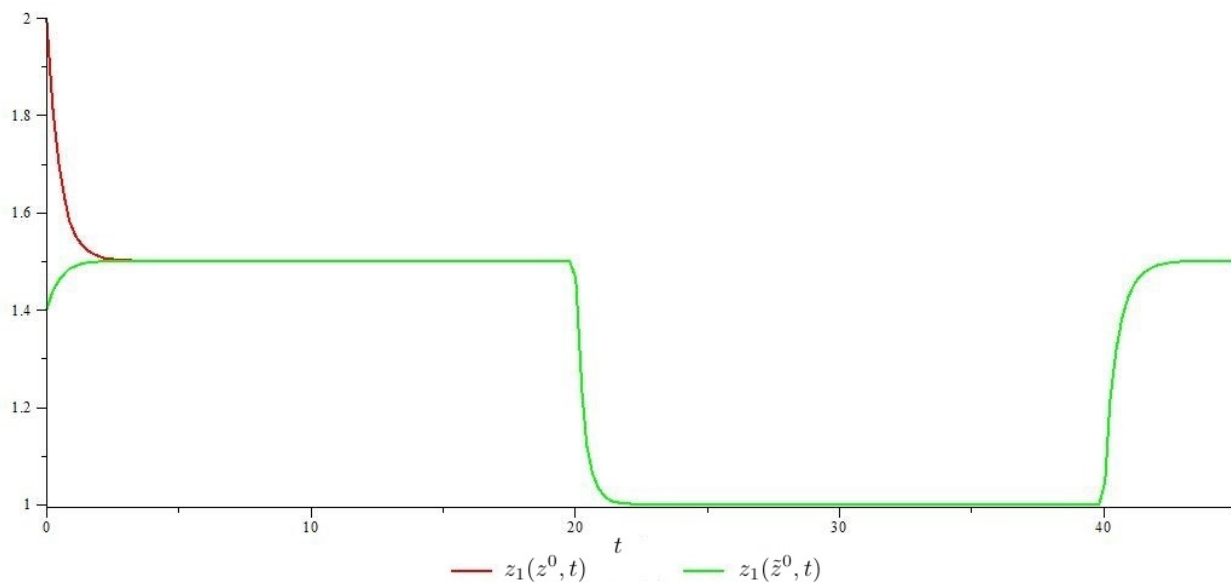


Рис. 1. Приближение компонент решений $z_1(z^0, t)$ и $z_1(\tilde{z}^0, t)$ системы (3)
 Fig. 1. The convergence of solutions components $z_1(z^0, t)$ and $z_1(\tilde{z}^0, t)$ of the system (3)

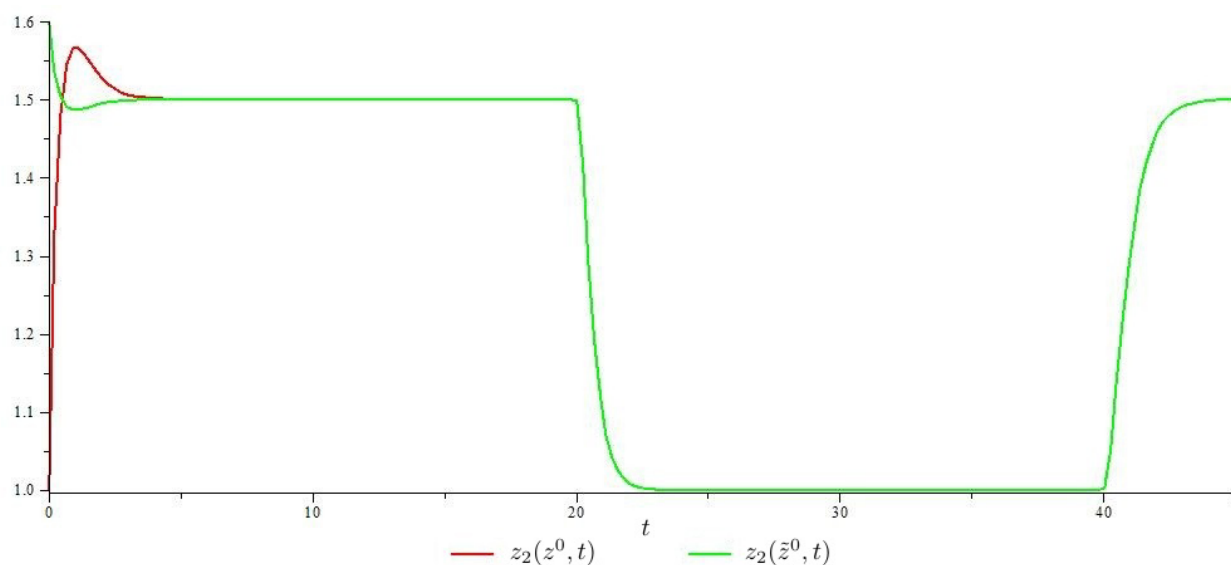


Рис. 2. Приближение компонент решений $z_2(z^0, t)$ и $z_2(\tilde{z}^0, t)$ системы (3)
 Fig. 2. The convergence of solutions components $z_2(z^0, t)$ and $z_2(\tilde{z}^0, t)$ of the system (3)

ЛИТЕРАТУРА

1. Вавилин В. А. Нелинейные модели биологической очистки и процессов самоочищения в реках. М.: Наука, 1983. 185 с.
2. Кириллов А. Н., Данилова И. В. Динамика процесса биоочистки при переменном входном потоке загрязнений: инвариантные множества и стабилизация // Труды Карельского научного центра РАН. 2020. № 7. С. 65–71. doi: 10.17076/mat1230
3. Кириллов А. Н., Сазонов А. М. Моделирование процесса биологической очистки сточных

вод на основе шумпетеровской динамики // Труды ИСА РАН. 2020. Т. 70, № 3. С. 24–28. doi: 10.14357/20790279200303

4. Brune D. Optimal control of the complete-mix activated sludge process // Environmental Technology Letters. 1985. Vol. 6. P. 467–476. doi: 10.1080/09593338509384365

5. Dupont R., Sinkjaer O. Optimization of wastewater treatment plants by means of computer models // Water Sci. Technol. 1994. Vol. 30. P. 181–190.

6. Grigorieva E. V., Khailov E. N. Minimization of pollution concentration on a given time interval for the wastewater cleaning plant // J. Control Sci. Eng. 2010. Vol. 2010. Art. 712794. doi: 10.1155/2010/712794

7. Henze M., Grady Jr. C.P.L., Gujer W., Marais G.v.R., Matsuo T. A general model for single-sludge activated sludge wastewater treatment systems // Water Res. 1987. Vol. 21. P. 505–515. doi: 10.1016/0043-1354(87)90058-3

8. Steffens M. A., Lant P. A., Newell R. B. A systematic approach for reducing biological wastewater treatment models // Water Res. 1997. Vol. 31. P. 590–606.

9. Tacacs I., Patry G. G., Nolasco D. A dynamic model of the clarification-thickening process // Water Res. 1991. Vol. 25. P. 1263–1271. doi: 10.1016/0043-1354(91)90066-Y

REFERENCES

1. Vavilin V. A. Nonlinear models of the biological wastewater treatment and the processes of the self-purification in rivers. Moscow: Nauka; 1983. 185 p. (In Russ.)

2. Danilova I. V., Kirillov A. N. Dynamics of the biological wastewater treatment process under variable pollution input flow: invariant sets and stabilization. *Trudy Karel'skogo nauchnogo tsentra RAN = Transactions of the Karelian Research Centre RAS*. 2020;7:67–71. (In Russ.). doi: 10.17076/mat1230

3. Kirillov A. N., Sazonov A. M. The modeling of the biological wastewater treatment process on the base of the Schumpeterian dynamics. *Trudy ISA RAN = Transactions of the ISA RAS*. 2020;70(3):24–28. (In Russ.). doi: 10.14357/20790279200303

4. Brune D. Optimal control of the complete-mix activated sludge process. *Environmental Technology Letters*. 1985;6:467–476. doi: 10.1080/09593338509384365

5. Dupont R., Sinkjaer O. Optimization of wastewater treatment plants by means of computer models. *Water Sci. Technol*. 1994;30:181–190.

6. Grigorieva E. V., Khailov E. N. Minimization of pollution concentration on a given time interval for the wastewater cleaning plant. *J. Control Sci. Eng*. 2010;2010:712794. P. 1–10. doi: 10.1155/2010/712794

7. Henze M., Grady Jr. C.P.L., Gujer W., Marais G.v.R., Matsuo T. A general model for single-sludge activated sludge wastewater treatment systems. *Water Res*. 1987;21:505–515. doi: 10.1016/0043-1354(87)90058-3

8. Steffens M. A., Lant P. A., Newell R. B. A systematic approach for reducing biological wastewater treatment models. *Water Res*. 1997;31:590–606.

9. Tacacs I., Patry G. G., Nolasco D. A dynamic model of the clarification-thickening process. *Water Res*. 1991;25:1263–1271. doi: 10.1016/0043-1354(91)90066-Y

Поступила в редакцию / received: 11.03.2026; принята к публикации / accepted: 29.05.2026.

Авторы заявляют об отсутствии конфликта интересов / The authors declare no conflict of interest.

СВЕДЕНИЯ ОБ АВТОРАХ:

Кириллов Александр Николаевич

д-р физ.-мат. наук, ведущий научный сотрудник

e-mail: kirillov@krc.karelia.ru

Сазонов Александр Михайлович

канд. физ.-мат. наук, младший научный сотрудник

e-mail: sazon-tb@mail.ru

CONTRIBUTORS:

Kirillov, Alexander

Dr. Sci. (Phys.-Math.), Leading Researcher

Sazonov, Alexander

Cand. Sci. (Phys.-Math.), Junior Researcher

УДК 004.01:006.72 (470.22)

ПОТЕНЦИАЛ В НЕКООПЕРАТИВНОЙ ИГРЕ УПРАВЛЕНИЯ ЗАДАНИЯМИ С ЛИНЕЙНЫМИ ЭКСТЕРНАЛИЯМИ

Н. Н. Кукушкина

*Институт прикладных математических исследований КарНЦ РАН,
ФИЦ «Карельский научный центр РАН» (ул. Пушкинская, 11,
Петрозаводск, Республика Карелия, Россия, 185910)*

Представлена и исследована теоретико-игровая математическая модель управления заданиями в вычислительной системе с линейными экстерналиями. Каждый игрок стремится максимизировать свою производительность в присутствии положительных экстерналий. Предложенная модель описывает управление заданиями в системе добровольных вычислений типа Desktop Grid, а экстерналии выражают обмен информацией в процессе решения научной задачи. Системы Desktop Grid задействуют вычислительные мощности неспециализированных компьютеров, объединенных сетью передачи данных и выполняющих вычисления в то время, когда они не заняты другой работой. В таких системах вычислительноемкая научная задача зачастую делится на несколько взаимосвязанных подзадач, выполняемых параллельно. Промежуточные результаты решения одной подзадачи способны помогать в решении остальных подзадач, оптимизируя или сокращая вычисления. В данной работе доказано, что игра является потенциальной в случаях однородных процессоров или заданий и одинаковых экстерналий, инвариантных или симметричных экстерналий. При этом в случаях с однородными процессорами или заданиями и одинаковыми экстерналиями или инвариантными экстерналиями равновесие по Нэшу единственно и глобально оптимально, в то время как случай с симметричными экстерналиями допускает множество равновесий по Нэшу в чистых стратегиях. Приводятся результаты вычислительных экспериментов по моделированию управления заданиями предложенным методом в сравнении с рядом популярных алгоритмов. Представленные результаты расширяют область применения классических моделей, доказывая существование равновесия и сходимости к нему как в задачах минимизации задержки, так и в задачах максимизации производительности. Найденный вид функции потенциала позволяет использовать методы глобальной оптимизации для поиска равновесий.

Ключевые слова: управление заданиями; задача покрытия машин; равновесие по Нэшу; потенциал; линейные экстерналии

Для цитирования: Кукушкина Н. Н. Потенциал в некооперативной игре управления заданиями с линейными экстерналиями // Труды Карельского научного центра РАН. 2026. № 6. С. 61–72. doi: 10.17076/mat2352

Финансирование. Финансовое обеспечение исследований осуществлялось из средств федерального бюджета на выполнение государственного задания КарНЦ РАН (Институт прикладных математических исследований КарНЦ РАН).

N. N. Kukushkina. POTENTIAL IN A NON-COOPERATIVE TASK SCHEDULING GAME WITH LINEAR EXTERNALITIES

Institute of Applied Mathematical Research, Karelian Research Centre, Russian Academy of Sciences (11 Pushkinskaya St., 185910 Petrozavodsk, Karelia, Russia)

This paper investigates a game-theoretic mathematical model of task scheduling in a computational system with linear externalities. Each player aims to maximize their performance in the presence of positive externalities. The proposed model describes task scheduling in a volunteer computing system of Desktop Grid type, while the externalities express the information exchange in the process of solving a scientific problem. Desktop Grid systems employ computational capacities of non-dedicated computers that are united by a data transmission network and perform computations in their idle time. In such systems, a computationally intensive scientific problem is often divided into several interconnected subproblems to be performed in parallel. The intermediate results of solving one subproblem can aid in solving other subproblems by optimizing or reducing the computations. Proof is provided that the game is a potential game in cases of homogeneous processors or tasks and equal externalities, invariant or symmetric externalities. In the cases with homogeneous processors or tasks and equal externalities or invariant externalities, the Nash equilibrium is unique and globally optimal, while the case with symmetric externalities allows multiple Nash equilibriums in pure strategies. We present the results of computational experiments modeling task scheduling using the proposed method in comparison with a number of popular algorithms. The provided results expand the applicability of classical models, proving the existence of an equilibrium and convergence to it in both delay minimization and performance maximization problems. The potential function form found here facilitates the use of global optimization methods to search for equilibriums.

Keywords: task scheduling; machine covering problem; Nash equilibrium; potential; linear externalities

For citation: Kukushkina N. N. Potential in a non-cooperative task scheduling game with linear externalities. *Trudy Karel'skogo nauchnogo tsentra RAN = Transactions of the Karelian Research Centre RAS*. 2026. No. 6. P. 61–72. doi: 10.17076/mat2352

Funding. The studies were funded from the federal budget through state assignment to the Karelian Research Centre RAS (Institute of Applied Mathematical Research, Karelian Research Centre RAS).

ВВЕДЕНИЕ

В современной теории расписаний и математической теории игр особое место занимают игры распределения ресурсов. В играх такого типа эгоистичные участники претендуют на ограниченные ресурсы, и стратегия каждого игрока определяется не только его действиями, но и действиями конкурентов. Математические модели игр распределения ресурсов применяются в различных прикладных областях: например, для управления трафиком в сетях передачи данных, распределения нагрузки в облачных сервисах, оптимизации логистических цепочек и т. д.

В частности, рассматриваются игры покрытия машин [13, 18], в которых эгоистичные участники стремятся оптимизировать использование ресурсов, распределяя свои задачи по машинам (процессорам) таким образом, что-

бы максимизировать минимальную задержку. При этом в классической модели функция выигрыша или затрат игрока зависит исключительно от состояния выбранного им ресурса, а не от действий игроков на других машинах.

В дополнение к классической модели рассматриваются также модели с экстерналиями – эффектами, выражающими влияние текущей загрузки каждого процессора на производительность каждого из остальных процессоров. В игре покрытия машин такое влияние обусловлено необходимостью координированного управления совокупностью вычислительных заданий с сопутствующим обменом данными между процессорами.

В общем случае в играх распределения ресурсов отрицательные экстерналии позволяют моделировать эффект конкуренции [1, 4, 8], а положительные – эффекты взаимовыгодно-

сти, такие как, например, обмен информацией между агентами, разделение затрат и др. [4].

В литературе целый ряд математических моделей берут начало из КР-модели [6] с параллельными каналами различной емкости. В работе [1] рассмотрена задача максимизации минимальной задержки по процессорам в присутствии линейных экстерналий. Доказано существование равновесия по Нэшу и показано, что в общем случае цена анархии конечна. При этом в общем случае равновесие по Нэшу может отсутствовать, а алгоритм лучших ответов – закидываться.

Рассматривались также задачи с экстерналиями в сетевой постановке [2, 4, 7]. В таких моделях взаимное влияние игроков друг на друга определяется топологией сети.

В данной работе исследована математическая модель управления заданиями с линейными экстерналиями. В отличие от работы [1], каждый игрок стремится максимизировать свою производительность в присутствии экстерналий, оказывающих положительный эффект. Модель описывает управление заданиями в системе добровольных вычислений. Экстерналии выражают обмен информацией в процессе решения научной задачи, и игроки, действуя эгоистично, в равновесии поступают оптимальным образом, чтобы решить общую задачу. Доказано, что при соблюдении ряда условий игра является потенциальной и, таким образом, в ней существует равновесие по Нэшу в чистых стратегиях, которого игроки достигают при помощи алгоритма лучших или наилучших ответов.

Предложенная модель описывает вычислительные системы, в которых большая задача разбита на несколько взаимосвязанных подзадач, что характерно для научного поиска. Получаемые в динамике результаты одной подзадачи косвенно помогают в решении остальных, например, сужая пространство поиска.

При этом рассматриваются системы добровольных вычислений, участники которых – не зависящие друг от друга агенты, стремящиеся максимизировать свой собственный выигрыш, отдавая свои мощности на решение той или иной задачи. Благодаря структуре математической модели агенты делят между собой пространство поиска в одном и том же блоке, а от остальных блоков тем временем поступает полезная информация, что увеличивает производительность системы. Таким образом, интенсивность поиска в соседних блоках имеет значение и позволяет игрокам оптимизировать собственные вычисления.

УПРАВЛЕНИЕ ЗАДАНИЯМИ В СИСТЕМАХ ДОБРОВОЛЬНЫХ ВЫЧИСЛЕНИЙ

Научные задачи тесно связаны с генерацией и анализом больших объемов данных, что требует специальных алгоритмов управления заданиями в современных гетерогенных вычислительных системах [12, 15]. Активно используются Desktop Grid – распределенные системы неспециализированных компьютеров, объединенных сетью передачи данных и выполняющих вычисления в то время, когда они не заняты другой работой. В Desktop Grid владельцы компьютеров добровольно предоставляют свои вычислительные мощности для решения той или иной научной задачи. На сегодняшний день существует множество проектов добровольных вычислений в различных областях науки; среди примеров можно отметить Einstein@home [3] (астрофизика), World Community Grid [17] (биология и медицина), PrimeGrid [11] (математика) и другие.

Системы типа Desktop Grid наиболее хорошо подходят для решения задач, требующих значительных вычислительных мощностей и при этом поддающихся разделению на большое число независимых рабочих заданий, каждое из которых можно за сравнительно небольшое время выполнить на обычном компьютере (задачи типа Bag-of-Tasks). Подобные задачи часто возникают в научных исследованиях: например, проверка свойств математических объектов в комбинаторике [14] и криптографии [19], проведение экспериментов *in silico* по большим базам данных [10, 16] и т. д.

Несмотря на то что в отдельных ситуациях в системах добровольных вычислений может возникать конкуренция за ресурсы сервера [5], в общем случае приход нового участника увеличивает производительность системы для всех: если участников станет больше, то научная задача будет решена быстрее.

При этом вычислительная научная задача зачастую логически подразделяется на несколько подзадач, которые могут выполняться параллельно друг с другом. Промежуточные результаты решения одной подзадачи способны помогать в решении остальных подзадач, оптимизируя или сокращая вычисления. Это отражает специфику научного поиска. Поясним эту идею на примере виртуального скрининга лекарств. Виртуальный скрининг – это перебор миллионов малых молекул *in silico* и выбор среди них кандидатов в лекарства, удовлетворяющих определенным условиям. Пусть библиотека молекул, среди которых ищутся кандидаты в лекарства, разделена на несколько кластеров и в каждом кластере

одновременно проводится виртуальный скрининг. Если в некотором кластере обнаружены бесперспективные (например, ввиду токсичности) молекулы, содержащие определенный структурный фрагмент, то аналогичные молекулы можно исключить из обработки и в других кластерах. Таким образом, обмен промежуточными данными и статистикой поиска способен ускорить работу всей системы в целом. При этом обнаружение конкретного целевого результата носит случайный характер и может быть редким событием.

На рис. 1 показана общая схема вычислительного процесса в Desktop Grid: научная задача делится на подзадачи, а подзадачи – на рабочие задания, каждое из которых может выполняться на любом вычислительном узле независимо от других рабочих заданий.

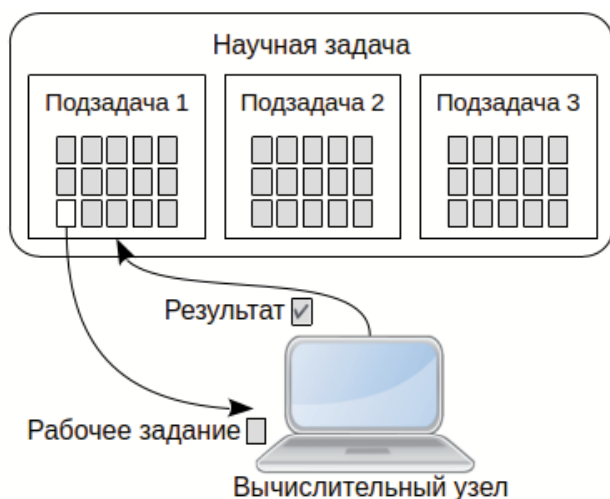


Рис. 1. Вычислительный процесс в Desktop Grid
Fig. 1. Computational process in a Desktop Grid

Далее рассмотрим математическую модель управления заданиями в Desktop Grid. Докажем, что в случаях однородных процессоров или заданий и одинаковых экстерналий, инвариантных или симметричных экстерналий игра является потенциальной и имеет равновесие по Нэшу в чистых стратегиях.

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ

Рассмотрим вычислительную систему $S = S(\mathcal{N}, p, e)$ со множеством заданий $\mathcal{N}$, вероятностями получения полезных результатов p и множеством экстерналий e . Множество заданий $\mathcal{N}$ разбито на n кластеров (для краткости будем далее называть их блоками) $N_1, \dots, N_n$. Обозначим как $N = \{1, 2, \dots, n\}$ множество

индексов блоков заданий; $N_i \cap N_k = \emptyset \forall i \neq k$, $i, k \in N$, $\bigcup_{i=1}^n N_i = \mathcal{N}$.

Для каждого блока N_i задана вероятность получения полезных результатов в данном блоке p_i , $0 < p_i \leq 1$. В контексте виртуального скрининга качество результата может выражаться, например, предсказанной энергией Гиббса связывания молекулы с мишенью (ккал/моль). Чем ниже отрицательное значение энергии, тем сильнее молекула теоретически взаимодействует с мишенью. Тогда полезным результатом может считаться хит – молекула с предсказанной энергией связывания ниже заданного порога. Вероятность получения полезных результатов p_i будет выше в блоках, содержащих молекулы небольшого размера, простой формы или схожих с известными лигандами для того же класса мишеней. Поскольку процесс виртуального скрининга носит стохастический характер, а молекулы внутри одного блока схожи между собой, можно оценить значение p_i , проверив небольшую выборку молекул из блока N_i .

Коэффициент e_{ik} , где $i, k \in N, i \neq k$, выражает вклад загрузки блока заданий N_k в производительность вычислений в блоке N_i .

Задания выполняются процессорами из множества $U = U(M, w)$, которые имеют скорости $w_1, w_2, \dots, w_m$ (количество выполняемых заданий в единицу времени), где $m = |M|$.

Обозначим как $W = \sum_{j=1}^m w_j$ суммарную скорость всех процессоров.

Каждый процессор из множества M выбирает блок, в котором будет выполнять задания.

В данной модели процессоры можно считать игроками. Стратегия игрока j – это выбор блока заданий l_j . Тогда профиль стратегий представим в виде вектора $L = (l_1, \dots, l_m)$. Обозначим суммарную мощность всех процессоров, выбравших блок заданий i , как $\delta_i(L) = \sum_{\substack{j \in M: \\ l_j = i}} w_j$. Производительность вычислений в блоке заданий i имеет вид

$$\lambda_i(L) = \delta_i(L)p_i + \sum_{k \neq i} e_{ik}\delta_k(L). \quad (1)$$

Выражение (1) представляет собой производительность любого из игроков, выбравших блок заданий N_i , поскольку получаемые общими усилиями полезные результаты имеют одинаковую ценность для всех игроков.

Таким образом, имеется некооперативная игра $\Gamma = \langle S(\mathcal{N}, p, e), U(M, w), \lambda \rangle$. Профиль стратегий является равновесием по Нэшу, ес-

ли в нем ни одному игроку не выгодно в одностороннем порядке отклониться от выбранной стратегии. Формально профиль стратегий L – равновесие по Нэшу тогда и только тогда, когда для любого игрока $j \in M$ выполняется неравенство $\lambda_j(L) \geq \lambda_i(L_{j:l_j \rightarrow i})$ для всех блоков заданий $i \in N$.

Будем полагать, что выполнены условия, вытекающие из физического смысла поведения вычислительной системы [1]:

$$e_{ik} > p_i \quad \forall i \neq k, \quad i, k \in N \quad (2)$$

$$e_{ki} > p_i \quad \forall i \neq k, \quad i, k \in N \quad (3)$$

$$\sum_{l \neq i} e_{il} \geq \sum_{l \neq k} e_{kl} \quad \forall i \neq k : p_i \leq p_k, \quad i, k \in N. \quad (4)$$

Условия (2) и (3) означают, что блоки заданий сильно взаимосвязаны и значимость их взаимодействия превышает вероятность получения полезного результата внутри одного блока изолированно. Условие (4) означает, что влияние результатов других блоков на более продуктивный блок оказывается выше, чем на менее продуктивный блок.

Известно, что в игре с конечным числом стратегий наличие потенциальной функции гарантирует существование равновесия по Нэшу в чистых стратегиях [9]. По определению потенциальной игры, если один игрок меняет стратегию и увеличивает свой выигрыш («лучший ответ»), то значение потенциальной функции тоже строго увеличивается. В точке максимума потенциала ни один игрок не может увеличить значение функции, изменив свою стратегию, а значит, не может увеличить и свой собственный выигрыш. Это и есть определение равновесия по Нэшу. При этом в точных потенциальных играх изменение значения потенциала в точности совпадает с изменением выигрыша отклонившегося игрока, во взвешенных – совпадает с точностью до множителя – веса игрока. Каждая точка локального максимума потенциальной функции является равновесием по Нэшу в чистых стратегиях.

Рассмотрим варианты игры Γ , в которых всегда существует равновесие по Нэшу.

Игра с однородными процессорами и одинаковыми экстерналиями

Рассмотрим игру $\Gamma = \langle S(\mathcal{N}, p, e), U(M, w), \lambda \rangle$. Пусть $w_1 = w_2 = \dots = w_m = w$ и $e_{ik} = e$ для всех i и k . Обозначим число игроков, выбравших блок i , как m_i . Тогда производительность в блоке i примет вид

$$\lambda_i(L) = w(m_i(p_i - e) + em).$$

Теорема 1. *Игра Γ с однородными процессорами и одинаковыми экстерналиями имеет точный потенциал*

$$\begin{aligned} P(L) &= \sum_{i=1}^n \sum_{k=1}^{m_i} wk(p_i - e) \\ &= \frac{w}{2} \sum_{i=1}^n (p_i - e)(m_i^2 + m_i). \end{aligned} \quad (5)$$

Доказательство. Предположим, что некоторый процессор $j \in M$ переходит с блока s в блок t и профиль стратегий меняется с L на L' . Изменение затрат процессора j составит

$$\Delta C_j = w(m_t + 1)(p_t - e) - wm_s(p_s - e).$$

Вычислим изменение значения потенциала.

$$\begin{aligned} \Delta P &= \sum_{k=1}^{m_t+1} wk(p_t - e) + \sum_{k=1}^{m_s-1} wk(p_s - e) \\ &\quad - \sum_{k=1}^{m_t} wk(p_t - e) - \sum_{k=1}^{m_s} wk(p_s - e) \\ &= w(m_t + 1)(p_t - e) - wm_s(p_s - e), \end{aligned}$$

что совпадает с выражением для ΔC_j . Следовательно, игра является потенциальной. $\square$

Следствие 1. *Непрерывное расширение функции $P(L)$ (5) на область $\mathbb{R}_+^n$ является строгой вогнутой функцией, т. е. квадратичная форма отрицательно определена в силу условия (2). Поэтому игра имеет единственное равновесие по Нэшу (с точностью до дискретности переменных m_i), которое соответствует глобальному максимуму функции $P(L)$ в точке m^* с координатами*

$$m_i^* = \frac{m + n/2}{(p_i - e) \sum_{k=1}^n \frac{1}{p_k - e}}.$$

Игра с однородными блоками заданий и одинаковыми экстерналиями

Рассмотрим игру $\Gamma = \langle S(\mathcal{N}, p, e), U(M, w), \lambda \rangle$. Пусть $p_1 = p_2 = \dots = p_n = p$ и $e_{ik} = e$ для всех i и k .

Теорема 2. *Игра Γ с однородными блоками заданий и одинаковыми экстерналиями имеет взвешенный потенциал*

$$P(L) = \frac{1}{2} (p - e) \sum_{i=1}^n \delta_i^2(L). \quad (6)$$

Доказательство. Предположим, что некоторый процессор $j \in M$ переходит с блока заданий s на блок заданий t и профиль стратегий меняется с L на L' . Изменение производительности процессора j составит

$$\begin{aligned}\Delta C_j &= -e\delta_t(L') + \delta_t(L')p + e\delta_s(L) - \delta_s(L)p \\ &= -e(\delta_t(L) + w_j) + (\delta_t(L) + w_j)p + e\delta_s(L) \\ &\quad - \delta_s(L)p = (p - e)\delta_t(L) - (p - e)\delta_s(L) \\ &\quad + (p - e)w_j = (p - e)(\delta_t(L) - \delta_s(L) + w_j).\end{aligned}$$

Вычислим изменение значения потенциала. Для этого заметим, что все слагаемые (6) для $i \neq s, t$ остаются неизменными, меняются лишь значения слагаемых для $i = s$ и $i = t$.

$$\begin{aligned}\Delta P &= \frac{1}{2}(p - e)(\delta_t^2(L') - \delta_t^2(L) + \delta_s^2(L') \\ &\quad - \delta_s^2(L)) = \frac{1}{2}(p - e)((\delta_t(L) + w_j)^2 - \delta_t^2(L) \\ &\quad + (\delta_s(L) - w_j)^2 - \delta_s^2(L)) = \frac{1}{2}(p - e) \\ &\quad \times (2\delta_t(L)w_j + w_j^2 - 2\delta_s(L)w_j + w_j^2) \\ &= w_j(p - e)(\delta_t(L) - \delta_s(L) + w_j),\end{aligned}$$

что с точностью до множителя совпадает с выражением для ΔC_j . Следовательно, игра является потенциальной. $\square$

Следствие 2. Непрерывное расширение функции $P(L)$ (6) на область $\mathbb{R}_+^n$ является строго вогнутой функцией как сумма квадратов с отрицательным множителем. Поэтому игра имеет единственное равновесие по Нэшу (с точностью до дискретности переменных δ_i), которое соответствует глобальному максимуму функции $P(L)$ в точке δ^* с координатами

$$\delta_1^* = \delta_2^* = \dots = \delta_n^* = \frac{W}{n}.$$

Игра с инвариантными экстерналиями

Рассмотрим игру $\Gamma = \langle S(\mathcal{N}, p, e), U(M, w), \lambda \rangle$. Пусть экстерналии инвариантны: для любого $i \in N$ влияние загрузки в блоке заданий N_i на производительность во всех остальных блоках одинаково, то есть $e_{ik} = e_{jk} = e_k \forall i, j, k \in N, k \neq i, k \neq j$.

Теорема 3. Игра Γ с инвариантными экстерналиями имеет взвешенный потенциал

$$P(L) = \sum_{i=1}^n \left(\frac{p_i - e_i}{2} \delta_i^2(L) \right.$$

$$\left. + \frac{p_i + e_i}{2} \sum_{k:l_k=i} w_k^2 \right). \quad (7)$$

Доказательство. Предположим, что некоторый процессор $j \in M$ переходит с блока s в блок t и профиль стратегий меняется с L на L' . Изменение затрат процессора j составит

$$\Delta C_j = \delta_s(L)(e_s - p_s) + \delta_t(L)(p_t - e_t) + w_j(p_t - e_s).$$

Вычислим изменение значения потенциала.

$$\begin{aligned}\Delta P &= \frac{p_s - e_s}{2}(\delta_s^2(L) - 2\delta_s(L)w_j + w_j^2) \\ &\quad + \frac{p_t - e_t}{2}(\delta_t^2(L) + 2\delta_t(L)w_j + w_j^2) \\ &\quad + \frac{p_s - e_s}{2} \sum_{\substack{k:l_k(L)=s, \\ k \neq j}} w_j^2 + \frac{p_t - e_t}{2} \sum_{\substack{k:l_k(L)=t, \\ k \neq j}} w_j^2 \\ &= (e_s - p_s)\delta_s(L)w_j + \frac{p_s - e_s}{2}w_j^2 - \frac{p_s + e_s}{2}w_j^2 \\ &\quad + (p_t - e_t)\delta_t(L)w_j + \frac{p_t - e_t}{2}w_j^2 + \frac{p_t + e_t}{2}w_j^2 \\ &= w_j(\delta_s(L)(e_s - p_s) + \delta_t(L)(p_t - e_t) + w_j(p_t - e_s)),\end{aligned}$$

что совпадает с выражением для ΔC_j . Следовательно, игра является потенциальной. $\square$

Следствие 3. Непрерывное расширение функции $P(L)$ (7) на область $\mathbb{R}_+^n$ является строго вогнутой функцией по переменным δ_i , т. к. квадратичная форма отрицательно определена в силу условия (2). Поэтому игра имеет единственное равновесие по Нэшу (с точностью до дискретности переменных δ_i), которое соответствует глобальному максимуму функции $P(L)$ в точке δ^* с координатами

$$\delta_i^* = \frac{W}{(p_i - e_i) \sum_{k=1}^n \frac{1}{p_k - e_k}}.$$

Игра с симметричными экстерналиями

Рассмотрим игру $\Gamma = \langle S(\mathcal{N}, p, e), U(M, w), \lambda \rangle$. Пусть экстерналии симметричны: $e_{ik} = e_{ki} \forall i, k \in N, k \neq i$.

Теорема 4. Игра Γ с симметричными экстерналиями имеет взвешенный потенциал

$$\begin{aligned}P(L) &= \sum_{i=1}^n \frac{p_i}{2} \left(\delta_i^2(L) + \sum_{j:l_j=i} w_j^2 \right) \\ &\quad + \sum_{i < k} e_{ik} \delta_i(L) \delta_k(L). \quad (8)\end{aligned}$$

Доказательство. Предположим, что некоторый процессор $j \in M$ переходит с блока s в блок t и профиль стратегий меняется с L на L' . Изменение затрат процессора j составит

$$\Delta C_j = \delta_s(L)(e_{ts} - p_s) + \delta_t(L)(p_t - e_{st}) + w_j(p_t - e_{ts}) + \sum_{k \neq s, t} \delta_k(L)(e_{tk} - e_{sk}).$$

Вычислим изменение значения потенциала.

$$\begin{aligned} \Delta P = & \frac{p_s}{2}(\delta_s(L) - w_j)^2 + \frac{p_s}{2} \sum_{j: l_j(L')=s} w_j^2 \\ & + \frac{p_t}{2}(\delta_t(L) + w_j)^2 + \frac{p_t}{2} \sum_{j: l_j(L')=t} w_j^2 \\ & + \sum_{i < k} e_{ik} \delta_i(L') \delta_k(L') - \frac{p_s}{2} \delta_s^2(L) - \frac{p_t}{2} \delta_t^2(L) \\ & - \frac{p_s}{2} \sum_{j: l_j(L)=s} w_j^2 - \frac{p_t}{2} \sum_{j: l_j(L)=t} w_j^2 \\ & - \sum_{i < k} e_{ik} \delta_i(L) \delta_k(L) = w_j(\delta_s(L)(e_{ts} - p_s) \\ & + \delta_t(L)(p_t - e_{st}) + w_j(p_t - e_{ts}) \\ & + \sum_{k \neq s, t} \delta_k(L)(e_{tk} - e_{sk})) \end{aligned}$$

что с точностью до множителя совпадает с выражением для ΔC_j . Следовательно, игра является потенциальной. $\square$

Следствие 4. В игре с симметричными экстремальными может существовать несколько различных равновесий по Нэшу, в каждом из которых производительности во всех задействованных блоках заданий равны (с точностью до дискретности переменных):

$$\lambda_1^* = \lambda_2^* = \dots = \lambda_n^*.$$

Доказательство. Запишем непрерывное расширение функции $P(L)$ (8) на область $\mathbb{R}_+^n$ в матричной форме:

$$P(L) \approx \frac{1}{2} \delta^T A \delta,$$

где элементы матрицы A : $A_{ik} = e_{ik}$, $A_{ii} = p_i$. Поскольку $e_{ik} > p_i$, собственные числа матрицы могут принимать как отрицательные, так и положительные значения, и может существовать несколько локальных экстремумов потенциала. Найдем точки экстремума, запи-

сав условия Лагранжа:

$$\begin{cases} p_1 \delta_1 + e_{12} \delta_2 + \dots + e_{1n} \delta_n = \eta \\ e_{21} \delta_1 + p_2 \delta_2 + \dots + e_{2n} \delta_n = \eta \\ \dots \\ e_{n1} \delta_1 + e_{n2} \delta_2 + \dots + p_n \delta_n = \eta \\ \sum_{i=1}^n \delta_i = W \end{cases}$$

Заметим, что в левой части каждого из первых n уравнений стоит формула производительности вычислений в блоке i , λ_i (1). Видно, что в точке экстремума потенциала производительности во всех задействованных блоках равны между собой: $\lambda_1 = \lambda_2 = \dots = \lambda_n = \eta$. $\square$

Из теорем 1–4 следует, что алгоритм лучших ответов A1 сходится к равновесию по Нэшу для любого из перечисленных вариантов игры управления заданиями Γ , причем в первых трех вариантах – к единственному равновесию, а в четвертом – к одному из равновесий, если их несколько. В записи алгоритма используем следующие структуры данных:

- массив $L = (l_1, \dots, l_m)$ хранит номера блоков, выбранных игроками $1, \dots, m$. Назовем его расписанием;
- $random_schedule(\Gamma)$ – это любое выбранное случайным образом расписание в игре Γ ;
- $L_{j:s \rightarrow t}$ – это расписание, полученное из расписания L переходом игрока j с блока s на блок t . Стратегии других игроков при этом не меняются.

Алгоритм управления заданиями

Алгоритм управления заданиями A1 начинает работу с расписания L , выбранного случайным образом. Далее в цикле по всем игрокам $j \in M$ проверяется существование лучшего ответа (то есть строго выгодного перехода $s \rightarrow t$) для текущего игрока, и если такой ответ найден, то игрок совершает его. Цикл лучших ответов продолжается до тех пор, пока в установившемся расписании ни один игрок не сможет никуда перейти с выгодой для себя. Алгоритм A1.

```

1:  $L \leftarrow random\_schedule(\Gamma)$ 
2:  $moveFound \leftarrow false$ 
3: loop
4:   for  $j \in M$  do
5:      $s \leftarrow L[j]$ 
6:     for  $t \in N, t \neq s$  do
7:       if  $C_j(L_{j:s \rightarrow t}) > C_j(L)$  then
8:          $L \leftarrow L_{j:s \rightarrow t}$ 

```

```

9:         moveFound  $\leftarrow$  true
10:        break
11:    end if
12: end for
13: if moveFound then
14:     break
15: end if
16: end for
17: if !moveFound then
18:     break
19: end if
20: end loop
21: return  $L$ 

```

ВЫЧИСЛИТЕЛЬНЫЕ ЭКСПЕРИМЕНТЫ

Пример 1. Рассмотрим игру Γ с инвариантными экстерналиями и разнородными игроками. Пусть $n = 10$, $m = 1000$, $e_1 = 0.3$, $e_2 = 0.32$, $e_3 = 0.33$, $e_4 = e_5 = e_6 = e_7 = 0.35$, $e_8 = 0.4$, $e_9 = 0.42$, $e_{10} = 0.45$; $p_1 = 0.01$, $p_2 = 0.025$, $p_3 = 0.05$, $p_4 = 0.075$, $p_5 = 0.1$, $p_6 = 0.125$, $p_7 = 0.15$, $p_8 = 0.175$, $p_9 = 0.2$, $p_{10} = 0.225$. Веса игроков сгенерированы случайным образом в диапазоне от 1 до 100.

На рис. 2 приведена динамика сходимости алгоритма A1 к равновесию по Нэшу в игре Γ .

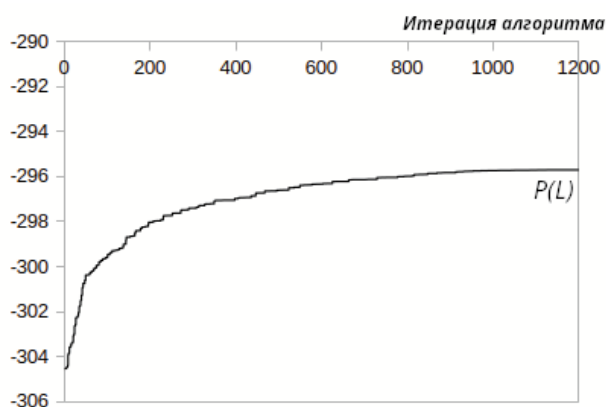


Рис. 2. Динамика сходимости алгоритма A1 к равновесию по Нэшу в игре с инвариантными экстерналиями ($n = 10$, $m = 1000$)

Fig. 2. Dynamics of convergence to the Nash equilibrium in a game with invariant externalities ($n = 10$, $m = 1000$)

В полученном равновесии по Нэшу

$$\delta = (4317, 4244, 4471, 4561, 5008, 6261, 5565, 5567, 5695, 5567),$$

что, с учетом дискретности переменных, совпадает с округленным аналитическим решением из следствия 3

$$\delta^* = (4318.8, 4245.6, 4473.0, 4554.4, 5009.8, 6262.2, 5566.4, 5566.4, 5692.9, 5566.4).$$

Заметим, что в играх с неоднородными блоками заданий и различными экстерналиями в равновесии по Нэшу, в зависимости от значений параметров модели, производительность вычислений в некотором блоке t может оказаться нулевой, так как ни одному игроку не выгодно выбрать этот блок. Таких блоков может быть и несколько (но не более n). Это обусловлено спецификой модели (1)–(4): при определенном соотношении параметров всем игрокам будет выгоднее работать в блоке заданий с высоким p_i , чем распределиться по разным блокам заданий и помогать друг другу через обмен информацией. В итоге задания блока t не будет выполнять никто. Чтобы подобной ситуации не возникало, можно действовать разными способами, например:

- 1) Выполнить фиксированное количество заданий из блоков, выбранных игроками, а затем запретить выбор этих блоков, оставив лишь не выбранные ранее блоки. Тогда результаты из всех блоков будут поступать последовательно, множество доступных блоков заданий будет меняться на каждом этапе, а в остальном игра останется неизменной.
- 2) Направить в каждый блок вычислительные мощности фиксированного объема (например, собственные компьютеры администраторов проекта добровольных вычислений) и запретить им переходить в другие блоки. Тогда поток результатов из всех блоков будет обеспечен базовыми мощностями $\delta_i^0 > 0$, суммарные мощности δ_i в формуле (1) будут вычисляться с учетом базовых мощностей, а в остальной игре останется неизменной.
- 3) Изменить параметры математической модели, увеличив вознаграждения участников за работу в непопулярных блоках или снизив их за работу в популярных блоках (увеличить p_t или уменьшить p_i). Значения параметров подобрать так, чтобы равновесное расписание заданий стало более сбалансированным.

Пример 2. Рассмотрим игру Γ с одинаковыми экстерналиями и однородными игроками. Пусть $n = 10$, $e = 0.23$, $w = 5$, а частоты получения полезных результатов p_i такие же, как в примере 1. На рис. 3 приведена зависимость числа блоков с $\delta_i(L) > 0$ в равновесии от числа игроков в игре Γ . Здесь начальное расписание

выбрано случайным образом и зафиксировано на протяжении всей серии экспериментов. Большинство игроков в равновесии выбирают блок 10 с минимальной разницей $|p_i - e|$. График показывает, что при малых m есть никем не выбранные блоки заданий, но при $m > 61$ выбраны все блоки заданий.

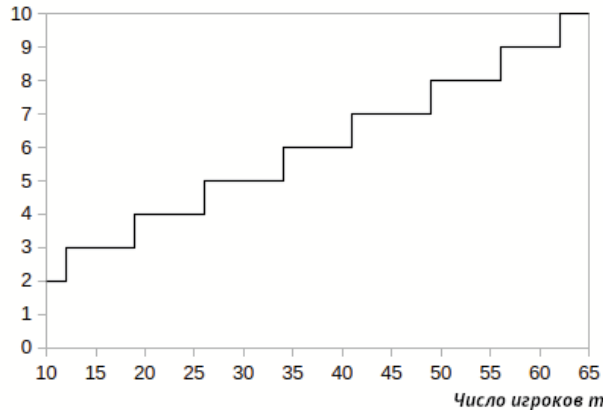


Рис. 3. Число блоков с $\delta_i > 0$ в равновесии в игре с одинаковыми экстерналиями и однородными игроками в зависимости от числа игроков m

Fig. 3. The number of the blocks with $\delta_i > 0$ in the equilibrium in the game with equal externalities and homogeneous players, depending on the number of players m

Пример 3. Рассмотрим игру Γ с симметричными экстерналиями, разнородными игроками и блоками заданий. Пусть $n = 3$, $m = 20$, $p_1 = 0.01$, $p_2 = 0.02$, $p_3 = 0.2$, $e_{12} = e_{21} = 0.5$, $e_{13} = e_{31} = 0.25$, $e_{23} = e_{32} = 0.25$. Веса игроков сгенерированы случайным образом в диапазоне от 1 до 500.

В общем случае равновесие по Нэшу может быть неоптимальным с точки зрения системы в целом, поскольку игроки преследуют эгоистические цели и действуют независимо друг от друга. В зависимости от целей вычислительного проекта (скорейшее получение всех результатов, скорейшее получение первых полезных результатов и т. д.) оптимальным может оказаться централизованное назначение заданий игрокам. Рассмотрим две меры эффективности системы и оценим эффективность равновесий по Нэшу.

Пусть $E_1(L) = \sum_{i=1}^n \lambda_i(L)$ – полнота охвата подзадач, которые кто-либо выполняет; $E_2(L) = \sum_{i=1}^n \delta_i(L) \lambda_i(L)$ – суммарная производительность всех вычислительных узлов.

Введем меры эффективности

$$E_1^*(L) = \frac{E_1(L)}{\max_L E_1(L)}, \quad E_2^*(L) = \frac{E_2(L)}{\max_L E_2(L)}.$$

Сравним поведение системы при различных алгоритмах управления заданиями:

- L_{NE} – расписание в равновесии по Нэшу, найденное при помощи алгоритма A1.
- L_1 – расписание, найденное последовательным «жадным» назначением процессоров на блоки заданий: каждый игрок выбирает для себя лучший вариант, исходя из текущего состояния системы.
- L_2 – расписание, найденное последовательной «жадной» максимизацией минимума: каждый игрок назначается в тот блок, где текущая производительность минимальна.
- L_3 – расписание, найденное алгоритмом Round Robin с учетом частот получения полезных результатов и экстерналий.
- L_4 – расписание, найденное случайным назначением процессоров на блоки.

Максимально возможная эффективность системы найдена полным перебором профилей стратегий и составляет $\max_L E_1(L) = 3\,956.76$, $\max_L E_2(L) = 6\,803\,728$. В таблице приведена эффективность системы, описанной в примере 3, усредненная по 1000 экспериментов.

Эффективность вычислительной системы при различных алгоритмах управления заданиями (пример 3)

Efficiency of the computational system with different task scheduling algorithms (example 3)

	L_{NE}	L_1	L_2	L_3	L_4
E_1^*	0.884	0.971	0.972	0.973	0.965
E_2^*	0.995	0.971	0.390	0.974	0.924

При этом в равновесиях по Нэшу ни один игрок не выбирает третий блок. Таким образом, эгоистичное поведение игроков приводит к неоптимальной полноте охвата блоков заданий, но обеспечивает максимальную суммарную производительность (это следует из вида функции потенциала). Если целью является гарантированное получение промежуточных результатов всех подзадач, необходимо использовать механизмы балансировки, описанные выше. Например, чтобы сбалансировать загрузку и улучшить эффективность в

равновесии, назначим в каждый блок фиксированные вычислительные мощности по $\delta_i^0 = 100$ единиц каждая. Скорректируем меру эффективности с учетом введенных мощностей:

$$E'_1(L) = E_1(L) - \sum_{i=1}^n (\delta_i^0 p_i + \sum e_{ik} \delta_k^0).$$

Теперь усредненная эффективность системы с эгоистичными игроками близка к единице: $E'_1(L_{NE}) = 0.9897$.

ЗАКЛЮЧЕНИЕ

В статье представлена и исследована математическая модель управления заданиями с линейными экстерналиями, где каждый игрок стремится максимизировать свою производительность в присутствии экстерналий, оказывающих положительный эффект. Экстерналии выражают обмен информацией в процессе решения научной задачи, например, в системе добровольных вычислений.

Доказано, что игра является потенциальной в случаях однородных процессоров или заданий и одинаковых экстерналий, инвариантных или симметричных экстерналий. Доказано также, что в случаях с однородными процессорами или заданиями и одинаковыми экстерналиями или инвариантными экстерналиями равновесие по Нэшу единственно и глобально оптимально, в то время как случай с симметричными экстерналиями допускает множество равновесий, что требует применения алгоритма лучших ответов. Приводятся результаты вычислительных экспериментов по моделированию управления заданиями.

Представленные результаты расширяют область применения классических моделей, доказывая существование равновесия и сходямость к нему как в задачах минимизации задержки, так и в задачах максимизации производительности. Найденный вид функции потенциала позволяет использовать методы глобальной оптимизации для поиска равновесий, что эффективнее прямого перебора стратегий.

ЛИТЕРАТУРА

1. *Chirkova J. V.* Maximizing the minimum processor load with linear externalities // Mathematical Optimization Theory and Operations Research: Recent Trends. MOTOR 2021: 20th International Conference. CCIS. Vol. 1476 / A. Strekalovsky, Yu. Kochetov, T. Gruzdeva, A. Orlov (eds.). Cham: Springer, 2021. P. 147–162. doi: 10.1007/978-3-030-86433-0_10
2. *Chirkova J. V.* Potential game in general transport network with symmetric externalities

// Mathematical Optimization Theory and Operations Research. MOTOR 2024. Lecture Notes in Computer Science. Vol. 14766 / A. Ereemeev, M. Khachay, Yu. Kochetov, V. Mazalov, P. Pardalos (eds.). Cham: Springer, 2024. P. 231–242. doi: 10.1007/978-3-031-62792-7_16

3. Einstein@Home. URL: <https://einsteinathome.org/> (дата обращения: 18.03.2026).

4. *Etesami S. R.* Maximizing social welfare subject to network externalities: a unifying submodular optimization approach // IEEE Trans. Network Sci. Eng. 2024. Vol. 11(5). P. 4860–4874. doi: 10.1109/TNSE.2024.3397188

5. *Ivashko E., Nikitina N.* Characterization of a desktop grid project as a queueing system // Supercomputing. RuSCDays 2024. Lecture Notes in Computer Science. Vol. 15407 / V. Voevodin, A. Antonov, D. Nikitenko (eds.). 2025. P. 32–43. doi: 10.1007/978-3-031-78462-0_3

6. *Koutsoupias E., Papadimitriou C.* Worst-case equilibria // STACS 1999. Lecture Notes in Computer Science. Vol. 1563 / C. Meinel, S. Tison (eds.). Heidelberg: Springer, 1999. P. 404–413. doi: 10.1007/3-540-49116-3_38

7. *Kuang Z., Mazalov V., Tang X., Zheng J.* Transportation network with externalities // J. Comput. Appl. Math. 2021. Vol. 382. Art. 113091. doi: 10.1016/j.cam.2020.113091

8. *Liao W., Wang L., Li J.* Congestion game with inter-cell interference for cell selection in heterogeneous cellular network // 2014 IEEE/CIC International Conference on Communications in China (ICCC). 2014. P. 603–608. doi: 10.1109/ICCCChina.2014.7008348

9. *Monderer D., Shapley L. S.* Potential games // Games and economic behavior. 1996. Vol. 14, iss. 1. P. 124–143. doi: 10.1006/game.1996.0044

10. *Nikitina N., Manzyuk M., Podlipnik Č., Jukić M.* Volunteer computing project SiDock@home for virtual drug screening against SARS-CoV-2 // IFIP Advances in Information and Communication Technology. 2021. Vol. 616. P. 23–34. doi: 10.1007/978-3-030-86582-5_3

11. PrimeGrid. URL: <https://www.primegrid.com/> (дата обращения: 18.03.2026).

12. *Quang B. T., Kim J. S., Rho S., Kim S., Kim S., Hwang S., Medernach E., Breton V.* A comparative analysis of scheduling mechanisms for virtual screening workflow in a shared resource environment // 2015 15th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing. 2015. P. 853–862. doi: 10.1109/CCGrid.2015.123

13. *Tan Z., Wan L., Zhang Q., Ren W.* Inefficiency of equilibria for the machine covering game on uniform machines // Acta Informatica. 2012. Vol. 49. P. 361–379. doi: 10.1007/s00236-012-0163-1

14. Vatutin E., Belyshev A., Nikitina N., Manzhuk M., Albertian A., Kurochkin I., Kripachev A., Pykhtin A. Diagonalization and canonization of Latin squares // *Supercomputing. RuSCDays 2023. Lecture Notes in Computer Science*. Vol. 14389 / V. Voevodin, S. Sobolev, M. Yakobovskiy, R. Shagaliyev (eds.). 2023. P. 48–61. doi: 10.1007/978-3-031-49435-2_4
15. Vivas A., Tchernykh A., Castro H. Trends, approaches, and gaps in scientific workflow scheduling: a systematic review // *IEEE Access*. 2024. Vol. 12. P. 182203–182231. doi: 10.1109/ACCESS.2024.3509218
16. Voelz V. A., Pande V. S., Bowman G. R. Folding@home: Achievements from over 20 years of citizen science herald the exascale era // *Biophys. J.* 2023. Vol. 122, iss. 14. P. 2852–2863. doi: 10.1016/j.bpj.2023.03.028
17. World Community Grid. URL: <https://www.worldcommunitygrid.org/> (дата обращения: 18.03.2026).
18. Wu Y., Cheng T. C. E., Ji M. Inefficiency of the Nash equilibrium for selfish machine covering on two hierarchical uniform machines // *Inf. Process. Lett.* 2015. Vol. 115. P. 838–844. doi: 10.1016/j.ipl.2015.06.005
19. Zaikin O. SAT-Based cryptanalysis: from parallel computing to volunteer computing // *Supercomputing. RuSCDays 2019. Communications in Computer and Information Science*. Vol. 1129 / V. Voevodin, S. Sobolev (eds.). 2019. P. 701–712. doi: 10.1007/978-3-030-36592-9_57

REFERENCES

1. Chirkova J. V. Maximizing the minimum processor load with linear externalities. *Mathematical Optimization Theory and Operations Research: Recent Trends. MOTOR 2021: 20th International Conference. CCIS*. Vol. 1476. Cham: Springer; 2021. P. 147–162. doi: 10.1007/978-3-030-86433-0_10
2. Chirkova J. V. Potential game in general transport network with symmetric externalities. *Mathematical Optimization Theory and Operations Research. MOTOR 2024. Lecture Notes in Computer Science*. Vol. 14766. Cham: Springer; 2024. P. 231–242. doi: 10.1007/978-3-031-62792-7_16
3. Einstein@Home. URL: <https://einsteinathome.org/> (accessed: 18.03.2026).
4. Etesami S. R. Maximizing social welfare subject to network externalities: a unifying submodular optimization approach. *IEEE Trans. Network Sci. Eng.* 2024;11(5):4860–4874. doi: 10.1109/TNSE.2024.3397188
5. Ivashko E., Nikitina N. Characterization of a desktop grid project as a queueing system. *Supercomputing. RuSCDays 2024. Lecture Notes in Computer Science*. Vol. 15407. 2025. P. 32–43. doi: 10.1007/978-3-031-78462-0_3

6. Koutsoupias E., Papadimitriou C. Worst-case equilibria. *STACS 1999. Lecture Notes in Computer Science*. Vol. 1563. Heidelberg: Springer; 1999. P. 404–413. doi: 10.1007/3-540-49116-3_38
7. Kuang Z., Mazalov V., Tang X., Zheng J. Transportation network with externalities. *J. Comput. Appl. Math.* 2021;382:113091. doi: 10.1016/j.cam.2020.113091
8. Liao W., Wang L., Li J. Congestion game with inter-cell interference for cell selection in heterogeneous cellular network. *2014 IEEE/CIC International Conference on Communications in China (ICCC)*. 2014. P. 603–608. doi: 10.1109/ICCChina.2014.7008348
9. Monderer D., Shapley L. S. Potential games. *Games and economic behavior*. 1996;14(1):124–143. doi: 10.1006/game.1996.0044
10. Nikitina N., Manzyuk M., Podlipnik Č., Jukić M. Volunteer computing project SiDock@home for virtual drug screening against SARS-CoV-2. *IFIP Advances in Information and Communication Technology*. 2021;616:23–34. doi: 10.1007/978-3-030-86582-5_3
11. PrimeGrid. URL: <https://www.primegrid.com/> (accessed: 18.03.2026).
12. Quang B. T., Kim J. S., Rho S., Kim S., Kim S., Hwang S., Medernach E., Breton V. A comparative analysis of scheduling mechanisms for virtual screening workflow in a shared resource environment. *2015 15th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*. 2015. P. 853–862. doi: 10.1109/CCGrid.2015.123
13. Tan Z., Wan L., Zhang Q., Ren W. Inefficiency of equilibria for the machine covering game on uniform machines. *Acta Informatica*. 2012;49:361–379. doi: 10.1007/s00236-012-0163-1
14. Vatutin E., Belyshev A., Nikitina N., Manzhuk M., Albertian A., Kurochkin I., Kripachev A., Pykhtin A. Diagonalization and canonization of Latin squares. *Supercomputing. RuSCDays 2023. Lecture Notes in Computer Science*. Vol. 14389. 2023. P. 48–61. doi: 10.1007/978-3-031-49435-2_4
15. Vivas A., Tchernykh A., Castro H. Trends, approaches, and gaps in scientific workflow scheduling: a systematic review. *IEEE Access*. 2024;12:182203–182231. doi: 10.1109/ACCESS.2024.3509218
16. Voelz V. A., Pande V. S., Bowman G. R. Folding@home: Achievements from over 20 years of citizen science herald the exascale era. *Biophys. J.* 2023;122(14):2852–2863. doi: 10.1016/j.bpj.2023.03.028
17. World Community Grid. URL: <https://www.worldcommunitygrid.org/> (accessed: 18.03.2026).
18. Wu Y., Cheng T. C. E., Ji M. Inefficiency of the Nash equilibrium for selfish machine covering on

two hierarchical uniform machines. *Inf. Process. Lett.* 2015;115:838–844. doi: 10.1016/j.ipl.2015.06.005

19. *Zaikin O.* SAT-Based cryptanalysis: from parallel computing to volunteer computing. *Super-*

computing. RuSCDays 2019. Communications in Computer and Information Science. Vol. 1129. 2019. P. 701–712. doi: 10.1007/978-3-030-36592-9_57

Поступила в редакцию / received: 31.03.2026; принята к публикации / accepted: 19.05.2026.
Автор заявляет об отсутствии конфликта интересов / The author declares no conflict of interest.

СВЕДЕНИЯ ОБ АВТОРЕ:

Кукушкина Наталия Николаевна

канд. техн. наук, старший научный сотрудник

e-mail: kukushkina@krc.karelia.ru

CONTRIBUTOR:

Kukushkina, Natalia

Cand. Sci. (Tech.), Senior Researcher

УДК 519.214.4+519.214.8

ЕЩЕ РАЗ ОБ ОБЪЕМАХ ДЕРЕВЬЕВ ЛЕСА ГАЛЬТОНА – ВАТСОНА

Ю. Л. Павлов

*Институт прикладных математических исследований КарНЦ РАН,
ФИЦ «Карельский научный центр РАН» (ул. Пушкинская, 11, Петрозаводск,
Республика Карелия, Россия, 185910)*

Рассматриваются леса Гальтона – Ватсона с N корневыми деревьями и n некорневыми вершинами. Предполагается, что такие леса образуются критическим ветвящимся процессом, в котором число прямых потомков каждой частицы имеет распределение

$$p_k = \frac{h(k+1)}{(k+1)^\tau}, \quad k = 0, 1, 2, \dots, \quad \tau \in (2, 3),$$

где $h(x)$ – медленно меняющаяся на бесконечности функция. Найдены предельные распределения числа деревьев заданного объема, если $N, n \rightarrow \infty$, $n/N \rightarrow \infty$.

Ключевые слова: лес Гальтона – Ватсона; объем дерева; предельные теоремы; число деревьев заданного объема

Для цитирования: Павлов Ю. Л. Еще раз об объемах деревьев леса Гальтона – Ватсона // Труды Карельского научного центра РАН. 2026. № 6. С. 73–81. doi: 10.17076/mat2296

Финансирование. Финансовое обеспечение исследований осуществлялось из средств федерального бюджета на выполнение государственного задания КарНЦ РАН (Институт прикладных математических исследований КарНЦ РАН).

Yu. L. Pavlov. MORE ON THE SIZES OF TREES IN THE GALTON – WATSON FOREST

Institute of Applied Mathematical Research, Karelian Research Centre, Russian Academy of Sciences (11 Pushkinskaya St., 185910 Petrozavodsk, Karelia, Russia)

Galton – Watson forests with N rooted trees and n nonroot vertices are considered. It is assumed that such forests are formed by a critical branching process in which the number of offsprings for each particle has the distribution

$$p_k = \frac{h(k+1)}{(k+1)^\tau}, \quad k = 0, 1, 2, \dots, \quad \tau \in (2, 3),$$

where the function $h(x)$ is slowly varying at infinity. Some limit distributions of the number of trees of a given size are found if $N, n \rightarrow \infty$, $n/N \rightarrow \infty$.

Key words: Galton–Watson forest; tree size; limit theorems; number of trees of a given size

For citation: Pavlov Yu. L. Once again on the sizes of trees in the Galton–Watson forest. *Trudy Karelskogo nauchnogo tsentra RAN = Transactions of the Karelian Research Centre RAS*. 2026. No. 6. P. 73–81. doi: 10.17076/mat2296

Funding. The studies were funded from the federal budget through state assignment to the Karelian Research Centre RAS (Institute of Applied Mathematical Research, Karelian Research Centre RAS).

ВВЕДЕНИЕ И ОСНОВНОЙ РЕЗУЛЬТАТ

В статьях [3, 5] рассматривался лес Гальтона–Ватсона $\mathfrak{F}_{N,n}$, состоящий из N упорядоченных корневых деревьев и содержащий $N + n$ вершин. Предполагалось, что такой лес образован начинающимся с N частиц критическим ветвящимся процессом G , а случайная величина ξ , равная числу прямых потомков каждой частицы, имеет распределение

$$p_k = \mathbf{P}\{\xi = k\} = \frac{h(k+1)}{(k+1)^\tau}, \quad (1)$$

где $k = 0, 1, 2, \dots$, $\tau \in (2, 3)$, а $h(x)$ – медленно меняющаяся на бесконечности функция. Поскольку процесс G критический, распределение (1) удовлетворяет условию

$$\mathbf{E}\xi = 1. \quad (2)$$

Кроме того, из (1) следует, что дисперсия ξ бесконечна. Далее будем считать, что для случайной величины ξ верно равенство (2).

Случайный лес $\mathfrak{F}_{N,n}$ состоит из множества траекторий процесса G таких, что общее число частиц, существовавших до вырождения процесса, равно $N + n$. Распределение вероятностей на этом множестве естественным образом генерируется процессом G . Объемом дерева называется число его вершин, включая корень.

Обозначим μ_r случайную величину, равную числу деревьев объема r в $\mathfrak{F}_{N,n}$. В [5] были впервые доказаны теоремы о предельном поведении μ_r при различном характере стремления N и n к бесконечности, в том числе при фиксированных или стремящихся к бесконечности r . Заметим, что в случае $N, n \rightarrow \infty$ так, что $n/N \rightarrow \infty$, предельные распределения μ_r удалось получить только при весьма медленном росте отношения n/N . В дальнейшем, в статье [3], была предпринята попытка существенно расширить область применения полученных в [5] результатов. Это удалось сделать, но условия, при которых верны доказанные в [3] теоремы, по-прежнему не были оптимальными.

Для того чтобы сформулировать один из полученных в [3] результатов, введем необходимые понятия и обозначения. Техника доказательств теорем статей [3, 5] и настоящей статьи в значительной степени опирается на использование свойств сопровождающего G вспомогательного ветвящегося процесса $G(\lambda)$. Для того чтобы описать процесс $G(\lambda)$, введем случайную величину $\xi(\lambda)$, имеющую распределение

$$p_k(\lambda) = \mathbf{P}\{\xi(\lambda) = k\} = \frac{\lambda^k p_k}{F(\lambda)}, \quad (3)$$

где $k = 0, 1, 2, \dots$, $0 < \lambda < 1$ и

$$F(z) = \sum_{k=0}^{\infty} p_k z^k. \quad (4)$$

Введем случайный процесс $G(\lambda)$, отличающийся от G только тем, что число непосредственных потомков каждой частицы имеет распределение (3), а не (1). Обозначим $m = \mathbf{E}\xi(\lambda)$. В качестве значения параметра λ распределения (3) далее будем использовать единственное решение уравнения $m = n/(N + n)$. Тогда из (4) следует, что

$$m = \frac{\lambda F'(\lambda)}{F(\lambda)} = \frac{n}{N + n}. \quad (5)$$

Очевидно, что процесс $G(\lambda)$, как и процесс G , состоит из N независимых начинающихся с одной частицы подпроцессов $G_1(\lambda), \dots, G_N(\lambda)$. Обозначим $\nu^{(1)}, \dots, \nu^{(N)}$ независимые случайные величины, равные числам частиц, существовавших в подпроцессах $G_1(\lambda), \dots, G_N(\lambda)$ соответственно до их вырождения.

Пусть $\sigma^2 = \mathbf{D}\xi(\lambda)$, $a = \mathbf{E}\nu^{(1)}$, $b^2 = \mathbf{D}\nu^{(1)}$. В [6, формулы (2.7)] показано, что

$$a = \frac{1}{1 - m}, \quad b^2 = \frac{\sigma^2}{(1 - m)^3}, \quad (6)$$

где m определено в (5). Обозначим

$$q_k = \mathbf{P}\{\nu^{(1)} = k\}, \quad k = 1, 2, \dots, \quad (7)$$

и пусть

$$\sigma_{rr}^2 = q_r \left(1 - q_r - \frac{(a-r)^2}{b^2} q_r \right). \quad (8)$$

Для случая $N, n \rightarrow \infty$ так, что $n/N \rightarrow \infty$, в [5, теорема 1] доказана следующая теорема.

Теорема 1. Пусть r – фиксированное натуральное число, а $N, n \rightarrow \infty$ так, что $n/N \rightarrow \infty$ и существуют такие положительные числа Δ, v , где $v < 1/2$, что

$$\left(\frac{n}{N} \right)^2 = O(N^v (1 - \lambda)^{1+\Delta}). \quad (9)$$

Тогда для целых положительных k равномерно относительно $u_r = (k - Nq_r)/(\sigma_{rr}\sqrt{N})$ в любом фиксированном конечном интервале

$$\mathbf{P}\{\mu_r = k\} = \frac{1 + o(1)}{\sigma_{rr}\sqrt{2\pi N}} e^{-u_r^2/2}.$$

Это утверждение остается справедливым и при $r \rightarrow \infty$, если

$$r = O\left(\frac{n}{N}\right).$$

В [3, с. 106] отмечается, что если $N, n \rightarrow \infty$, то $\lambda \uparrow 1$ и условие (9) может быть выполнено, только если $n = O(N^{7/6})$. Это значит, что теорема 1 справедлива, если отношение n/N стремится к бесконечности очень медленно. Более точно необходимую скорость можно себе представить, если разность $1 - \lambda$ в (9) оценить, используя лемму 4 из [3], согласно которой

$$1 - \lambda \sim \left(-\frac{N}{n\Gamma(2-\tau)h(1/(1-\lambda))} \right)^{1/(\tau-2)}, \quad (10)$$

где $\Gamma(2-\tau)$ – значение гамма-функции в точке $2-\tau$. Заметим также, что в конкретных ситуациях проверка условия (9) не всегда является легкой задачей.

В статье [3] была предпринята попытка получить результат, аналогичный теореме 1, но при более простых условиях. Вместо условия (9) предполагалось, что $N, n \rightarrow \infty$, $n/N \rightarrow \infty$ и

$$\frac{n}{N^{\tau-1}} \rightarrow 0. \quad (11)$$

Кроме того, в [3] медленно меняющаяся функция $h(x)$ из распределения (1) подчинялась условию

$$h(x) \rightarrow D, \quad (12)$$

где $D > 0$ – некоторая постоянная. В [3] доказана следующая теорема.

Теорема 2. Пусть r – фиксированное натуральное число, а $N, n \rightarrow \infty$ так, что $n/N \rightarrow \infty$ и выполнены условия (11), (12). Тогда для целых положительных k равномерно относительно $v_r = (k - Nq_r)/\sqrt{Nq_r(1 - q_r)}$ в любом фиксированном конечном интервале

$$\mathbf{P}\{\mu_r = k\} = \frac{1 + o(1)}{\sqrt{2\pi Nq_r(1 - q_r)}} e^{-v_r^2/2}.$$

Это утверждение остается справедливым и при $r \rightarrow \infty$, если

$$r = O\left(\left(\frac{n}{N}\right)^{(\tau-1)/(\tau-2)}\right).$$

Основным результатом настоящей статьи является доказанная ниже теорема 3 о предельном поведении случайных величин μ_r без условия (12). В теореме 3 предполагается, что r – фиксированное число. Доказать подобный результат при $r \rightarrow \infty$ пока не удалось, и автор надеется, что эта задача в дальнейшем будет решена.

Теорема 3. Пусть r – фиксированное натуральное число, а $N, n \rightarrow \infty$ так, что $n/N \rightarrow \infty$ и существует $\gamma > 0$ такое, что

$$\frac{n}{N^{\tau-1-\gamma}} \rightarrow 0. \quad (13)$$

Тогда для целых положительных k равномерно относительно $w_r = (k - Nq_r)/\sqrt{Nq_r(1 - q_r)}$ в любом фиксированном конечном интервале

$$\mathbf{P}\{\mu_r = k\} = \frac{1 + o(1)}{\sqrt{2\pi Nq_r(1 - q_r)}} e^{-w_r^2/2}.$$

Статья состоит из трех разделов. В следующем разделе получены вспомогательные утверждения (леммы 1–6), с помощью которых в третьем разделе доказывается теорема 3.

ВСПОМОГАТЕЛЬНЫЕ РЕЗУЛЬТАТЫ

Приводимые ниже леммы формулируются и доказываются подобно вспомогательным результатам статьи [3]. Поэтому в некоторых местах мы ограничимся ссылками на полученные в [3, 5] промежуточные выражения, подробно описывая только выкладки, в которых новое условие (13) заменяет (11) и (12).

Метод доказательства теоремы 3 в значительной степени опирается на использование так называемой обобщенной схемы размещения частиц по ячейкам, введенной и подробно изученной В. Ф. Колчиным [2]. Обозначим $\nu_1(\mathfrak{F}), \dots, \nu_N(\mathfrak{F})$ случайные величины, равные объемам деревьев рассматриваемого

леса Гальтона–Ватсона, имеющих корни с номерами $1, \dots, N$ соответственно. Пусть ν_N означает число частиц, существовавших в процессе $G(\lambda)$ за все время его эволюции, следовательно, $\nu_N = \nu^{(1)} + \dots + \nu^{(N)}$. Известно [5, равенство (3.1)], что для двух наборов случайных величин $\nu_1(\mathfrak{F}), \dots, \nu_N(\mathfrak{F})$ и $\nu^{(1)}, \dots, \nu^{(N)}$ выполнены условия обобщенной схемы размещения и верно следующее утверждение.

Лемма 1. *Справедливо равенство*

$$\begin{aligned} & \mathbf{P}\{\nu_1(\mathfrak{F}) = k_1, \dots, \nu_N(\mathfrak{F}) = k_N\} \\ &= \mathbf{P}\{\nu^{(1)} = k_1, \dots, \nu^{(N)} = k_N | \nu_N = N + n\}. \end{aligned}$$

Заметим, что, как это обычно встречается при использовании обобщенной схемы для решения различных комбинаторных задач, утверждение леммы 1 справедливо при любых значениях параметра $\lambda, 0 < \lambda \leq 1$. Поэтому при исследовании случайных лесов в ряде случаев целесообразно использовать ветвящийся процесс $G(\lambda)$, выбирая значение λ , наиболее удобное для проведения необходимых вычислений.

Пусть $\nu_r^{(1)}, \dots, \nu_r^{(N)}$ означают независимые случайные величины такие, что

$$\mathbf{P}\{\nu_r^{(i)} = k\} = \mathbf{P}\{\nu^{(i)} = k | \nu^{(i)} \neq r\}, \quad (14)$$

где $k = 1, 2, \dots, i = 1, \dots, N$. Обозначим $\zeta_S^{(r)} = \nu_r^{(1)} + \dots + \nu_r^{(S)}$, где $S \leq N$. В [2, гл. I, §2, лемма 1] показано, что из леммы 1 и (14) следует такой результат.

Лемма 2. *Справедливо равенство*

$$\begin{aligned} \mathbf{P}\{\mu = k\} &= \binom{N}{k} q_r^k (1 - q_r)^{N-k} \\ &\times \frac{\mathbf{P}\{\zeta_{N-k}^{(r)} = N + n - kr\}}{\mathbf{P}\{\nu_N = N + n\}}. \end{aligned}$$

Далее рассмотрим предельное поведение суммы ν_N . Обозначим $\psi_N(t)$ характеристическую функцию случайной величины $(\nu_N - aN)/(b\sqrt{N})$.

Лемма 3. *Пусть $N, n \rightarrow \infty$ так, что $n/N \rightarrow \infty$ и выполнено условие (13). Тогда $\psi_N(t) \rightarrow e^{-t^2/2}$ при любом фиксированном t .*

Доказательство. Обозначим $\varphi(t)$ характеристическую функцию случайной величины $\nu^{(1)}$. В [5, формула (4.3)] показано, что

$$\ln \psi_N(t) = -\frac{t^2}{2} + \frac{t^3}{6b^3\sqrt{N}} Q\left(\frac{t}{b\sqrt{N}}\right), \quad (15)$$

где

$$|Q(t)| \leq 2 \max_{|s| \leq |t|} |(\ln \varphi(s))_s''|. \quad (16)$$

Из (15) видим, что для доказательства леммы достаточно установить соотношение

$$\frac{t^3}{6b^3\sqrt{N}} Q\left(\frac{t}{b\sqrt{N}}\right) \rightarrow 0. \quad (17)$$

Используя (3), (5), (6) и применяя тауберову теорему [6, гл. XIII, §5, теорема 5], находим, что

$$b^2 \asymp \left(\frac{n}{N}\right)^3 (1 - \lambda)^{\tau-3} h \left(\frac{1}{1 - \lambda}\right). \quad (18)$$

В [5, формула (4.8)] был найден явный вид выражения $(\ln \varphi(t))_t'''$. Поскольку это выражение весьма громоздко, здесь оно не приводится. Кроме того, там найдена и оценка модуля этого выражения [5, формула (4.9)]. Эту оценку нетрудно уточнить, если к содержащимся в [5, формула (4.8)] производным производящей функции распределения (3) опять, как и при получении (18), применить тауберову теорему. Тогда

$$|(\ln \varphi(t))_t'''| \asymp \left(\frac{n}{N}\right)^5 (1 - \lambda)^{2\tau-6} h \left(\frac{1}{1 - \lambda}\right). \quad (19)$$

Отсюда и из (10), (16), (18) находим, что

$$\begin{aligned} & \frac{1}{b^3\sqrt{N}} Q\left(\frac{t}{b\sqrt{N}}\right) \\ &= O\left(\left(\frac{n}{N^{\tau-1}} h^{5-2\tau} \left(\frac{1}{1 - \lambda}\right)\right)^{1/(2\tau-4)}\right) \end{aligned}$$

и из (13) следует (17), что и завершает доказательство леммы 3. $\square$

В следующей лемме доказывается, что на самом деле имеет место локальная сходимость распределений сумм ν_N к нормальному закону.

Лемма 4. *Пусть выполнены условия леммы 3. Тогда для натуральных l равномерно относительно $u = (l - N - n)/(b\sqrt{N})$ в любом фиксированном конечном интервале*

$$\mathbf{P}\{\nu_N = l\} = \frac{1 + o(1)}{b\sqrt{2\pi N}} e^{-u^2/2}.$$

Доказательство. Следуя стандартной схеме доказательства локальных предельных теорем (например, [1, теорема 4.2.1]), представим вероятность $\mathbf{P}\{\nu_N = k\}$ по формуле обращения в виде интеграла

$$\mathbf{P}\{\nu_N = k\} = (2\pi b\sqrt{N})^{-1} \int_{-\pi b\sqrt{N}}^{\pi b\sqrt{N}} e^{-iut} \psi_N(t) dt. \quad (20)$$

Поскольку

$$(2\pi)^{-1/2} e^{-u^2/2} = (2\pi)^{-1} \int_{-\infty}^{\infty} \exp\{-iut - t^2/2\} dt, \quad (21)$$

используя (20), (21), представим разность

$$R = 2\pi(b\sqrt{N}\mathbf{P}\{\nu_N = l\} - (2\pi)^{-1/2} e^{-u^2/2})$$

в виде суммы $R = I_1 + I_2 + I_3 + I_4 + I_5$, где

$$I_1 = \int_{-A}^A e^{-iut} (\psi_N(t) - e^{-t^2/2}) dt,$$

$$I_2 = \int_{A < |t| \leq \varepsilon L} e^{-iut} \psi_N(t) dt,$$

$$I_3 = \int_{\varepsilon L < |t| \leq \varepsilon b\sqrt{N}} e^{-iut} \psi_N(t) dt,$$

$$I_4 = \int_{\varepsilon b\sqrt{N} < |t| \leq \pi b\sqrt{N}} e^{-iut} \psi_N(t) dt,$$

$$I_5 = - \int_{A < |t|} \exp\{-iut - t^2/2\} dt, \quad (22)$$

$$L = \left| b^3 \sqrt{N} Q^{-1} \left(\frac{1}{b\sqrt{N}} \right) \right|, \quad (23)$$

а положительные постоянные A и ε будут выбраны позднее.

Очевидно, что лемма 4 справедлива, если $R \rightarrow 0$. Для доказательства этого соотношения рассмотрим отдельно каждый из перечисленных в (22) интегралов. Из леммы 3 следует, что $I_1 \rightarrow 0$ при любом фиксированном A . Понятно, что I_5 можно сделать сколь угодно малым выбором достаточно большого A . Хорошо известно, что в области интегрирования I_4 справедливо неравенство

$$\left| \varphi \left(\frac{t}{b\sqrt{N}} \right) \right| \leq e^{-C_1},$$

здесь и далее символы $C_1, C_2, \dots$ означают некоторые положительные постоянные. Отсюда

$$|\psi_N(t)| \leq e^{-C_1 N}$$

и из (10), (13), (18) находим, что $I_4 \rightarrow 0$. Из разложения (15) и из (23) следует, что при достаточно малом ε в области интегрирования I_2

$$|\psi_N(t)| \leq e^{-C_3 t^2},$$

следовательно, интеграл I_2 можно сделать сколь угодно малым выбором достаточно большого A .

Осталось рассмотреть I_3 . Для получения необходимой оценки будем следовать доказательству леммы 4 из [2, гл. II, §4]. Обозначим $U(z)$ и $F_\lambda(z)$ производящие функции случайных величин $\nu^{(1)}$ и $\xi(\lambda)$ соответственно. Хорошо известно [2, гл. II, §4, лемма 1], что

$$U(z) = z F_\lambda(U(z)).$$

Поскольку $\lambda < 1$, функция $F_\lambda(z)$ имеет все производные и при $|z| \leq 1$ она может быть представлена следующим образом:

$$F_\lambda(z) = 1 + F'_\lambda(1)(z-1) + \frac{F''_\lambda(1)(1+\gamma(z))}{2}(z-1)^2,$$

где

$$F'_\lambda(1) = \frac{n}{N+n},$$

а $\gamma(z) \rightarrow 0$ при $z \rightarrow 1$. Отсюда и из предыдущего равенства

$$U(z) = z \left(1 + F'_\lambda(1)(U(z)-1) + \frac{F''_\lambda(1)(1+\gamma(z))}{2}(U(z)-1)^2 \right).$$

Разрешая это квадратное уравнение относительно $U(z)$ и выбирая ветвь корня из условия $0 \leq U(z) \leq 1$ для действительных z , $0 \leq z \leq 1$, получаем

$$U(z) = 1 - A(N, n)(1 + o(1)) - \sqrt{A^2(N, n) - \frac{2(z-1)(1+o(1))}{zF''_\lambda(1)}}(1 + o(1)), \quad (24)$$

где

$$A(N, n) = \frac{z-1}{zF''_\lambda(1)} - \frac{N}{(N+n)F''_\lambda(1)}.$$

Положим

$$z = \exp \left\{ \frac{it}{b\sqrt{N}} \right\} = \cos \left(\frac{t}{b\sqrt{N}} \right) + i \sin \left(\frac{t}{b\sqrt{N}} \right). \quad (25)$$

Нетрудно видеть, что интеграл I_3 можно представить в виде суммы $I_3 = I'_3 + I''_3$, где в области интегрирования I'_3 верно соотношение $z-1 = o(N/n)$ и $N/n = O(z-1)$ в области интегрирования I''_3 .

Рассмотрим I'_3 . Из (24)

$$U(z) = 1 + \frac{N(1+o(1))}{(N+n)F''_\lambda(1)} \left(1 - \frac{n}{N}(z-1) - \sqrt{1 - 2 \left(\frac{N+n}{N}\right)^2 F''_\lambda(1)(z-1)(1+o(1))}\right). \quad (26)$$

Обозначим для краткости

$$B = B(N, n) = 2 \left(\frac{N+n}{N}\right)^2 F''_\lambda(1)(1+o(1)).$$

Используя (3), (4) и опять применяя тауберову теорему, находим, что

$$F''_\lambda(1) \sim \frac{\Gamma(3-\tau)}{(1-\lambda)^{3-\tau}} h\left(\frac{1}{1-\lambda}\right). \quad (27)$$

Ясно, что

$$\begin{aligned} & \sqrt{1 - B(z-1)} \\ &= \sqrt{1 + B \left(1 - \cos\left(\frac{t}{b\sqrt{N}}\right)\right) - iB \sin\left(\frac{t}{b\sqrt{N}}\right)}. \end{aligned}$$

Отсюда

$$\begin{aligned} & \sqrt{1 - B(z-1)} \\ &= R^{1/2} \left(\cos\left(\frac{\alpha + 2\pi k}{2}\right) + i \sin\left(\frac{\alpha + 2\pi k}{2}\right) \right), \end{aligned} \quad (28)$$

где $k = 0, 1$.

$$\begin{aligned} R &= \left| 1 + B \left(1 - \cos\left(\frac{t}{b\sqrt{N}}\right)\right) - iB \sin\left(\frac{t}{b\sqrt{N}}\right) \right| \\ &= \left| 1 + 2B(1+B) \left(1 - \cos\left(\frac{t}{b\sqrt{N}}\right)\right) \right|^{1/2}, \end{aligned}$$

$$\alpha = \arctg \frac{\operatorname{Im}(W)}{1 + \operatorname{Re}(W)},$$

$$W = B \left(1 - \cos\left(\frac{t}{b\sqrt{N}}\right)\right) - iB \sin\left(\frac{t}{b\sqrt{N}}\right).$$

Обозначим

$$D = D(N, n) = \frac{N}{(N+n)F''_\lambda(1)}.$$

Тогда из (26), (28)

$$\begin{aligned} & \varphi\left(\frac{t}{b\sqrt{N}}\right) = 1 - D(1+o(1)) \\ & \times \left(R^{1/2} \cos\left(\frac{\alpha}{2} + \pi k\right) - 1\right) \end{aligned}$$

$$\begin{aligned} & + \frac{n}{N} \left(\cos\left(\frac{t}{b\sqrt{N}}\right) - 1 \right) \\ & - iD(1+o(1)) \left(R^{1/2} \sin\left(\frac{\alpha}{2} + \pi k\right) \right. \\ & \left. + \frac{n}{N} \sin\left(\frac{t}{b\sqrt{N}}\right) \right). \end{aligned}$$

Далее, для того чтобы с помощью этого равенства оценить $|\varphi(t/(b\sqrt{N}))|$, соберем вместе (10), (16), (18), (19), (23), (26), (27), (28). Учтывая, что $|t/(b\sqrt{N})| > \varepsilon L$, $t/(b\sqrt{N}) \rightarrow 0$ и проводя непосредственные вычисления, хотя и громоздкие, но элементарные, получаем, что $D \rightarrow 0$ и существует $C_3 > 0$ такое, что

$$R^{1/2} \cos\left(\frac{\alpha}{2} + \pi k\right) > 1 + C_3 h^\rho \left(\frac{1}{1-\lambda}\right),$$

где $\rho = 2$, если функция $h(x)$ ограничена, и $\rho = 1/2$, если $h(x) \rightarrow \infty$ при $x \rightarrow \infty$. Тогда

$$\left| \varphi\left(\frac{t}{b\sqrt{N}}\right) \right|^2 < 1 - C_4 D h^\rho \left(\frac{1}{1-\lambda}\right),$$

следовательно,

$$\left| \varphi\left(\frac{t}{b\sqrt{N}}\right) \right| \leq \exp \left\{ -C_5 \frac{N}{nF''_\lambda(1)} h^\rho \left(\frac{1}{1-\lambda}\right) \right\}. \quad (29)$$

Поэтому

$$|I'_3| \leq 2b\sqrt{N} \exp \left\{ -C_5 \frac{N^2}{nF''_\lambda(1)} h^\rho \left(\frac{1}{1-\lambda}\right) \right\}.$$

Из этого неравенства, (10), (13), (18) и (27) следует, что $I'_3 \rightarrow 0$.

Пусть теперь $N/n = O(z-1)$ и рассмотрим I''_3 . Из (24) получаем, что

$$U(z) = 1 - \sqrt{-\frac{2(z-1)}{F''_\lambda(1)}} (1+o(1)).$$

Отсюда и из (25)

$$\begin{aligned} U(z) &= 1 - \sqrt{-\frac{2it}{b\sqrt{N}F''_\lambda(1)}} (1+o(1)) \\ &= 1 - \sqrt{\frac{|t|}{b\sqrt{N}F''_\lambda(1)}} \left(1 - i\frac{t}{|t|}\right) (1+o(1)). \end{aligned}$$

Это значит, что

$$\left| \varphi\left(\frac{t}{b\sqrt{N}}\right) \right| \leq \exp \left\{ -C_6 \sqrt{\frac{|t|}{b\sqrt{N}F''_\lambda(1)}} \right\}.$$

Из этого неравенства и из (10), (16), (18), (23), (27) следует, что в области интегрирования I_3'' для любых $\delta > 0$

$$\left| \varphi \left(\frac{t}{b\sqrt{N}} \right) \right| \leq \exp \left\{ -C_7 \left(\frac{N}{n} \right)^{(1+\delta)/(\tau-2)} \right\}. \quad (30)$$

Тогда

$$|I_3''| \leq 2\varepsilon b\sqrt{N} \times \exp \left\{ -C_7 \left(\frac{N^{\tau-1+\delta}}{n^{1+\delta}} \right)^{1/(\tau-2)} \right\}.$$

Отсюда и из (13), (18) находим, что $I_3'' \rightarrow 0$ при достаточно малых $\delta > 0$. $\square$

Для того чтобы описать предельное поведение $\zeta_S^{(r)}$, введем обозначения

$$a_r = \mathbf{E}\nu_r^{(1)} = \frac{a - rq_r}{1 - q_r},$$

$$b_r^2 = \mathbf{D}\nu_r^{(1)} = \frac{b^2}{(1 - q_r)^2} \left(1 - q_r - \frac{(a - r)^2}{b^2} q_r \right), \quad (31)$$

где a и b определены в (6). Пусть $\psi_S^{(r)}(t)$ означает характеристическую функцию случайной величины $(\zeta_S^{(r)} - Sa_r)/(b_r\sqrt{S})$.

Лемма 5. Пусть выполнены условия теоремы 3 и

$$S = N(1 - q_r)(1 + o(1)). \quad (32)$$

Тогда $\psi_S^{(r)}(t) \rightarrow e^{-t^2/2}$ при любом фиксированном t .

Доказательство. Обозначим $\varphi_r(t)$ характеристическую функцию случайной величины $\nu_r^{(1)}(t)$. Как и в доказательстве леммы 7 из [3],

$$\ln \psi_S^{(r)}(t) = -\frac{t^2}{2} + \frac{t^3}{6b^3\sqrt{S}} Q_r \left(\frac{t}{b_r\sqrt{S}} \right), \quad (33)$$

где

$$|Q_r(t)| \leq 2 \max_{|s| \leq |t|} |(\ln \varphi_r(s))_s'''|. \quad (34)$$

Согласно (33), достаточно показать, что

$$\frac{t^3}{6b^3\sqrt{S}} Q_r \left(\frac{t}{b_r\sqrt{S}} \right) \rightarrow 0. \quad (35)$$

Понятно, что

$$\varphi_r(t) = \frac{\varphi(t) - q_r e^{itr}}{1 - q_r}, \quad (36)$$

где q_r определено в (7). В [4, гл. 2, §3, формула (3) и гл. 3, §2, формула (3)] были найдены

явные выражения для $(\ln \varphi(t))_t'''$ и $(\ln \varphi_r(t))_t'''$ соответственно. Указанные формулы здесь не приводятся, поскольку они весьма громоздки. С помощью этих формул путем проведения необходимых вычислений можно показать (см. также [5, формула (5.23)]), что

$$|(\ln \varphi_r(t))_y'''| \asymp |(\ln \varphi(t))_t'''| + r^3 q_r. \quad (37)$$

В [5, формула (5.9)] установлено, что для фиксированных r

$$0 < C_8 \leq q_r \leq C_9 < 1. \quad (38)$$

Теперь покажем, что

$$b_r^2 \asymp b^2. \quad (39)$$

Из (31) следует, что для доказательства (39) достаточно установить, что

$$\frac{(a - r)^2}{b^2} q_r \rightarrow 0. \quad (40)$$

Используя (5) и (18), находим, что

$$\frac{a^2}{b^2} \leq C_{10}(1 - \lambda)^{3-\tau-\varepsilon} \frac{N}{n} \times \left(\left(\frac{1}{1 - \lambda} \right)^\varepsilon h \left(\frac{1}{1 - \lambda} \right) \right)^{-1}.$$

Отсюда, учитывая свойства медленно меняющихся на бесконечности функций и выбирая, например, $\varepsilon = (3 - \tau)/2$, при $n/N \rightarrow \infty$ и $\lambda \uparrow 1$ получим

$$\frac{a^2}{b^2} q_r \rightarrow 0. \quad (41)$$

Это соотношение, вместе с (38), показывает, что (40), а вместе с ним и (39), верно при фиксированных r . Тогда из (17), (18), (34), (37), (39) следует (35). $\square$

Замечание. Из (8), (40) и (41) следует, что $\sigma_{rr}^2 \sim q_r(1 - q_r)$. Это значит, что утверждения теорем 1 и 3 эквивалентны.

Теперь покажем, что для $\zeta_S^{(r)}$ имеет место и локальная сходимость.

Лемма 6. Пусть верно (32) и выполнены условия теоремы 3. Тогда для натуральных l равномерно относительно $y_r = (l - Sa_r)/(b_r\sqrt{S})$ в любом фиксированном конечном интервале

$$\mathbf{P}\{\zeta_S^{(r)} = l\} = \frac{1 + o(1)}{b_r\sqrt{2\pi S}} e^{-y_r^2/2}.$$

Доказательство. Следуя доказательству леммы 5 из [3], представим вероятность $\mathbf{P}\{\zeta_S^{(r)} = l\}$ по формуле обращения в виде интеграла

$$\mathbf{P}\{\zeta_S^{(r)} = l\} = \frac{1}{2\pi b_r \sqrt{S}} \int_{-\pi b_r \sqrt{S}}^{\pi b_r \sqrt{S}} e^{-iy_r t} \psi_S^{(r)}(t) dt.$$

Используя это равенство и (21), представим разность

$$R_r = 2\pi(b_r \sqrt{S} \mathbf{P}\{\zeta_S^{(r)} = l\} - (2\pi)^{-1/2} e^{-y_r^2/2})$$

в виде суммы пяти интегралов

$$\begin{aligned} I_1 &= \int_{-A}^A e^{-iy_r t} (\psi_S^{(r)}(t) - e^{-t^2/2}) dt, \\ I_2 &= \int_{A < |t| \leq \varepsilon \varepsilon L_r} e^{-iy_r t} \psi_S^{(r)}(t) dt, \\ I_3 &= \int_{\varepsilon L_r < |t| \leq \varepsilon b_r \sqrt{S}} e^{-iy_r t} \psi_S^{(r)}(t) dt, \\ I_4 &= \int_{\varepsilon b_r \sqrt{S} < |t| \leq \pi b_r \sqrt{S}} e^{-iy_r t} \psi_S^{(r)}(t) dt, \\ I_5 &= - \int_{A < |t|} \exp\{-iy_r t - t^2/2\} dt, \end{aligned} \quad (42)$$

$$L_r = b_r^3 \sqrt{S} Q_r^{-1} \left(\frac{1}{b_r \sqrt{S}} \right), \quad (43)$$

а положительные постоянные A и ε будут выбраны позднее.

Понятно, что если $R_r \rightarrow 0$, то лемма 6 верна. В силу леммы 5 при любом фиксированном A верно соотношение $I_1 \rightarrow 0$. Нетрудно проверить, собирая (16), (18), (19), (23), (37), (42), (43), что интегралы I_2, I_4, I_5 оцениваются так же, как и одноименные интегралы в доказательстве леммы 4.

Рассмотрим I_3 . Из (36) следует, что

$$|\varphi_r(t)| = |\varphi(t)| \left| \frac{1 - \varphi^{-1}(t) q_r e^{itr}}{1 - q_r} \right|. \quad (44)$$

В области интегрирования I_3 при фиксированных r величину $tr/(b_r \sqrt{S})$ можно сделать сколь угодно малой выбором достаточно малого ε . Используя это и равенство (27), несложно показать, что

$$\left| 1 - \varphi^{-1} \left(\frac{t}{b_r \sqrt{S}} \right) q_r \exp \left\{ \frac{itr}{b_r \sqrt{S}} \right\} \right| \leq 1 - q_r.$$

Отсюда и из (44)

$$\left| \varphi_r \left(\frac{t}{b_r \sqrt{S}} \right) \right| \leq \left| \varphi \left(\frac{t}{b_r \sqrt{S}} \right) \right|$$

и I_3 с помощью (29), (30) оценивается так же, как одноименный интеграл в доказательстве леммы 4. $\square$

ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ 3

В условиях теоремы 3

$$k = Nq_r + w_r \sqrt{Nq_r(1 - q_r)} \quad (45)$$

и

$$N - k = N(1 - q_r)(1 + o(1)). \quad (46)$$

Из леммы 4

$$\mathbf{P}\{\nu_N = N + n\} \sim (b\sqrt{2\pi N})^{-1}. \quad (47)$$

В утверждении леммы 6 положим $l = N + n - kr$ и $S = N - k$. Тогда, учитывая (18), (39), (45), (46), находим, что

$$\mathbf{P}\{\zeta_{N-k}^{(r)} = N + n - kr\} \sim (b\sqrt{2\pi N})^{-1}.$$

Отсюда и из (47)

$$\frac{\mathbf{P}\{\zeta_{N-k}^{(r)} = N + n - kr\}}{\mathbf{P}\{\nu_N = N + n\}} \rightarrow 1. \quad (48)$$

Известно, что если $Nq_r(1 - q_r) \rightarrow \infty$, то для биномиальных вероятностей справедливо нормальное приближение:

$$\begin{aligned} &\binom{N}{k} q_r^k (1 - q_r)^{N-k} \\ &= \frac{1 + o(1)}{\sqrt{2\pi Nq_r(1 - q_r)}} \exp \left\{ -\frac{(k - Nq_r)^2}{2Nq_r(1 - q_r)} \right\} \end{aligned} \quad (49)$$

равномерно относительно

$$\frac{k - Nq_r}{\sqrt{Nq_r(1 - q_r)}}$$

в любом фиксированном конечном интервале. Из (38) следует, что в условиях теоремы 3 $Nq_r(1 - q_r) \rightarrow \infty$. Поэтому из леммы 2, (48) и (49) получаем утверждение теоремы 3.

Автор выражает глубокую благодарность рецензенту за очень полезные замечания.

ЛИТЕРАТУРА

1. Ибрагимов И. А., Линник Ю. В. Независимые и стационарно связанные величины. М.: Наука, 1965. 524 с.
2. Колчин В. Ф. Случайные отображения. М.: Наука, 1984. 208 с.

3. Павлов Ю. Л. Замечание об объемах деревьев леса Гальтона–Ватсона // Дискретная математика. 2025. Т. 37, вып. 4. С. 102–117. doi: 10.4213/dm1872

4. Павлов Ю. Л. Случайные леса. Петрозаводск: Карел. науч. центр РАН, 1996. 259 с.

5. Павлов Ю. Л., Чеплюкова И. А. Объемы деревьев случайного леса и конфигурационные графы // Труды МИАН им. В. А. Стеклова. 2022. Т. 316. С. 298–315. doi: 10.4213/tm4216

6. Феллер В. Введение в теорию вероятностей и ее применения. Т. 2. М.: Мир, 1984. 738 с.

REFERENCES

1. Ibragimov I. A., Linnik Yu. V. Independent and stationary sequences of random variables. Groningen:

Wolters-Noordho; 1971. 438 p.

2. Kolchin V. F. Random mappings. New York: Springer; 1986. 206 p.

3. Pavlov Yu. L. A remark on the Galton-Watson forest tree sizes. *Diskretnaya matematika = Discrete Mathematics and Applications*. 2025;37(4):102–117. (In Russ.). doi: 10.4213/dm1872

4. Pavlov Yu. L. Random forests. Utrecht: VSP; 2000. 122 p. doi: 10.1515/9783110941975

5. Pavlov Yu. L., Cheplyukova I. A. Sizes of trees in a random forest and configuration graphs. *Proceedings of the Steklov Institute of Mathematics*. 2022;316:280–297. doi: 10.1134/S0081543822010205

6. Feller W. An introduction to probability theory and its applications. Vol. II. New York: John Wiley; 1971. 704 p.

Поступила в редакцию / received: 26.01.2026; принята к публикации / accepted: 08.06.2026.
Автор заявляет об отсутствии конфликта интересов / The author declares no conflict of interest.

СВЕДЕНИЯ ОБ АВТОРЕ:

Павлов Юрий Леонидович

д-р физ.-мат. наук, профессор, главный научный сотрудник

e-mail: pavlov@krc.karelia.ru

CONTRIBUTOR:

Pavlov, Yury

Dr. Sci. (Phis.-Math.), Professor, Chief Researcher

УДК 519.214.4+519.214.8

О СКОРОСТИ СХОДИМОСТИ К НОРМАЛЬНОМУ ЗАКОНУ РАСПРЕДЕЛЕНИЙ ЧИСЛА ДЕРЕВЬЕВ ЗАДАННОГО ОБЪЕМА В ЛЕСЕ ГАЛЬТОНА – ВАТСОНА

Ю. Л. Павлов

*Институт прикладных математических исследований КарНЦ РАН,
ФИЦ «Карельский научный центр РАН» (ул. Пушкинская, 11, Петрозаводск,
Республика Карелия, Россия, 185910)*

Рассматриваются леса Гальтона–Ватсона, порожденные однородным критическим ветвящимся процессом с N начальными частицами, в котором число прямых потомков каждой частицы является случайной величиной ξ с распределением

$$p_k = \frac{h(k+1)}{(k+1)^\tau}, \quad k = 0, 1, 2, \dots, \quad \tau \in (2, 3),$$

где $h(x)$ – медленно меняющаяся на бесконечности функция. Предполагается, что число некорневых вершин случайного леса равно n . Пусть μ_r – число деревьев с r вершинами. Доказаны теоремы о скорости сходимости распределений μ_r к нормальному закону при $N, n, r \rightarrow \infty$ и $n/N \leq C < \infty$.

Ключевые слова: лес Гальтона–Ватсона; объем дерева; предельные теоремы; скорость сходимости; число деревьев заданного объема

Для цитирования: Павлов Ю. Л. О скорости сходимости к нормальному закону распределений числа деревьев заданного объема в лесу Гальтона–Ватсона // Труды Карельского научного центра РАН. 2026. № 6. С. 82–88. doi: 10.17076/mat2302

Финансирование. Финансовое обеспечение исследований осуществлялось из средств федерального бюджета на выполнение государственного задания КарНЦ РАН (Институт прикладных математических исследований КарНЦ РАН).

Yu. L. Pavlov. ON THE RATE OF NORMAL APPROXIMATION FOR THE DISTRIBUTION OF THE NUMBER OF TREES OF A GIVEN SIZE IN A GALTON – WATSON FOREST

Institute of Applied Mathematical Research, Karelian Research Centre, Russian Academy of Sciences (11 Pushkinskaya St., 185910 Petrozavodsk, Karelia, Russia)

The article studies Galton–Watson forests formed by a critical branching process starting with N particles. It is assumed that the total number of descendants of the initial particles is n until the process extinction. The number of offsprings of each particle has the distribution

$$p_k = \frac{h(k+1)}{(k+1)^\tau}, \quad k = 0, 1, 2, \dots, \quad \tau \in (2, 3),$$

where the function $h(x)$ is slowly varying at infinity. Let μ_r be the number of trees with r vertices. Theorems are proven for the rate of convergence of μ_r distributions to the normal law if $N, n, r \rightarrow \infty$ and $n/N \leq C < \infty$.

Key words: Galton–Watson forest; tree size; limit theorems; rate of convergence; number of trees of a given size

For citation: Pavlov Yu. L. On the rate of normal approximation for the distribution of the number of trees of a given size in a Galton–Watson forest. *Trudy Karelskogo nauchnogo tsentra RAN = Transactions of the Karelian Research Centre RAS*. 2026. No. 6. P. 82–88. doi: 10.17076/mat2302

Funding. The studies were funded from the federal budget through state assignment to the Karelian Research Centre RAS (Institute of Applied Mathematical Research, Karelian Research Centre RAS).

ВВЕДЕНИЕ

В статьях [3, 4] рассматривались случайные леса Гальтона–Ватсона с N упорядоченными корневыми деревьями и n некорневыми вершинами. Предполагалось, что такие леса образованы начинающимися с N частиц критическими ветвящимися процессами, в которых случайная величина ξ , равная числу непосредственных потомков любой частицы, имеет распределение

$$p_k = \mathbf{P}\{\xi = k\} = \frac{h(k+1)}{(k+1)^\tau}, \quad (1)$$

где $k = 0, 1, 2, \dots, \tau \in (2, 3)$, а $h(x)$ – медленно меняющаяся на бесконечности функция. Пусть G – такой ветвящийся процесс Гальтона–Ватсона. Генерируемому им случайному лесу соответствует множество траекторий процесса G , в которых число всех потомков начальных частиц равно n , а распределение вероятностей на этом множестве естественным образом индуцируется процессом G . В [4] также кратко сообщается о причинах интереса к ветвящимся процессам вида G и к образуемым ими лесам Гальтона–Ватсона. Там же даны ссылки на литературу, более подробно освещающую это направление исследований.

Поскольку процесс G критический, справедливо равенство

$$\mathbf{E}\xi = 1. \quad (2)$$

Объемом дерева случайного леса называется число его вершин, включая корень. В [4] изучалось предельное поведение случайных величин μ_r , равных числам деревьев леса объема r при всех возможных соотношениях между стремящимися к бесконечности параметрами N и n . При этом рассматривался также и случай $r \rightarrow \infty$. В частности, были найдены предельные распределения μ_r , когда $n/N \leq$

$C < \infty$. Техника доказательств теорем статьи [4] и настоящей статьи в значительной степени опирается на использование свойств сопровождающего G вспомогательного ветвящегося процесса $G(\lambda)$. Для того чтобы описать процесс $G(\lambda)$, введем случайную величину $\xi(\lambda)$, имеющую распределение

$$p_k(\lambda) = \mathbf{P}\{\xi(\lambda) = k\} = \frac{\lambda^k p_k}{F(\lambda)}, \quad k = 0, 1, 2, \dots, \quad (3)$$

где $0 < \lambda < 1$ и

$$F(\lambda) = \sum_{k=0}^{\infty} p_k \lambda^k. \quad (4)$$

Введем ветвящийся процесс $G(\lambda)$, отличающийся от G только тем, что число прямых потомков каждой частицы процесса $G(\lambda)$ имеет распределение (3), а не (1), как в G . В качестве значения параметра λ распределения (3) используется единственное решение уравнения

$$m = \mathbf{E}\xi(\lambda) = \frac{\lambda F'(\lambda)}{F(\lambda)} = \frac{n}{N+n}. \quad (5)$$

Процесс $G(\lambda)$, как и G , распадается на N независимых начинающихся с одной частицы подпроцессов $G_1(\lambda), \dots, G_N(\lambda)$. Обозначим $\nu^{(1)}, \dots, \nu^{(N)}$ независимые одинаково распределенные случайные величины, равные числам частиц, существовавших в подпроцессах $G_1(\lambda), \dots, G_N(\lambda)$ соответственно до их вырождения. Иногда будет удобнее вместо $\nu^{(1)}$ писать ν без указания верхнего индекса. Пусть $\sigma^2 = \mathbf{D}\xi(\lambda)$, $a = \mathbf{E}\nu$, $b^2 = \mathbf{D}\nu$. В [4, формулы (2.7)] показано, что

$$a = \frac{1}{1-m}, \quad b^2 = \frac{\sigma^2}{(1-m)^3}. \quad (6)$$

Обозначим

$$q_k = \mathbf{P}\{\nu = k\}, \quad k = 1, 2, \dots, \quad (7)$$

и пусть

$$\sigma_{rr}^2 = q_r \left(1 - q_r - \frac{(a-r)^2}{b^2} q_r \right). \quad (8)$$

Следующие две теоремы доказаны в [4, теоремы 1, 2]. В них установлена локальная сходимость распределений μ_r к предельным законам в случае $r \rightarrow \infty$.

Теорема 1. Пусть $N, n, r \rightarrow \infty$ и выполнено одно из следующих условий.

1. $n/N \rightarrow 0$, $n^2/N \rightarrow \infty$, $Nq_r \rightarrow \infty$.
2. $0 < C_1 \leq n/N \leq C_2 < \infty$, $r = o(\ln N)$.

Тогда для целых неотрицательных k равномерно относительно $u_r = (k - Nq_r)/(\sigma_{rr}\sqrt{N})$ в любом фиксированном конечном интервале

$$\mathbf{P}\{\mu_r = k\} = \frac{1 + o(1)}{\sigma_{rr}\sqrt{2\pi N}} e^{-u_r^2/2}.$$

Теорема 2. Пусть $N, n, r \rightarrow \infty$ так, что $n/N \leq C_2 < \infty$. Тогда для целых неотрицательных k равномерно относительно $u_r = (k - Nq_r)/(\sqrt{Nq_r})$ в любом фиксированном конечном интервале

$$\mathbf{P}\{\mu_r = k\} = \frac{1}{k!} (Nq_r)^k \exp\{-Nq_r\} (1 + o(1)).$$

Из теорем 1 и 2 следует, что в некоторых случаях распределения μ_r одновременно сближаются с нормальным и с пуассоновским распределениями. В таких случаях имеет смысл оценить близость распределений μ_r к предельным законам, то есть найти скорость сходимости. Это позволит выяснить, какое из предельных распределений и при каких условиях осуществляет более точную аппроксимацию.

Целью настоящей статьи является оценка скорости локальной сходимости распределений μ_r к нормальному закону в условиях теорем 1 и 2. Статья состоит из четырех разделов. В следующем разделе формулируется основной результат в виде теоремы 3. В третьем разделе приводятся вспомогательные утверждения (леммы 2–6), с помощью которых в последнем разделе доказывается теорема 3. Символы $C_1, C_2, \dots$ означают некоторые положительные постоянные.

ОСНОВНОЙ РЕЗУЛЬТАТ

Прежде чем сформулировать теорему 3, являющуюся основным результатом статьи, представим несколько соотношений в надежде, что они позволят читателю лучше понять условия и утверждение этой теоремы. Сначала приведем следующую лемму, представляющую собой часть леммы 2 из статьи [4].

Лемма 1. Пусть $N, n \rightarrow \infty$. Справедливы следующие утверждения.

1. Если $0 < C_1 \leq n/N \leq C_2 < \infty$,
то $0 < C_3 \leq \lambda \leq C_4 < 1$.
2. Если $n/N \rightarrow 0$,
то $\lambda = \frac{np_0}{Np_1} (1 + o(1))$.

Из (4), (5) и леммы 1 нетрудно получить, что функция $\lambda/F(\lambda)$ при $0 \leq \lambda \leq 1$ монотонно возрастает от 0 до 1, поэтому если $n/N \leq C_2 < \infty$, то

$$\frac{\lambda}{F(\lambda)} \leq C_5 < 1. \quad (9)$$

Из (7), формул (5.6), (5.7), (5.15), (5.17) статьи [4] и простейших свойств медленно меняющихся функций вытекает, что при достаточно больших r и сколь угодно малых $\delta > 0$

$$\begin{aligned} \frac{1}{r^{\tau/(\tau-1)+\delta}} \left(\frac{\lambda}{F(\lambda)} \right)^{r-1} &\leq q_r \\ &\leq \frac{1}{r^{\tau/(\tau-1)-\delta}} \left(\frac{\lambda}{F(\lambda)} \right)^{r-1}. \end{aligned} \quad (10)$$

Выражения (9) и (10) позволяют оценить, при каких соотношениях между стремящимися к бесконечности N, n, r выполнено условие

$$Nq_r \rightarrow \infty. \quad (11)$$

Рассмотрим поведение дисперсии b^2 . Предположим, что $0 < C_1 \leq n/N \leq C_2 < \infty$. Тогда из (3)–(6) и первого утверждения леммы 1 следует, что

$$0 < C_6 \leq b^2 \leq C_7 < \infty. \quad (12)$$

Пусть теперь $n/N \rightarrow 0$. Тогда из (3)–(6) и второго утверждения леммы 1 находим, что

$$b^2 \asymp \frac{n}{N}. \quad (13)$$

Из (9), (10) очевидным образом следует, что если $r \rightarrow \infty$, то

$$q_r \rightarrow 0, \quad r^2 q_r \rightarrow 0. \quad (14)$$

Отсюда и из (8), (12) находим, что если $0 < C_1 \leq n/N \leq C_2 < \infty$, то при $r \rightarrow \infty$

$$\sigma_{rr}^2 \sim q_r. \quad (15)$$

Пусть $n/N \rightarrow 0$. Из утверждения 2 леммы 1, (10) и (13) видим, что при любом фиксированном $k > 0$ и любом $\delta > 0$

$$\begin{aligned} \frac{r^k q_r}{b^2} &\leq C_8 r^{k-\tau/(\tau-1)+\delta} \left(\frac{\lambda}{F(\lambda)} \right)^{r-2} \\ &= C_8 \exp \left\{ (r-2) \ln \left(\frac{\lambda}{F(\lambda)} \right) \right. \\ &\quad \left. \times \left(1 + \left(k - \frac{\tau}{\tau-1} + \delta \right) \frac{\ln r}{(r-2) \ln(\lambda/F(\lambda))} \right) \right\}. \end{aligned}$$

Легко видеть, что последнее выражение стремится к нулю, поэтому

$$\frac{r^k q_r}{b^2} \rightarrow 0 \quad (16)$$

и из (8) опять получаем (15).

Обозначим

$$v_r = \frac{k - Nq_r}{\sqrt{Nq_r}}.$$

Учитывая (15), нетрудно получить, что утверждение теоремы 1 можно изложить следующим образом. Для целых неотрицательных k равномерно относительно v_r в любом фиксированном конечном интервале

$$\mathbf{P}\{\mu_r = k\} = \frac{1 + o(1)}{\sqrt{2\pi Nq_r}} e^{-v_r^2/2}.$$

Теперь сформулируем основной результат статьи.

Теорема 3. Пусть $N, n, r \rightarrow \infty$ так, что $n/N \leq C_2 < \infty$, $n^3/N^2 \rightarrow \infty$ и выполнено условие (11). Пусть также

$$v_r = \frac{k - Nq_r}{\sqrt{Nq_r}} = O(1). \quad (17)$$

Тогда

$$\begin{aligned} \mathbf{P}\{\mu_r = k\} &= \frac{1}{\sqrt{2\pi Nq_r}} e^{-v_r^2/2} \\ &\times \left(1 + \frac{3v_r - v_r^3}{6\sqrt{Nq_r}} + \frac{r^2 q_r}{2b^2} (1 - v_r^2) \right. \\ &\quad \left. + O\left(\frac{Nr q_r}{n} + \frac{N}{n^{3/2}} + \frac{1}{Nq_r} \right) \right). \end{aligned}$$

ВСПОМОГАТЕЛЬНЫЕ УТВЕРЖДЕНИЯ

Введем вспомогательные независимые случайные величины $\nu_r^{(1)}, \dots, \nu_r^{(N)}$ такие, что

$$\mathbf{P}\{\nu_r^{(i)} = k\} = \mathbf{P}\{\nu^{(i)} = k | \nu^{(i)} \neq r\}, \quad (18)$$

где $i = 1, \dots, N$, $k = 1, 2, \dots$. Пусть $\nu_N = \nu^{(1)} + \dots + \nu^{(N)}$, $\zeta_S^{(r)} = \nu_r^{(1)} + \dots + \nu_r^{(S)}$, где $S \leq N$. В основе доказательства теоремы 1 лежит следующая лемма [4, лемма 1].

Лемма 2. Справедливо равенство

$$\begin{aligned} \mathbf{P}\{\mu_r = k\} &= \binom{N}{k} q_r^k (1 - q_r)^{N-k} \\ &\times \frac{\mathbf{P}\{\zeta_{N-k}^{(r)} = N + n - kr\}}{\mathbf{P}\{\nu_N = N + n\}}. \end{aligned}$$

В [4, лемма 3] показано, что распределения сумм $(\nu_N - Na)/(b\sqrt{N})$ слабо сходятся к стандартному нормальному закону. Поскольку, в силу (5), параметр λ распределения (3) зависит от растущих N и n , суммы ν_N образуют схемы серий. В работе [2] найдены условия, при выполнении которых из слабой сходимости распределений серий сумм независимых решетчатых случайных величин к нормальному закону следует локальная сходимость. Поэтому в [4] соответствующее доказательство для ν_N было проведено путем проверки найденных в [2] условий. Задача оценки скорости локальной сходимости к нормальному закону распределений серий сумм решалась в теореме 5 статьи [2]. Для того чтобы сформулировать соответствующий результат, введем обозначение:

$$H(\nu) = \inf_{1/4 \leq d \leq 1/2} H(\nu, d), \quad (19)$$

где

$$H(\nu, d) = \mathbf{E}\langle \nu^* d \rangle,$$

ν^* означает симметризованную случайную величину ν , $\langle x \rangle$ – расстояние от x до ближайшего целого числа.

Поскольку локальная теорема для ν_N в условиях теоремы 1 доказана [4, лемма 4], из [2, теорема 5 и следствие из нее] вытекает такое уточнение этого результата.

Лемма 3. Пусть выполнены условия теоремы 1 и $u = (l - N - n)/(b\sqrt{N})$. Тогда для натуральных l

$$\left| \mathbf{P}\{\nu_N = l\} - \frac{1}{b\sqrt{2\pi N}} e^{-u^2/2} \right| \leq \Delta,$$

где

$$\Delta \leq C_9 \frac{\mathbf{E}\nu^3}{Nb^2 H(\nu)}.$$

В [4, равенство (5.7)] показано, что

$$q_r = \frac{\lambda^{r-1}}{rF^r(\lambda)} \mathbf{P}\{\xi_1 + \dots + \xi_r = r-1\},$$

где $r = 1, 2, \dots$, а $\xi_1, \dots, \xi_r$ – независимые одинаково распределенные по закону (1) случайные величины. Отсюда, из (5) и леммы 1 следует, что если $n/N \leq C_2 < \infty$, то

$$0 < C_{10} \leq \mathbf{E}\nu^3 \leq C_{11} < \infty. \quad (20)$$

Рассмотрим $H(\nu)$. Согласно лемме 1 из [2],

$$D(\nu, d) \leq H(\nu, d) \leq 4D(\nu, d), \quad (21)$$

где

$$D(\nu, d) = \inf_{a \in \mathbb{R}} \mathbf{E}\langle(\nu - a)d\rangle^2.$$

Расстояние от $(\nu - a)d$ до ближайшего целого при любых a и d не превосходит единицы. Следовательно,

$$D(\nu, d) \leq 1. \quad (22)$$

В лемме 2 [2] доказано, что при $1/4 \leq d \leq 1/2$

$$D(\nu, d) \geq \frac{d^2}{4} \sum_{k=1}^{\infty} \min\{\mathbf{P}\{\nu = k\}, \mathbf{P}\{\nu = k+1\}\}. \quad (23)$$

Лемма 4. Пусть выполнены условия теоремы 3, $u = (l - N - n)/(b\sqrt{N})$, где l – натуральное число. Тогда

$$\mathbf{P}\{\nu_N = l\} = \frac{1}{b\sqrt{2\pi N}} e^{-u^2/2} + O\left(\frac{N}{n^2}\right).$$

Доказательство. Поскольку

$$\mathbf{P}\{\nu = 1\} = p_0(\lambda), \quad \mathbf{P}\{\nu = 2\} = p_1(\lambda)p_0(\lambda),$$

из (3), (23) и леммы 1 видим, что при $d \geq 1/4$

$$D(\nu, d) \geq \frac{1}{64} p_0(\lambda)p_1(\lambda) > 0. \quad (24)$$

Если $0 < C_1 \leq n/N \leq C_2 < \infty$, то из (24) следует, что

$$D(\nu, d) \geq C_{12} > 0.$$

Отсюда и из (19), (21), (22), (24)

$$0 < C_{12} \leq H(\nu) \leq 4. \quad (25)$$

Из леммы 3, соотношений (12), (20), (25) следует утверждение леммы 4 при $0 < C_1 \leq n/N \leq C_2 < \infty$.

Пусть $n/N \rightarrow 0$. Из (3), (21), (24) и второго утверждения леммы 1 получаем, что

$$H(\nu) \geq C_{13} \frac{n}{N}. \quad (26)$$

Отсюда, из леммы 3 и соотношений (13), (20) опять приходим к утверждению леммы 4. $\square$

Локальная сходимость распределений $\zeta_{N-k}^{(r)}$ при $r \rightarrow \infty$ установлена в [4, лемма 6]. Пусть

$$S = N(1 - q_r)(1 + o(1)). \quad (27)$$

Покажем, что для сумм вида $\zeta_S^{(r)}$ верно утверждение, подобное лемме 4 для ν_N . Обозначим

$$a_r = \mathbf{E}\nu_r^{(1)} = \frac{a - rq_r}{1 - q_r},$$

$$b_r^2 = \mathbf{D}\nu_r^{(1)} = \frac{b^2}{(1 - q_r)^2} \left(1 - q_r - \frac{(a - r)^2}{b^2} q_r\right). \quad (28)$$

Если $N, n \rightarrow \infty$ так, что $n/N \leq C_2 < \infty$, то, используя (18), для $\mathbf{E}(\nu_r^{(1)})^3$ нетрудно получить оценку, аналогичную (20):

$$0 < C_{14} \leq \mathbf{E}(\nu_r^{(1)})^3 \leq C_{15} < \infty. \quad (29)$$

Из (14), (16) и (28) видим, что если $r \rightarrow \infty$, то $b_r^2 \sim b^2$. Но нам потребуется более точное соотношение между этими дисперсиями. Пусть $0 < C_1 \leq n/N \leq C_2 < \infty$. Тогда, используя (1)–(6) и первое утверждение леммы 1, нетрудно обнаружить, что

$$0 < C_{16} \leq a, b^2 \leq C_{17} < \infty. \quad (30)$$

Из (28), (30) теперь находим, что

$$b_r^2 = b^2 \left(1 - \frac{r^2}{b^2} q_r + O\left(rq_r \frac{N}{n}\right)\right). \quad (31)$$

Если $n/N \rightarrow 0$, то из (5), (6) и второго утверждения леммы 1 следует, что $a \rightarrow 1$, и из (13), (28) опять приходим к (31).

Следуя идее получения оценок (25), (26), нетрудно проверить, что они остаются в силе и для $\nu_r^{(1)}$. Это значит, что в случае $0 < C_1 \leq n/N \leq C_2 < \infty$

$$0 < C_{18} \leq H(\nu_r^{(1)}) \leq 4, \quad (32)$$

а если $n/N \rightarrow 0$, то

$$H(\nu_r^{(1)}) \geq C_{19} \frac{n}{N}. \quad (33)$$

Учитывая (27), (29), (32), (33) и заменяя в лемме 3 N на S , a на a_r и b^2 на b_r^2 , приходим к следующему результату.

Лемма 5. Пусть выполнены условия теоремы 3, условие (27) и $w_r = (l - Sa_r)/(b_r\sqrt{S})$, где l – натуральное число. Тогда

$$\mathbf{P}\{\zeta_S^{(r)} = l\} = \frac{1}{b_r\sqrt{2\pi S}} e^{-w_r^2/2} + O\left(\frac{S}{n^2}\right).$$

Для доказательства теоремы 3 нам требуется теорема 1 из [1, глава II, §5], в которой дана оценка скорости сближения биномиальных вероятностей с нормальным законом.

Лемма 6. Если $N \rightarrow \infty$, $(1 + v^6)/(Npq) \rightarrow 0$, где $0 < p < 1$, $q = 1 - p$, $v = (k - Nq)/\sqrt{Npq}$, то

$$\binom{N}{k} q^k (1 - q)^{N-k} = \frac{1}{\sqrt{2\pi Npq}} e^{-v^2/2} \times \left(1 + \frac{p - q}{6\sqrt{Npq}} (3v - v^3) + O\left(\frac{1 + v^6}{Npq}\right)\right).$$

ДОКАЗАТЕЛЬСТВО ОСНОВНОГО РЕЗУЛЬТАТА

Доказательство теоремы 3 опирается на использование леммы 2. Полагая $l = N + n$ в лемме 4, получаем, что

$$\mathbf{P}\{\nu_N = N + n\} = \frac{1}{b\sqrt{2\pi N}} + O\left(\frac{N}{n^2}\right). \quad (34)$$

Из (17) следует, что

$$k = Nq_r + v_r\sqrt{Nq_r}, \quad (35)$$

поэтому

$$N - k = N(1 - q_r) \left(1 - \frac{v_r}{1 - q_r} \sqrt{\frac{q_r}{N}}\right). \quad (36)$$

Из (14) и (36) видим, что выполнены условия леммы 5. Полагая в ней $l = N + n - kr$, $S = N - k$ и используя (6), (28), (31), (35), (36), прямыми вычислениями устанавливаем, что

$$w_r^2 = \frac{v_r^2 r^2 q_r}{b^2} + O\left(rq_r \frac{N}{n}\right). \quad (37)$$

Отсюда и из (16), (17)

$$e^{-w_r^2/2} = 1 - \frac{v_r^2 r^2 q_r}{2b^2} + O\left(rq_r \frac{N}{n}\right). \quad (38)$$

Таким образом, из (37), (38) и леммы 5 вытекает, что

$$\mathbf{P}\{\zeta_{N-k}^{(r)} = N + n - kr\} = \frac{1}{b_r\sqrt{2\pi(N-k)}}$$

$$\times \left(1 - \frac{v_r^2 r^2 q_r}{2b^2} + O\left(rq_r \frac{N}{n}\right)\right) + O\left(\frac{N}{n^2}\right). \quad (39)$$

Используя (13), (16), (31), (34), (39), не очень сложно обнаружить, что

$$\frac{\mathbf{P}\{\zeta_{N-k}^{(r)} = N + n - kr\}}{\mathbf{P}\{\nu_N = N + n\}} = 1 + \frac{r^2 q_r}{2b^2} (1 - v_r^2) + O\left(rq_r \frac{N}{n} + \frac{N}{n^{3/2}}\right). \quad (40)$$

Теперь обратимся к лемме 6. Полагая в ней $q = q_r$ и учитывая (14), (17), приходим к равенству

$$\binom{N}{k} q_r^k (1 - q_r)^{N-k} = \frac{1}{\sqrt{2\pi Nq_r(1 - q_r)}} e^{-v^2/2} \times \left(1 + \frac{1 - 2q_r}{6\sqrt{Nq_r(1 - q_r)}} (3v - v^3) + O\left(\frac{1}{Nq_r}\right)\right), \quad (41)$$

при этом

$$v = v_r \left(1 + \frac{q_r}{2} + O(q_r^2)\right). \quad (42)$$

С помощью (17), (42) находим, что

$$\frac{1}{\sqrt{2\pi Nq_r(1 - q_r)}} e^{-v^2/2} = \frac{1}{\sqrt{2\pi Nq_r}} e^{-v_r^2/2} \times \left(1 + \frac{1 - v_r^2}{2} q_r + O(q_r^2)\right), \quad (43)$$

$$1 + \frac{1 - 2q_r}{6\sqrt{Nq_r(1 - q_r)}} (3v - v^3) + O\left(\frac{1}{Nq_r}\right) = 1 + \frac{3v_r - v_r^3}{6\sqrt{Nq_r}} + O\left(\sqrt{\frac{q_r}{N}} + \frac{1}{Nq_r}\right). \quad (44)$$

Из (41)–(44) следует равенство

$$\binom{N}{k} q_r^k (1 - q_r)^{N-k} = \frac{1}{\sqrt{2\pi Nq_r}} e^{-v_r^2/2} \times \left(1 + \frac{3v_r - v_r^3}{6\sqrt{Nq_r}} + \frac{1 - v_r^2}{2} q_r + O\left(q_r^2 + \sqrt{\frac{q_r}{N}} + \frac{1}{Nq_r}\right)\right).$$

Отсюда, из (16), (40) и леммы 2 нетрудно получить утверждение теоремы 3.

Автор выражает глубокую благодарность рецензенту за выявленные ошибки и полезные рекомендации.

ЛИТЕРАТУРА

1. Колчин В. Ф., Севастьянов Б. А., Чистяков В. П. Случайные размещения. М.: Наука, 1976. 223 с.
2. Мухин А. В. Локальные предельные теоремы для решетчатых случайных величин // Теория вероятностей и ее применения. 1991. Т. 36, вып. 4. С. 660–674.
3. Павлов Ю. Л. Об объемах деревьев леса Гальтона–Ватсона с бесконечной дисперсией в критическом случае // Дискретная математика. 2024. Т. 36, вып. 2. С. 33–49. doi: 10.4213/dm1813
4. Павлов Ю. Л., Чеплюкова И. А. Объемы деревьев случайного леса и конфигурационные графы // Труды МИАН им. В. А. Стеклова. 2022. Т. 316. С. 298–315. doi: 10.4213/tm4216

Поступила в редакцию / received: 31.01.2026; принята к публикации / accepted: 26.05.2026.
Автор заявляет об отсутствии конфликта интересов / The author declares no conflict of interest.

СВЕДЕНИЯ ОБ АВТОРЕ:

Павлов Юрий Леонидович

д-р физ.-мат. наук, профессор, главный научный сотрудник

e-mail: pavlov@krc.karelia.ru

REFERENCES

1. Kolchin V. F., Sevastianov B. A., Chistiakov V. P. Random allocations. New York: Wiley; 1978. 262 p.
2. Mukhin A. B. Local limit theorems for lattice random variables. *Theory of Probability and its Applications*. 1992;36(4):698–713. doi: 10.1137/1136086
3. Pavlov Yu. L. On the sizes of trees in a Galton–Watson forest with infinite variance in the critical case. *Diskretnaya matematika = Discrete Mathematics and Applications*. 2024;36(2):33–49. (In Russ.). doi: 10.4213/dm1813
4. Pavlov Yu. L., Cheplyukova I. A. Sizes of trees in a random forest and configuration graphs. *Proceedings of the Steklov Institute of Mathematics*. 2022;316:280–297. doi: 10.1134/S0081543822010205

CONTRIBUTOR:

Pavlov, Yury

Dr. Sci. (Phis.-Math.), Professor, Chief Researcher

УДК 339.5:519.17

ПРИМЕНЕНИЕ ИНДЕКСА ДРУЖБЫ И ФИЛЬТРА ДИСПАРИТЕТА ДЛЯ АНАЛИЗА МИРОВОЙ ТОРГОВОЙ СЕТИ

А. А. Печников

*Институт прикладных математических исследований КарНЦ РАН,
ФИЦ «Карельский научный центр РАН» (ул. Пушкинская, 11, Петрозаводск,
Республика Карелия, Россия, 185910)*

Ориентированные графы и сети, имеющие небольшой масштаб, высокую плотность и распределение весов дуг с тяжелым хвостом, являются интересными объектами исследований в силу указанных особенностей. Одним из ярких примеров таких сетей является Всемирная торговая сеть экспортно-импортных торговых потоков между странами, плотность которой превышает 90 %. При анализе сети, как правило, используются базовые характеристики, традиционно применяемые в теории графов и сетевом анализе. В работе предлагается подход, основанный на совмещении «парадокса дружбы», используемого для оценки локальной значимости участников сети, и функции диспаритета, применяемого для отсеивания слабых связей на основе статистической значимости. Полученные результаты и их содержательная интерпретация позволяют глубже понять сложную структуру, скрытую за малой размерностью таких сетей.

Ключевые слова: модели экспорта и импорта; теория графов; сетевой анализ; индекс дружбы; функция диспаритета

Для цитирования: Печников А. А. Применение индекса дружбы и фильтра диспаритета для анализа мировой торговой сети // Труды Карельского научного центра РАН. 2026. № 6. С. 89–99. doi: 10.17076/mat2312

Финансирование. Финансовое обеспечение исследований осуществлялось из средств федерального бюджета на выполнение государственного задания КарНЦ РАН (Институт прикладных математических исследований КарНЦ РАН).

A. A. Pechnikov. ANALYSIS OF THE WORLD TRADE NETWORK USING THE FRIENDSHIP INDEX AND DISPARITY FILTER

Institute of Applied Mathematical Research, Karelian Research Centre, Russian Academy of Sciences (11 Pushkinskaya St., 185910 Petrozavodsk, Karelia, Russia)

Directed graphs and networks that have a small scale, high density, and a heavy-tailed distribution of edge weights are interesting research objects due to these specified characteristics. A prominent example of such networks is the World Trade Network of export-import flows between countries, whose density exceeds 90 %. When analyzing a network, researchers typically rely on basic

characteristics traditionally used in graph theory and network analysis. This paper proposes an approach based on combining the “friendship paradox”, used to assess the local significance of network participants, with the disparity filter, applied to discard weak ties based on statistical significance. The obtained results and their substantive interpretation allow for a deeper understanding of the complex structure hidden behind the small dimensionality of such networks.

Keywords: export-import models; graph theory; network analysis; friendship index; disparity filter

For citation: Pechnikov A. A. Analysis of the world trade network using the friendship index and disparity filter. *Trudy Karelskogo nauchnogo tsentra RAN = Transactions of the Karelian Research Centre RAS*. 2026. No. 6. P. 89–99. doi: 10.17076/mat2312

Funding. The studies were funded from the federal budget through state assignment to the Karelian Research Centre RAS (Institute of Applied Mathematical Research, Karelian Research Centre RAS).

ВВЕДЕНИЕ

Экспортно-импортные торговые потоки являются одними из ключевых объектов анализа в эмпирических исследованиях международной торговли, позволяющих количественно оценивать интенсивность двусторонних отношений и структурные свойства глобальной торговой системы в целом. Сетевой анализ и теория графов прочно вошли в инструментарий экономических исследований в России и за рубежом. Поскольку ключевые вопросы международной торговли, такие как экспорт и импорт, являются основным объектом анализа в данной статье, представим краткий обзор современных публикаций, посвященных применению этих методов.

Классической работой в области применения сетевого анализа и теории графов по изучению международной торговли считается статья [5]. Эта статья заложила методологическую основу для нового направления в экономических исследованиях, продемонстрировав силу сетевого подхода для понимания глобальных торговых взаимосвязей, их устойчивости и эволюции, что подтверждается ее активным цитированием в последующих работах.

В частности, в ней сформулировано определение Всемирной торговой сети (International Trade Network, ВТС) как композиции двух составляющих. Первая составляющая — это граф $G = G(V, L)$, где $V = \{1, 2, \dots, n\}$ — множество вершин (стран), а L — множество дуг, то есть ориентированных связей (торговых потоков) между парами вершин, идущих от страны-экспортера i к стране-импортеру j в случае наличия такой экспортно-импортной операции. Вторая составляющая — это пара (W, P) , включающая дополнительную информацию о соответствующих характеристиках

вершин и связей между ними. Положительные элементы $w_{ij} \in W$ выступают в качестве диадных весов на G , модифицируя его исходную бинарную структуру и преобразуя простой направленный граф во взвешенный, где w_{ij} указывает на силу связи между страной i и страной j (например, объемы экспорта). Элементы $p_i \in P$, в свою очередь, включают специфические для страны значения (например, доход, население, географическое местоположение).

При анализе сети используются базовые характеристики, традиционно применяемые в теории графов и сетевом анализе: степень вершины, полустепени захода/исхода, коэффициент кластеризации, плотность графа. Для оценки положения вершины относительно ее окрестности или всей сети используются меры центральности по степени, близости и по посредничеству. Сетевой анализ применяется в сочетании с гравитационной моделью международной торговли [4], основанной на предположении, что величина взаимодействия участников пропорциональна произведению показателей их значимости (например, объему ВВП) и обратно пропорциональна расстоянию между ними. Результаты показывают, что глобальный уровень торговой интеграции продолжает расти, за исключением некоторых регионов, существует сильная неоднородность в выборе странами торговых партнеров, а члены Всемирной торговой организации связаны между собой теснее, чем остальной мир.

В работе [8] исследованы взаимосвязи между первыми 50 странами мира по объему экспорта с помощью программы для анализа социальных сетей Gephi. Сеть построена по данным Всемирного банка об экспортных потоках между странами за период 2010–2020 годов, что позволяет оценить долгосрочные тренды. Применены стандартные метрики сетево-

го анализа – центральность по степени, степени близости и посредничества, а также анализ модулярности для выявления сообществ. Исследование подтверждает, что глобальная торговая система остается высоко неоднородной и иерархической, но ее центр тяжести смещается в Азию.

В работе российских авторов с использованием сетевых методов проводится анализ структуры экспорта и импорта на примере рынка агропромышленных товаров для определения факторов, влияющих на торговые потоки между государствами, поскольку, как говорится в статье, «...статистический анализ мировых торговых отношений не всегда позволяет выявить все существенные аспекты сотрудничества» [1, с. 421]. В исследовании используются упомянутые ранее стандартные показатели с акцентом на определение важности вершин посредством центральностей. По данным Всемирного банка был построен и проанализирован граф для 10 наиболее крупных мировых экспортеров/импортеров сельскохозяйственной продукции и их 20 крупнейших партнеров по рынку агропромышленных товаров. Полученные характеристики центральности этого графа затем используются в регрессионной модели импорта. Подтверждено предположение о значимости показателя центральности по собственному вектору, демонстрирующего положение страны в мировой торговой сети.

Несколько иной подход с использованием теоретико-графовой модели международной торговли, рассмотренный в работе [2], позволяет решить задачу о том, каковы должны быть соотношения госбюджетов торгующих между собой стран, чтобы не было существенного дефицита торгового баланса каждой из стран-участниц. Определяется понятие структурной матрицы торговли A , где ее элементы a_{ij} – часть госбюджета, которую j -я страна тратит на закупки товаров i -й страны. С применением теории графов сформулирована и доказана теорема, позволяющая определить существование и единственность точного сбалансированного решения для общего случая неотрицательной структурной матрицы торговли (под сбалансированностью международной торговли понимается бездефицитность торгового бюджета каждого участника). Результат демонстрируется на условном примере для шести участников.

Сети международной торговли, вне зависимости от количества участников и принципов их объединения в одно исследуемое множество, имеют достаточно небольшие разме-

ры, хотя и могут обладать чертами сложных сетей (например, гетерогенностью) [13]. В данном случае содержательно гетерогенность сети подчеркивает как неравномерное распределение торговых партнеров, так и неравномерное распределение суммарных объемов экспорта и импорта по странам, что означает наличие большого количества слабых связей.

Для отсеивания слабых связей на основе статистической значимости, а не просто по абсолютному или относительному порогу, может быть использован фильтр диспаратета (disparity filter). Метод был предложен в работе [15], и в ней же получил название «фильтр диспаратета». Несколько ранее этими же авторами данный метод, еще не получивший своего названия, был применен в исследовании вопросов международной торговли [16], чтобы отсеять статистически незначимые связи и оставить только «каркас» (backbone) – наиболее важные для глобальной архитектуры торговые потоки. Основной вывод заключается в том, что глобальная торговая система, несмотря на свою кажущуюся плотность, на самом деле держится на небольшом количестве стратегически важных «магистралей» (highways), которые и образуют ее устойчивую архитектуру, а фильтр диспаратета – это инструмент, который позволяет объективно выявить указанные магистрали. В нашем исследовании мы будем использовать этот подход для удаления слабых связей при сохранении сильной связности сети.

Анализ публикаций показывает, что в современных исследованиях международных торговых сетей для определения важности вершин-участников преимущественно используются взвешенные расширения классических показателей центральности (влиятельность, близость, степень посредничества). Возможно, это связано с некоей «стандартизацией» – результаты сравнимы между разными исследованиями.

В нашем исследовании предлагается иной подход, основанный на так называемом парадоксе дружбы. Парадокс дружбы (англ. friendship paradox) – это явление, которое в 1991 году обнаружил социолог Scott L. Feld, изучая социальные сети, и которое состоит в том, что, как правило, у большинства людей друзей меньше, чем в среднем у их друзей [7]. Заметим, что это было еще до появления Facebook, ВКонтакте и им подобных, и поэтому термин «социальные сети» имел точное социологическое определение как устойчивая структура социальных отношений между отдельными индивидами или группами [3]. Ло-

кальным сетевым показателем, количественно оценивающим выполнение парадокса дружбы для отдельной вершины, является индекс дружбы (friendship index) как мера зависимости вершины от смежных с ней вершин. Применение индекса дружбы позволяет, с одной стороны, проверить выполнение парадокса дружбы для узлов сети, а с другой – интерпретировать его значение как характеристику относительной значимости вершины.

Материалы, изложенные в статье, представляют собой результаты решения двух задач с использованием реальных данных по экспорту-импорту – определения значимости вершин графа ВТС с использованием индекса дружбы и определения сообществ графа ВТС с применением фильтра диспаритета – с проведением сравнения с традиционными методами и содержательной интерпретацией результатов. Полученные результаты предлагают еще один подход к пониманию взаимосвязей в торговых сетях, что позволяет глубже изучить их сложную структуру, скрытую за малой размерностью таких сетей.

ГРАФ ВСЕМИРНОЙ ТОРГОВОЙ СЕТИ

Приведем описание используемой далее теоретико-графовой модели ВТС, в основном следуя работе [5] с адаптацией к нашим задачам.

Граф ВТС – это граф $G = (V, E, W)$, где

- V – множество вершин, соответствующих участникам внешнеторговых операций, $|V| = n$, идентификаторами вершин являются англоязычные названия участников в стандарте UN Comtrade [18] (USA, Russian Federation, Côte d'Ivoire, Türkiye),
- E – множество дуг (v_i, v_j) , то есть ориентированных связей (торговых потоков) между парами вершин, идущими от экспортера v_i к импортеру v_j в случае наличия такой экспортно-импортной операции, $|E| = m$,
- W – множество весов дуг, $w_{ij} \in W$ указывает на силу связи между v_i и v_j как объема экспорта.

Реальные данные получены с использованием базы данных Comtrade Организации Объединенных Наций (UN Comtrade Database, [18]), содержащей подробную глобальную годовую и ежемесячную статистику торговли в мире. За 2020 год найдено 236 участников экспортно-импортных торговых операций по всем категориям товаров (материальным продуктам), без учета услуг.

Выбор 2020 года в качестве анализируемого достаточно условен. Дополнительным фактором такого выбора является пандемия COVID-19, позволяющая посмотреть на адаптацию взвешенной ориентированной сети к экстремальному внешнему воздействию.

Заметим, что в UN Comtrade Database объектом статистики является «таможенная территория» (customs area), что для нашего исследования не имеет принципиального значения, поэтому мы для простоты будем говорить «страна». Результаты запросов к базе данных были сведены в таблицу, содержащую построчную информацию по экспорту за 2020 год в следующей форме:

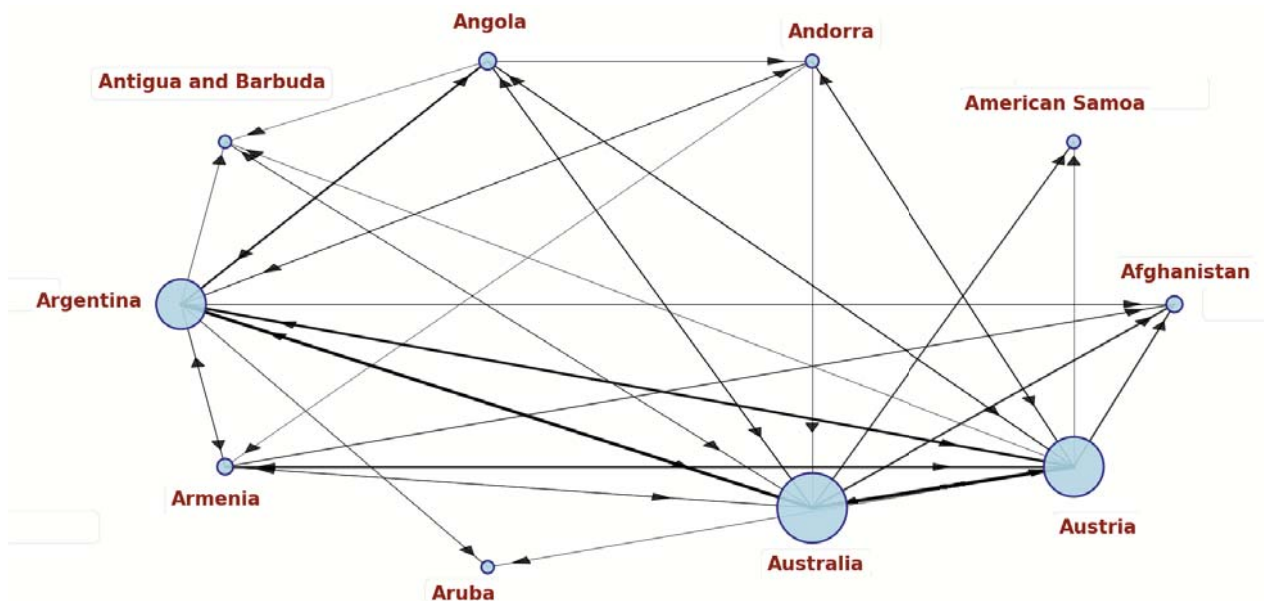
<страна-экспортер><страна-импортер><экспорт в млн US\$>

Граф ВТС, построенный по этим данным, содержит 236 вершин и 22 868 дуг. Он состоит из одной компоненты (слабой) связности и при этом содержит 70 «висячих» вершин, соответствующих 70 странам, которые только импортируют товары. Небольшой фрагмент графа для 10 первых по алфавиту стран из множества вершин V приведен на рисунке; здесь даже при наличии таких экзотических участников, как Антигуа и Барбуда и Аруба, обнаруживается высокая плотность 0,422.

Далее будет рассматриваться ориентированная версия парадокса дружбы, для которой требуется сильно связный граф, в котором каждая вершина имеет входящие и исходящие дуги [9], поэтому «висячие» вершины были удалены из графа G . Удаление 70 «висячих» вершин порождает граф $G_{166} = (V, E, W)$, у которого 166 вершин и 18 877 дуг. Граф обладает сильной связностью и высокой плотностью, равной 0,689, и является базовым объектом исследований.

ИНДЕКС ДРУЖБЫ И ФУНКЦИЯ ДИСПАРИТЕТА ПРИМЕНИТЕЛЬНО К ГРАФУ ВТС

Подход, основанный на использовании индекса дружбы и фильтра диспаритета, применительно к графам ВТС в кратком виде можно описать так. Индекс дружбы в данном случае рассматривается как показатель «значимости» вершин графа, дополняющий классические показатели (показатели центральности, Page Rank) и добавляющий к анализу структуры анализ качества связей. Фильтр диспаритета для графов, как уже было сказано, – это статистический метод для выделения значимого каркаса графа за счет отбрасывания несущественных связей. Совместное при-



Фрагмент графа, соответствующий реальной BTC за 2020 год
A fragment of the graph corresponding to the actual ITN for 2020

менение индекса дружбы и фильтра диспаритета позволяет выявлять наиболее значимые вершины не только на исходном графе, но и на его каркасе, выявляя, таким образом, устойчивое ядро, не зависящее от слабых связей.

В терминологии теории графов парадокс дружбы в своей первоначальной постановке утверждает, что в простом связном неориентированном графе средняя степень вершин, смежных с данной, не меньше, чем степень самой вершины. Равенство достигается только в случае, если в графе степени всех вершин равны [6]. Количественным показателем для измерения парадокса дружбы является индекс дружбы.

Процесс развития исследований парадокса дружбы от базовой постановки к его более общим формам, учитывающим веса ребер, ориентацию и произвольные атрибуты вершин, систематизирован в [6].

В работе [9] вместо термина «индекс дружбы» с учетом более общей постановки используется термин «отклонение дружбы» (friendship bias) $\Delta_i^{(r)}$, определяемое для вершины i и произвольного показателя центральности r как

$$\Delta_i^{(r)} = (1/d_i) \sum_{j \sim i} r_j - r_i, \quad (1)$$

где

- d_i – это исходящая степень вершины i ,
- $j \sim i$ – множество всех вершин j , смежных с вершиной i ,
- r_i – показатель центральности вершины i .

Используем формулу (1) как основу для вычисления индекса дружбы вершин графа BTC с учетом того, что граф BTC является взвешенным ориентированным графом, следовательно, надо учесть, какой индекс дружбы рассматривается – по импорту или по экспорту. Важным фактором является выбор показателя, используемого в качестве атрибутов вершин. В этой работе мы остановимся на взвешенной степени (weighted degree), которую принято включать в семейство показателей центральности, в особенности когда это касается ориентированных графов [14]. Учтем также, что этот показатель очень легко интерпретируется содержательно: исходящая степень – это суммарный экспорт страны, а входящая степень – ее суммарный импорт.

Введем следующие обозначения для формулы индекса дружбы по экспорту для графа $G = (V, E, W)$:

$\{1, 2, \dots, n\}$ – множество индексов (порядковых номеров), соответствующих странам множества V ,

Δ_i^{exp} – индекс дружбы по экспорту для i -й вершины,

W – матрица смежности графа G , построенная по множеству весов дуг W ,

$w_{ij} \geq 0$ – вес дуги (i, j) (объем экспорта из страны i в страну j),

d_i^{out} – исходящая степень вершины i , вычисляемая по матрице E ,

$N^{out}(i)$ – множество конечных вершин для дуг, у которых начальной вершиной является i .

Тогда $r_i = \sum_{j \neq i} w_{ij}$ и формулу индекса дружбы по экспорту можно записать следующим образом:

$$\Delta_i^{exp} = 1/d_i^{out} \sum_{j \in N^{out}(i)} r_j - r_i.$$

Случай, когда $\Delta_i^{exp} < 0$, означает, что экспортная привлекательность страны i выше, чем средняя экспортная привлекательность стран, которые являются импортерами для i , и напротив, когда $\Delta_i^{exp} > 0$, страны, в которые она экспортирует, в среднем более привлекательны.

Формула для Δ_i^{exp} может быть записана по аналогии с учетом того, что рассматривается входящая степень d_i^{in} , суммарный импорт в качестве r_i и подмножество $N^{in}(i)$. Интерпретация в этом случае аналогична случаю с экспортом: отрицательное значение Δ_i^{imp} означает большую привлекательность страны i по импорту, и наоборот.

На примере графа G_{166} и его фрагмента можно было убедиться в такой особенности графа ВТС, как его высокая плотность. При этом распределение весов дуг крайне асимметричное с тяжелым хвостом: от наибольшего значения 452492,9 до минимального 0,000001 (в миллионах US\$).

Для того чтобы «проредить» очень плотные взвешенные ориентированные графы, естественным кажется задание некоего порога с последующим удалением всех дуг, веса которых ниже этого порога. Очевидным недостатком такого подхода является игнорирование относительной значимости дуги в контексте ее локального окружения, возможное удаление критически важных дуг для вершин малой инцидентности, а значит, потенциального разрушения иерархической организации графа.

Подход, основанный на фильтре диспаритета [15], для каждой вершины графа проверяет статистическую гипотезу о том, что веса ее инцидентных дуг распределены случайным образом. Нулевая гипотеза предполагает, что после выделения минимального ненулевого веса на каждую существующую дугу оставшийся суммарный вес вершины распределяется равномерно между всеми ее дугами. Для каждой дуги вычисляется р-значение – вероятность получить в рамках этой нулевой гипотезы вес, больший или равный наблюдаемому, по формуле

$$pvalue(i, j) = (1 - w_{ij}/s_i)^{k_i - 1},$$

где

- w_{ij} – вес дуги из вершины i в вершину j ,
- s_i – суммарный вес всех дуг, инцидентных вершине i (сила вершины),

- k_i – степень вершины i (количество инцидентных дуг).

Дуги со значением $pvalue$ выше заданного порога α считаются статистически незначимыми и удаляются. Выбор порога α во многом зависит от целей исследования: выделение ядра графа, очистка от явного «шума», сравнение структуры графа при разных значениях порога.

Фильтр диспаритета в работе [15] рассматривается применительно к неориентированным графам, а его расширение на ориентированные графы описано в приложении [17]. Здесь предлагается формировать каркас графа путем сохранения всех входящих и исходящих связей, которые превосходят порог значимости хотя бы для одной из двух вершин на концах дуг, с отбрасыванием остальных. Более жесткий критерий заключается в том, что порог значимости должен быть превзойден для обеих вершин.

Отличие мягкого и жесткого критериев заключается, в частности, в том, что в первом случае сохраняется ориентированная архитектура графа, а во втором сохраняются только взаимно значимые дуги, что сильно снижает асимметрию. В данной работе далее будет использоваться мягкий критерий, как сохраняющий связность при меньшем количестве дуг. При этом порог значимости задается пользователем из содержательных соображений и, по сути, представляет собой «порог фильтрации».

Опишем работу фильтра диспаритета в виде алгоритма в соответствии с подходом, изложенным в [17], использующего мягкий критерий сохранения дуг.

Исходные данные

1. Взвешенная матрица смежности ориентированного графа A размерности $n \times n$, где a_{ij} – вес дуги, $a_{ij} = 0$ означает отсутствие дуги, $a_{ii} = 0$ по определению.

2. Задаваемый пользователем порог фильтрации α ($0 < \alpha < 1$).

Шаг 1: Предварительный расчет локальных характеристик всех вершин

1. k_i^{out} – исходящая степень узла i , $k_i^{out} := |\{j : a_{ij} > 0\}|$,

2. s_i^{out} – исходящая сила узла i , $s_i^{out} := \sum_{j=1}^n a_{ij}$,

3. k_i^{in} – входящая степень узла i , $k_i^{in} := |\{j : a_{ji} > 0\}|$,

4. s_i^{in} – входящая сила узла i , $s_i^{in} := \sum_{j=1}^n a_{ji}$.

Шаг 2: Расчет р-значений для каждой дуги

1. p_{ij}^{out} – значимость дуги со стороны вершины-источника i .

Если $k_i^{out} \leq 1$ или $s_i^{out} = 0$, то $p_{ij}^{out} := 1$,

иначе $p_{ij}^{out} := (1 - a_{ij}/s_i^{out})^{(k_i^{out}-1)}$.

2. p_{ij}^{in} – значимость дуги со стороны вершины-приемника j .

Если $k_j^{in} \leq 1$ или $s_j^{in} = 0$, то $p_{ij}^{in} := 1$,

иначе $p_{ij}^{in} := (1 - a_{ij}/s_j^{in})^{(k_j^{in}-1)}$.

Шаг 3: Применение мягкого критерия и сравнение с порогом α

Для каждой дуги (i, j) вычисляется p_{ij}^{gentle} по формуле $p_{ij}^{gentle} := \min(p_{ij}^{out}, p_{ij}^{in})$.

Для каждой дуги (i, j) : если $p_{ij}^{gentle} < \alpha$, то a_{ij} в матрице A сохраняется, иначе $a_{ij} := 0$.

Результат

По матрице A построена матрица $A^{backbone}$, являющаяся взвешенной матрицей смежности графа-каркаса исходного графа.

Алгоритмы вычисления индекса дружбы, фильтра диспаритета и использующихся далее классических мер центральности реализованы на языке Python (версия 3.8.2).

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ И АНАЛИЗ

В этом разделе приведен ряд характерных результатов более широкого исследования, де-

монстрирующих возможности применения индекса дружбы и функции диспаритета на примере графа ВТС.

Как говорилось ранее, базовым объектом исследования является граф $G_{166} = (V, E, W)$. Напомним некоторые его характеристики: 166 вершин, 18 877 дуг, крайне асимметричное распределение с тяжелым хвостом весов множества дуг, сильная связность, высокая плотность 0,689.

Для графа G_{166} были вычислены значения Page Rank, степени близости (closeness centrality) и индексов дружбы по экспорту и по импорту. В результате 32 вершины имеют значения Δ_i^{exp} и Δ_i^{imp} меньше нуля (причем это один и тот же список вершин, как для экспорта, так и для импорта). Это означает сильную иерархию торговой сети и соответствует выводу о том, что Всемирная торговая сеть — «... это сильно неоднородная структура, где небольшое число стран занимают центральное положение, тогда как подавляющее большинство стран остаются на периферии» [5, р. 1417].

В таблице для графа G_{166} приведены первые 15 наиболее значимых стран по каждому из показателей (названия стран в колонках PageRank и Closeness упорядочены по убыванию показателя, а в колонках Δ_{exp} и Δ_{imp} — по возрастанию).

Значения показателей важности для 15 стран
Importance indicator values for 15 countries

PageRank	Closeness	Δ_{exp}	Δ_{imp}
USA	China	China	China
China	Hong Kong SAR	USA	USA
Germany	Germany	Germany	Germany
United Kingdom	USA	Japan	Japan
France	Japan	Netherlands	Netherlands
Netherlands	Rep. of Korea	Hong Kong SAR	Hong Kong SAR
Canada	Mexico	Rep. of Korea	Italy
Italy	Canada	Italy	Rep. of Korea
Japan	Australia	France	France
Hong Kong SAR	Netherlands	Canada	Mexico
Mexico	Brazil	United Kingdom	Canada
Belgium	Switzerland	Mexico	United Kingdom
Switzerland	France	Singapore	Singapore
India	Poland	Russian Federation	Russian Federation
Spain	Italy	Switzerland	Switzerland

Один из первых моментов в таблице, бросающихся в глаза читателю, не посвященному глубоко в проблемы мировой торговли, заключается в наличии такого субъекта, как

Hong Kong SAR (Специальный административный район Гонконг). Изучение публикаций подтверждает правильность результата: Гонконг является, говоря по-простому, перевалоч-

ным хабом для Китая [10], 90 % экспорта Гонконга – это реэкспорт, а не то, что производится здесь. Заметим также, что Гонконг стоит в рейтинге Closeness выше, чем США, и возможное объяснение этому факту заключается как в связях с Китаем, так и в географии – это центр Восточной Азии, быстрорастущего региона.

Три страны – Китай, США и Германия – находятся в первой четверке по всем четырем показателям. В терминах случайного блуждания по сети высокий рейтинг Page Rank можно интерпретировать как указание на страны, «где деньги оседают», степень близости показывает, насколько сложно или легко стране проводить экспортно-импортные операции в рамках торговой сети, а индексы дружбы показывают относительное положение среди ближайших соседей по этим операциям. Следовательно, можно сделать вывод о том, что Китай, США и Германия – глобальные лидеры, которые доминируют и локально, что соответствует [19, р. 58].

Как уже сказано, Всемирная торговая сеть – это очень неоднородная структура. Поэтому на следующем шаге исследования было сформировано множество, содержащее 71 страну, каждая из которых имеет суммарный внешнеторговый оборот за 2020 год больше 30 млрд \$. На долю этих стран приходится 94 % глобального экспорта и импорта, образуя, таким образом, ядро мировой торговой сети. Граф G_{71} , построенный из графа G_{166} путем удаления остальных 95 вершин, не вошедших в ядро, и смежных с ними дуг, содержит 4917 дуг, состоит из одной компоненты сильной связности и имеет очень высокую плотность 0,989.

Для графа G_{71} также были вычислены показатели PageRank, Closeness и индексов дружбы по экспорту и по импорту. Теперь в графе G_{71} уже только 22 вершины имеют значения Δ_i^{exp} и Δ_i^{imp} меньше нуля (опять-таки это один и тот же список вершин как для экспорта, так и для импорта).

Если построить таблицу, аналогичную представленной в этой работе, то самое существенное изменение, которое произошло, заключается в том, что Германия по Closeness, Δ_{exp} и Δ_{imp} поднялась на вторые позиции, оставшись на третьей по PageRank. Позиции по PageRank понятны – уход небольших по экспорту/импорту стран не повлиял на глобальную привлекательность. Рост показателя близости говорит о том, что в G_{71} у Германии исчезли длинные цепочки связей, а на рост показателей индексов дружбы повлиял уход

небольших по экспорту/импорту стран, которые тянули средний показатель вниз.

Несмотря на то что страны, вошедшие в качестве вершин в граф G_{71} , имеют значительные показатели суммарного экспортно-импортного оборота, это не избавило граф от очень большой разницы в весах дуг: от наибольшего значения 452492,9 (как и ранее) до минимального 0,00028 (в миллионах \$). Минимальное значение по сравнению с графом G_{166} выросло в 280 раз, но остается в миллиарды раз меньше максимального.

Применим функцию диспаратета к графу G_{71} в попытке оставить только те дуги, которые действительно важны для функционирования системы. Вначале было найдено минимальное пороговое значение $\alpha = 0,00018$, сохраняющее сильную связность графа. Такое низкое значение α говорит о том, что сохранившиеся дуги должны быть доминирующими для одного из своих концов, то есть это связи, через которые проходит почти весь экспорт (или импорт) для этой вершины. Граф, полученный при этом значении α , назовем каркасом графа G_{71} и обозначим $G_{71}^{backbone}$.

В графе $G_{71}^{backbone}$ при 71 вершине и сохранении сильной связности осталось всего 227 дуг (плотность 0,046). Вычисление показателей PageRank, Closeness, Δ_{exp} и Δ_{imp} оставляет трех прежних лидеров – Китай, США и Германию – по всем показателям, кроме близости. По показателю Closeness Китай остался безусловным лидером, в то время как Германия опустилась на 7-е место, а США на 12-е.

Вполне возможно, что такой результат может быть объяснен объективной экономической остановкой, вызванной пандемией: Китай остался лидером, потому что он был единственным крупным игроком, который в 2020 году работал «как обычно», в отличие от США и Германии (но это утверждение требует дополнительных исследований).

Задачи нахождения сообществ в ориентированных графах с очень высокой плотностью, как в случае с графом ВТС, имеют большое значение для выявления блоков наиболее интенсивно торгующих стран. Известный алгоритм Лихта–Ньюмана (Leicht–Newman) нахождения сообществ в ориентированном графе [11] как обобщение на взвешенные графы реализован в библиотеке Python в виде функции `nx.modularity` [12], которую мы используем далее. Здесь для оценки качества разбиения на сообщества используется мера модулярности Q , учитывающая общую сумму весов в графе и каждом сообществе.

Считается, что $0,2 \leq Q \leq 0,3$ означает приемлемое значение для выявления сообществ, а $0,3 < Q \leq 0,5$ – выявление значимой структуры сообществ.

Известно, что в очень плотных графах выделение сообществ часто не дает содержательных результатов, поскольку модулярность таких графов невысока. Поэтому мы провели эксперимент с уже известными нам сильно связными графами G_{71} и $G_{71}^{backbone}$, то есть с очень плотным графом и графом с очень низкой плотностью, полученным из первого с использованием функции диспаритета.

В первом случае Q равно 0,283, а во втором 0,443, то есть в графе $G_{71}^{backbone}$ выявлена значимая структура сообществ. Некоторые результаты и их возможная содержательная интерпретация показывают перспективность такого подхода.

В графе G_{71} получены пять сообществ; перечислим их, присвоив каждому достаточно условное название:

«Европейское торговое ядро с расширением» (ЕТЯР, 28 стран): 22 страны ЕС (Австрия, Бельгия, Болгария, Хорватия, Чехия, Дания, Финляндия, Франция, Германия, Греция, Венгрия, Ирландия, Италия, Люксембург, Нидерланды, Польша, Португалия, Румыния, Словакия, Словения, Испания, Швеция), 2 страны – кандидаты в ЕС (Турция, Сербия), 3 европейских страны не ЕС (Норвегия, Швейцария, Великобритания) и Марокко;

«Азиатско-Тихоокеанское ядро с расширением» (АТЯР, 24 страны): Азия, 16 стран (Камбоджа, Китай, Гонконг, Индия, Индонезия, Иран, Япония, Малайзия, Мьянма, Пакистан, Филиппины, Катар, Корея, Сингапур, Таиланд, Вьетнам), Латинская Америка, 4 страны (Аргентина, Бразилия, Чили, Перу), Океания, 2 страны (Австралия, Новая Зеландия), Африка, 2 страны (Нигерия, ЮАР);

«Постсоветское пространство» (7 стран): Беларусь, Казахстан, Латвия, Литва, Россия, Украина и Эстония;

«Американский континент и Израиль» (7 стран): Израиль, Канада, Колумбия, Мексика, Панама, США и Эквадор;

«Страны Аравийского полуострова и Египет» (5 стран): Египет, Кувейт, ОАЭ, Оман, Саудовская Аравия.

В графе $G_{71}^{backbone}$ шесть сообществ (26, 25, 9, 5, 4 и 2 участника), их условные названия сохранились с некоторыми количественными изменениями и добавилось сообщество «Балтия» из двух стран, о чем речь пойдет ниже.

В графе $G_{71}^{backbone}$ из 28 стран ЕТЯР Ирландия и Великобритания ушли в «Американский

континент и Израиль». Марокко же устойчиво присутствует в обоих вариантах ЕТЯР – ее связи с Европой сильны настолько, что не зависят от порога отсека.

В графе $G_{71}^{backbone}$ к АТЯР добавился Египет, из чего следует, что сообщества при удалении слабых связей могут и прирастать, а не только сжиматься. Возможно, Египет – пример страны, реальная торговая ориентация которой на Азию и БРИКС была скрыта за формальными слабыми связями с арабским миром и Европой.

Приведем пример уменьшения сообщества. Сообщество «Постсоветское пространство» в графе G_{71} показывает, что торговые связи разрушить сложнее, чем политические, а география и инфраструктура все еще работали в 2020 году.

Но в графе $G_{71}^{backbone}$ происходит разделение на два сообщества: Латвию и Литву («Балтия») и Беларусь, Казахстан, Россию, Украину и Эстонию в «Постсоветском пространстве». Возможное объяснение: Латвия и Литва имеют сильные торговые связи между собой, превосходящие связи с Европой и Россией, в то время как у Эстонии в 2020 году осталось больше «сильных» постсоветских связей благодаря транзиту и логистике.

ЗАКЛЮЧЕНИЕ

Ориентированные графы и сети, имеющие небольшой масштаб, высокую плотность и распределение весов дуг с тяжелым хвостом, являются интересными объектами исследований в силу указанных особенностей. В статье предлагается подход для исследования таких сетей, основанный на совмещении индекса дружбы, используемого для оценки локальной значимости участников сети, и функции диспаритета, применяемой для отсеивания слабых связей. Подход демонстрируется на примере Всемирной торговой сети экспортно-импортных потоков между странами в 2020 году, для которой построены несколько графов.

Начальный граф с 236 участниками экспортно-импортных операций построен по данным базы данных Comtrade Организации Объединенных Наций и содержит 22 868 дуг. Граф является связным, но не сильно связным, поскольку содержит 70 «висячих» вершин, соответствующих странам, которые только импортируют товары.

Удаление 70 «висячих» вершин порождает граф, у которого 166 вершин и 18 877 дуг, обладающий сильной связностью и высокой плотностью, равной 0,689.

Далее в базовом графе были оставлены только те вершины (и связывающие их дуги), у которых суммарный годовой торговый оборот был больше 30 млрд\$. В этом графе сохранился 71 вершина и 4917 дуг, он сильно связный с плотностью 0,989.

Применение функции диспаритета с целью «очистки» графа от «слабых» дуг при сохранении его сильной связности порождает граф-каркас, содержащий 71 вершину и 227 дуг (плотность 0,046).

Далее для трех графов (графов с 166 и 71 вершинами и графа-каркаса) были вычислены показатели значимости: PageRank, Closeness centrality и индексы дружбы по экспорту и по импорту.

Ожидаемый результат заключается в том, что для первых двух графов Китай, США и Германия являются глобальными лидерами (по PageRank, Closeness centrality), которые доминируют и локально (по индексам дружбы) во всех трех случаях. Менее ожидаемый результат – очень высокие места Гонконга, в некоторых случаях обходящего Германию, что находит содержательное объяснение. При этом Германия по индексам дружбы во втором графе поднимается на вторые позиции в мире. В графе-каркасе по индексам дружбы сохраняется лидерство Китая, США и Германии, а вот по показателю близости Германия и США значительно теряют позиции.

Таким образом, индексы дружбы представляют собой устойчивые показатели значимости стран-вершин, стабильные для лидеров во всех трех графах.

Кроме того, для графа с 71 вершиной и его графа-каркаса были построены разбиения на сообщества и проведен их сравнительный анализ с попытками содержательной интерпретации. В статье в качестве характерных примеров приведены три сообщества («Европейское торговое ядро с расширением», «Азиатско-Тихоокеанское ядро с расширением» и «Постсоветское пространство»). Один из значимых результатов заключается в том, что при уменьшении плотности графа с сохранением «сильных» дуг сообщества могут не только уменьшаться, но и увеличиваться. Интересны также варианты переходов стран из одних крупных сообществ в другие и сохранение экономических связей при разрушении политических.

Полученные результаты и их содержательная интерпретация позволяют глубже понять сложную структуру, скрытую за малой размерностью таких сетей.

ЛИТЕРАТУРА

1. Лапинова С. А., Аникина А. И., Ошарин А. М. Анализ структур экспорта и импорта с использованием сетевых методов (на примере рынка агропромышленных товаров) // Вестник Санкт-Петербургского университета. Экономика. 2020. Т. 36, вып. 3. С. 421–454. doi: 10.21638/spbu05.2020.304
2. Медведева Т. А. Алгоритмы теории графов в модели международной торговли // Advanced Engineering Research (Rostov-on-Don). 2010. Т. 10, № 6. С. 838–843.
3. Сушко В. А. Метод социометрии и анализ социальных сетей. М.: ИД КДУ, 2017. 310 с.
4. Anderson J. E., van Wincoop E. Gravity with Gravitas: a solution to the border puzzle // Am. Econ. Rev. 2003. Vol. 93, no. 1. P. 170–192. doi: 10.1257/00028280321455214
5. De Benedictis L., Tajoli L. The World Trade Network // The World Economy. 2011. Vol. 34, iss. 8. P. 1417–1454. doi: 10.1111/j.1467-9701.2011.01360.x
6. Evtushenko A., Kleinberg J. A comprehensive generalization of the Friendship Paradox to weights and attributes // Sci. Rep. 2024. Vol. 14. Art. 13730. doi: 10.1038/s41598-024-63167-9
7. Feld S. L. Why your friends have more friends than you do // Am. J. Sociol. 1991. Vol. 96, no. 6. P. 1464–1477.
8. Gönçer-Demiral D., İnce-Yenilmez M. Network analysis of international export pattern // Social Network Analysis and Mining. 2022. Vol. 12, iss. 1. Art. 156. doi: 10.21203/rs.3.rs-1538734/v1
9. Hazra R. S., Verbitskiy E. The generalized Friendship Paradox for spectral centralities. arXiv preprint. 2026. URL: <https://arXiv:2507.13059> (дата обращения: 02.02.2026).
10. Hong Kong SAR Ranks 10th Globally in Terms of Connectedness. URL: <https://www.dhl.com/hk-en/home/press/press-archive/2024/hong-kong-sar-ranks-10th-globally-in-terms-of-connectedness-reveals-dhl-global-connectedness-report-2024.html> (дата обращения: 02.02.2026).
11. Leicht E. A., Newman M. E. J. Community structure in directed networks // Phys. Rev. Lett. 2008. Vol. 100, no. 11. Art. 118703. doi: 10.1103/PhysRevLett.100.118703
12. Modularity. URL: <https://networkx.org/documentation/networkx-2.8.4/reference/algorithms/generated/networkx.algorithms.community.quality.modularity.html> (дата обращения: 16.02.2026).
13. Newman M. E. J. The structure and function of complex networks // SIAM Rev. 2003. Vol. 45, iss. 2. P. 167–256. doi: 10.1137/S003614450342480
14. Savić M., Ivanović M., Jain L. C. Complex networks in software, knowledge, and social systems. Cham, Switzerland: Springer, 2019. 331 p. doi: 10.1007/978-3-319-91196-0

15. Serrano M. Á., Boguñá M., Vespignani A. Extracting the multiscale backbone of complex weighted networks // *Proceedings of the National Academy of Sciences*. 2009. Vol. 106, no. 16. P. 6483–6488. doi: 10.1073/pnas.0808904106

16. Serrano M. Á., Boguñá M., Vespignani A. Patterns of dominant flows in the world trade web // *J. Econ. Interact. Coord.* 2007. Vol. 2. P. 111–124. doi: 10.1007/s11403-007-0026-y

17. Serrano M. Á., Boguñá M., Vespignani A. Supporting information. URL: http://www.pnas.org/cgi/data/0808904106/DCSupplemental/SI_Appendix (дата обращения: 06.02.2026).

18. Trade Data. URL: <https://comtradeplus.un.org/TradeFlow> (дата обращения: 02.02.2026)

19. World Trade Statistical Review 2021. URL: https://www.wto.org/english/res_e/statis_e/wts2021_e/wts2021_e.pdf (дата обращения: 02.02.2026).

REFERENCES

1. Lapinova S. A., Anikina A. I., Osharin A. M. Analysis of export and import structures using network methods (using the example of the agro-industrial goods market). *Vestnik Sankt-Petersburgskogo Universiteta. Ekonomika = St. Petersburg University Journal of Economic Studies*. 2020;36(3):421–454. doi: 10.21638/spbu05.2020.304 (In Russ.)

2. Medvedeva T. A. Graph theory algorithms in the International Trade model. *Advanced Engineering Research (Rostov-on-Don)*. 2010;10(6):838–843. (In Russ.)

3. Sushko V. A. Sociometry method and social network analysis. Moscow: KDU Publ. House; 2017. 310 p. (In Russ.)

4. Anderson J. E., van Wincoop E. Gravity with Gravitas: a solution to the border puzzle. *Am. Econ. Rev.* 2003;93(1):170–192. doi: 10.1257/000282803321455214

5. De Benedictis L., Tajoli L. The World Trade Network. *The World Economy*. 2011;34(8):1417–1454. doi: 10.1111/j.1467-9701.2011.01360.x

6. Evtushenko A., Kleinberg J. A comprehensive generalization of the Friendship Paradox to weights and attributes. *Sci. Rep.* 2024;14:13730. doi: 10.1038/s41598-024-63167-9

7. Feld S. L. Why your friends have more friends than you do. *Am. J. Sociol.* 1991;96(6):1464–1477.

8. Gönçer-Demiral D., İnce-Yenilmez M. Network analysis of international export pattern. *Social Network Analysis and Mining*. 2022;12(1):156. doi: 10.21203/rs.3.rs-1538734/v1

9. Hazra R. S., Verbitskiy E. The generalized Friendship Paradox for spectral centralities. arXiv preprint. 2026. URL: <https://arXiv:2507.13059> (accessed: 02.02.2026).

10. Hong Kong SAR Ranks 10th Globally in Terms of Connectedness. URL: <https://www.dhl.com/hk-en/home/press/press-archive/2024/hong-kong-sar-ranks-10th-globally-in-terms-of-connectedness-reveals-dhl-global-connectedness-report-2024.html> (accessed: 02.02.2026).

11. Leicht E. A., Newman M. E. J. Community structure in directed networks. *Phys. Rev. Lett.* 2008;100(11):118703. doi: 10.1103/PhysRevLett.100.118703

12. Modularity. URL: <https://networkx.org/documentation/networkx-2.8.4/reference/algorithms/generated/networkx.algorithms.community.quality.modularity.html> (accessed: 16.02.2026).

13. Newman M. E. J. The structure and function of complex networks. *SIAM Rev.* 2003;45(2):167–256. doi: 10.1137/S003614450342480

14. Savić M., Ivanović M., Jain L. C. Complex networks in software, knowledge, and social systems. Cham, Switzerland: Springer; 2019. 331 p. doi: 10.1007/978-3-319-91196-0

15. Serrano M. Á., Boguñá M., Vespignani A. Extracting the multiscale backbone of complex weighted networks. *Proceedings of the National Academy of Sciences*. 2009;106(16):6483–6488. doi: 10.1073/pnas.0808904106

16. Serrano M. Á., Boguñá M., Vespignani A. Patterns of dominant flows in the world trade web. *J. Econ. Interact. Coord.* 2007;2:111–124. doi: 10.1007/s11403-007-0026-y

17. Serrano M. Á., Boguñá M., Vespignani A. Supporting information. URL: http://www.pnas.org/cgi/data/0808904106/DCSupplemental/SI_Appendix (accessed: 06.02.2026).

18. TradeData. URL: <https://comtradeplus.un.org/TradeFlow> (accessed: 02.02.2026)

19. World Trade Statistical Review 2021. URL: https://www.wto.org/english/res_e/statis_e/wts2021_e/wts2021_e.pdf (accessed: 02.02.2026).

Поступила в редакцию / received: 18.02.2026; принята к публикации / accepted: 18.05.2026.

Автор заявляет об отсутствии конфликта интересов / The author declares no conflict of interest.

СВЕДЕНИЯ ОБ АВТОРЕ:

Печников Андрей Анатольевич

д-р техн. наук, ведущий научный сотрудник
e-mail: pechnikov@krc.karelia.ru

CONTRIBUTOR:

Pechnikov, Andrey

Dr. Sci. (Tech.), Leading Researcher

УДК 004.85

ДЕРЕВЬЯ РЕШЕНИЙ С ЛИНЕЙНОЙ КОМБИНАЦИЕЙ ПРИЗНАКОВ

А. А. Рогов*, К. А. Кулаков, Н. Д. Москин

*Институт математики и информационных технологий,
Петрозаводский государственный университет (пр. Ленина, 33,
Петрозаводск, Республика Карелия, Россия, 185910), *rogov@petrsu.ru*

Дерево решений — это один из алгоритмов машинного обучения, у которого есть ряд преимуществ (хорошая интерпретация, возможность работы с разными типами данных, автоматическое формирование правил и пр.). В некоторых задачах, где набор данных ограничен, а признаки связаны между собой (например, в задаче атрибуции текстов), можно при построении условий в узлах модели использовать не одиночные признаки, а их линейные комбинации. С учетом этого в статье изложено развитие метода «дерево решений», которое заключается в том, что к старым k признакам добавляются новые признаки как линейные комбинации двух исходных: $x_{ij} = \alpha x_i \pm (1 - \alpha)x_j$, где $i, j = 1, \dots, k$ и параметр $\alpha \in [0, 1]$. Для проведения численных экспериментов на примере задачи атрибуции текстов выполнена реализация построения линейных комбинаций признаков в информационной системе СМАЛТ («Статистические методы анализа литературного текста»). Результаты классификации дореволюционных текстов из журналов «Время», «Эпоха» и еженедельника «Гражданин» с использованием разных видов n -грамм показали, что данное улучшение повышает точность метода, при этом несущественно снижает интерпретацию полученного результата.

Ключевые слова: машинное обучение; дерево решений; линейная комбинация признаков; атрибуция текстов; n -граммы; информационная система СМАЛТ

Для цитирования: Рогов А. А., Кулаков К. А., Москин Н. Д. Деревья решений с линейной комбинацией признаков // Труды Карельского научного центра РАН. 2026. № 6. С. 100–105. doi: 10.17076/mat2314

A. A. Rogov*, K. A. Kulakov, N. D. Moskin. DECISION TREES WITH A LINEAR COMBINATION OF ATTRIBUTES

*Institute of Mathematics and Information Technologies, Petrozavodsk State University (33 Lenin St., 185910 Petrozavodsk, Karelia, Russia), *rogov@petrsu.ru*

A decision tree is a machine learning algorithm featuring a number of advantages (good interpretability, the ability to work with different types of data, automatic rule generation, etc.). In some tasks, where the data set is limited and attributes are interrelated (e.g., in the text attribution task), linear combinations of attributes can be used instead of single attributes when constructing conditions at the model's nodes. Taking this into account, the article presents an advancement of the decision

tree method, which consists in supplementing k old attributes with new attributes as linear combinations of two original ones: $x_{ij} = \alpha x_i \pm (1 - \alpha)x_j$, where $i, j = 1, \dots, k$ and the parameter $\alpha \in [0, 1]$. To conduct numerical experiments, the construction of linear combinations of attributes was implemented in the SMALT information system («Statistical Methods for the Analysis of Literary Text») for the case of the text attribution task. The results of classifying pre-1917 texts from the «Vremya» and «Epokha» magazines, and the weekly edition «Grazhdanin» using different types of n -grams showed that this improvement increases the accuracy of the method, but does not significantly reduce the interpretation of the outcome.

Key words: machine learning; decision tree; linear combination of attributes; text attribution; n -grams; SMALT information system

For citation: Rogov A. A., Kulakov K. A., Moskin N. D. Decision trees with a linear combination of attributes. *Trudy Karelskogo nauchnogo tsentra RAN = Transactions of the Karelian Research Centre RAS*. 2026. No. 6. P. 100–105. doi: 10.17076/mat2314

ВВЕДЕНИЕ

Дерево решений – это алгоритм машинного обучения, основанный на правиле «если <условие>, то <результат>». Представляет собой иерархическую структуру, состоящую из узлов и листьев, где каждый узел – это проверка условия, а листья – конечные результаты. Сама идея создания и дальнейшего развития моделей дерева решений появилась в середине XX века и была развита в работах К. Ховеленда «Компьютерное моделирование мышления» [10] и Е. Ханта «Эксперименты по индукции» [11]. Данным методом решаются следующие задачи: классификация данных, прогнозирование и выявление значимых факторов. Сегодня деревья решений используются в финансовой сфере, медицине, клиентском сервисе, промышленности и других областях [2, 5, 8]. Они помогают прогнозировать риски, диагностировать заболевания, оптимизировать процессы и принимать решения на основе больших объемов данных. Деревья решений можно применять в задачах атрибуции текстов (например, при решении вопросов их принадлежности тому или иному автору) [7]. Однако при этом не всегда удается добиться желаемого качества распознавания, так как объем текстов может быть весьма ограничен.

Этот метод обладает рядом преимуществ:

- простота интерпретации результатов (под интерпретацией результатов в машинном обучении понимают объяснение того, как модель пришла к своим прогнозам и какие факторы на них повлияли. Это нужно, чтобы сделать «черный ящик» модели понятным для людей: аналитиков, бизнеса, регуляторов);
- работа с разными типами данных;

- не требует нормализации данных;
- автоматическое формирование правил;
- быстрая скорость обучения (порядок $O(n \cdot \ln(n) \cdot d)$, где n – число объектов, d – число признаков).

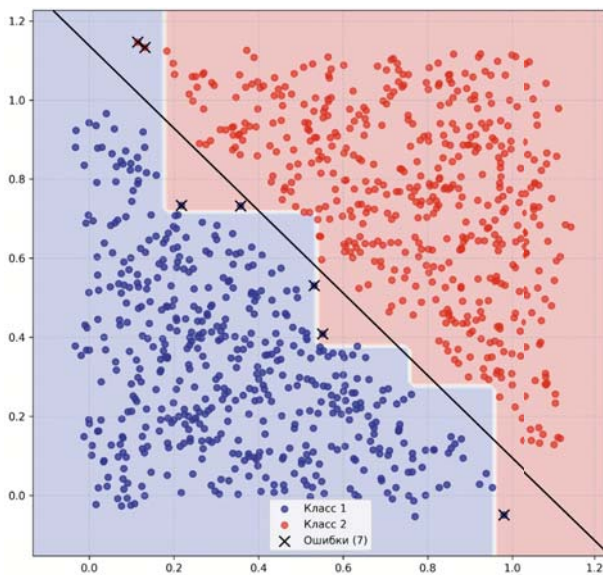
Как правило, при делении анализируемых данных в узле, содержащем <условие>, по двум веткам используется разделитель в виде простого правила, которое сравнивает некоторое значение признака с заданным порогом. Деревья решений с более сложными условиями рассмотрены в [1, 12]. Помимо одномерных (*univariate decision trees*) существует описание многомерных деревьев решений (*multivariate decision trees*). Во втором случае условие содержит более чем один признак, а именно лучшую линейную комбинацию признаков. Для ее поиска могут использоваться нейронная сеть и градиентный метод. В данной работе показано, что можно обойтись более простыми способами.

К основным существенным недостаткам указанного метода следует отнести проблему точности. Для борьбы с этой проблемой предлагают использовать набор деревьев (леса деревьев решений) [7]. В этом случае ухудшается интерпретация результатов или она становится вообще невозможной. В данной статье предлагается развитие метода «дерево решений», увеличивающее точность классификации за счет аугментации исходных данных.

МАТЕМАТИЧЕСКОЕ ОПИСАНИЕ ДЕРЕВА РЕШЕНИЙ

В качестве исходных данных для поставленной задачи используются размеченные пары $\{(z_l, y_l)\}_{l=1}^n$, где n – число объектов, z_l –

признаковое описание l -го объекта (точка в k -мерном пространстве $z_l = (x_1^l, x_2^l, \dots, x_k^l)$), а y_l – метка класса. На искусственно сгенерированном авторами демонстрационном примере (рис.) видно, что дерево решений в качестве классификатора строит прямоугольные границы решений (все расщепления перпендикулярны той или иной оси), при этом допущено 7 ошибок. На самом деле это простой линейно сепарабельный набор данных. То есть существует линейная разделяющая прямая, которая непараллельна осям координат (на рисунке пример такой прямой изображен сплошной линией). Если ее использовать в качестве классификатора, то ошибок не будет.



Графическое представление результата работы алгоритма «дерево решений»

Graphical representation of the result of the «decision tree» algorithm

Линейный классификатор даст всего одно условие (точка на плоскости выше или ниже этой линии), а классификатор, основанный на дереве решений, вынуждает проводить в худшем случае проверку 9 условий, что существенно удлиняет объяснение того, какая точка к какому классу относится, т. е. ухудшает интерпретацию.

Предлагаемый подход позволяет при той же точности уменьшить глубину построения решения классификатора, а значит, сохранить главное достоинство – простоту объяснения результатов разбиения (подробнее о проблемах, возникающих при построении моделей машинного обучения, см. в [3]). Для этого увеличим размерность обучающей выборки следующим образом. Добавим к старым k признакам новые признаки как линейные комби-

нации двух исходных: $x_{ij} = \alpha x_i \pm (1 - \alpha)x_j$, где $\alpha \in [0, 1]$. Этот подход был впервые применен в работе [9]. Если α равно 0 или 1, получаем старые признаки. При этом параметр α для каждой пары признаков будет свой.

В качестве значений α можно брать равномерные значения из интервала $(0, 1)$. Пусть их будет m . Тогда общее количество новых признаков будет равно $2mk(k - 1)$. Если добавить исходные признаки, то их будет не больше, чем $2mk(k - 1) + k$ с учетом различных значений α . Как отмечено выше, сложность данного алгоритма линейно зависит от числа признаков, значит, для данного метода увеличение размерности числа признаков не критично, так как существующие программные реализации построения деревьев решений способны обрабатывать массивы данных высокой размерности, а также выполнять параллельные вычисления.

ЧИСЛЕННЫЙ ЭКСПЕРИМЕНТ

Для оценки данного подхода рассмотрим задачу атрибуции текстов и фрагментов текстов из журналов «Время», «Эпоха» и еженедельника «Гражданин» под редакторством Ф. М. Достоевского [4]. Известно, что ряд анонимных публикаций принадлежат Ф. М. Достоевскому, а также он в ходе редактуры вносил правки в статьи других авторов, поэтому перед филологами стоит задача классификации текстов на две группы: «тексты, написанные Ф. М. Достоевским» и «тексты других авторов».

Определим анализируемый текст как последовательность из N фрагментов, каждый фрагмент состоит из M слов. Для унификации с каждым словом сопоставим признак части речи (атрибут), обозначим их число за k . В результате получаем N фрагментов, каждый из которых состоит из M значений атрибутов.

Рассмотрим следующие способы признакового описания фрагмента текста:

1. Униграммы частей речи. Построим вектор $z_l = (x_1^l, x_2^l, \dots, x_k^l)$, где x_i^l – частота встречаемости атрибута i в фрагменте l (признак i), $n = 1, \dots, N$. В результате получаем для текста набор из N векторов размерности k .

2. Линейные комбинации униграмм. Построим вектор $z_l^\alpha = (x_{11}^l, y_{11}^l, x_{12}^l, y_{12}^l, \dots, x_{kk}^l, y_{kk}^l)$, где $x_{ij}^l = \alpha x_i^l + (1 - \alpha)x_j^l$, $y_{ij}^l = \alpha x_i^l - (1 - \alpha)x_j^l$ – линейные комбинации частоты встречаемости униграмм i и j в фрагменте n . Таким образом получаем для текста набор из N векторов размерности $k \times (k - 1) \times 2$ для одного значения α .

В результате получаем набор сопоставимых векторов одинаковой размерности.

Для проведения численного эксперимента выполнена реализация построения линейных комбинаций признаков в информационной системе СМАЛТ («Статистические методы анализа литературного текста») [7]. Подсистема состоит из модуля управления деревьями решений (Tree app), интерфейса пользователя (Tree UI) и сервиса для их построения (Tree worker service).

Модуль управления реализован в виде Django-приложения и предоставляет функциональность по добавлению, изменению, удалению и просмотру деревьев решений и алгоритм проверки текста в дереве решений. Для построения дерева решений необходимо указать следующие параметры:

1. название исследования;
2. перечень текстов первого автора(ов);
3. перечень текстов второго автора(ов);
4. размер блока слов;
5. глубина дерева решений;
6. используемые признаки, описание которых изложено ранее.

Сервис построения деревьев решений реализован в виде самостоятельного Python-приложения, запускается в виде демона и реализует алгоритм подготовки данных и построения деревьев решений.

Алгоритм построения деревьев решений состоит из следующих шагов:

1. Определение перечня признаков. В системе реализована разметка текстов с использованием 23 атрибутов [6]. При построении дерева решений пользователь может исключить некоторые из них из рассмотрения.
2. Построение частотных матриц для перечней текстов. Последовательность текстов разбивается на блоки слов заданной длины. Для каждого блока вычисляется распределение частот встречаемости признаков в виде строки матрицы. Число столбцов определяется числом признаков.
3. Расчет частотных характеристик линейных комбинаций атрибутов.
4. Построение деревьев решений. Для построения используется модуль RandomForestClassifier пакета sklearn. Для получения наилучших решений в настройках метода указывается параметр числа попыток ($n_estimators$) не

менее 100 или числа строк матриц, если их больше; глубина дерева решений (max_depth) и отключение построения попыток на меньшем наборе данных ($bootstrap = False$). После построения выполняется оценка точности работы дерева решений на исходных данных (методы $predict()$ и $accuracy_score()$). Так как построение деревьев решений выполняется случайным образом, то выполняется 10 запусков с выбором наилучшего результата по точности.

5. Сохранение результата алгоритма. В базу данных информационной системы сохраняется в бинарном виде дерево решений и статистики.

Для классификации текста выполняется разбиение текста на блоки, расчет частотных характеристик каждого блока и для каждого полученного вектора выполняется оценка принадлежности блока к первой или второй группе текстов с помощью выбранного дерева решений.

Ключевой особенностью расчета частотных характеристик является построение блока текстов. На вход подаются тексты в виде потока слов. Слово пропускается, если его атрибут не входит в перечень требуемых. Блок считается завершенным, если число обработанных слов совпадает с размером блока.

Для проведения эксперимента использовались следующие параметры:

- размер блока – 200 слов;
- глубина дерева решений – 5 узлов;
- первая группа текстов – 50 текстов, подписанных Ф. М. Достоевским, из журналов «Время», «Эпоха» и еженедельника «Гражданин» [6];
- вторая группа текстов – 60 текстов, подписанных другими авторами (А. А. Григорьев, Я. П. Полонский, М. М. Достоевский, Н. Н. Прокопович, К. К. Сунгуров, Н. Н. Страхов, В. П. Мещерский, К. П. Победоносцев, В. Ф. Пуцкович), из журналов «Время», «Эпоха» и еженедельника «Гражданин» [6];
- число блоков – 873;
- используемые атрибуты – существительное, прилагательное, числительное, местоимение, глагол, причастие, деепричастие, наречие, категория состояния, частица, предлог, союз, модальное слово, междометие, звукоподражательное слово, иностранное слово, вводное слово.

В ходе эксперимента выполнялась оценка точности деревьев решений при заданных параметрах (подробнее о методике деления текстов на обучающую и тестовую выборку можно ознакомиться в [7]). Результаты экспериментов представлены в таблице. Показатели «точность» (*precision*) (или «точность положительных предсказаний», эта метрика отвечает на вопрос: «Из всех случаев, когда модель

сказала “это положительный класс”, сколько раз она была права?») и «F1-мера» (гармоническое среднее между точностью и полнотой). Полнота показывает, какую долю реальных положительных случаев модель сумела найти. Они широко используются в сфере искусственного интеллекта (подробнее о метриках качества моделей машинного обучения см. в [3]).

Результаты экспериментов (оценка точности деревьев решений при заданных параметрах)
Experimental results (evaluation of the accuracy of decision trees for given parameters)

Признаки Attributes	размер вектора vector size	точность (сред.) accuracy (avg.)	точность (макс.) accuracy (max.)	F1-мера F1-score
Униграммы частей речи Unigrams of parts of speech	17	0,816	0,821	0,824
Линейные комбинации униграмм с $\alpha = 0; 0,25; 0,5$ Linear combinations of unigrams with $\alpha = 0; 0,25; 0,5$	1632	0,853	0,856	0,863

Анализируя таблицу, можно сделать вывод, что использование классификатора на основе линейных комбинаций униграмм показывает более высокую точность по сравнению с классификатором на основе униграмм. Если построить доверительные интервалы для средних значений точности на уровне значимости 0,001, то они будут соответственно равны (0,815; 0,817) и (0,852; 0,854). Можно считать, что различия между показателями в таблице являются статистически значимыми.

Полученные деревья решений можно посмотреть по ссылке на информационную систему СМАЛТ [6].

ЗАКЛЮЧЕНИЕ

В работе предлагается развитие метода «дерево решений», которое можно использовать в том числе для задачи атрибуции текстов. При построении условий в узлах модели предлагается использовать не одиночные признаки, а линейные комбинации признаков. Данное усовершенствование при фиксированной глубине позволяет добиться большей точности, а при заданной точности уменьшает число узлов в классификационном дереве решений, а значит, улучшает интерпретацию полученного результата. Заметим, что предлагаемая аугментация признаков для построения дерева решений для примера, показанного на рисунке, приведет к линейной классификации.

ЛИТЕРАТУРА

1. Девяткин Д. А. Построение ансамблей деревьев решений с использованием линейных и нели-

нейных разделителей: Дис... канд. физ.-мат. наук. М.; 2022. 115 с.

2. Жаркова О. С., Шаропин К. А., Сеидова А. С., Берестнева Е. В., Осадчая И. А. Построение систем поддержки принятия решений в медицине на основе деревьев решений // Современные наукоемкие технологии. 2016. № 6. С. 33–37.

3. Жерон О. Прикладное машинное обучение с помощью Scikit-Learn, Keras и TensorFlow: концепции, инструменты и техники для создания интеллектуальных систем. СПб.: Диалектика, 2020. 1040 с.

4. Захаров В. Н. Задачи атрибуции анонимных статей в изданиях Достоевского // Известный Достоевский. 2022. № 9(3). С. 5–24. doi: 10.15393/j10.art.2022.6221

5. Иванов В. В., Акишина Е. П., Приказчикова А. С. Деревья и леса решений в задаче классификации кредитных организаций // Письма в журнал «Физика элементарных частиц и атомного ядра». 2025. № 22(1). С. 62–74.

6. Информационная система СМАЛТ («Статистические методы анализа литературного текста») [Электронный ресурс]. URL: http://smalt.karelia.ru/pysmalt/research/r_tree/list (дата обращения: 10.02.2026).

7. Rogov A. A., Abramov P. V., Burchneva D. D., Zakharova O. V., Kulakov K. A., Lebedev A. A., Moskin N. D., Otlivanichik A. V., Savinov E. D., Sidоров Ю. В. Проблема атрибуции в журналах «Время», «Эпоха» и еженедельнике «Гражданин». Петрозаводск: Острова, 2021. 391 с.

8. Франчук В. В., Михайличенко Б. В., Франчук М. В. Применение метода дерева решений в судебно-медицинской экспертной практи-

ке при анализе «врачебных дел» // Судебно-медицинская экспертиза. 2020. № 63(1). С. 9–14. doi: 10.17116/sudmed2020630119

9. Friedman J. H. A recursive partitioning decision rule for nonparametric classification // IEEE Transactions on Computers. 1977. No. 26(4). P. 404–408. doi: 10.1109/TC.1977.1674849

10. Hovland C. I. Computer simulation of thinking // Am. Psychol. 1960. No. 15(11). P. 687–693. doi: 10.1037/h0044165

11. Hunt E. B., Marin J., Stone P. J. Experiments in induction. New York: Academic Press, 1966. 247 p.

12. Kotsiantis S. B. Decision trees: a recent overview // Artif. Intell. Rev. 2013. No. 39(4). P. 261–283. doi: 10.1007/s10462-011-9272-4

REFERENCES

1. Devyatkin D. A. Constructing ensembles of decision trees using linear and nonlinear separators: PhD (Cand. of Phys.-Math.) thesis. Moscow; 2022. 115 p. (In Russ.)

2. Zharkova O. S., Sharopin K. A., Seidova K. S., Berestneva E. V., Osadchaya I. A. Construction of decision support systems in medicine based on decision tree. *Sovremennye naukoemkie tekhnologii = Modern Science-Intensive Technologies*. 2016;6:33–37. (In Russ.)

3. Geron A. Hands-on machine learning with Scikit-Learn, Keras and Tensor Flow: concepts, tools, and techniques to build intelligent systems. St. Petersburg: Dialektika; 2020. 1040 p. (In Russ.)

4. Zakharov V. N. Problems of attribution of anonymous articles in Dostoevsky's publications. *Neizvestnyi Dostoevskii = The Unknown Dostoevsky*. 2022; 9(3):5–24. (In Russ.). doi: 10.15393/j10.art.2022.6221

5. Ivanov V. V., Akishina E. P., Prikazchikova A. S. Decision trees and forests in the problem of classification of credit institutions. *Pis'ma v zhurnal «Fizika elementarnykh chastits i atomnogo yadra» = Letters to the journal 'Physics of Elementary Particles and Atomic Nuclei'*. 2025;22(1):62–74. (In Russ.)

6. SMALT information system («Statistical Methods for the Analysis of Literary Text») (In Russ.). URL: http://smalt.karelia.ru/pysmalt/research/r_tree/list (accessed: 10.02.2026).

7. Rogov A. A., Abramov R. V., Buchneva D. D., Zaharova O. V., Kulakov K. A., Lebedev A. A., Moskin N. D., Otlivanchik A. V., Savinov E. D., Sidorov Yu. V. The problem of attribution in the magazines 'Time', 'Epoch' and the weekly 'Citizen'. Petrozavodsk: Ostrova; 2021. 391 p. (In Russ.)

8. Franchuk V. V., Mikhaylichenko B. V., Franchuk M. V. Application of the decision tree method in forensic medical practice in the analysis of medical cases. *Sudebno-meditsinskaya ekspertiza = Forensic Medical Expertise*. 2020;63(1):9–14. (In Russ.). doi: 10.17116/sudmed2020630119

9. Friedman J. H. A recursive partitioning decision rule for nonparametric classification. *IEEE Transactions on Computers*. 1977;26(4):404–408. doi: 10.1109/TC.1977.1674849

10. Hovland C. I. Computer simulation of thinking. *Am. Psychol.* 1960;15(11):687–693. doi: 10.1037/h0044165

11. Hunt E. B., Marin J., Stone P. J. Experiments in induction. New York: Academic Press; 1966. 247 p.

12. Kotsiantis S. B. Decision trees: a recent overview. *Artif. Intell. Rev.* 2013;39(4):261–283. doi: 10.1007/s10462-011-9272-4

Поступила в редакцию / received: 23.02.2026; принята к публикации / accepted: 17.06.2026.

Авторы заявляют об отсутствии конфликта интересов / The authors declare no conflict of interest.

СВЕДЕНИЯ ОБ АВТОРАХ:

Рогов Александр Александрович
д-р техн. наук, заведующий кафедрой
e-mail: rogov@petrsu.ru

Кулаков Кирилл Александрович
канд. физ.-мат. наук, доцент
e-mail: kulakov@cs.petrsu.ru

Москин Николай Дмитриевич
д-р техн. наук, профессор
e-mail: moskin@petrsu.ru

CONTRIBUTORS:

Rogov, Aleksandr
Dr. Sci. (Tech.), Head of Department

Kulakov, Kirill
Cand. Sci. (Phys.-Math.), Associate Professor

Moskin, Nikolai
Dr. Sci. (Tech.), Professor

УДК 519.179.4

О ГЛОБАЛЬНОМ КЛАСТЕРНОМ КОЭФФИЦИЕНТЕ УСЛОВНОГО КОНФИГУРАЦИОННОГО ГРАФА

И. А. Чеплюкова

*Институт прикладных математических исследований КарНЦ РАН,
ФИЦ «Карельский научный центр РАН» (ул. Пушкинская, 11, Петрозаводск,
Республика Карелия, Россия, 185910)*

Рассматриваются конфигурационные графы с N вершинами. Степени вершин графа являются независимыми одинаково распределенными случайными величинами, распределение которых удовлетворяет условию: при $k \rightarrow \infty$

$$p_k = \mathbf{P}\{\xi = k\} \sim \frac{h(k)}{k^{\tau+1}}, \quad \tau > 0,$$

где случайная величина ξ равна степени любой вершины графа, а $h(x)$ обозначает медленно меняющуюся на бесконечности функцию. Исследуется глобальный кластерный коэффициент условного конфигурационного графа при различных значениях параметра $\tau > 0$. Получены предельные теоремы для этой характеристики при стремлении числа вершин к бесконечности.

Ключевые слова: конфигурационный граф; кластерный коэффициент; предельные теоремы

Для цитирования: Чеплюкова И. А. О глобальном кластерном коэффициенте условного конфигурационного графа // Труды Карельского научного центра РАН. 2026. № 6. С. 106–113. doi: 10.17076/mat2367

Финансирование. Финансовое обеспечение исследований осуществлялось из средств федерального бюджета на выполнение государственного задания КарНЦ РАН (Институт прикладных математических исследований КарНЦ РАН).

I. A. Cheplyukova. ON THE GLOBAL CLUSTERING COEFFICIENT OF A CONDITIONAL CONFIGURATION GRAPH

*Institute of Applied Mathematical Research, Karelian Research Centre, Russian
Academy of Sciences (11 Pushkinskaya St., 185910 Petrozavodsk, Karelia, Russia)*

We consider configuration graphs with N vertices. The vertex degrees are independent equally distributed random variables whose distribution satisfies the following condition: as $k \rightarrow \infty$

$$p_k = \mathbf{P}\{\xi = k\} \sim \frac{h(k)}{k^{\tau+1}},$$

where the random value η is equal to the degree of any vertex of the graph and

the function $h(x)$ is slowly varying (in the Karamata sense). The global clustering coefficient of a conditional configuration graph is investigated for different values of the parameter $\tau > 0$. Limit theorems are obtained for this characteristic as the number of vertices tends to infinity.

Keywords: configuration graph; clustering coefficient; limit theorems

For citation: Cheplyukova I. A. On the global clustering coefficient of a conditional configuration graph. *Trudy Karelskogo nauchnogo tsentra RAN = Transactions of the Karelian Research Centre RAS*. 2026. No. 6. P. 106–113. doi: 10.17076/mat2367

Funding. The studies were funded from the federal budget through state assignment to the Karelian Research Centre RAS (Institute of Applied Mathematical Research, Karelian Research Centre RAS).

ВВЕДЕНИЕ И ОСНОВНОЙ РЕЗУЛЬТАТ

В настоящее время большое внимание уделяется исследованию случайных графов, которые получили широкое применение при моделировании сложных сетей коммуникаций, таких как Интернет (например, [5, 6]). Одной из таких моделей является конфигурационный граф. Как показали многочисленные наблюдения за реальными сетями, степени узлов можно рассматривать как независимые одинаково распределенные случайные величины, при этом число вершин степени, не меньшей чем k , при достаточно больших k пропорционально $1/k^\alpha$, где α – некоторый параметр, $\alpha > 0$. Это значит, что распределение случайной величины η , равной степени любой вершины, при больших значениях k можно определить соотношением

$$\mathbf{P}\{\eta \geq k\} \sim \frac{h(k)}{k^\alpha}, \quad \alpha > 0,$$

где $h(x)$ – медленно меняющаяся на бесконечности функция (например, [4]).

Процесс построения конфигурационного графа можно описать следующим образом. Обозначим через N число вершин графа. Через $\xi_1, \xi_2, \dots, \xi_N$ обозначим случайные величины, равные степеням вершин графа с номерами $1, 2, \dots, N$ соответственно. Значения случайных величин $\xi_1, \xi_2, \dots, \xi_N$ равны числу полуребер, инцидентных конкретной вершине, т. е. ребер, для которых смежные вершины еще не определены [8]. Все полуребра считаются различными и, соединяясь друг с другом попарно и равновероятно, образуют ребра. Поскольку сумма степеней всех вершин графа должна быть четной, при необходимости вводится вспомогательная вершина единичной степени. В [8] замечено, что эта дополнительная вершина не влияет на асимптотические свойства графа при $N \rightarrow \infty$. Поэтому

в дальнейшем мы будем рассматривать степени только основных вершин. Нетрудно понять, что такая конструкция допускает появление и петель, и кратных ребер.

В современной литературе, посвященной случайным графам, большое внимание уделяется исследованию их структуры (см., например, [5]). Кластерный коэффициент является одной из важных характеристик в теории сетей, он количественно определяет степень, в которой узлы в сети имеют тенденцию группироваться вместе. Проще говоря, он измеряет, насколько связаны друг с другом соседи узла как локально для данного узла, так и для сети в целом. Обычно выделяют глобальный, локальный и средний кластерные коэффициенты (см., например, [7], [6]).

Далее мы будем рассматривать глобальный кластерный коэффициент случайного графа. Введем необходимые определения. Под «глобальным кластерным коэффициентом» будем понимать определение кластерного коэффициента для всей сети, предложенное в [5]. Сначала напомним определение глобального кластерного коэффициента графа, затем – определение для случайного графа. Следуя [5], глобальным кластерным коэффициентом графа называется отношение утроенного числа различных треугольников к числу пар смежных вершин. Пусть $G = G(V, E)$ – случайный граф, где V обозначает множество вершин графа, а E – множество ребер. Согласно [5] (соотношение (4.7.1)) глобальным кластерным коэффициентом C_G случайного графа будем называть

$$C_G = \frac{\mathbf{E}(\Delta_G)}{\mathbf{E}(W_G)},$$

где

$$\Delta_G = \sum_{i,j,k \in V} \mathbf{I}\{\text{вершины } i, j, k \text{ образуют треугольник}\},$$

$$W_G = \sum_{i,j,k \in V} \mathbf{I}\{\text{ребра } (i,j) \text{ и } (j,k) \text{ смежные}\},$$

$\mathbf{I}\{A\}$ обозначает индикатор события A . Заметим, что число треугольников между тремя вершинами может быть больше одного, если существуют кратные ребра. Учитывая особенности конфигурационных графов, в [7] была получена формула для нахождения глобального кластерного коэффициента конфигурационного графа:

$$C_G = \frac{\left(\sum_{k=1}^{\infty} k q_k\right)^2}{N \sum_{k=1}^{\infty} k \mathbf{P}\{\eta = k\}},$$

где

$$q_k = (k+1) \mathbf{P}\{\eta = k+1\} / \sum_{k=1}^{\infty} k \mathbf{P}\{\eta = k\}.$$

В данной работе исследуется предельное поведение глобального кластерного коэффициента конфигурационного графа с N вершинами, степени вершин которого являются независимыми одинаково распределенными случайными величинами с распределением, удовлетворяющим условию: при $k \rightarrow \infty$

$$p_k = \mathbf{P}\{\xi_i = k\} = \mathbf{P}\{\xi = k\} \sim \frac{h(k)}{k^{\tau+1}}, \quad (1)$$

где $\tau > 0$, $i = 1, \dots, N$, а $h(x)$ обозначает медленно меняющуюся на бесконечности функцию.

В [2] рассматривается асимптотика глобального кластерного коэффициента конфигурационного графа, степени вершин которого имеют распределение

$$\mathbf{P}\{\xi_i = k\} = \frac{h(k)}{k^{\tau+1}}, \quad k = 1, 2, \dots, \tau \in (2, 3).$$

Кроме того, в [2] было доказано следующее утверждение (лемма) о максимальной степени вершины конфигурационного графа. Пусть

$$\xi_{(N)} = \max_{1 \leq i \leq N} \xi_i.$$

Лемма. Пусть $N \rightarrow \infty$, z – фиксированное положительное число, а u – решение уравнения

$$u = h\left(\left(\frac{Nu}{\tau z}\right)^{1/\tau}\right). \quad (2)$$

Тогда

$$\mathbf{P}\left\{\xi_{(N)} < \left(\frac{Nu}{\tau z}\right)^{1/\tau}\right\} \rightarrow e^{-z}.$$

Используя свойства медленно меняющейся на бесконечности функции и (2), нетрудно видеть, что для любого β , $0 < \beta < 1$, справедливо

$$N^{-\beta/(1+\beta)} < u < N^{\beta/(1-\beta)}. \quad (3)$$

Учитывая это, можно рассматривать не все множество конфигурационных графов, а ограничиться его подмножеством, в котором степень любой вершины не превосходит $(Nu/(\tau z))^{1/\tau}$, где u определено в (2), z – некоторое фиксированное положительное число. В данной работе исследуется предельное поведение глобального кластерного коэффициента такого условного конфигурационного графа при всех значениях параметра распределения $\tau > 0$ и стремлении числа вершин графа к бесконечности.

Сформулируем основные результаты данной работы. Пусть

$$m_\tau = \mathbf{E}\xi.$$

Через $C_1, C_2, \dots$ будем обозначать некоторые положительные постоянные.

Теорема 1. Пусть $N \rightarrow \infty$ и выполнено одно из условий:

1. дисперсия случайной величины ξ конечна;
2. $\tau \searrow 2$, $u/(\tau - 2) \rightarrow \gamma$, где $0 \leq \gamma < \infty$;
3. $\tau = 2$, $u \ln(Nu) \rightarrow \gamma$, где $0 \leq \gamma < \infty$;
4. $\tau \nearrow 2$, $u^{2/\tau} N^{(2-\tau)/\tau} / (2 - \tau) \rightarrow \gamma$, где $0 \leq \gamma < \infty$,

тогда с вероятностью сколь угодно близкой к единице

$$C_G \asymp \frac{1}{N}.$$

Теорема 2. Пусть $N \rightarrow \infty$. Тогда с вероятностью сколь угодно близкой к единице справедливы следующие утверждения.

1. Если $\tau \searrow 2$ и $u/(\tau - 2) \rightarrow \infty$, то

$$C_G \asymp \frac{u^2}{N(\tau - 2)^2}.$$

2. Если $\tau = 2$ и $u \ln(Nu) \rightarrow \infty$, то

$$C_G \sim \frac{e^{-z}}{4Nm_2^3} u^2 \ln^2(uN).$$

3. Если $\tau \nearrow 2$ и $u^{2/\tau} N^{(2-\tau)/\tau} / (2 - \tau) \rightarrow \infty$, то

$$C_G \sim \frac{e^{-z} u^{4/\tau} N^{(4-3\tau)/\tau}}{m_2^3 (2 - \tau)^2}.$$

4. Если $1 < C_1 \leq \tau \leq C_2 < 2$, то

$$C_G \sim \frac{e^{-z}(z\tau)^{2(\tau-2)/\tau}}{m_\tau^3(2-\tau)^2} u^{4/\tau} N^{(4-3\tau)/\tau}.$$

5. Если $\tau \searrow 1$, $u/(1-\tau) \rightarrow \gamma$, где $0 \leq \gamma < \infty$, то

$$C_G \asymp N^{(4-3\tau)/\tau} u^{4/\tau}.$$

6. Если $\tau \searrow 1$, $u/(\tau-1) \rightarrow \infty$, то

$$C_G \asymp (\tau-1)^3 (Nu)^{(4-3\tau)/\tau}.$$

7. Если $\tau = 1$, $m_1 < \infty$, то

$$C_G \sim \frac{e^{-z}}{z^2 m_1^3} u^4 N.$$

8. Если $\tau = 1$, $u \ln(Nu) \rightarrow \gamma$, где $0 \leq \gamma < \infty$, то

$$C_G \asymp u^4 N.$$

9. Если $\tau = 1$, $u \ln(Nu) \rightarrow \infty$, то

$$C_G \sim \frac{e^{-z} Nu}{z^2 \ln^3(Nu)}.$$

10. Если $\tau \nearrow 1$, $u^{1/\tau} N^{(1-\tau)/\tau} / (1-\tau) \rightarrow \gamma$, где $0 \leq \gamma < \infty$, то

$$C_G \asymp u^{4/\tau} N^{(4-3\tau)/\tau}.$$

11. Если $\tau \nearrow 1$, $u^{1/\tau} N^{(1-\tau)/\tau} / (1-\tau) \rightarrow \infty$, то

$$C_G \sim \frac{e^{-z}(1-\tau)^3}{z^2} (Nu)^{1/\tau}.$$

12. Если $0 < C_3 \leq \tau \leq C_4 < 1$, то

$$C_G \sim \frac{e^{-z}(1-\tau)^3}{(2-\tau)^2(z\tau)^{(1+\tau)/\tau}} (Nu)^{1/\tau}.$$

Доказательство. Используя формулу (1), получим, что для глобального кластерного коэффициента C_G условного графа справедлива нижеуказанная формула:

$$C_G = \frac{\left(\sum_{k=1}^{\infty} k q'_k \right)^2}{N m'_\tau}, \quad (4)$$

где

$$p'_k = \mathbf{P} \left\{ \xi = k \mid \xi_i \leq \left(\frac{Nu}{\tau z} \right)^{1/\tau}, i = 1, \dots, N \right\},$$

$$q'_k = (k+1)p'_{k+1}/m'_\tau, \quad k = 0, 1, 2, \dots,$$

$$m'_\tau = \sum_{k=1}^{\infty} k p'_k.$$

Обозначим

$$X = X(z, \tau, N) = \left(\frac{Nu}{z\tau} \right)^{1/\tau}. \quad (5)$$

Из (1), (4) и (5) находим, что для достаточно большого A

$$C_G = \frac{(\mathbf{P} \{ \xi \leq X \})^N}{N} \frac{\left(\sum_{k=1}^X k(k+1)p_{k+1} \right)^2}{\left(\sum_{k=1}^X k p_k \right)^3} = \frac{(\mathbf{P} \{ \xi \leq X \})^N}{N} \quad (6)$$

$$\times \frac{\left(\sum_{k=1}^A k(k+1)p_{k+1} + \sum_{k=A+1}^X k(k+1) \frac{h(k+1)}{(k+1)^{\tau+1}} \right)^2}{\left(\sum_{k=1}^X k p_k \right)^3}.$$

Используя (1), (5), [3, гл. VIII, § 9, теорема 1, а] и свойства медленно меняющейся функции, получаем, что при $N \rightarrow \infty$

$$\mathbf{P} \{ \xi \leq X \} = 1 - \mathbf{P} \{ \xi > X \}$$

$$= 1 - (1 + o(1)) \int_X^\infty y^{-\tau-1} h(y) dy$$

$$= 1 - \frac{h(X)(1 + o(1))}{\tau X^\tau} = 1 + o(1).$$

Отсюда и из (2), (5), (6) следует, что

$$C_G = \frac{e^{-z}(1 + o(1))}{N} \quad (7)$$

$$\times \frac{\left(\sum_{k=1}^A k(k+1)p_{k+1} + \sum_{k=A+1}^X k(k+1) \frac{h(k+1)}{(k+1)^{\tau+1}} \right)^2}{\left(\sum_{k=1}^X k p_k \right)^3}.$$

Докажем теорему 1. Утверждение пункта 1 теоремы 1 следует из [1, теорема 1, п. 1].

Используя известные свойства медленно меняющихся функций, из (1) находим, что при достаточно больших l и любого $\varepsilon > 0$

$$\sum_{k>l} \frac{h(k)}{k^\tau} \leq \sum_{k>l} \frac{1}{k^{\tau-\varepsilon}}$$

$$\sim \int_l^\infty x^{-\tau+\varepsilon} dx \leq C_5 l^{-\tau+\varepsilon+1}. \quad (8)$$

Пусть выполнены условия пункта 2 теоремы 1. Из (1) и (8) получаем, что

$$\sum_{k=1}^X kp_k = \sum_{k=1}^\infty kp_k - \sum_{k>X} kp_k \sim m_2. \quad (9)$$

Используя определение медленно меняющейся функции, нетрудно показать, что при достаточно больших A

$$\begin{aligned} \sum_{A+1}^X \frac{h(k+1)}{(k+1)^{\tau-1}} &\sim \int_A^X \frac{h(x)}{x^{\tau-1}} dx \\ &= X^{2-\tau} \int_{A/X}^1 \frac{h(tX)}{t^{\tau-1}} dt \\ &\sim X^{2-\tau} h(X) \int_{A/X}^1 \frac{1}{t^{\tau-1}} dt \\ &= \frac{h(X)}{\tau-2} (-X^{2-\tau} + A^{2-\tau}). \end{aligned} \quad (10)$$

Учитывая, что

$$\sum_{A+1}^X \frac{h(k+1)}{(k+1)^{\tau-1}} > 0 \quad (11)$$

и условие $\tau \searrow 2$, нетрудно видеть, что дробь $h(X)/(\tau-2)$ положительна, тогда из (10) следует, что

$$\sum_{A+1}^X \frac{h(k+1)}{(k+1)^{\tau-1}} \asymp \left(\frac{h(X)}{\tau-2} \right). \quad (12)$$

Из (2), (5), (7), (9), (12) условий пункта 2 находим, что

$$C_G \asymp 1/N.$$

Пусть выполнены условия пункта 3 теоремы 1. Из (7) находим, что при достаточно больших A

$$\begin{aligned} C_G &\sim \frac{e^{-z}}{N} \left(\sum_{k=1}^A k(k+1)p_{k+1} \right. \\ &+ \left. \sum_{k=A+1}^X \frac{h(k+1)}{(k+1)} - \sum_{k=A+1}^X \frac{h(k+1)}{(k+1)^2} \right)^2 \\ &\quad \times \left(\sum_{k=1}^X kp_k \right)^{-3}. \end{aligned} \quad (13)$$

Используя хорошо известную формулу:

$$\sum_{k=1}^{[l]} \frac{1}{k} = \ln l + c + \varepsilon(l), \quad (14)$$

где c – постоянная Эйлера – Маскерони, $\varepsilon(l) \sim 1/(2l)$, можно показать, что при $N \rightarrow \infty$ и достаточно большом A

$$\begin{aligned} \sum_{k=A+1}^X \frac{h(k+1)}{(k+1)} &= \sum_{k=A+2}^{X+1} \frac{h(k)}{k} \\ &= \sum_{k=(A+2)/X}^1 \frac{h(kX)}{kX} \sim h(X) \sum_{k=A+2}^X \frac{1}{k} \\ &\sim h(X) \ln(X). \end{aligned} \quad (15)$$

Из (2), (5), (8), (9), (13), (15) находим, что при $N \rightarrow \infty$

$$C_G \asymp 1/N.$$

Пусть выполнены условия пункта 4 теоремы 1. Из (10), (11) и условия $\tau \nearrow 2$ получаем, что

$$\sum_{k=A+1}^X \frac{h(k+1)}{(k+1)^{\tau-1}} \asymp \frac{X^{2-\tau} h(X)}{2-\tau}. \quad (16)$$

Тогда отсюда и из (2), (5), (7), (9) следует, что при $N \rightarrow \infty$

$$C_G \asymp 1/N.$$

Значит, теорема 1 доказана.

Докажем теорему 2. Если выполнены условия пункта 1, то, учитывая, что в этом случае

$$\sum_{A+1}^X \frac{h(k+1)}{(k+1)^{\tau-1}} \rightarrow \infty,$$

из (2), (5), (7), (9), (12) получаем, что

$$C_G \asymp \left(\frac{u^2}{N(\tau-2)^2} \right).$$

Рассмотрим случай $\tau = 2$. Учитывая условия пункта 2 теоремы 2 и соотношение (15), при $N \rightarrow \infty$ выполняется соотношение

$$\sum_{A+1}^X \frac{h(k+1)}{(k+1)} \rightarrow \infty, \quad (17)$$

тогда из (2), (5), (8), (9), (13), (15) и (17) находим, что при $N \rightarrow \infty$

$$C_G \sim \frac{e^{-z}}{Nm_2^3} h^2(X) \ln^2(X) \sim \frac{e^{-z}}{4Nm_2^3} u^2 \ln^2(uN).$$

Если выполнены условия пункта 3, то из (2), (5), (7)–(9) и (16) следует, что при $N \rightarrow \infty$

$$C_G \sim \frac{e^{-z}}{m_2^3 N (2-\tau)^2} (X^{2-\tau} h(X))^2 \sim \frac{e^{-z} u^{4/\tau} N^{(4-3\tau)/\tau}}{4m_2^3 (2-\tau)^2 z^2}.$$

Рассмотрим случай, когда $1 < C_1 \leq \tau \leq C_2 < 2$. Из леммы 1.3.2 и утверждения 1.5.8 [4] получаем, что при $N \rightarrow \infty$

$$\sum_{k=A+1}^X \frac{h(k+1)}{(k+1)^{\tau-1}} = \int_A^X \frac{h(x)}{x^{\tau-1}} dx (1+o(1)) = X^{2-\tau} h(X) / (2-\tau)(1+o(1)). \quad (18)$$

Соотношение (9) выполняется и в этом случае. Из (2), (5), (7)–(9) и (18) следует, что при $N \rightarrow \infty$

$$C_G \sim \frac{e^{-z}}{m_\tau^3 N (2-\tau)^2} (X^{2-\tau} h(X))^2 = \frac{e^{-z}}{m_\tau^3 (2-\tau)^2} (z\tau)^{2(\tau-2)/\tau} u^{4/\tau} N^{(4-3\tau)/\tau}.$$

Пусть выполнены условия пункта 5. Тогда, используя лемму 1.3.2 и утверждение 1.5.8 [4], получаем, что при $N \rightarrow \infty$

$$\sum_{k=A+1}^X k p_k = \sum_{k=A+1}^X \frac{h(k)}{k^\tau} \asymp \frac{h(X)}{1-\tau}. \quad (19)$$

Отсюда, из (2), (5) и условия $u/(1-\tau) \rightarrow \gamma$ находим, что

$$\begin{aligned} \sum_{k=1}^X k p_k &= \sum_{k=1}^A k p_k + \sum_{k=A+1}^X k p_k \\ &= \sum_{k=1}^A k p_k + \frac{h(X)}{1-\tau} \neq 0. \end{aligned}$$

Тогда из (2), (5), (7) и (18) следует, что при $N \rightarrow \infty$

$$C_G \asymp N^{(4-3\tau)/\tau} u^{4/\tau}.$$

Пусть выполнены условия пункта 6. Тогда нетрудно видеть, что при достаточно больших A

$$\sum_{k=1}^X k p_k = \sum_{k=1}^A k p_k + \sum_{k=A+1}^X \frac{h(k)}{k^\tau}. \quad (20)$$

Из (2), (5), (19) следует, что

$$\sum_{k=A+1}^X \frac{h(k)}{k^\tau} \rightarrow \infty.$$

Тогда из (2), (5), (7), (18)–(20) находим, что при $N \rightarrow \infty$

$$\begin{aligned} C_G &\asymp \frac{(\tau-1)^3}{N h^3(X)} \left(\frac{X^{2-\tau} h(X)}{2-\tau} - \frac{h(X)}{\tau-1} \right)^2 \\ &= \frac{(\tau-1)^3}{N h(X)} \left(\frac{X^{2-\tau}}{2-\tau} - \frac{1}{\tau-1} \right)^2 \\ &\sim \frac{(\tau-1)^3}{N u} \left(\frac{(Nu)^{(2-\tau)/\tau}}{z} - \frac{1}{\tau-1} \right)^2. \quad (21) \end{aligned}$$

Согласно определению медленно меняющейся функции, несложно видеть, что при $N \rightarrow \infty$

$$\begin{aligned} \sum_{k=A}^X \frac{h(k)}{k^\tau} &= \sum_{A/X \leq t \leq 1} \frac{h(tX)}{(tX)^\tau} \\ &\sim h(X) \sum_{k=A}^X \frac{1}{k^\tau} < h(X) \sum_{k=A}^X \frac{1}{k}. \end{aligned}$$

Тогда из (14) получаем, что

$$\sum_{k=A}^X \frac{h(k)}{k^\tau} < C_6 h(X) \ln X.$$

Отсюда и из (2), (19) следует, что

$$\frac{1}{\tau-1} < C_7 \ln(uN).$$

Учитывая последнее соотношение и (21), получаем, что

$$C_G \asymp (\tau-1)^3 (Nu)^{(4-3\tau)/\tau}.$$

Пусть $\tau = 1$ и $m_1 < \infty$. Тогда из (2), (5), (7) находим, что при $N \rightarrow \infty$

$$C_G \sim \frac{e^{-z}}{N m_1^3} h^2(X) X^2 = \frac{e^{-z}}{z^2 m_1^3} u^4 N.$$

Пусть выполнены условия пункта 8. Тогда из (1) и (15) получаем, что при достаточно большом A

$$\begin{aligned} \sum_1^X k p_k &\sim \sum_1^A k p_k + \sum_{A+1}^X \frac{h(k)}{k} \\ &\sim \sum_1^A k p_k + h(X) \ln X. \quad (22) \end{aligned}$$

Отсюда, из (2), (5), (7) и условия пункта 8 находим, что при $N \rightarrow \infty$

$$C_G \asymp \frac{1}{N} h^2(X) X^2 \asymp u^4 N.$$

Если выполнены условия пункта 9, то из (1), (2), (5), (7), (14) и (22) нетрудно получить, что при достаточно большом A и $N \rightarrow \infty$

$$\begin{aligned} C_G &\sim \frac{e^{-z}}{N} h^{-3}(X) \ln^{-3}(X) (h(X)X \\ &\quad - h(X) \ln(X))^2 \\ &= \frac{e^{-z} X^2}{N h(X) \ln^3(X)} = \frac{e^{-z} Nu}{z^2 \ln^3(Nu)}. \end{aligned}$$

Пусть $0 < C_3 \leq \tau < C_4 < 1$. Согласно утверждению 1.5.8 и лемме 1.3.2 [4] для достаточно больших A и $N \rightarrow \infty$ справедливы соотношения (18) и

$$\sum_{A+1}^X \frac{h(k)}{k^\tau} = h(X) \frac{X^{1-\tau}}{1-\tau} (1 + o(1)). \quad (23)$$

Отсюда и из (2), (5), (7), (18) находим, что при $N \rightarrow \infty$

$$\begin{aligned} C_G &\sim \frac{e^{-z}(1-\tau)^3}{N h^3(X) X^{3(1-\tau)}} \left(\frac{h(X) X^{2-\tau}}{2-\tau} \right. \\ &\quad \left. - \frac{h(X) X^{1-\tau}}{1-\tau} \right)^2 \sim \frac{e^{-z}(1-\tau)^3}{N} \frac{X^{1+\tau}}{(2-\tau)^2 h(X)} \\ &= \frac{e^{-z}(1-\tau)^3}{(2-\tau)^2 (z\tau)^{(1+\tau)/\tau}} (Nu)^{1/\tau}. \end{aligned}$$

Если выполнены условия пункта 10, то соотношение (23) справедливо. Тогда из (2), (5), (7), (18), (23) и условий пункта 10 находим, что при $N \rightarrow \infty$

$$C_G \asymp \frac{1}{N} \left(\frac{h(X) X^{2-\tau}}{2-\tau} \right)^2 \sim u^{4/\tau} N^{(4-3\tau)/\tau}.$$

Пусть выполнены условия пункта 11. Равенство (23) выполняется и в этом случае. Кроме того, используя свойства медленно меняющихся функций и переходя от суммирования к интегрированию, несложно видеть, что для $\varepsilon > 0$

$$\begin{aligned} \sum_{A+1}^X \frac{h(k)}{k^\tau} &< \sum_{A+1}^X \frac{1}{k^{\tau-\varepsilon}} = \int_A^X \frac{1}{x^{\tau-\varepsilon}} dx (1 + o(1)) \\ &= \frac{X^{1-\tau+\varepsilon}}{1-\tau+\varepsilon} (1 + o(1)). \end{aligned} \quad (24)$$

Кроме того, можно заметить, что при выполнении условий пункта 11 из (23) и (24) следует, что для любого $\varepsilon > 0$

$$1 - \tau > h(X) X^{-\varepsilon}.$$

Из (24) получаем, что

$$\begin{aligned} &\frac{h(X) X^{2-\tau}}{(2-\tau) \sum_A^X h(k)/k^\tau} \\ &> C_9 \frac{h(X) X^{2-\tau}}{(2-\tau) X^{-\tau+1+\varepsilon}} = C_9 \frac{h(X) X^{1-\varepsilon}}{2-\tau} \rightarrow \infty. \end{aligned}$$

Отсюда и из (2), (5), (7), (18), (23) следует, что при $N \rightarrow \infty$

$$\begin{aligned} C_G &\sim \frac{e^{-z}}{N} \frac{X^{4-2\tau} h^2(X)}{(2-\tau)^2} \left(1 - \frac{(2-\tau) \sum_A^X h(k)/k^\tau}{h(X) X^{2-\tau}} \right)^2 \\ &\quad \times \left(\frac{X^{1-\tau} h(X)}{1-\tau} \right)^3 \\ &\sim \frac{e^{-z}(1-\tau)^3 X^{1+\tau}}{N h(X)} \sim \frac{e^{-z}(1-\tau)^3 (Nu)^{1/\tau}}{z^2}. \end{aligned}$$

Теперь теорема 2 доказана полностью.

Несложно заметить, что из утверждений теорем 1 и 2, их доказательств и (3) вытекает следствие.

Следствие. *Справедливы следующие утверждения.*

1. Если выполнено одно из условий:

- $\tau \geq 2$;
- $\tau \rightarrow 2$;
- $1 < C_1 \leq \tau \leq C_2 < 2$,
 $u^{4/\tau} N^{(4-3\tau)/\tau} \rightarrow 0$,

то $C_G \rightarrow 0$;

2. Если выполнено одно из условий:

- $0 < \tau \leq 1$;
- $\tau \rightarrow 1$;
- $1 < C_1 \leq \tau \leq C_2 < 2$,
 $u^{4/\tau} N^{(4-3\tau)/\tau} \rightarrow \infty$,

то $C_G \rightarrow \infty$

3. Если $1 < C_1 \leq \tau \leq C_2 < 2$, $u^{4/\tau} N^{(4-3\tau)/\tau} \rightarrow \text{const} > 0$, то C_G стремится к некоторой положительной константе.

Такое поведение глобального кластерного коэффициента можно объяснить тем, что при малых значениях параметра τ в графе образуется много ребер, а значит, много треугольников. Поскольку u зависит от функции $h(x)$, N , τ , поведение кластерного коэффициента C_G также зависит от этих параметров. Такая особенность кластерного коэффициента конфигурационного графа отмечалась еще в работе [7], где в качестве медленно меняющейся функции $h(x)$ рассматривался нормирующий множитель $1/\zeta(\tau)$.

ЛИТЕРАТУРА

1. Павлов Ю. Л. Об асимптотике кластерного коэффициента конфигурационного графа с неизвестным распределением степеней вершин // Информатика и ее применения. 2019. Т. 13, № 3. С. 3–19. doi: 10.14357/19922264190302
2. Павлов Ю. Л. О глобальном кластерном коэффициенте Интернет-графа // Труды Карельского научного центра РАН. 2023. № 4. С. 50–53. doi: 10.17076/mat1765
3. Феллер В. Введение в теорию вероятностей и ее применения. Т. 2. М.: Мир, 1984. 738 с.
4. Bibgman N. H., Goldie C. M., Teugels J. L. R. Regular variation. Cambridge: Cambridge Univ. Press, 1987. 513 p.
5. Hofstad R. Random graphs and complex networks. Vol. 1. Cambridge: Cambridge Univ. Press, 2017. 337 p. doi: 10.1017/9781316779422
6. Newman M. E. J. Networks. An introduction. Oxford: Oxford Univ. Press, 2010. 772 p.
7. Newman M., E., J. The structure and function of complex networks // SIAM Rev. 2003. Vol. 45, no. 2. P. 167–256. doi: 10.1137/S003614450342480

8. Reittu Y., Norros I. On the power-law random graph model of massive data networks // Perform. Eval. 2004. Vol. 55, no. 1–2. P. 3–23. doi: 10.1016/S0166-5316(03)00097-X

REFERENCES

1. Pavlov Yu. L. On the asymptotics of clustering coefficient in a configuration graph with unknown distribution of vertex degrees. *Informatika i ee primeneniya = Computer Science and its Applications*. 2019;13(3):3–19. (In Russ.). doi: 10.14357/19922264190302
2. Pavlov Yu. L. On the Internet-graph clustering coefficient. *Trudy Karelskogo nauchnogo tsentra RAN = Transactions of the Karelian Research Centre RAS*. 2023;4:50–53. (In Russ.). doi: 10.17076/mat1765
3. Feller W. An introduction to probability theory and its applications. Vol. 2. New York: John Wiley; 1991. 704 p.
4. Bibgman N. H., Goldie C. M., Teugels J. L. R. Regular variation. Cambridge: Cambridge Univ. Press; 1987. 513 p.
5. Hofstad R. Random graphs and complex networks. Vol. 1. Cambridge: Cambridge Univ. Press; 2017. 337 p. doi: 10.1017/9781316779422
6. Newman M. E. J. Networks. An introduction. Oxford: Oxford Univ. Press; 2010. 772 p.
7. Newman M. E. J. The structure and function of complex networks. *SIAM Rev.* 2003;45(2):167–256. doi: 10.1137/S003614450342480
8. Reittu Y., Norros I. On the power-law random graph model of massive data networks. *Perform. Eval.* 2004;55(1–2):3–23. doi: 10.1016/S0166-5316(03)00097-X

Поступила в редакцию / received: 09.04.2026; принята к публикации / accepted: 08.06.2026.
Автор заявляет об отсутствии конфликта интересов / The author declares no conflict of interest.

СВЕДЕНИЯ ОБ АВТОРЕ:

Чеплюкова Ирина Александровна

канд. физ.-мат. наук, старший научный сотрудник

e-mail: chia@krc.karelia.ru

CONTRIBUTOR:

Cheplyukova, Irina

Cand. Sci. (Phys.-Math.), Senior Researcher

УДК 519.179.4

ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ ДЛЯ ЧИСЛА ВЕРШИН ЗАДАННОЙ СТЕПЕНИ В УСЛОВНОМ КОНФИГУРАЦИОННОМ ГРАФЕ

И. А. Чеплюкова

*Институт прикладных математических исследований КарНЦ РАН,
ФИЦ «Карельский научный центр РАН» (ул. Пушкинская, 11, Петрозаводск,
Республика Карелия, Россия, 185910)*

Рассматриваются конфигурационные графы с N вершинами. Степени вершин являются независимыми одинаково распределенными случайными величинами, распределение которых удовлетворяет следующему условию: при $k \rightarrow \infty$

$$p_k = \mathbf{P}\{\eta = k\} \sim \frac{h(k)}{k^g}, \quad 2 < g < 3,$$

где $h(k)$ – медленно меняющаяся на бесконечности функция. Изучаются случайные графы при условии, что сумма степеней всех вершин равна n . Найдены предельные распределения числа вершин заданной степени в таком условном графе при $N, n \rightarrow \infty$ так, что $h(N)n^2N^{(4-3g)/(g-1)} \geq C > 0$.

Ключевые слова: конфигурационный граф; степень вершины; предельное распределение

Для цитирования: Чеплюкова И. А. Предельные теоремы для числа вершин заданной степени в условном конфигурационном графе // Труды Карельского научного центра РАН. 2026. № 6. С. 114–126. doi: 10.17076/mat2327

Финансирование. Финансовое обеспечение исследований осуществлялось из средств федерального бюджета на выполнение государственного задания КарНЦ РАН (Институт прикладных математических исследований КарНЦ РАН).

I. A. Cheplyukova. LIMIT THEOREMS FOR THE NUMBER OF VERTICES WITH A GIVEN DEGREE IN A CONDITIONAL CONFIGURATION GRAPH

Institute of Applied Mathematical Research, Karelian Research Centre, Russian Academy of Sciences (11 Pushkinskaya St., 185910 Petrozavodsk, Karelia, Russia)

We consider configuration graphs with N vertices. The vertex degrees are independent equally distributed random variables whose distribution satisfies the following condition: as $k \rightarrow \infty$

$$p_k = \mathbf{P}\{\eta = k\} \sim \frac{h(k)}{k^g}, \quad 2 < g < 3,$$

where $h(k)$ is a slowly varying function. We study the random graphs under the

conditions that the sum of vertex degrees is n . The limit distributions of the number of vertices with a given degree in the conditional graph are obtained as $N, n \rightarrow \infty$ and $h(N)n^2N^{(4-3g)/(g-1)} \geq C > 0$ are obtained.

Key words: configuration graph; vertex degree; limit distribution

For citation: Cheplyukova I. A. Limit theorems for the number of vertices with a given degree in a conditional configuration graph. *Trudy Karelskogo nauchnogo tsentra RAN = Transactions of the Karelian Research Centre RAS*. 2026. No. 6. P. 114–126. doi: 10.17076/mat2327

Funding. The studies were funded from the federal budget through state assignment to the Karelian Research Centre RAS (Institute of Applied Mathematical Research, Karelian Research Centre RAS).

ВВЕДЕНИЕ И ОСНОВНЫЕ РЕЗУЛЬТАТЫ

Объектом исследования в данной работе являются конфигурационные графы. Случайные графы широко используются при моделировании различных сложных сетей коммуникаций, таких как социальные, биологические, транспортные сети и Интернет [13, 15, 16]. Наблюдения за реальными сетями [11, 14] показали, что все они обладают некоторыми общими свойствами, и эти свойства должны быть отражены в их моделях. Например, степени вершин можно рассматривать как независимые одинаково распределенные случайные величины, при этом число вершин степени, не меньшей чем k , при достаточно больших k пропорционально $1/k^{\tau-1}$, где $\tau > 1$. Это значит, что распределение случайной величины η , равной степени любой вершины, при больших значениях k можно определить равенством

$$P\{\eta \geq k\} = \frac{h(k)}{k^{\tau-1}}, \quad \tau > 1,$$

где $h(k)$ – медленно меняющаяся на бесконечности функция.

Впервые конфигурационные графы были введены в работе [12]. Опишем процесс построения этих графов. Обозначим через N число вершин конфигурационного графа, степени вершин которого являются независимыми случайными величинами. Значения этих случайных величин равны числу полуребер, инцидентных данной вершине, т. е. ребер, для которых смежные вершины еще не определены [17]. Все полуребра считаются различными и, соединяясь друг с другом попарно и равновероятно, образуют ребра. Поскольку сумма степеней вершин графа должна быть четной, при необходимости вводится вспомогательная вершина единичной степени. В [17] замечено, что эта дополнительная вершина не влияет на асимптотические свойства графа при $N \rightarrow \infty$. Поэтому в дальнейшем мы будем рассматривать степени только основных вершин.

Нетрудно понять, что такая конструкция допускает появление и петель, и кратных ребер.

Исследованию структуры конфигурационных графов посвящено большое число работ (например, [13, 15, 16]). Есть ряд статей, в которых изучается предельное поведение числовых характеристик конфигурационных графов при условии ограничения на число ребер [2, 4–7, 9]. Например, в статье [7] впервые рассматривались такие конфигурационные графы при условии, что сумма степеней всех вершин известна и равна n . В [4] и [6] изучались конфигурационные графы при более естественном предположении ограничения числа ребер сверху. Исследование таких условий графов представляет определенный интерес, поскольку на практике иногда удается оценить число связей в реальных сетях.

В [2–7, 9] главное внимание уделяется предельному поведению таких важных числовых характеристик степенной структуры условных графов, как максимальная степень и число вершин заданной степени. В частности, в статье [9] изучались конфигурационные графы, распределение степеней вершин которых имеют вид:

$$p_k = \frac{h(k)}{k^g \Sigma(1, g)}, \quad k = 1, 2, \dots, \quad g > 1,$$

где

$$\Sigma(x, y) = \sum_{k=1}^{\infty} x^k \frac{h(k)}{k^y},$$

при условии, что число ребер известно и равно n и найдены предельные распределения указанных характеристик в случаях, когда $n, N \rightarrow \infty$, $1 < C_1 \leq n/N \leq C_2 < \mathbf{E}\eta$. (Здесь и далее $C, C_1, C_2, \dots$ обозначают некоторые положительные постоянные.) А в [3] решена аналогичная задача для конфигурационных графов при достаточно слабых ограничениях на хвост распределения степени вершины:

$$p_k \sim \frac{d}{k^g (\ln k)^h}, \quad d > 0, \quad g > 1, \quad h \geq 0, \quad k \rightarrow \infty.$$

В [2] изучены конфигурационные графы, распределение степеней вершин которых имеет вид:

$$p_k = \frac{h(k)}{k^g}, \quad k = 1, 2, \dots, \quad g > 3,$$

при условии, что число ребер ограничено сверху и найдены предельные распределения для максимальной степени вершины и числа вершин заданной степени. В основе доказательств этих результатов лежит обобщенная схема размещения частиц по ячейкам, введенная и изученная В. Ф. Колчиным [1], или ее аналог [10].

Пусть далее η обозначает целочисленную положительную случайную величину, распределение которой удовлетворяет условию: при $k \rightarrow \infty$

$$p_k = \mathbf{P}\{\eta = k\} \sim \frac{h(k)}{k^g}, \quad 2 < g < 3. \quad (1)$$

Рассмотрим конфигурационный граф с N вершинами, степени вершин которого являются независимыми одинаково распределенными случайными величинами с таким же распределением, как у случайной величины η . В данном исследовании значение параметра g принадлежит интервалу $(2, 3)$, выбор таких значений параметра обусловлен тем, что, как показывают многочисленные наблюдения, такие значения параметра g являются наиболее подходящими для большого числа реальных сетей. Из (1) следует, что при $2 < g < 3$ распределение степеней вершин имеет конечное математическое ожидание и бесконечную дисперсию. В настоящей работе исследуется указанный случайный граф при условии, что сумма степеней всех вершин известна и равна n , для таких конфигурационных графов мы получили предельные теоремы для числа вершин заданной степени при $n, N \rightarrow \infty$ так, что $h(N)n^2N^{(4-3g)/(g-1)} \geq C > 0$. В статье [8] уже рассматривались такие условные конфигурационные графы и была доказана предельная теорема для максимальной степени вершины в вышеописанной области изменения параметров n, N, g . Сформулируем основные результаты настоящей работы. Для этого введем необходимые обозначения. Пусть $F(x)$ обозначает функцию распределения целочисленной случайной величины η , значения которой положительны и удовлетворяют условию (1). Положим для $x > 0$

$$\kappa(x) = 1 - F(x) + F(-x)$$

и определим нормирующие постоянные B_N как нижнюю границу тех x , для которых

$$\kappa(x-0) \geq (N)^{-1} \geq \kappa(x+0).$$

Через μ_r обозначим случайную величину, равную числу вершин графа степени r . Тогда справедливы следующие утверждения.

Теорема 1. Пусть $n, N \rightarrow \infty$, $h(N)n^2N^{(4-3g)/(g-1)} \geq C > 0$, $Nh(r)r^{-g} \rightarrow \infty$. Тогда равномерно относительно целых неотрицательных k таких, что $u_r = \frac{k-Np_r}{\sqrt{Np_r(1-p_r)}}$ лежит в любом фиксированном конечном интервале

$$\mathbf{P}\{\mu_r = k\} = \frac{1 + o(1)}{\sqrt{2\pi Np_r(1-p_r)}} e^{-u_r^2/2}.$$

Теорема 2. Пусть $n, N, r \rightarrow \infty$, $h(N)n^2N^{(4-3g)/(g-1)} \geq C > 0$ и выполнено одно из условий:

1. $r\sqrt{N}/B_N \rightarrow 0$;
2. $r\sqrt{N}/B_N \geq C_1$, $Nh(r)/r^g \rightarrow 0$.

Тогда равномерно относительно целых неотрицательных k таких, что $(k - Np_r)/\sqrt{Np_r}$ лежит в любом фиксированном конечном интервале

$$\mathbf{P}\{\mu_r = k\} = \frac{(Np_r)^k}{k!} e^{-Np_r} (1 + o(1)).$$

Замечание. В замечании 1 к теореме 1 статьи [8] показано, что при $N \rightarrow \infty$ верно

$$\frac{Nh(B_N)}{(g-1)B_N^{g-1}} \rightarrow 1 \quad (2)$$

и для достаточно большого N и любого сколь угодно малого $\varepsilon > 0$ выполнено

$$C_2 N^{1/(g-1+\varepsilon)} < B_N < C_3 N^{1/(g-1-\varepsilon)}. \quad (3)$$

ДОКАЗАТЕЛЬСТВА ВСПОМОГАТЕЛЬНЫХ УТВЕРЖДЕНИЙ И ТЕОРЕМ

Обозначим через $\xi_1, \xi_2, \dots, \xi_N$ случайные величины, равные степеням вершин графа с номерами $1, 2, \dots, N$ соответственно. Пусть $\eta_1, \dots, \eta_N$ – независимые одинаково распределенные случайные величины, которые распределены так же, как и случайная величина η . Тогда, используя (1), для целых $k_1, \dots, k_N$ таких, что $k_1 + \dots + k_N = n$, выполнено равенство

$$\begin{aligned} & \mathbf{P}\{\xi_1 = k_1, \dots, \xi_N = k_N\} \\ &= \mathbf{P}\{\eta_1 = k_1, \dots, \eta_N = k_N | \eta_1 + \dots + \eta_N = n\}. \end{aligned} \quad (4)$$

Следовательно, для двух наборов случайных величин $(\xi_1, \dots, \xi_N)$ и $(\eta_1, \dots, \eta_N)$ выполнены условия обобщенной схемы размещения частиц по ячейкам [1].

Введем вспомогательные независимые одинаково распределенные случайные величины $\eta_1^{(r)}, \dots, \eta_N^{(r)}$, распределение которых имеет вид:

$$\mathbf{P}\{\eta_i^{(r)} = k\} = \mathbf{P}\{\eta_i = k | \eta_i \neq r\}, \quad (5)$$

где $k = 1, 2, \dots$, $i = 1, 2, \dots, N$. Пусть

$$\zeta_N = \eta_1 + \dots + \eta_N, \quad \zeta_N^{(r)} = \eta_1^{(r)} + \dots + \eta_N^{(r)}.$$

В [1] доказано, что для случайной величины μ_r следствием из равенства (4) является следующее утверждение.

Лемма 1. *Справедливо равенство*

$$\begin{aligned} & \mathbf{P}\{\mu_r = k\} \\ &= \binom{N}{k} p_r^k (1 - p_r)^{N-k} \frac{\mathbf{P}\{\zeta_{N-k}^{(r)} = n - kr\}}{\mathbf{P}\{\zeta_N = n\}}. \end{aligned}$$

Согласно утверждению леммы 1, для получения предельных результатов для μ_r нам необходимо знать предельные поведения сумм независимых случайных величин $\zeta_{N-k}^{(r)}$, ζ_N и асимптотику биномиальной вероятности. Для суммы ζ_N можно использовать лемму 2 [8]. Сформулируем ее ниже в виде леммы 2.

Лемма 2. *Пусть $N, n \rightarrow \infty$, $2 < g < 3$, $(n - MN)/B_N \rightarrow \infty$. Тогда*

$$\mathbf{P}\{\zeta_N = n\} \sim \frac{Nh(n - MN)}{(n - MN)^g},$$

где $M = \mathbf{E}\eta$.

Докажем, что для суммы $\zeta_N^{(r)}$ справедливо следующее утверждение.

Лемма 3. *Пусть $N, n \rightarrow \infty$, $h(N)n^2N^{(4-3g)/(g-1)} \geq C > 0$ и выполнено одно из условий:*

1. $r\sqrt{N}/B_N \rightarrow 0$;
2. $r\sqrt{N}/B_N \geq C_1$, $Nh(r)/r^g \rightarrow 0$.

Тогда при $S = N(1 - p_r)(1 + o(1))$ равномерно относительно целых неотрицательных k , для которых $(k - Np_r)/\sqrt{Np_r(1 - p_r)}$ лежит в любом фиксированном конечном интервале

$$\mathbf{P}\{\zeta_S^{(r)} = n - kr\} = \frac{h(n)N(1 + o(1))}{n^g}.$$

Доказательство. Величина r может принимать и фиксированные значения, и стремиться к бесконечности. Будем одновременно рассматривать эти два случая, учитывая, что при $r \rightarrow \infty$ случайные величины $\eta_1^{(r)}, \dots, \eta_N^{(r)}$ образуют схему серий. Из (5) нетрудно получить, что

$$m_r = \mathbf{E}\eta_1^{(r)} = \frac{M - rp_r}{1 - p_r}. \quad (6)$$

Введем независимые одинаково распределенные случайные величины

$$\xi_j^{(r)} = \eta_j^{(r)} - m_r, \quad j = 1, \dots, N, \quad (7)$$

и их сумму $\tilde{\zeta}_S^{(r)} = \xi_1^{(r)} + \dots + \xi_S^{(r)} = \zeta_S^{(r)} - Sm_r$. Обозначим $l = n - kr - Sm_r$ и

$$\gamma = \left(\frac{S^{1/(g-1)}h(S)}{l} \right)^{1/3}. \quad (8)$$

Согласно условиям леммы нетрудно видеть, что $\gamma \rightarrow 0$. Тогда вероятность $\mathbf{P}\{\zeta_S^{(r)} = n - kr\}$ можно представить в виде

$$\begin{aligned} \mathbf{P}\{\zeta_S^{(r)} = n - kr\} &= \mathbf{P}\{\tilde{\zeta}_S^{(r)} = l\} \\ &= P_1(l) + SP_2(l) + P_3(l), \end{aligned} \quad (9)$$

где

$$\begin{aligned} P_1(l) &= \mathbf{P}\{\tilde{\zeta}_S^{(r)} = l, \xi_j^{(r)} \leq \gamma l, j = 1, \dots, S\}, \\ P_2(l) &= \mathbf{P}\{\tilde{\zeta}_S^{(r)} = l, \xi_j^{(r)} \leq \gamma l, j = 1, \dots, S-1, \\ &\quad \xi_S^{(r)} > \gamma l\}, \\ P_3(l) &= \mathbf{P}\{\tilde{\zeta}_S^{(r)} = l, \bigcup_{i \neq j} \{\xi_i^{(r)} > \gamma l, \xi_j^{(r)} > \gamma l\}\}. \end{aligned}$$

Ниже мы покажем, что основной вклад в сумму, стоящую в правой части соотношения (9), дает второе слагаемое.

Оценим первое слагаемое $P_1(l)$. Положим

$$R(\omega) \quad (10)$$

$$= \sum_{j \leq \gamma l + m_r} \exp\{(j - m_r)\omega\} \mathbf{P}\{\xi_1^{(r)} = j - m_r\}.$$

Введем вспомогательные независимые случайные величины $\xi_i^{(r)}(\gamma)$, $i = 1, \dots, S$, такие, что

$$\mathbf{P}\{\xi_i^{(r)}(\gamma) = j - m_r\} \quad (11)$$

$$= \frac{\mathbf{P}\{\xi_i^{(r)} = j - m_r\} \exp\{(j - m_r)/(\gamma l)\}}{R(1/(\gamma l))},$$

где $1 \leq j \leq \gamma l + m_r$. Тогда $P_1(l)$ можно представить в виде

$$P_1(l) = R^S \left(\frac{1}{\gamma l} \right) \exp \left\{ -\frac{1}{\gamma} \right\} \mathbf{P} \{ \zeta_S^{(r)}(\gamma) = l \}, \quad (12)$$

где $\zeta_S^{(r)}(\gamma) = \xi_1^{(r)}(\gamma) + \dots + \xi_S^{(r)}(\gamma)$.

Рассмотрим вероятность $\mathbf{P} \{ \zeta_S^{(r)}(\gamma) = l \}$. Обозначим через $\varphi(t)$, $\varphi_1^{(r)}(t)$, $\tilde{\varphi}(t)$ и $\varphi_\gamma(t)$ характеристические функции случайных величин η_1 , $\eta_1^{(r)}$, $\xi_1^{(r)}$ и $\xi_1^{(r)}(\gamma)$ соответственно. Используя (10) и (11), нетрудно получить, что

$$\varphi_\gamma(t) = R \left(it + \frac{1}{\gamma l} \right) / R \left(\frac{1}{\gamma l} \right). \quad (13)$$

Согласно формуле обращения, для вероятности $\mathbf{P} \{ \zeta_S^{(r)}(\gamma) = l \}$ справедливо неравенство

$$\mathbf{P} \{ \zeta_S^{(r)}(\gamma) = l \} \leq \frac{C_4}{B_S} \int_{-\pi B_S}^{\pi B_S} \left| \varphi_\gamma^S \left(\frac{t}{B_S} \right) \right| dt. \quad (14)$$

Используя (12), (13) и (14), нетрудно показать, что при $0 < \varepsilon < 1$

$$P_1(l) \leq \frac{C_4}{B_S} e^{-1/\gamma} \left(\int_{|t| \leq \varepsilon B_S} \left| R \left(\frac{it}{B_S} + \frac{1}{\gamma l} \right) \right|^S dt + \int_{\varepsilon B_S < |t| \leq \pi B_S} \left| R \left(\frac{it}{B_S} + \frac{1}{\gamma l} \right) \right|^S dt \right). \quad (15)$$

Покажем, что интегралы правой части (15) ограничены. Учитывая, что при $0 \leq y \leq 1$ справедливы соотношения

$$e^y = 1 + y + \delta(y), \quad e^{-y} = 1 - y + \delta(y), \quad (16)$$

где $\delta(y) \leq y^2$, из (7), (10) и (16) находим, что

$$\begin{aligned} & \left| R \left(it + \frac{1}{\gamma l} \right) \right| \\ &= \left| \sum_{1 \leq j \leq \gamma l + m_r} \exp \left\{ (j - m_r) \left(it + \frac{1}{\gamma l} \right) \right\} \right. \\ & \quad \left. \times \mathbf{P} \{ \xi_1^{(r)} = j - m_r \} \right| \\ &\leq \left| \sum_{1 \leq j \leq \gamma l + m_r} \exp \{ itj \} \exp \left\{ \frac{j - m_r}{\gamma l} \right\} \right| \end{aligned}$$

$$\begin{aligned} & \times \mathbf{P} \{ \xi_1^{(r)} = j - m_r \} \Big| \\ &\leq \left| \sum_{1 \leq j \leq \gamma l + m_r} \exp \{ itj \} \left(1 + \frac{j - m_r}{\gamma l} + \left(\frac{j - m_r}{\gamma l} \right)^2 \right) \mathbf{P} \{ \xi_1^{(r)} = j - m_r \} \right| \\ &\leq \left| \sum_{j \geq 1} \exp \{ itj \} \mathbf{P} \{ \eta_1^{(r)} = j \} - \sum_{j > \gamma l + m_r} \exp \{ itj \} \mathbf{P} \{ \eta_1^{(r)} = j \} \right. \\ & \quad \left. + \sum_{1 \leq j \leq \gamma l + m_r} \exp \{ itj \} \frac{j - m_r}{\gamma l} \mathbf{P} \{ \xi_1^{(r)} = j - m_r \} \right. \\ & \quad \left. + \frac{1}{(\gamma l)^2} \sum_{1 \leq j \leq \gamma l + m_r} (j - m_r)^2 \mathbf{P} \{ \eta_1^{(r)} = j \} \right. \\ & \quad \leq \left| \varphi_1^{(r)}(t) \right| + \sum_{j > \gamma l + m_r} \mathbf{P} \{ \eta_1^{(r)} = j \} \\ & \quad + \frac{1}{\gamma l} \sum_{j > \gamma l + m_r} (j - m_r) \mathbf{P} \{ \xi_1^{(r)} = j - m_r \} \\ & \quad \left. + \frac{1}{(\gamma l)^2} \sum_{1 \leq j \leq \gamma l + m_r} (j - m_r)^2 \mathbf{P} \{ \eta_1^{(r)} = j \} \right|. \quad (17) \end{aligned}$$

Используя свойства медленно меняющейся функции: для любого достаточно малого $\varepsilon > 0$

$$\lim_{x \rightarrow \infty} h(x)x^{-\varepsilon} = 0, \quad \lim_{x \rightarrow \infty} h(x)x^\varepsilon = \infty, \quad (18)$$

из (1) нетрудно получить, что при достаточно больших k и любом достаточно малом $\varepsilon > 0$

$$\begin{aligned} & \sum_{i > k} p_i \leq C_5 k^{-g+1+\varepsilon}, \\ & \sum_{i > k} (i - M) p_i \leq C_6 k^{-g+2+\varepsilon}, \\ & \sum_{i \leq k} (i - M)^2 p_i \leq C_7 k^{-g+3+\varepsilon}. \end{aligned} \quad (19)$$

Отсюда нетрудно видеть, что справедливы следующие соотношения:

$$\sum_{j > l\gamma} \mathbf{P} \{ \eta_1^{(r)} = j \} = \frac{1}{1 - p_r} \sum_{j > l\gamma} \mathbf{P} \{ \eta_1 = j, \eta_1 \neq r \}$$

$$\leq C_8 \sum_{j>l\gamma} p_k \leq C_9 (l\gamma)^{-g+1+\varepsilon}, \quad (20)$$

$$\sum_{j>\gamma l+m_r} (j-m_r) \mathbf{P}\{\xi_1^{(r)} = j-m_r\} \leq C_{10} (l\gamma)^{-g+2+\varepsilon} \quad (21)$$

и

$$\sum_{j<l\gamma} (j-m_r)^2 \mathbf{P}\{\eta_1^{(r)} = j\} \leq C_{11} (l\gamma)^{-g+3+\varepsilon}. \quad (22)$$

Тогда из (17), (20)–(22) вытекает, что

$$\left| R \left(it + \frac{1}{\gamma l} \right) \right| \leq \left| \varphi_1^{(r)}(t) \right| + C_{12} (l\gamma)^{-g+1+\varepsilon}, \quad (23)$$

где

$$\varphi_1^{(r)}(t) = \frac{\varphi(t) - p_r e^{itr}}{1 - p_r}. \quad (24)$$

Согласно условиям леммы и (8) выполняется соотношение

$$\frac{\gamma l}{S} = h^{1/3}(S) l^{2/3} S^{(4-3g)/(3(g-1))} \geq C_{13} > 0, \quad (25)$$

тогда найдется такое $\varepsilon > 0$, что при $N \rightarrow \infty$

$$(\gamma l)^{-g+1+\varepsilon} = o(1/S). \quad (26)$$

В статье [8] (см. соотношение (2.30)) показано, что в окрестности нуля логарифм характеристической функции $\eta_1 - M$ можно представить в виде:

$$\begin{aligned} \ln \varphi_{\eta_1-M}(t) &= \Gamma(2-g) \cos \frac{\pi(g-1)}{2} \frac{h(1/|t|)}{g-1} |t|^{g-1} \\ &\times \left(1 - i \frac{t}{|t|} \operatorname{tg} \frac{\pi(g-1)}{2} \right) \\ &= \Gamma(1-g) h \left(\frac{1}{|t|} \right) |t|^{g-1} \\ &\times \cos \frac{\pi(g-1)}{2} \left(1 - i \frac{t}{|t|} \operatorname{tg} \frac{\pi(g-1)}{2} \right), \quad (27) \end{aligned}$$

где $\Gamma(z)$ – значение гамма-функции в точке z , а $\varphi_{\eta_1-M}(t)$ – характеристическая функция случайной величины $\eta_1 - M$. Отсюда и (24) нетрудно показать, что при $rt \rightarrow 0$

$$\begin{aligned} \varphi_1^{(r)}(t) &= \frac{1}{1-p_r} (\varphi(t) - p_r e^{itr}) \\ &= 1 + \frac{it}{1-p_r} (M - rp_r) \end{aligned}$$

$$\begin{aligned} &+ \frac{\Gamma(1-g)}{1-p_r} h \left(\frac{1}{|t|} \right) |t|^{g-1} \cos \frac{\pi(g-1)}{2} \\ &\times \left(1 - i \frac{t}{|t|} \operatorname{tg} \frac{\pi(g-1)}{2} \right) + O((rt)^2). \quad (28) \end{aligned}$$

Из (2), (6) и (28) следует, что для любого фиксированного t , $N \rightarrow \infty$ и $r\sqrt{S}/B_S \leq C_{14}$

$$\begin{aligned} \ln \left(\varphi_1^{(r)} \left(\frac{t}{B_S} \right) \right)^S &= itm_r S/B_S \\ &+ \frac{S\Gamma(1-g)}{1-p_r} h \left(\frac{B_S}{|t|} \right) \cos \frac{\pi(g-1)}{2} \frac{|t|^{g-1}}{B_S^{g-1}} \\ &\times \left(1 - i \frac{t}{|t|} \operatorname{tg} \frac{\pi(g-1)}{2} \right) + O \left(\frac{r^2 S}{B_S^2} \right) \\ &= itm_r S/B_S \\ &- \frac{\Gamma(2-g)}{1-p_r} |t|^{g-1} \cos \frac{\pi(g-1)}{2} \\ &\times \left(1 - i \frac{t}{|t|} \operatorname{tg} \frac{\pi(g-1)}{2} \right) + O \left(\frac{r^2 S}{B_S^2} \right) \quad (29) \end{aligned}$$

(здесь рассматривается та ветвь логарифмической функции, для которой $\ln 1 = 0$). Тогда из (3), (23), (26) и (29), используя свойства характеристических функций, находим, что при любом фиксированном t и $r\sqrt{S}/B_S \leq C_{14}$

$$\begin{aligned} \left| R \left(\frac{it}{B_S} + \frac{1}{\gamma l} \right) \right|^S &\leq C_{15} \left| \varphi_1^{(r)} \left(\frac{t}{B_S} \right) \right|^S \\ &\times (1 + C_{16} (\gamma l)^{-g+1+\varepsilon})^S \leq C_{17} \exp\{-C_{18} |t|^{g-1}\}. \end{aligned}$$

Следовательно, при $N \rightarrow \infty$, $r\sqrt{S}/B_S \leq C_{14}$ первый интеграл в правой части (15) ограничен.

Рассмотрим первый интеграл правой части (15) при $r\sqrt{S}/B_S \rightarrow \infty$. Согласно условию леммы в этом случае выполняется соотношение

$$Sh(r)/r^g \rightarrow 0. \quad (30)$$

Покажем, что и в этом случае интеграл также будет ограничен. Если $r\sqrt{S}/B_S \rightarrow \infty$, тогда из (3) следует, что $r \rightarrow \infty$. Из (24) получаем, что

$$\begin{aligned} \left(\varphi_1^{(r)} \left(\frac{t}{B_S} \right) \right)^S &= \varphi^S \left(\frac{t}{B_S} \right) (1-p_r)^{-S} \\ &\times \left(1 - p_r \exp \left\{ \frac{itr}{B_S} \right\} \varphi^{-1} \left(\frac{t}{B_S} \right) \right)^S. \quad (31) \end{aligned}$$

Из (1) нетрудно видеть, что при достаточно больших N

$$S \ln(1 - p_r) = O(Nh(r)/r^g). \quad (32)$$

Покажем, что при любых фиксированных t , $N \rightarrow \infty$ справедливо

$$\begin{aligned} S \ln \left(1 - p_r \exp \left\{ \frac{itr}{B_S} \right\} \varphi^{-1} \left(\frac{t}{B_S} \right) \right) \\ = O(Sh(r)/r^g). \end{aligned} \quad (33)$$

Учитывая, что при $r \rightarrow \infty$

$$\left| p_r \exp \left\{ \frac{itr}{B_S} \right\} \varphi^{-1} \left(\frac{t}{B_S} \right) \right| \leq C_{18} p_r,$$

и раскладывая логарифмическую функцию в окрестности точки 0 в ряд Тейлора, получим, что при любых фиксированных t и $N \rightarrow \infty$

$$\begin{aligned} \ln \left(1 - p_r \exp \left\{ \frac{itr}{B_S} \right\} \varphi^{-1} \left(\frac{t}{B_S} \right) \right) \\ = -p_r \exp \left\{ \frac{itr}{B_S} \right\} \varphi^{-1} \left(\frac{t}{B_S} \right) \\ - \left(p_r \exp \left\{ \frac{itr}{B_S} \right\} \varphi^{-1} \left(\frac{t}{B_S} \right) \right)^2 / 2 + \dots \quad (34) \\ = -p_r \exp \left\{ \frac{itr}{B_S} \right\} \varphi^{-1} \left(\frac{t}{B_S} \right) \\ \times \left(1 + \left(p_r \exp \left\{ \frac{itr}{B_S} \right\} \varphi^{-1} \left(\frac{t}{B_S} \right) \right) / 2 + \dots \right). \end{aligned}$$

Несложно видеть, что

$$\begin{aligned} \left| -p_r \exp \left\{ \frac{itr}{B_S} \right\} \varphi^{-1} \left(\frac{t}{B_S} \right) \right| \\ \times \left(1 + \left(p_r \exp \left\{ \frac{itr}{B_S} \right\} \varphi^{-1} \left(\frac{t}{B_S} \right) \right) / 2 + \dots \right) \\ \leq p_r \left| \varphi^{-1} \left(\frac{t}{B_S} \right) \right| \left(1 + p_r \left| \varphi^{-1} \left(\frac{t}{B_S} \right) \right| \right. \\ \left. + \left(p_r \left| \varphi^{-1} \left(\frac{t}{B_S} \right) \right| \right)^2 + \dots \right) \\ = p_r \left| \varphi^{-1} \left(\frac{t}{B_S} \right) \right| / \left| 1 - p_r \left| \varphi^{-1} \left(\frac{t}{B_S} \right) \right| \right|. \end{aligned}$$

Отсюда и из (3), (34) следует, что при любых фиксированных t и $N \rightarrow \infty$

$$\ln \left(1 - p_r \exp \left\{ \frac{itr}{B_S} \right\} \varphi^{-1} \left(\frac{t}{B_S} \right) \right)$$

$$\begin{aligned} = O \left(p_r \left| \varphi^{-1} \left(\frac{t}{B_S} \right) \right| / \left| 1 - p_r \left| \varphi^{-1} \left(\frac{t}{B_S} \right) \right| \right| \right) \\ = O(p_r), \end{aligned}$$

тогда отсюда и из (1) вытекает справедливость равенства (33). Тогда из (23), (26), (30)–(33) и используя свойства характеристических функций, находим, что при любом фиксированном t

$$\begin{aligned} \left| R \left(\frac{it}{B_S} + \frac{1}{\gamma l} \right) \right|^S \leq \left| \varphi_1^{(r)} \left(\frac{t}{B_S} \right) \right|^S \\ \times (1 + C_{19}(\gamma l)^{-g+1+\varepsilon})^S \leq C_{20} \left| \varphi \left(\frac{t}{B_S} \right) \right|^S. \end{aligned} \quad (35)$$

Из (2) и (27) получаем, что при любом фиксированном t и $N \rightarrow \infty$

$$\begin{aligned} \ln \varphi_{\eta_1-M}^S(t/B_S) = -\Gamma(2-g)|t|^{g-1} \cos \frac{\pi(g-1)}{2} \\ \times \left(1 - i \frac{t}{|t|} \operatorname{tg} \frac{\pi(g-1)}{2} \right) + o(1). \end{aligned}$$

Отсюда и из (35) находим, что при выполнении условий леммы в случае, когда $r\sqrt{S}/B_S \rightarrow \infty$ и любом фиксированном t

$$\left| R \left(\frac{it}{B_S} + \frac{1}{\gamma l} \right) \right|^S \leq C_{21} \exp \{ -C_{22}|t|^{g-1} \},$$

откуда и вытекает, что при $r\sqrt{S}/B_S \rightarrow \infty$ первый интеграл в (15) тоже ограничен.

Для оценки второго интеграла в (15) можно использовать известное свойство характеристических функций решетчатых распределений с единичным максимальным шагом, согласно которому при $\varepsilon < |t| \leq \pi$ справедливо неравенство

$$|\varphi_1^{(r)}(t)| \leq e^{-C_{23}}.$$

Отсюда и из (8), (23), (26) получаем, что при $N \rightarrow \infty$

$$\begin{aligned} \int_{\varepsilon B_S < |t| \leq \pi B_S} \left| R \left(\frac{it}{B_S} + \frac{1}{\gamma l} \right) \right|^S dt \\ \leq \int_{\varepsilon B_S < |t| \leq \pi B_S} \left(\left| \varphi_1^{(r)} \left(\frac{t}{B_S} \right) \right| + C_{24}(\gamma l)^{-g+1+\varepsilon} \right)^S dt \\ \leq C_{25} \int_{\varepsilon B_S < |t| \leq \pi B_S} e^{-C_{26}S} dt. \end{aligned}$$

Следовательно, второй интеграл в (15) стремится к 0 при $N \rightarrow \infty$. Учитывая, что оба интеграла в равенстве (15) ограничены, находим, что

$$P_1(l) \leq C_{27} B_S^{-1} e^{-1/\gamma}. \quad (36)$$

Покажем, что

$$P_1(l) = o\left(\frac{Sh(l)}{l^g}\right). \quad (37)$$

Используя (3), получаем, что для достаточно большого N и достаточно малого $\varepsilon > 0$

$$\begin{aligned} C_{28} S^{\varepsilon(g-2-\varepsilon)/(g-1+\varepsilon)} &> \left(\frac{S}{B_S}\right)^\varepsilon \\ &> C_{29} S^{\varepsilon(g-2+\varepsilon)/(g-1+\varepsilon)}. \end{aligned}$$

Следовательно, по свойствам медленно меняющейся функции для любого $\varepsilon > 0$ выполнено соотношение

$$\lim_{S \rightarrow \infty} S^\varepsilon h(B_S) = \infty. \quad (38)$$

Согласно (2) справедливо соотношение $B_S \asymp S^{1/(g-1)} h^{1/(g-1)}(B_S)$, где запись $f_1(x) \asymp f_2(x)$ означает, что $f_1(x) = O(f_2(x))$ и $f_2(x) = O(f_1(x))$ для положительных функций $f_1(x)$ и $f_2(x)$. Отсюда и из (8) нетрудно получить, что для достаточно малого положительного ε

$$\frac{\exp\{-1/\gamma\} n^g}{B_S Sh(l)} \quad (39)$$

$$\asymp \exp\left\{-\frac{1}{\gamma}\right\} \left(\frac{1}{\gamma}\right)^{3(g+\varepsilon)} \frac{h^{g+\varepsilon}(S) S^{\varepsilon/(g-1)}}{h^{1/(g-1)}(B_S) h(l) l^\varepsilon}.$$

Поскольку $n/N \rightarrow \infty$ при выполнении условий леммы, из (2), (8), (18), (38), (39) можно показать, что при $N \rightarrow \infty$

$$\frac{\exp\{-1/\gamma\} l^g}{B_S Sh(l)} \rightarrow 0.$$

Отсюда и из (36), учитывая условия леммы, вытекает равенство (37).

Рассмотрим $P_2(l)$. Нетрудно видеть, что $P_2(l)$ можно представить в виде суммы:

$$\begin{aligned} P_2(l) &= \sum_{j > \gamma l + m_r} \mathbf{P}\{\xi_S^{(r)} = j - m_r\} \\ &\times \mathbf{P}\{\tilde{\zeta}_{S-1}^{(r)} = l - j + m_r, \xi_i^{(r)} \leq \gamma l, \\ &\quad i = 1, \dots, S-1\}. \end{aligned} \quad (40)$$

Введем вспомогательные независимые одинаково распределенные случайные величины $\xi_{i,\gamma l}^{(r)}$, $i = 1, \dots, S-1$, для которых

$$\mathbf{P}\{\xi_{1,\gamma l}^{(r)} = k\}$$

$$= \mathbf{P}\{\eta_1 = k - m_r \mid \eta_1 \neq r, \eta_1 \leq \gamma l\}, \quad (41)$$

и их сумму $\zeta_{S-1,\gamma l}^{(r)} = \xi_{1,\gamma l}^{(r)} + \dots + \xi_{S-1,\gamma l}^{(r)}$. Из (5), (7), (20), (25), (26) и (41) получаем, что

$$\begin{aligned} \mathbf{P}\{\tilde{\zeta}_{S-1}^{(r)} = l - j + m_r, \xi_i^{(r)} \leq \gamma l, i = 1, \dots, S-1\} \\ = \left(\mathbf{P}\{\xi_1^{(r)} \leq \gamma l\}\right)^{S-1} \mathbf{P}\{\zeta_{S-1,\gamma l}^{(r)} = l - j + m_r\} \\ = \left(1 - \sum_{k > \gamma l} \mathbf{P}\{\xi_1^{(r)} = k\}\right)^{S-1} \\ \times \mathbf{P}\{\zeta_{S-1,\gamma l}^{(r)} = l - j + m_r\} \\ = \left(1 - \sum_{k > \gamma l} \mathbf{P}\{\eta_1^{(r)} = k + m_r\}\right)^{S-1} \\ \times \mathbf{P}\{\zeta_{S-1,\gamma l}^{(r)} = l - j + m_r\} \\ = \mathbf{P}\{\zeta_{S-1,\gamma l}^{(r)} = l - j + m_r\} (1 + o(1)). \end{aligned}$$

Тогда, учитывая, что случайные величины $\xi_{1,\gamma l}^{(r)}, \dots, \xi_{S-1,\gamma l}^{(r)}$ могут принимать отрицательные значения, из (40) получаем, что $P_2(l)$ можно представить в виде:

$$P_2(l) = \sum_{i=1}^3 R_i(l) (1 + o(1)), \quad (42)$$

где

$$R_i(l) = \sum_{J_i} \mathbf{P}\{\xi_S^{(r)} = j\} \mathbf{P}\{\zeta_{S-1,\gamma l}^{(r)} = l - j\},$$

$$J_1 = \{j : \gamma l < j < \gamma^{1/2} l\};$$

$$J_2 = \{j : \gamma^{1/2} l \leq j < l - \gamma^{1-\Delta} l\};$$

$$J_3 = \{j : l - \gamma^{1-\Delta} l \leq j \leq l + m_r(S-1)\}$$

при достаточно малом положительном Δ .

Покажем, что основной вклад в $P_2(l)$ дает $R_3(l)$. Из (1), (5) и (7) при выполнении условий леммы находим, что при $j \in J_3$

$$\mathbf{P}\{\xi_S^{(r)} = j\} = h(l) (1 + o(1)) / ((1 - p_r) l^g).$$

Отсюда и из (42) получаем, что

$$R_3(l) = \frac{h(l)(1 + o(1))}{(1 - p_r)l^g} \quad (43)$$

$$\times \mathbf{P} \left\{ -m_r(S - 1) \leq \zeta_{S-1, \gamma l}^{(r)} \leq \gamma^{1-\Delta} l \right\}.$$

Пусть $g(x)$ обозначает плотность устойчивого закона с показателем $g - 1$ и характеристической функцией

$$f(t) = \exp \left\{ -\Gamma(2 - g) \cos(\pi(g - 1)/2) |t|^{g-1} \right. \\ \left. \times \left(1 - i \frac{t}{|t|} \operatorname{tg}(\pi(g - 1)/2) \right) \right\}. \quad (44)$$

Через $\varphi_{\gamma l, r}(t)$ обозначим характеристическую функцию случайных величин $\xi_{1, \gamma l}^{(r)}$. Из (1), (5), (20), (26) и (41) нетрудно видеть, что в окрестности нуля (при $t \rightarrow 0$) при достаточно большом N и достаточно малом положительном ε

$$\begin{aligned} \varphi_{\gamma l, r}(t) &= \sum_k e^{itk} \mathbf{P} \{ \xi_{1, \gamma l}^{(r)} = k \} \\ &= \frac{\tilde{\varphi}(t) - \sum_{k > \gamma l} e^{itk} \mathbf{P} \{ \xi_1^{(r)} = k \}}{1 - \mathbf{P} \{ \xi_1^{(r)} > \gamma l \}} \\ &= \frac{\tilde{\varphi}(t) - \sum_{k > \gamma l} e^{itk} \mathbf{P} \{ \xi_1^{(r)} = k \}}{1 - \mathbf{P} \{ \eta_1^{(r)} > \gamma l + m_r \}} \\ &= \tilde{\varphi}(t) (1 + O((\gamma l)^{1-g+\varepsilon})) \\ &= e^{-itm_r} \varphi_1^{(r)}(t) (1 + O((\gamma l)^{1-g+\varepsilon})) \\ &= e^{-itm_r} \varphi_1^{(r)}(t) (1 + o(1/S)). \end{aligned} \quad (45)$$

Рассмотрим случай, когда $r\sqrt{S}/B_S \rightarrow 0$. Используя равенство (28), нетрудно показать, что для любого фиксированного t

$$\begin{aligned} &\ln \left(\varphi_1^{(r)} \left(\frac{t(1 - p_r)^{1/(g-1)}}{B_S} \right) \right)^S \\ &= itm_r S \frac{(1 - p_r)^{1/(g-1)}}{B_S} - \Gamma(2 - g) \cos \frac{\pi(g - 1)}{2} \\ &\times |t|^{g-1} \left(1 - i \frac{t}{|t|} \operatorname{tg} \frac{\pi(g - 1)}{2} \right) + O \left(\frac{r^2 S}{B_S^2} \right). \end{aligned}$$

Отсюда и из (45) следует, что при $S \rightarrow \infty$ и $r\sqrt{S}/B_S \rightarrow 0$ распределение случайной величины

$$\frac{\zeta_{S, \gamma l}^{(r)} - m_r S}{B_S / (1 - p_r)^{1/(g-1)}}$$

слабо сходится к устойчивому закону с показателем $g - 1$ и характеристической функцией (44). Значит, при $S \rightarrow \infty$ и $r\sqrt{S}/B_S \rightarrow 0$ справедливо

$$\begin{aligned} &\mathbf{P} \{ -m_r(S - 1) \leq \zeta_{S-1, \gamma l}^{(r)} \leq \gamma^{1-\Delta} l \} \\ &= \mathbf{P} \left\{ \frac{-2m_r(S - 1)}{B_{S-1} / (1 - p_r)^{1/(g-1)}} \right. \\ &\leq \left. \frac{\zeta_{S-1, \gamma l}^{(r)} - m_r(S - 1)}{B_{S-1} / (1 - p_r)^{1/(g-1)}} \leq \frac{\gamma^{1-\Delta} l - m_r(S - 1)}{B_{S-1} / (1 - p_r)^{1/(g-1)}} \right\} \\ &\sim \int_{-\infty}^{\infty} g(x) dx = 1, \end{aligned} \quad (46)$$

где $g(x)$ обозначает плотность устойчивого закона с показателем $g - 1$ и характеристической функцией (44). Из (3), (43) и (46) находим, что при $S, l \rightarrow \infty$ и $r\sqrt{S}/B_S \rightarrow 0$

$$R_3(l) \sim \frac{h(l)}{(1 - p_r)l^g}. \quad (47)$$

Рассмотрим $R_3(l)$ при $r\sqrt{S}/B_S \geq C_{14} > 0$. Учитывая, что в этом случае $r \rightarrow \infty$, нетрудно проверить, что соотношения (31), (32), (33) выполняются. Из условия леммы (30) и соотношений (31), (32), (33) получаем, что при $S \rightarrow \infty$

$$\left(\varphi_1^{(r)} \left(\frac{t}{B_S} \right) \right)^S = \varphi^S \left(\frac{t}{B_S} \right) (1 + o(1)).$$

Отсюда и из (2), (3), (27), (45) следует, что при $S \rightarrow \infty$ и $r\sqrt{S}/B_S \geq C_{14} > 0$ распределение случайной величины

$$\frac{\zeta_{S, \gamma l}^{(r)} - m_r S}{B_S}$$

слабо сходится к устойчивому закону с показателем $g - 1$ и характеристической функцией (44). Значит, и для случая, когда $r/B_S \geq C_{15} > 0$, соотношение (47) сохраняет свою силу.

Рассмотрим $R_1(l)$. Из (1), (5), (7) и (42) находим, что

$$R_1(l) \leq \frac{C_{30} h(\gamma l)}{(\gamma l)^g} \mathbf{P} \{ \zeta_{S-1, \gamma l}^{(r)} > l - l\gamma^{1/2} \}. \quad (48)$$

Для оценки вероятности $\mathbf{P} \{ \zeta_{S-1, \gamma l}^{(r)} > l - l\gamma^{1/2} \}$ будем использовать неравенство Чебышева.

Сделаем оценку первого и второго момента случайной величины $\xi_{1,\gamma l}^{(r)}$. Из (1), (5), (7), (21), (22), (41), учитывая, что $\mathbf{E}\xi_1^{(r)} = 0$, получаем, что для достаточно малого $\varepsilon > 0$

$$\begin{aligned} |\mathbf{E}\xi_{1,\gamma l}^{(r)}| &= \left| \sum_{k \leq \gamma l} k \frac{\mathbf{P}\{\eta_1^{(r)} = k - m_r\}}{1 - (1 - p_r)^{-1} \sum_{k > \gamma l} p_k} \right| \\ &\leq \left| \sum_{1-m_r \leq k \leq \gamma l} k \mathbf{P}\{\xi_1^{(r)} = k\} \right| (1 + C_{31}(\gamma l)^{-g+1+\varepsilon}) \\ &= \sum_{k > \gamma l} k \mathbf{P}\{\xi_1^{(r)} = k\} (1 + C_{31}(\gamma l)^{-g+1+\varepsilon}) \\ &\leq C_{32} \sum_{k > \gamma l} (k - m_r) p_k \leq C_{33}(\gamma l)^{2+\varepsilon-g}. \quad (49) \end{aligned}$$

и

$$\begin{aligned} \mathbf{E} \left(\xi_{1,\gamma l}^{(r)} \right)^2 &= \sum_{1-m_r \leq k} k^2 \mathbf{P}\{\xi_{1,\gamma l}^{(r)} = k\} \\ &\leq C_{34} \sum_{1-m_r \leq k \leq \gamma l} k^2 \mathbf{P}\{\xi_1^{(r)} = k\} \\ &\leq C_{35} \sum_{1-m_r \leq k \leq \gamma l} (k - m_r)^2 p_k \\ &\leq C_{36}(\gamma l)^{3+\varepsilon-g}. \quad (50) \end{aligned}$$

Применяя неравенство Чебышева, из (49) и (50) находим, что при достаточно больших l

$$\begin{aligned} \mathbf{P}\{\zeta_{S-1,\gamma l}^{(r)} > l - l\gamma^{1/2}\} \\ &\leq \frac{C_{37}}{l^2} \mathbf{E} \left(\xi_{1,\gamma l}^{(r)} + \dots + \xi_{S-1,\gamma l}^{(r)} \right)^2 \\ &\leq \frac{C_{38}}{l^2} ((S-1)(l\gamma)^{3+\varepsilon-g} \\ &\quad + (S-1)(S-2)(\gamma l)^{-2g+4+2\varepsilon}) \\ &\leq C_{39} S \gamma^{-g+3+\varepsilon} l^{-g+1+\varepsilon}. \quad (51) \end{aligned}$$

Из (8), (48) и (51) получаем, что при $S \rightarrow \infty$ и достаточно малом $\varepsilon > 0$

$$\begin{aligned} R_1(l) &= O \left(\frac{h(\gamma l)}{(\gamma l)^g} S l^{-g+1+\varepsilon} \gamma^{-g+3+\varepsilon} \right) \\ &= O \left(\frac{h(l)}{l^g} \frac{h(\gamma l)}{h(l)} S^A l^B h^D(S) \right), \quad (52) \end{aligned}$$

где

$$A = \frac{1}{3} \left(1 + \frac{\varepsilon + 1}{g - 1} \right), \quad B = -\frac{1}{3}g + \frac{2}{3}\varepsilon,$$

$$D = -\frac{2}{3}g + 1 + \frac{1}{3}\varepsilon.$$

Используя (53), можно показать, что $-B - A > 0$ при $2 < g < 3$. Тогда из (18) и (25) следует, что при $N \rightarrow \infty$

$$\frac{h(\gamma l)}{h(l)} S^A l^B h^D(S) \rightarrow 0.$$

Отсюда и из (52) получаем, что

$$R_1(l) = o \left(\frac{h(l)}{l^g} \right). \quad (53)$$

Рассмотрим $R_2(l)$. Из (1), (5), (7), (8) и (42) следует, что

$$\begin{aligned} R_2(l) &\leq \frac{C_{40} h(\gamma^{1/2} l)}{(\gamma^{1/2} l)^g} \mathbf{P}\{\zeta_{S-1,\gamma l}^{(r)} > \gamma^{1-\Delta} l\} \\ &\leq \frac{C_{41} h(\gamma^{1/2} l)}{(\gamma^{1/2} l)^g} \mathbf{P}\{\zeta_{S-1,\gamma l}^{(r)} > \gamma l\}. \quad (54) \end{aligned}$$

Оценка вероятности $\mathbf{P}\{\zeta_{S-1,\gamma l}^{(r)} > \gamma l\}$ осуществляется аналогично тому, как получена оценка для вероятности $\mathbf{P}\{\zeta_{S-1,\gamma l}^{(r)} > l - l\gamma^{1/2}\}$. Применяя неравенство Чебышева, из (49) и (50) находим, что при $S \rightarrow \infty$ и достаточно малом $\varepsilon > 0$

$$\begin{aligned} \mathbf{P}\{\zeta_{S-1,\gamma l}^{(r)} > \gamma l\} &\leq \frac{C_{42}}{(\gamma l)^2} S (l\gamma)^{3+\varepsilon-g} \\ &\leq C_{43} S (\gamma l)^{-g+1+\varepsilon}. \end{aligned}$$

Отсюда и из (18), (54) получаем, что для достаточно больших S, l и достаточно малом $\varepsilon > 0$

$$\begin{aligned} R_2(l) &\leq C_{44} \left(\gamma^{1/2} l \right)^{-g+\varepsilon} \mathbf{P}\{\zeta_{S-1,\gamma l}^{(r)} > \gamma l\} \\ &\leq C_{45} \frac{S(\gamma l)^{-g+1+\varepsilon}}{(\gamma^{1/2} l)^{g-\varepsilon}} \\ &= C_{45} \frac{h(l)}{l^g} \left(\frac{1}{h(l)} S l^{1-g+2\varepsilon} \gamma^{1-3g/2+3\varepsilon/2} \right). \quad (55) \end{aligned}$$

Покажем, что при $S \rightarrow \infty$

$$\frac{1}{h(l)} S l^{1-g+2\varepsilon} \gamma^{1-3g/2+3\varepsilon/2} \rightarrow 0. \quad (56)$$

Из (8) находим, что

$$\frac{1}{h(l)} S l^{1-g+2\varepsilon} \gamma^{1-3g/2+3\varepsilon/2} = \frac{1}{h(l)} l^A S^B h^D(S),$$

где

$$A = 2/3 - g/2 + 3\varepsilon/2, \quad B = \frac{3g/2 - 2 + 3\varepsilon/2}{3(g-1)},$$

$$D = \frac{1 - 3g/2 + 3\varepsilon/2}{3}.$$

Нетрудно видеть, что в рассматриваемой области изменения параметра g значение $-A - B > 0$. Тогда, согласно свойствам медленно меняющейся функции, справедливо соотношение

$$\begin{aligned} & \frac{1}{h(l)} S l^{1-g+2\varepsilon} \gamma^{1-3g/2+3\varepsilon/2} \\ &= \frac{1}{h(l) n^{(-A-B)/2}} \left(\frac{l}{S} \right)^B \frac{h^D(S)}{l^{(-A-B)/2}} \rightarrow 0, \end{aligned}$$

значит, соотношение (56) верно. Из (55) и (56) вытекает, что

$$R_2(l) = o\left(\frac{h(l)}{l^g}\right). \quad (57)$$

Из (42), (47), (53) и (57) находим, что

$$P_2(l) = \frac{h(l)(1 + o(l))}{(1 - p_r)l^g}. \quad (58)$$

Осталось рассмотреть $P_3(l)$. Из (9) нетрудно видеть, что

$$P_3(l) = \binom{S}{2} \sum_{k > 2(\gamma l + m_r)} \mathbf{P}\{\xi_{S-1}^{(r)} + \xi_S^{(r)} = k - 2m_r,$$

$$\xi_i^{(r)} > \gamma l, i = S-1, S\}$$

$$\times \mathbf{P}\{\tilde{\zeta}_{S-2}^{(r)} = l - k + 2m_r\}. \quad (59)$$

Из (5), (7), (19) получаем, что при достаточно большом N и достаточно малом $\varepsilon > 0$

$$\begin{aligned} & \mathbf{P}\{\xi_{S-1}^{(r)} + \xi_S^{(r)} = k - 2m_r, \xi_i^{(r)} > \gamma l, i = S-1, S\} \\ &= \sum_J \mathbf{P}\{\xi_{S-1}^{(r)} = j - m_r\} \mathbf{P}\{\xi_S^{(r)} = k - m_r - j\} \end{aligned}$$

$$\leq C_{46} \frac{h(\gamma l)}{(\gamma l)^g} \mathbf{P}\{\xi_S^{(r)} > \gamma l\}$$

$$\leq C_{47} h(\gamma l) (\gamma l)^{-2g+1+\varepsilon}, \quad (60)$$

где

$$J = \{j : \gamma l + m_r < j < k - \gamma l - m_r\}.$$

Из соотношений (59) и (60) вытекает, что

$$\begin{aligned} P_3(l) &\leq C_{48} S^2 h(\gamma l) (\gamma l)^{-2g+1+\varepsilon} \mathbf{P}\{\tilde{\zeta}_{S-2}^{(r)} > l - 2l\gamma\} \\ &\leq C_{49} S^2 h(\gamma l) (\gamma l)^{-2g+1+\varepsilon}. \end{aligned} \quad (61)$$

Используя (8), несложно показать, что

$$\frac{S^2 h(\gamma l)}{(\gamma l)^{2g-1-\varepsilon}} = \frac{Sh(l)}{l^g}$$

$$\times \left(\frac{h(\gamma l)}{h(l) h^{(2g-1-\varepsilon)/3}(S)} \left(\frac{S}{l} \right)^{\frac{g-2+\varepsilon}{3g-3}} \frac{1}{l^\omega} \right), \quad (62)$$

где

$$\omega = \frac{1}{3(g-1)} ((g-2)^2 - \varepsilon(2g+1)).$$

Учитывая, что $l/S \rightarrow \infty$ и $\omega > 0$, из (18), (61) и (62) получаем, что при $N \rightarrow \infty$

$$P_3(l) = o\left(\frac{Sh(l)}{l^g}\right). \quad (63)$$

Тогда утверждение леммы следует из соотношений (9), (37), (58), (63) и условий леммы 3. $\square$

Теперь можно перейти к доказательству основных теорем. Учитывая, что $N - k = N(1 - p_r)(1 + o(1))$, из лемм 2 и 3 находим, что

$$\frac{\mathbf{P}\{\zeta_{N-k}^{(r)} = n - kr\}}{\mathbf{P}\{\zeta_N = n\}} \rightarrow 1. \quad (64)$$

Согласно нормальному приближению биномиального распределения, при $Nh(r)/r^g \rightarrow \infty$ выполняется равенство

$$\begin{aligned} & \binom{N}{k} p_r^k (1 - p_r)^{N-k} \\ &= \frac{1 + o(1)}{\sqrt{2\pi N p_r (1 - p_r)}} \exp \left\{ -\frac{(k - N p_r)^2}{N p_r (1 - p_r)} \right\}. \end{aligned}$$

Тогда утверждение теоремы 1 следует из (64) и леммы 1.

Применяя пуассоновское приближение биномиального распределения при $p_r \rightarrow 0$, лемму 1 и равенство (64), получаем теорему 2.

ЛИТЕРАТУРА

1. Колчин В. Ф. Случайные графы. М.: Физматлит, 2000. 256 с.
2. Павлов Ю. Л. О структуре условных конфигурационных графов с ограниченным числом ребер // Труды Карельского научного центра РАН. 2025. № 4. С. 54–59. doi: 10.17076/mat2041
3. Павлов Ю. Л. Условные конфигурационные графы со случайным параметром степенного распределения степеней // Математический сборник. 2018. Т. 209, № 2. С. 120–137. doi: 10.4213/sm8832
4. Павлов Ю. Л., Хворостянская Е. В. О предельных распределениях степеней вершин конфигурационных графов с ограниченным числом ребер // Математический сборник. 2016. Т. 207, № 3. С. 93–110. doi: 10.4213/sm8512
5. Павлов Ю. Л., Чеплюкова И. А. Об асимптотике степенной структуры конфигурационных графов с ограничениями на число ребер // Дискретная математика. 2018. Т. 30, № 1. С. 77–94. doi: 10.4213/dm1445
6. Павлов Ю. Л., Чеплюкова И. А. Предельные распределения числа вершин заданной степени конфигурационного графа с ограниченным числом ребер // Теория вероятностей и ее применения. 2021. Т. 66, № 3. С. 468–486. doi: 10.4213/tpv5332
7. Павлов Ю. Л., Чеплюкова И. А. Случайные графы Интернет-типа и обобщенная схема размещения // Дискретная математика. 2008. Т. 20, № 3. С. 3–18. doi: 10.4213/dm1008
8. Чеплюкова И. А. О максимальной степени вершины условного конфигурационного графа // Теория вероятностей и ее применения. 2026. Т. 71, № 1. С. 78–94. doi: 10.4213/tpv5776
9. Чеплюкова И. А. О предельных распределениях степеней вершин конфигурационного графа // Труды Карельского научного центра РАН. 2015. № 10. С. 123–130. doi: 10.17076/mat138
10. Чупрунов А. Н., Фазекаш И. Аналог обобщенной схемы размещения. Предельные теоремы для числа ячеек заданного объема // Дискретная математика. 2012. Т. 24, № 1. С. 140–158. doi: 10.4213/dm1178
11. Barabasi L. A., Albert R. Emergence of scaling in random networks // Science. 1999. No. 286. P. 509–512. doi: 10.1126/science.286.5439.509
12. Bollobas B. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs // European J. Combin. 1980. Vol. 1, no. 4. P. 311–316. doi: 10.1016/S01195-6698(80)80030-8
13. Durrett R. Random graph dynamics. Cambridge: Cambridge Univ. Press, 2007. 233 p. doi: 10.1017/CBO9780511546594
14. Faloutsos C., Faloutsos P., Faloutsos M. On power-law relationship of the Internet topology

// Comput. Commun. Rev. 1999. No. 29. P. 251–262. doi: 10.1145/316194.316229

15. Hofstad R. Random graphs and complex networks. Vol. 1. Cambridge: Cambridge Univ. Press, 2017. 337 p. doi: 10.1017/9781316779422
16. Newman M. E. J. Networks. An introduction. Oxford: Oxford Univ. Press, 2010. 772 p.
17. Reittu Y., Norros I. On the power-law random graph model of massive data networks // Perform. Eval. 2004. Vol. 55, no. 1–2. P. 3–23. doi: 10.1016/S0166-5316(03)00097-X

REFERENCES

1. Kolchin V. F. Random graphs. Cambridge: Cambridge Univ. Press; 1999. 256 p. doi: 10.1017/CBO9780511721342
2. Pavlov Yu. L. On the structure of conditional configuration graphs with a bounded number of edges. *Trudy Karelskogo nauchnogo tsentra RAN = Transactions of the Karelian Research Centre RAS.* 2025;4:54–59. (In Russ.). doi: 10.17076/mat2041
3. Pavlov Yu. L. Conditional configuration graphs with discrete power-law distribution of vertex degrees. *Sbornik: Mathematics.* 2018;209(2):258–275. doi: 10.1070/SM8832
4. Pavlov Yu. L., Khvorostyanskaya E. V. On the limit distributions of the degrees of vertices in configuration graphs with a bounded number of edges. *Sbornik: Mathematics.* 2016;207(3):400–417. doi: 10.4213/sm8512
5. Pavlov Yu. L., Cheplyukova I. A. On the asymptotics of degree structure of configuration graphs with bounded number of edges. *Discrete Math. Appl.* 2019;29(4):219–132. doi: 10.1515/dma-2019-0020
6. Pavlov Yu. L., Cheplyukova I. A. Limit distributions of the number of vertices of a given degree in a configuration graph with bounded number of edges. *Theory of Probability and its Applications.* 2021;66(3):376–390. doi: 10.1137/S0040585X97T990460
7. Pavlov Yu. L., Cheplyukova I. A. Random graphs of Internet type and the generalised allocation scheme. *Discrete Math. Appl.* 2008;18(5):447–463. doi: 10.1515/DMA.2008.033
8. Cheplyukova I. A. On the maximum vertex degree of a conditional configuration graph. *Theory of Probability and its Applications.* 2026;71(1):78–94. (In Russ.). doi: 10.4213/tpv5776
9. Cheplyukova I. A. On limit distributions of vertex degrees in a configuration graph. *Trudy Karelskogo nauchnogo tsentra RAN = Transactions of the Karelian Research Centre RAS.* 2015;10:123–130. (In Russ.). doi: 10.17076/mat138
10. Chuprunov A. N., Fazekas I. An analogue of the generalised allocation scheme: limit theorems for the number of cells containing a given number of

particles. *Discrete Math. Appl.* 2012;22(1):101–122. doi: 10.1515/dma-2012-008

11. *Barabasi L. A., Albert R.* Emergence of scaling in random networks. *Science*. 1999;286:509–512. doi: 10.1126/science.286.5439.509

12. *Bollobas B.* A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.* 1980;1(4):311–316. doi: 10.1016/S01195-6698(80)80030-8

13. *Durrett R.* Random graph dynamics. Cambridge: Cambridge Univ. Press; 2007. 221 p. doi: 10.1017/CBO9780511546594

14. *Faloutsos C., Faloutsos P., Faloutsos M.* On power-law relationship of the Internet topology. *Comput. Commun. Rev.* 1999;29:251–262. doi: 10.1145/316194.316229

15. *Hofstad R.* Random graphs and complex networks. Vol. 1. Cambridge: Cambridge Univ. Press; 2017. 337 p. doi: 10.1017/9781316779422

16. *Newman M. E. J.* Networks. An introduction. Oxford: Oxford Univ. Press; 2010. 772 p.

17. *Reittu Y., Norros I.* On the power-law random graph model of massive data networks. *Perform. Eval.* 2004;55(1–2):3–23. doi: 10.1016/S0166-5316(03)00097-X

Поступила в редакцию / received: 13.03.2026; принята к публикации / accepted: 22.05.2026.
Автор заявляет об отсутствии конфликта интересов / The author declares no conflict of interest.

СВЕДЕНИЯ ОБ АВТОРЕ:

Чеплюкова Ирина Александровна

канд. физ.-мат. наук, старший научный сотрудник

e-mail: chia@krc.karelia.ru

CONTRIBUTOR:

Cheplyukova, Irina

Cand. Sci. (Phys.-Math.), Senior Researcher

УДК 519.21

О СХОДИМОСТИ ЧИСЕЛ ЯЧЕЕК ЗАДАННОГО ОБЪЕМА К ЭМПИРИЧЕСКОМУ ПРОЦЕССУ

А. Н. Чупрунов*, К. Н. Яковлев

*Факультет прикладной математики, физики и информационных технологий,
Чувашский государственный университет (ул. Университетская, 38,
Чебоксары, Чувашская Республика, Россия, 428034), *achuprunov@mail.ru*

Пусть r, r_1 – целые неотрицательные числа, $r < r_1$, $\eta_1, \dots, \eta_N$ – обобщенная схема размещения $(r_1 - r)n + rN$ частиц по N ячейкам, определенная независимыми случайными величинами $\xi_1, \dots, \xi_N$, которые имеют распределение степенного ряда, определенное рядом $\sum_{k=0}^{\infty} \frac{b_k \beta^k}{k!}$, случайный процесс $X_{n,N}^{\{r_1\}}(t) = \sum_{i=1}^{[tN]} I_{\{\eta_i=r_1\}}$, $0 \leq t \leq 1$, F_n – эмпирический процесс с параметром n . Показано, что если n, r, r_1 – фиксированные числа и выполнено условие $Ar(r_1)$: $b_k = 0, k < r, b_r > 0, b_k = 0, r < k < r_1, b_{r_1} > 0$, то при $N \rightarrow \infty$ случайные процессы $X_{n,N}^{\{r_1\}}$ сходятся по распределению в пространстве Скорохода к случайному процессу nF_n .

Ключевые слова: обобщенная схема размещения; функциональная предельная теорема; пространство Скорохода; броуновский мост; эмпирический процесс

Для цитирования: Чупрунов А. Н., Яковлев К. Н. О сходимости чисел ячеек заданного объема к эмпирическому процессу // Труды Карельского научного центра РАН. 2026. № 6. С. 127–131. doi: 10.17076/mat2285

Финансирование. Финансовое обеспечение исследования осуществлялось из средств федерального бюджета на выполнение государственного задания Чувашского государственного университета.

A. N. Chuprunov*, K. N. Iakovlev. ON THE CONVERGENCE OF GIVEN VALUE CELLS NUMBERS TO AN EMPIRICAL PROCESS

*Faculty of Applied Mathematics, Physics and Information Technologies, Chuvash State University (38 Universitetskaya St., 428034 Cheboksary, Chuvash Republic, Russia), *achuprunov@mail.ru*

Let r, r_1 be integer nonnegative numbers such that $r < r_1$. Let $\eta_1, \dots, \eta_N$ be a general allocation scheme of $(r_1 - r)n + rN$ particles by N cells, defined by independent random variables $\xi_1, \dots, \xi_N$ which have power series distribution defined by the series $\sum_{k=0}^{\infty} \frac{b_k \beta^k}{k!}$. Denote: the random process $X_{n,N}^{\{r_1\}}(t) = \sum_{i=1}^{[tN]} I_{\{\eta_i=r_1\}}$, $0 \leq t \leq 1$, F_n is an empirical process with the parameter n . It proved that if n, r and r_1 are fixed numbers and the condition $Ar(r_1)$: $b_k = 0, k < r, b_r > 0, b_k = 0, r < k < r_1, b_{r_1} > 0$, is valid, then as $N \rightarrow \infty$ the random processes $X_{n,N}^{\{r_1\}}$ converge in distribution in Skorokhod space to nF_n .

Key words: general allocation scheme; functional limit theorem; Skorokhod space; Brounian bridge; empirical process

For citation: Chuprunov A. N., Iakovlev K. N. On the convergence of given value cells numbers to an empirical process. *Trudy Karelskogo nauchnogo tsentra RAN = Transactions of the Karelian Research Centre RAS*. 2026. No. 6. P. 127–131. doi: 10.17076/mat2285

Funding. The studies were funded from the federal budget through state assignment to the Chuvash State University.

ВВЕДЕНИЕ

Мы будем использовать понятие обобщенной схемы размещения, введенное В. Ф. Колчиным в [5]. Неотрицательные целочисленные случайные величины $\eta_1, \dots, \eta_N$ образуют обобщенную схему размещения n частиц по N ячейкам, если существуют такие независимые одинаково распределенные неотрицательные целочисленные случайные величины $\xi_1, \xi_2, \dots, \xi_N$, что их совместное распределение имеет вид

$$\begin{aligned} & \mathbf{P}\{\eta_1 = k_1, \dots, \eta_N = k_N\} \\ &= \mathbf{P}\left\{\xi_1 = k_1, \dots, \xi_N = k_N \mid \sum_{i=1}^N \xi_i = n\right\}, \end{aligned}$$

где $k_1, k_2, \dots, k_N$ – такие неотрицательные целые числа, что $k_1 + k_2 + \dots + k_N = n$.

Различные схемы дискретной теории вероятностей, такие как случайные леса, случайные перестановки, случайные размещения, урновые схемы, являются частными случаями обобщенной схемы размещения [6, 7].

Пространством Скорохода $D[0, 1]$ называется множество функций, определенных на отрезке $[0, 1]$, имеющих пределы слева и непрерывных справа, наделенное метрикой

$$\rho(x, y) = \inf_{\lambda} \max \left\{ \sup_{t \in [0, 1]} |x(\lambda(t)) - y(t)|, \sup_{t \in [0, 1]} |\lambda(t) - t| \right\}, \quad x, y \in D[0, 1].$$

Здесь инфимум берется по всем строго возрастающим непрерывным отображениям отрезка $[0, 1]$ на $[0, 1]$ (см. [1, 2]).

Будем использовать следующие обозначения: $\xrightarrow{D}$ – сходимость по распределению в пространстве Скорохода, W_0 – броуновский мост, $\mathbf{N}_0$ – множество неотрицательных целых чисел.

Пусть

$$B(\beta) = \sum_{k=0}^{\infty} \frac{b_k}{k!} \beta^k, \beta \in (0, R), \quad (1)$$

ряд с неотрицательными коэффициентами, хотя бы один из которых не равен нулю, имеющий радиус сходимости $R > 0$. Случайные величины $\xi_i = \xi_i(\beta)$, $1 \leq i \leq N$, имеют распределение степенного ряда, если их распределение имеет вид

$$p_k = p_k(\beta) = \mathbf{P}\{\xi_i = k\} = \frac{b_k \beta^k}{k! B(\beta)}, \quad k \in \mathbf{N}_0, \quad (2)$$

где число $\beta \in (0, R)$. В [3, 4] изучалась сходимость сумм $S_N = S_N(\beta) = \sum_{i=1}^N \xi_i(\beta)$, $N \in \mathbf{N}$.

Пусть r, r_1 – целые неотрицательные числа такие, что $r < r_1$. Для последовательности коэффициентов ряда (1) мы будем использовать свойство $Ar(r_1)$:

$$b_0 = 0, \dots, b_{r-1} = 0, \quad b_r > 0,$$

$$b_{r+1} = 0, \dots, b_{r_1-1} = 0, \quad b_{r_1} > 0.$$

Свойство $A0(r_1)$ введено В. Ф. Колчиным и А. В. Колчиным в [4].

Случайные величины $\xi_i = \xi_i(\beta)$, $1 \leq i \leq N$, имеющие распределение степенного ряда, не являются константами тогда и только тогда, когда для некоторых $r < r_1$ выполнено свойство $Ar(r_1)$. И, рассматривая случайные величины с распределением степенного ряда, мы будем предполагать, что для некоторых $r < r_1$ выполнено свойство $Ar(r_1)$.

Замечание. Пусть $\eta_1, \dots, \eta_N$ – обобщенная схема размещения n частиц по N ячейкам со случайными величинами $\xi_1, \xi_2, \dots, \xi_N$, имеющими распределение степенного ряда. Тогда

$$\begin{aligned} & \mathbf{P}\{\eta_1 = k_1, \dots, \eta_N = k_N\} \\ &= \frac{\prod_{i=1}^N \frac{b_{k_i}}{(k_i)!}}{\sum_{\{k'_i \in \mathbf{N}_0, 1 \leq i \leq N, k'_1 + k'_2 + \dots + k'_N = n\}} \prod_{i=1}^N \frac{b_{k'_i}}{(k'_i)!}}, \end{aligned}$$

где $k_1, k_2, \dots, k_N$ – такие неотрицательные целые числа, что $k_1 + k_2 + \dots + k_N = n$. Поэтому в действительности распределение схемы $\eta_1, \dots, \eta_N$ не зависит от β .

Случайный процесс

$$F_n(t) = \frac{1}{n} \sum_{i=1}^n I_{\{\zeta_i < t\}}, \quad 0 \leq t \leq 1,$$

где ζ_i , $1 \leq i \leq n$ – независимые случайные величины, имеющие равномерное распределение на интервале $[0, 1]$, называется эмпирическим процессом.

Пусть r – целое неотрицательное число. В работе мы будем изучать сходимость по распределению в пространстве Скорохода случайных процессов

$$X_{n,N}^{\{r\}}(t) = \sum_{i=1}^{[tN]} I_{\{\eta_i=r\}}, \quad 0 \leq t \leq 1,$$

к эмпирическому процессу, где обобщенная схема размещения $\eta_1, \dots, \eta_N$ определяется случайными величинами $\xi_1, \xi_2, \dots, \xi_N$, имеющими распределение степенного ряда. Заметим, что случайная величина $X_{n,N}^{\{r\}}(t)$ – число ячеек среди первых $[tN]$ ячеек, содержащих r частиц, в обобщенной схеме размещения n частиц по N ячейкам.

В работе [8] изучалась сходимость по распределению случайных процессов

$$X_{n,N}(t) = \sum_{i=1}^{[tN]} \eta_i, \quad 0 \leq t \leq 1.$$

В частности, в [8] доказана теорема.

Теорема А. Пусть $\eta_1, \dots, \eta_N$ – обобщенная схема размещения rn частиц по N ячейкам, где r – фиксированное натуральное число, со случайными величинами $\xi_1, \xi_2, \dots, \xi_N$, имеющими распределение степенного ряда со свойством $A0(r)$. Если n фиксировано, а $N \rightarrow \infty$, то

$$\frac{X_{n,N}}{r} \xrightarrow{D} nF_n. \quad (3)$$

В настоящей работе получены аналоги теоремы А для случайных процессов $X_{n,N}^{\{r\}}$. Их доказательства основаны на доказательстве близости в пространстве Скорохода распределений случайных процессов $rX_{n,N}^{\{r\}}$ и $X_{n,N}$, что дает возможность использовать теорему А.

В следующем примере приведен случай, когда случайные процессы $rX_{n,N}^{\{r\}}$ и $X_{n,N}$ совпадают.

Пример. Рассмотрим обобщенную схему размещения rn частиц по N ячейкам со случайными величинами $\xi_i = rI_{A_i}$, $1 \leq i \leq N$, где I_{A_i} – индикаторы независимых равновероятных событий A_i таких, что $0 < \mathbf{P}(A_i) < 1$. Случайные величины ξ_i имеют распределение

степенного ряда, определенное функцией $B(\beta) = b_0 + \frac{b_r}{r!} \beta^r$. Вероятность $\mathbf{P}\{\xi_1 + \xi_2 + \dots + \xi_N = rn\} = 0$ при $n > N$. Поэтому формула (1) не имеет смысла. При $n = N$ $\eta_i = r$, $1 \leq i \leq N$. Таким образом, этот случай также не представляет интереса. Пусть $n < N$. Тогда каждая ячейка содержит 0 или r частиц и $rX_{n,N}^{\{r\}}(t)$ – количество частиц, содержащихся в первых $[tN]$ ячейках. То есть случайные процессы $rX_{n,N}^{\{r\}}$ и $X_{n,N}$ совпадают. Поэтому, по теореме 7 из [8], если n фиксировано, а $N \rightarrow \infty$, то

$$X_{n,N}^{\{r\}} \xrightarrow{D} nF_n.$$

Будем считать β таким, что $\mathbf{E}\xi_i(\beta) = \frac{rn}{N}$, т. е. $rp_r(\beta) = \frac{rn}{N}$. Так как $\frac{n}{N} < 1$, это уравнение имеет решение, и оно единственно. Тогда дисперсия случайной величины $\xi_i(\beta)$ равна $\sigma^2(\beta) = r^2 \frac{n}{N} (1 - \frac{n}{N})$. Рассмотрим случайные процессы

$$Y_{n,N}^{\{r\}}(t) = \frac{X_{n,N}^{\{r\}}(t) - [tN] \frac{n}{N}}{\sqrt{n(1 - \frac{n}{N})}}, \quad 0 \leq t \leq 1,$$

и

$$Y_{n,N}(t) = \frac{X_{n,N}(t) - [tN] \mathbf{E}\xi_i(\beta)}{\sqrt{N}\sigma(\beta)}, \quad 0 \leq t \leq 1.$$

Тогда

$$Y_{n,N}^{\{r\}}(t) = \frac{rX_{n,N}^{\{r\}}(t) - [tN] \frac{rn}{N}}{\sqrt{Nr^2 \frac{n}{N} (1 - \frac{n}{N})}} = Y_{n,N}(t),$$

где $0 \leq t \leq 1$. То есть случайные процессы $Y_{n,N}^{\{r\}}$ и $Y_{n,N}$ совпадают. Поэтому, по теореме 4 из [8], если $n, N \rightarrow \infty$ так, что $n(1 - \frac{n}{N}) \rightarrow \infty$, то

$$Y_{n,N}^{\{r\}} \xrightarrow{D} W_0.$$

ОСНОВНЫЕ РЕЗУЛЬТАТЫ

Основным результатом работы является следующая теорема.

Теорема 1. Пусть случайные величины $\eta_1, \dots, \eta_N$ образуют обобщенную схему размещения rn частиц по N ячейкам, где r, n – фиксированные натуральные числа, со случайными величинами $\xi_1, \xi_2, \dots, \xi_N$, имеющими распределение степенного ряда со свойством $A0(r)$. Тогда

$$X_{n,N}^{\{r\}} \xrightarrow{D} nF_n \quad \text{при } N \rightarrow \infty. \quad (4)$$

Справедливо следующее обобщение теоремы 1.

Теорема 2. Пусть случайные величины $\eta_1, \dots, \eta_N$ образуют обобщенную схему размещения $(r_1 - r)n + rN$ частиц по N ячейкам, где n, r и r_1 – фиксированные натуральные числа, со случайными величинами $\xi_1, \xi_2, \dots, \xi_N$, имеющими распределение степенного ряда со свойством $Ar(r_1)$. Тогда

$$X_{n,N}^{\{r_1\}} \xrightarrow{D} nF_n \text{ при } N \rightarrow \infty. \quad (5)$$

Мы будем рассматривать случайные процессы

$$X_{n,N}^c(t) = X_{n,N}(t) - r[Nt], \quad 0 \leq t \leq 1.$$

Справедливо следующее обобщение теоремы А.

Теорема А'. Пусть случайные величины $\eta_1, \dots, \eta_N$ образуют обобщенную схему размещения $(r_1 - r)n + rN$ частиц по N ячейкам, где n, r и r_1 – фиксированные натуральные числа, со случайными величинами $\xi_1, \xi_2, \dots, \xi_N$, имеющими распределение степенного ряда со свойством $Ar(r_1)$. Тогда

$$\frac{X_{n,N}^c}{r_1 - r} \xrightarrow{D} nF_n \text{ при } N \rightarrow \infty. \quad (6)$$

Доказательства

При доказательстве теоремы 1 мы будем использовать следующую лемму.

Лемма. Пусть $0 < C < R$, $C < 1$, $0 < \lambda < \infty$, k – неотрицательное целое число. Пусть условие $A0(r)$ выполнено для некоторого $r \in \mathbf{N}$, $\beta = \beta(N)$ такое, что $Np_r(\beta) = \lambda$. Тогда при $N \rightarrow \infty$

$$\mathbf{P}\{S_N = rk\} = e^{-\lambda} \frac{\lambda^k}{k!} + Q_N,$$

где

$$|Q_N| \leq C_1 N \beta^{r+1}, \quad \beta < C, \\ C_1 = \left(\frac{b_r}{r!b_0} \right)^2 C^{r-1} + \frac{1}{b_0} \sum_{n=r+1}^{\infty} \frac{b_n C^{n-r+1}}{n!}.$$

Лемма является следствием леммы 2.4 из [9].

Доказательство теоремы 1. Пусть $\beta = \beta_N$ такое, что $Np_r(\beta) = 1$, $N \in \mathbf{N}$. Тогда по лемме получаем

$$\mathbf{P}\{\xi_1 + \xi_2 + \dots + \xi_N = rn\} = \frac{e^{-1}}{n!} (1 + o(1)).$$

Учитывая, что $\beta = O\left(\frac{1}{N^{1/r}}\right)$, из (2) заключаем, что

$$\mathbf{P}\{\xi_i > r\} = O\left(\frac{1}{N^{\frac{r+1}{r}}}\right).$$

Следовательно,

$$\begin{aligned} & \mathbf{P}\left\{\sup_{t \in [0,1]} |X_{n,N}(t) - rX_{n,N}^{\{r\}}(t)| > 0\right\} \\ &= \mathbf{P}\left\{\sup_{t \in [0,1]} \sum_{i=1}^{[tN]} \eta_i I_{\{\eta_i > r\}} > 0\right\} \\ &= \frac{\mathbf{P}\left\{\sup_{t \in [0,1]} \sum_{i=1}^{[tN]} \xi_i I_{\{\xi_i > r\}} > 0, \sum_{i=1}^N \xi_i = rn\right\}}{\mathbf{P}\{\sum_{i=1}^N \xi_i = rn\}} \\ &\leq (1 + o(1))en! \sum_{i=1}^N \mathbf{P}\{\xi_i > r\} \\ &= (1 + o(1))en!N\mathbf{P}\{\xi_i > r\} \\ &= (1 + o(1))en!NO\left(\frac{1}{N^{\frac{r+1}{r}}}\right) = o(1). \quad (7) \end{aligned}$$

Так как

$$\rho(x, y) \leq \sup_{t \in [0,1]} |x(t) - y(t)|, \quad x, y \in D[0, 1],$$

то

$$\begin{aligned} & \mathbf{P}\left\{\rho(X_{n,N}, rX_{n,N}^{\{r\}}) < \varepsilon\right\} \\ &> \mathbf{P}\left\{\sup_{t \in [0,1]} |X_{n,N}(t) - rX_{n,N}^{\{r\}}(t)| < \varepsilon\right\}. \end{aligned}$$

Отсюда и из (7) находим, что для достаточно малого $\varepsilon > 0$

$$\mathbf{P}\left\{\rho(X_{n,N}, rX_{n,N}^{\{r\}}) < \varepsilon\right\} \rightarrow 1 \text{ при } N \rightarrow \infty. \quad (8)$$

Из (8) и (3) следует (4). $\square$

Доказательство теоремы 2. Заметим, что распределения случайных величин $\xi'_i = \xi_i - r$, $1 \leq i \leq N$, удовлетворяют свойству $A0(r_1 - r)$. Рассмотрим $\eta'_1, \dots, \eta'_N$ – обобщенную схему размещения $(r_1 - r)n$ частиц по N ячейкам, определенную случайными величинами $\xi'_i = \xi_i - r$, $1 \leq i \leq N$. Обозначим случайные векторы:

$$\xi = (\xi_1, \dots, \xi_N), \quad \xi' = (\xi'_1, \dots, \xi'_N),$$

$$\mathbf{k} = (k_1, \dots, k_N), \quad \mathbf{r} = (r, \dots, r).$$

Имеем

$$\begin{aligned} & \mathbf{P}\{\eta'_1 = k_1, \dots, \eta'_N = k_N\} \\ &= \mathbf{P}\left\{\xi' = \mathbf{k} \mid \sum_{i=1}^N \xi'_i = (r_1 - r)n\right\} \\ &= \mathbf{P}\left\{\xi - \mathbf{r} = \mathbf{k} \mid \sum_{i=1}^N \xi_i = (r_1 - r)n + rN\right\} \\ &= \mathbf{P}\{\eta_1 - r = k_1, \dots, \eta_N - r = k_N\}, \end{aligned}$$

где $k_1, k_2, \dots, k_N$ – такие неотрицательные целые числа, что $k_1 + k_2 + \dots + k_N = (r_1 - r)n$. Итак, распределения случайных векторов

$$(\eta'_1, \dots, \eta'_N), \quad (\eta_1 - r, \dots, \eta_N - r)$$

совпадают. Следовательно, совпадают распределения случайных процессов $X_{n,N}^{\{r_1\}}$ и $X_{n,N}^{\{rr_1\}} = \sum_{i=1}^{[Nt]} I_{\{\eta'_i=r_1-r\}}, t \in [0, 1]$. По теореме 1

$$X_{n,N}^{\{rr_1\}} \xrightarrow{D} nF_n \text{ при } N \rightarrow \infty.$$

Поэтому справедливо (5). $\square$

Доказательство теоремы А'. Из совпадения распределений случайных векторов $(\eta'_1, \dots, \eta'_N), (\eta_1 - r, \dots, \eta_N - r)$ следует, что распределения случайных процессов $X_{n,N}^c$ и $X'_{n,N} = \sum_{i=1}^{[Nt]} \eta'_i, t \in [0, 1]$, совпадают. По теореме А

$$\frac{X'_{n,N}}{r_1 - r} \xrightarrow{D} nF_n \text{ при } N \rightarrow \infty.$$

Поэтому справедливо (6). $\square$

ЛИТЕРАТУРА

1. *Афанасьев В. И.* Случайные блуждания и ветвящиеся процессы. М.: МИАН, 2007. 188 с.
2. *Биллингсли П.* Сходимость вероятностных мер. М.: Физматлит, 1977. 352 с.
3. *Колчин А. В.* Предельные теоремы для обобщенной схемы размещения // Дискретная математика. 2003. Т. 18, вып. 4. С. 148–157. doi: 10.4213/dm224
4. *Колчин А. В., Колчин В. Ф.* О переходе распределений сумм независимых одинаково распределенных случайных величин с одной решетки на другую в обобщенной схеме размещения // Дискретная математика. 2006. Т. 18, вып. 4. С. 113–127. doi: 10.4213/dm76
5. *Колчин В. Ф.* Один класс предельных теорем для условных распределений // Литовский математический сборник. 1968. Т. 8, вып. 1. С. 53–63. doi: 10.15388/LMJ.1968.20181
6. *Колчин В. Ф.* Случайные графы. М.: Физматлит, 2000. 256 с.

7. *Павлов Ю. Л.* Случайные леса. Петрозаводск: Карел. науч. центр РАН, 1996. 259 с.

8. *Чупрунов А. Н., Фазекаш И.* Принцип инвариантности для чисел частиц в ячейках обобщенной схемы размещения // Дискретная математика. 2023. Т. 35, вып. 3. С. 81–99. doi: 10.4213/dm1738

9. *Chuprunov A. N., Fazekas I.* Poisson limit theorems for the generalized allocation scheme // Ann. Univ. Sci. Budapest, Sect. Comp. 2019. Vol. 49. P. 77–96.

REFERENCES

1. *Afanasyev V. I.* Random walks and branching processes. Moscow: MIAN; 2007. 158 p. (In Russ.)
2. *Billingsley P.* Convergence of probability measures. New York: John Wiley & Sons; 1968. 253 p. doi: 10.1002/9780470316962
3. *Kolchin A. V.* On limit theorems for the generalised allocation scheme. *Discrete Math. Appl.* 2003;13(6):627–636. doi: 10.1515/156939203322733336
4. *Kolchin A. V., Kolchin V. F.* On transition of distributions of sums of independent identically distributed random variables from one lattice to another in the generalised allocation scheme. *Discrete Math. Appl.* 2006;16(6):527–540. doi: 10.1515/156939206779218023
5. *Kolchin V. F.* A certain class of limit theorems for conditional distributions. *Litovskii matematicheskii sbornik = Lithuanian Mathematical Journal.* 1968;8(1):53–63. (In Russ.). doi: 10.15388/LMJ.1968.20181
6. *Kolchin V. F.* Random Graphs. Cambridge: Cambridge University Press; 1999. 252 p. doi: 10.1017/CBO9780511721342
7. *Pavlov Yu. L.* Random Forests. Utrecht: VSP; 2000. 122 p. doi: 10.1515/9783110941975
8. *Chuprunov A. N., Fazekas I.* Invariance principle for numbers of particles in cells of a general allocation scheme. *Discrete Math. Appl.* 2023;35(3):81–99. (In Russ.). doi: 10.4213/dm1738
9. *Chuprunov A. N., Fazekas I.* Poisson limit theorems for the generalized allocation scheme. *Ann. Univ. Sci. Budapest, Sect. Comp.* 2019;49:77–96.

Поступила в редакцию / received: 16.01.2026; принята к публикации / accepted: 27.04.2026.

Авторы заявляют об отсутствии конфликта интересов / The authors declare no conflict of interest.

СВЕДЕНИЯ ОБ АВТОРАХ:

Чупрунов Алексей Николаевич

д-р физ.-мат. наук, профессор
e-mail: achuprunov@mail.ru

Яковлев Кирилл Николаевич

студент
e-mail: ktq0x@yandex.ru

CONTRIBUTORS:

Chuprunov, Aleksey

Dr. Sci. (Phys.-Math.), Professor

Iakovlev, Kirill

Student

УДК 519.115:519.2

ВЕРОЯТНОСТНЫЕ МОДЕЛИ РАЗМЕЩЕНИЯ ПО ЯЧЕЙКАМ СЛУЧАЙНОГО ЧИСЛА ЧАСТИЦ

Н. Ю. Энатская

*Московский институт электроники и математики,
Национальный исследовательский университет «Высшая школа экономики»
(ул. Таллинская, 34, Москва, Россия, 123458)*

В схемах S размещения r частиц по n различимым ячейкам изучаются их размещения в выделенных m ячейках – схем S^* , для которых проводится анализ новым перечислительным методом (ПМ) по расширенным направлениям перечислительной комбинаторики: нахождения числа исходов и на основе построения модели их неповторного нумерованного перечисления – решения для них задачи нумерации в прямой и обратной постановках по установлению взаимно-однозначного соответствия между их видами и номерами и введения управляемого вероятностного распределения на множестве исходов схемы.

Ключевые слова: выделенные ячейки; случайное число частиц

Для цитирования: Энатская Н. Ю. Вероятностные модели размещения по ячейкам случайного числа частиц // Труды Карельского научного центра РАН. 2026. № 6. С. 132–138. doi: 10.17076/mat2271

N. Yu. Enatskaya. PROBABILISTIC RANDOM NUMBER OF PARTICLES CELL PLACEMENT MODELS

*National Research University Higher School of Economics, Moscow Institute of
Electronics and Mathematics (34 Tallinskaya St., 123458 Moscow, Russia)*

In schemes S for placing r particles in n distinguishable cells, their placement in selected m cells – schemes S^* is studied, for which the analysis is carried out by a new enumerative method (EM) in extended areas of enumerative combinatorics: finding the number of outcomes and, based on the construction of a model of their non-repetitive numbered enumeration, solving the numbering problem for them in direct and inverse statements to establish a one-to-one correspondence between their types and numbers, and introducing a controlled probabilistic distribution on the set of outcomes of the scheme.

Key words: selected cells; random number of particles

For citation: Enatskaya N. Yu. Probabilistic random number of particles cell placement models. *Trudy Karel'skogo nauchnogo tsentra RAN = Transactions of the Karelian Research Centre RAS*. 2026. No. 6. P. 132–138. doi: 10.17076/mat2271

ВВЕДЕНИЕ

В [2] проведен асимптотический анализ классической схемы равновероятного размещения различных частиц по различным ячейкам. В [1] проводится асимптотический анализ поведения числа ячеек с фиксированными количествами частиц при размещении заданного или случайного чисел различных частиц по ячейкам.

Здесь в доасимптотической области значений параметров строятся вероятностные модели с другой интерпретацией случайного числа частиц при разных их качествах по различимости и условиям их размещения. Рассматривается схема S^* заполнения m выделенных ячеек среди n различных ячеек начальной схемы S размещения по ним r различных частиц, все результаты анализа для которой считаем известными [3].

При варьировании качества частиц по их различимости и введении ограничения на уровни заполнения ячеек в начальных схемах S наряду с непосредственным анализом с учетом специфики схем S^* предлагается общий алгоритмический прием пересчета результатов анализа схем S в соответствующих схемах S^* для перечисления исходов схем, решения задачи нумерации и нахождения вероятностных распределений их исходов. В частности, в условиях равновероятных исходов каждой начальной схемы S определяется вероятностное распределение исходов в соответствующих схемах S^* с просчетом на числовых примерах. Результаты проведенного анализа представлены алгоритмическими процедурами и аналитическими формулами.

Перечислим схемы S^* с указанием их различий и видов их исходов в условиях размещения по различным ячейкам в схемах $A-D$:

A – схема сочетаний – частицы неразличимы и ячейка вмещает одну частицу, $r \leq n$; вид исхода – набор номеров непустых ячеек в порядке ячеек;

B – схема сочетаний с повторением – частицы неразличимы и ячейка вмещает все частицы; вид исхода – уровни заполнения ячеек в порядке ячеек;

C – схема размещений – частицы различимы и ячейка вмещает одну частицу, $r \leq n$; вид исхода – набор номеров частиц в ячейках в порядке ячеек;

D – схема размещений с повторением – частицы различимы и ячейка вмещает все частицы; вид исхода – наборы номеров частиц в ячейках в порядке ячеек.

Схема S^* в условиях размещения схемы A была рассмотрена в [3, п. 2.23], а здесь приво-

дится в связи с изучением остальных ранее не рассматриваемых по аналогичным направлениям схем размещения частиц по части ячеек и для полноты изложения, т. е. для логической завершенности рассмотрения данного класса схем и установления аналогии их анализа.

Цель статьи – в дополнении к анализу схемы S^* в условиях размещения частиц схемы A доказательства теорем 2–4 для остальных схем в условиях $B-D$ класса S^* о численностях и вероятностях их исходов.

Каждый исход нашей изучаемой схемы S^* соответствует группе исходов схемы S с совпадающими с ее исходом частями исходов в первых m ячейках. Будем называть общий алгоритмический прием пересчета результатов анализа схемы S для схемы S^* методом группированных исходов (МГИ), рассмотрением которого предварим получение конкретных результатов по всем приведенным условиям схем $A-D$ размещения частиц с учетом их специфики.

ПМ – перечислительный метод – основной используемый здесь метод доасимптотического исследования моделей по приведенным в аннотации направлениям (подробно изложен в [3]); он состоит в построении вероятностных моделей на формирующих исходы их итерационных случайных процессов.

Основная конструкция комбинаторных схем – это обобщенная схема последовательных действий (ОПД); исследована в [3, п. 1.9] и предлагается здесь для перечисления исходов и решения задачи нумерации (ЗН). При использовании результатов других ранее изученных схем ограничимся ссылками на источники с доказанными результатами: схема сочетаний и сводящаяся к ней схема сочетаний с повторением [3, пп. 1.3 и 1.1.1.5]; схема размещений [3, п. 1.4]; схема размещений с повторением [3, пп. 1.1 и 2.16].

Постановка задачи. Провести доасимптотический анализ схем $B-D$ по ПМ в дополнение к анализу ранее изученной [3, п. 2.23] схемы A класса схем S^* по установлению численностей и вероятностей их исходов.

Актуальность исследований комбинаторных схем размещения частиц по ячейкам с различными ограничениями связана как с их широкими приложениями, так и с перспективным конструированием анализа новых все усложняющихся, включая изучаемые здесь, комбинаторных схем в нейросетях. В нашей постановке задачи анализ схем $A-D$ дает получение результатов ПМ для размещения частиц в фиксированной части ячеек по решению ана-

логичных задач для их размещения по всем ячейкам.

1. МГИ – ОБЩИЙ ПРИЕМ ПЕРЕСЧЕТА РЕЗУЛЬТАТОВ АНАЛИЗА СХЕМЫ S ДЛЯ СХЕМЫ S^*

МГИ – метод группированных данных – метод пересчета результатов анализа начальной схемы для схемы с исходами, соответствующими ее группированным исходам, является новым и рассматривается здесь.

Считаем схему S изученной по всем направлениям ПМ.

1.1. Перечисление и число исходов схемы S^*

Для перечисления исходов схемы S^* выпишем части (результаты заполнения первых t ячеек) исходов схемы S в порядке их перечисления с отбраковкой повторов. Тогда N_{S^*} – число исходов схемы S^* находится как число групп исходов схемы S с разными заполнениями первых t ячеек. Нумерацию исходов схемы S^* производим в порядке их появления.

1.2. Задача нумерации в схеме S^*

Решение ЗН состоит из решения двух задач: **ПЗН** – прямой задачи нумерации нахождения вида исхода схемы по его номеру и **ОЗН** – обратной задачи нумерации нахождения номера исхода схемы по его виду [3, п. 1.1].

Среди исходов схемы S в порядке их нумерации зафиксируем номера первых исходов групп совпадающих заполнений первых t ячеек, т. е. номера всех разных исходов схемы S^* . Пусть это будут номера $\bar{A} = (A_1, \dots, A_{N_{S^*}})$.

Порядок решения ПЗН

Пусть дан номер N^* исхода схемы S^* . Требуется найти его вид R^* .

1) Среди $\bar{A} = (A_1, \dots, A_{N_{S^*}})$ находим соответствующий данному номеру N^* номер исхода схемы S – это номер A_{N^*} ;

2) по решенной ПЗН в схеме S по номеру исхода A_{N^*} определяем его вид R ;

3) из R выписываем по результату заполнения первых t ячеек соответствующий вид искомого исхода схемы S^* .

Порядок решения ОЗН

Пусть дан вид R^* исхода схемы S^* . Требуется найти его номер N^* .

1) По решенной ОЗН в схеме S по последовательным номерам ее исходов $\bar{A}$ определяем их виды;

2) исход схемы S^* данного вида R^* последовательно сравниваем с видами исходов шага 1) по заполнению первых t ячеек схемы S с

номера $\bar{A} = (A_1, \dots, A_{N_{S^*}})$ до совпадения – тогда порядковый номер первого такого совпадения есть искомым номер N^* исхода схемы S^* .

2. ПРЯМОЙ АНАЛИЗ МОДЕЛЕЙ СХЕМ S^* В УСЛОВИЯХ РАЗМЕЩЕНИЯ СХЕМ $A-D$

Число исходов каждой из схем S^* в условиях размещения в схемах $A-D$ определяется значениями случайных чисел частиц схем S в первых t ячейках и вычисляется числом групп исходов схемы S с совпадающими заполнениями этих t ячеек суммированием по всем возможным значениям количеств частиц в первых t ячейках всех способов их размещения в условиях каждой конкретной схемы S .

Для анализа схем проводятся изучения конструкции графов перечисления их исходов. Они описываются в терминах их пучковых структур [3, п. 1.1], где

итерация – шаг перехода в графе перечисления исходов схемы к следующему этапу (шагу) их перечисления;

пучок в графе процесса – совокупность переходов из каждого исхода каждой итерации;

размер пучка на каждой итерации – число переходов из каждого исхода итерации;

пучковая структура итерации – перечисление размеров пучков итерации;

пучковая структура графа – перечисление размеров пучков всех итераций.

Процедуры перечисления исходов схем S^* в каждой из схем $A-D$ представлены двумя итерациями (действиями): 1) фиксацией варьируемого допустимого значения $k = \bar{t}, \bar{T}$ числа частиц в первых t ячейках в условиях схем $A-D$ и 2) перечислением $N_{S^*}(k)$ исходов с k частицами среди выделенных t ячеек, т. е. с заменой параметров схем S – чисел частиц r и ячеек n на зафиксированное число k частиц и t ячеек схем S^* . Таким образом, пучковая структура графа перечисления исходов на первой итерации есть размер одного пучка значений $k = \bar{t}, \bar{T}$, равный $\bar{d}^{(0)} = (T - t + 1)$, а на второй итерации – $(T - t + 1)$ пучков размерами $\bar{d}^{(1)}$ по $N_{S^*}(k)$, зависящими от k , откуда следует, что эта схема представляет собой схему ОПД с известной из [3, п. 1.9] по данной пучковой структуре процедурой перечисления исходов и решения ЗН.

ЗН для исходов схем S^* в каждой из схем $A-D$ решается по схеме ОПД [3, п. 1.9], для чего требуется информация о пучковой структуре графа перечисления исходов схемы в соответствии с описанными итерациями их перечисления.

Вероятностные распределения исходов схем S^* определяются условиями размещения частиц в схемах $A-D$ и будут изучаться отдельно.

Приведем анализ схем S^* в условиях схем $A-D$ при равновероятных исходах начальных схем S (в отличие от их любых распределений в алгоритмическом подходе к анализу схем S^* по МГИ в п. 1).

2.1. Анализ схемы S^* в условиях схемы $S = A$

Число исходов схемы $A : N_A = C_n^r$. Пусть X – случайное число частиц в выделенных m ячейках. При равновероятном распределении исходов начальной схемы A – схемы сочетаний вероятность каждого исхода $p_A = 1/C_n^r$.

Теорема 1. В схеме S^* в условиях схемы $S = A$ число исходов

$$N_{S^*} = \sum_{k=t}^T N_{S^*}(k) = \sum_{k=t}^T C_m^k, \quad (1)$$

где $N_{S^*}(k)$ – число исходов схемы S^* при фиксированном значении k , $t = \max(0, r - n + m) \leq k \leq \min(r, m) = T$, с вероятностью исходов при каждом k

$$p(k) = C_{n-m}^{r-k} / C_n^r. \quad (2)$$

Доказательство. В условиях схемы $S = A$ число ее исходов $N_A = C_n^r$ может быть представлено в виде суммы чисел $N_A(k)$ вариантов размещения частиц со всеми возможными значениями их количеств $X = k$ среди первых m ячеек: $N_A = C_n^r = \sum_{k=t}^T N_A(k) = \sum_{k=t}^T C_m^k C_{n-m}^{r-k}$, где допустимые из значений $X = k$ определяются из условий $t = \max(0, r - n + m) \leq k \leq \min(r, m) = T$, каждое слагаемое $N_A(k) = C_m^k C_{n-m}^{r-k}$ – число исходов схемы A при $X = k$, а исходы схемы $S = A$ при каждом значении $X = k$ представляют собой группы ее исходов, не меняющих исходы схемы S^* , численностями C_{n-m}^{r-k} . Отсюда получаем для числа N_{S^*} таких групп формулу (1).

В схеме A разные исходы определяются разными наборами непустых ячеек среди n ячеек в их порядке, а в схеме S^* – среди первых m ячеек. Тогда в схеме S^* вероятности ее исходов будут получаться сложением вероятностей исходов соответствующих групп размерами по C_{n-m}^{r-k} равновероятных исходов схемы A , что приводит к результату (2).

Пучковая структура графа перечисления исходов схемы S^* в виде последовательностей размеров пучков i -й итерации ($i = \overline{0, 1}$) при

$i = 0$ это $\bar{d}^{(0)} = (T - t + 1)$, при $i = 1$ это $\bar{d}^{(1)} = (C_m^t, \dots, C_m^T)$.

Проверка на распределение вероятностей исходов схем A и S^* дает результаты – для схемы $A : C_n^r p_A = \sum_{k=t}^T C_m^k C_{n-m}^{r-k} p_A = 1$, а для схемы $S^* : \sum_{k=t}^T C_m^k p(k) = \sum_{k=t}^T C_m^k (C_{n-m}^{r-k} / C_n^r) = 1$. $\square$

Пример 1. Пусть $n = 10$, $m = 5$, $r = 7$. Тогда $2 \leq k \leq 5$, $N_A = C_{10}^7 = 120$ или через $N_A(k)$ при $N_A(k) = C_5^k C_5^{7-k}$ получаем $N_A = C_5^2 C_5^5 + C_5^3 C_5^4 + C_5^4 C_5^3 + C_5^5 C_5^2 = 10 + 50 + 50 + 10 = 120$, $N_{S^*} = C_5^2 + C_5^3 + C_5^4 + C_5^5 = 26$; $\bar{d}^{(0)} = (4)$, $\bar{d}^{(1)} = (C_5^2, C_5^3, C_5^4, C_5^5) = (10, 10, 5, 1)$;

распределение вероятностей в схеме A :

$$\bar{p} = (1/120, \dots, 1/120);$$

распределение вероятностей по (2) $\{p_j(k)\}$ в схеме S^* с числами ее исходов совпадающих вероятностей по C_m^k , т. е. соответственно:

$$C_5^2 = 10, C_5^3 = 10, C_5^4 = 5, C_5^5 = 1,$$

и по (2):

$$p_1(2) = \dots = p_{10}(2) = C_5^5 / C_{10}^7 = 1/120;$$

$$p_{11}(3) = \dots = p_{20}(3) = C_5^4 / C_{10}^7 = 5/120;$$

$$p_{21}(4) = \dots = p_{25}(4) = C_5^3 / C_{10}^7 = 10/120;$$

$$p_{26}(5) = C_5^2 / C_{10}^7 = 1/120,$$

откуда проверка на распределение дает верный результат

$$(1/120) \cdot 10 + (5/120) \cdot 10 + (10/120) \cdot 5 + (1/120) \cdot 1 = 1.$$

Замечание. Элементы приведенного здесь доказательства теоремы 1 содержатся в [3, пп. 2.23.2 и 2.23.3], но теорема 1 в данном здесь виде не формулировалась.

2.2. Анализ схемы S^* в условиях схемы $S = B$

Число исходов схемы $B : N_B = C_{n+r-1}^{m-1}$. Пусть X – случайное число частиц в выделенных m ячейках. При равновероятном распределении исходов начальной схемы B – схемы сочетаний вероятность каждого исхода $p_B = 1/C_{n+r-1}^{m-1}$.

Теорема 2. В схеме S^* в условиях схемы $S = B$ число исходов

$$N_{S^*} = \sum_{k=0}^r N_{S^*}(k) = \sum_{k=0}^r C_{m+k-1}^{m-1}, \quad (3)$$

где $N_{S^*}(k)$ – число исходов схемы S^* при фиксированном значении k , $0 \leq k \leq r$, с вероятностью исходов при каждом k

$$p(k) = C_{n-m+r-k-1}^{m-m-1} / C_{n+r-1}^{m-1}. \quad (4)$$

Доказательство. В условиях схемы $S = B$ число ее исходов $N_B = C_{n+r-1}^{m-1}$ может быть

представлено в виде суммы чисел $N_B(k)$ вариантов размещения частиц со всеми возможными значениями их количеств $X = k$ среди первых m ячеек: $N_B = C_{n+r-1}^{m-1} = \sum_{k=0}^r N_B(k) = \sum_{k=0}^r C_{m+k-1}^{m-1} C_{n-m+r-k-1}^{m-1}$, где все указанные в сумме значения $X = k$ допустимы, каждое слагаемое $N_B(k) = C_{m+k-1}^{m-1}$ – число исходов схемы B при $X = k$, а исходы схемы $S = B$ при каждом значении $X = k$ представляют собой группы ее исходов, не меняющих исходы схемы S^* , численностями $C_{n-m+r-k-1}^{m-1}$. Отсюда получаем для числа N_{S^*} таких групп формулу (3).

В схеме B разные исходы определяются разными наборами уровней заполнения ячеек среди n ячеек в их порядке, а в схеме S^* – среди первых m ячеек. Тогда в схеме S^* вероятности ее исходов будут получаться сложением вероятностей исходов соответствующих групп размерами по $C_{n-m+r-k-1}^{m-1}$ равновероятных исходов схемы B , что приводит к результату (4).

Пучковая структура графа перечисления исходов схемы S^* в виде последовательностей размеров пучков i -й итерации ($i = \overline{0,1}$) при $i = 0$ это $\bar{d}^{(0)} = (r+1)$, при $i = 1$ это $\bar{d}^{(1)} = (C_{m-1}^{m-1}, \dots, C_{m+r-1}^{m-1})$.

Проверка на распределение вероятностей исходов схем B и S^* дает результаты – для схемы B : $C_{n+r-1}^{m-1} p_B = \sum_{k=0}^r C_{m+k-1}^{m-1} C_{n-m+r-k-1}^{m-1} p_B = 1$, а для схемы S^* : $\sum_{k=0}^r C_{m+k-1}^{m-1} p(k) = \sum_{k=0}^r C_{m+k-1}^{m-1} C_{n-m+r-k-1}^{m-1} / C_{n+r-1}^{m-1} = 1$. $\square$

Пример 2. Пусть $n = 5$, $m = 2$, $r = 3$. Тогда $0 \leq k \leq 3$, $N_B = C_7^4 = 35$ или через $N_B(k)$ при $N_B(k) = C_{1+k}^1 C_{5-k}^2$ получаем $N_B = C_1^1 C_5^2 + C_2^1 C_4^2 + C_3^1 C_3^2 + C_4^1 C_2^2 = 35$; $N_{S^*} = C_{2+0-1}^{2-1} + C_{2+1-1}^{2-1} + C_{2+2-1}^{2-1} + C_{2+3-1}^{2-1} = 1 + 2 + 3 + 4 = 10$, $\bar{d}^{(0)} = (r+1) = (4)$, $\bar{d}^{(1)} = (C_1^1, C_2^1, C_3^1, C_4^1) = (1, 2, 3, 4)$; распределение вероятностей в схеме B : $\bar{p} = (1/35, \dots, 1/35)$;

распределение вероятностей по (4) $\{p_j(k)\}$ в схеме S^* с числами ее исходов совпадающих вероятностей по C_{m+k-1}^{m-1} , т. е. соответственно: $C_1^1 = 1, C_2^1 = 2, C_3^1 = 3, C_4^1 = 4$, и по (4):

$$\begin{aligned} p_1(0) &= C_5^2/35 = 10/35; \\ p_2(1) &= p_3(1) = C_4^2/35 = 6/35; \\ p_4(2) &= p_5(2) = p_6(2) = C_3^2/35 = 3/35; \\ p_7(3) &= p_8(3) = p_9(3) = p_{10}(3) = 4/35, \end{aligned}$$

откуда проверка на распределение дает верный результат

$$(10/35) \cdot 1 + (6/35) \cdot 2 + (3/35) \cdot 3 + (1/35) \cdot 4 = 1.$$

2.3. Анализ схемы S^* в условиях схемы $S = C$

Число исходов схемы $C : N_C = A_n^r$. Пусть X – случайное число частиц в выделенных m ячейках. При равновероятном распределении исходов начальной схемы C – схемы размещений вероятность каждого исхода $p_C = 1/A_n^r$.

Теорема 3. В схеме S^* в условиях схемы $S = C$ число исходов

$$N_{S^*} = \sum_{k=t}^T N_{S^*}(k) = \sum_{k=t}^T A_m^k C_r^k, \quad (5)$$

где $N_{S^*}(k)$ – число исходов схемы S^* при фиксированном значении k , $t = \max(0, r - n + m) \leq k \leq \min(r, m) = T$, с вероятностью исходов при каждом k

$$p(k) = A_{n-m}^{r-k} / A_n^r, \quad (6)$$

где $A_n^r = n(n-1) \cdots (n-r+1)$ – число исходов схемы размещений из n элементов по r .

Доказательство. В условиях схемы $S = C$ число ее исходов $N_C = A_n^r$ может быть представлено в виде суммы чисел $N_C(k)$ вариантов размещения частиц со всеми возможными значениями их количеств $X = k$ среди первых m ячеек: $N_C = \sum_{k=t}^T N_C(k) = \sum_{k=t}^T C_r^k A_m^k A_{n-m}^{r-k}$, где допустимые из значений $X = k$ определяются из условий $t = \max(0, r - n + m) \leq k \leq \min(r, m) = T$, каждое слагаемое $N_C(k) = C_r^k A_m^k A_{n-m}^{r-k}$ – число исходов схемы C при $X = k$, а исходы схемы $S = C$ при каждом значении $X = k$ представляют собой группы ее исходов, не меняющих исходы схемы S^* , численностями A_{n-m}^{r-k} . Отсюда получаем для числа N_{S^*} таких групп формулу (5).

В схеме C разные исходы определяются разными номерами частиц в непустых ячейках среди n ячеек в их порядке, а в схеме S^* – среди первых m ячеек. Тогда в схеме S^* вероятности ее исходов будут получаться сложением вероятностей исходов соответствующих групп размерами по A_{n-m}^{r-k} равновероятных исходов схемы C , что приводит к результату (6).

Пучковая структура графа перечисления исходов схемы S^* в виде последовательностей размеров пучков i -й итерации ($i = \overline{0,1}$) при $i = 0$ это $\bar{d}^{(0)} = (T - t + 1)$, при $i = 1$ это $\bar{d}^{(1)} = (A_m^t C_r^t, \dots, A_m^T C_r^T)$.

Проверка на распределение вероятностей исходов схем C и S^* дает результаты – для схемы C : $A_n^r p_C = \sum_{k=t}^T C_r^k A_m^k A_{n-m}^{r-k} p_C = 1$,

а для схемы S^* : $\sum_{k=t}^T C_r^k A_m^k p(k) = \sum_{k=t}^T C_r^k A_m^k A_{n-m}^{r-k} / A_n^r = A_n^r p = 1$. $\square$

Пример 3. Пусть $n = 5$, $m = 2$, $r = 3$. Тогда $0 \leq k \leq 2$, $N_C = A_5^3 = 60$, или через $N_C(k) = C_3^k A_2^k A_3^{3-k}$ получаем $N_C = C_3^0 A_2^0 A_3^3 + C_3^1 A_2^1 A_3^2 + C_3^2 A_2^2 A_3^1 = 6 + 36 + 18 = 60$, $N_{S^*} = A_2^0 C_3^0 + A_2^1 C_3^1 + A_2^2 C_3^2 = 1 + 6 + 6 = 13$, $\bar{d}^{(0)} = (r + 1) = (3)$, $\bar{d}^{(1)} = (A_3^0 C_2^0, A_3^1 C_2^1, A_3^2 C_2^2) = (1, 6, 6)$; распределение вероятностей в схеме $C : \bar{p} = (1/60, \dots, 1/60)$; распределение вероятностей по (6) $\{p_j(k)\}$ в схеме S^* с числами ее исходов совпадающих вероятностей по $A_m^k C_r^k$, т. е. соответственно: $A_3^0 C_2^0 = 1$, $A_3^1 C_2^1 = 6$, $A_3^2 C_2^2 = 6$, и по (6): $p_1(0) = A_3^3/60 = 6/60$; $p_2(1) = \dots = p_7(1) = A_3^2/60 = 6/60$; $p_8(2) = \dots = p_{13}(2) = A_3^1/60 = 3/60$, откуда проверка на распределение дает верный результат $(6/60) \cdot 1 + (6/60) \cdot 6 + (3/60) \cdot 6 = 1$.

2.4. Анализ схемы S^* в условиях схемы $S = D$

Число исходов схемы $D : N_D = n^r$. Пусть X – случайное число частиц в выделенных m ячейках. При равновероятном распределении исходов начальной схемы D – схемы размещений с повторением вероятность каждого исхода $p_D = 1/n^r$.

Теорема 4. В схеме S^* в условиях схемы $S = D$ число исходов

$$N_{S^*} = \sum_{k=0}^r N_{S^*}(k) = \sum_{k=0}^r C_r^k m^k, \quad (7)$$

где $N_{S^*}(k)$ – число исходов схемы S^* при фиксированном значении k , $0 \leq k \leq r$, с вероятностью исходов при каждом k

$$p_j(k) = (n - m)^{r-k} / n^r, \quad j = \overline{1, N_{S^*}}. \quad (8)$$

Доказательство. В условиях схемы $S = D$ число ее исходов $N_D = n^r$ может быть представлено в виде суммы чисел $N_D(k)$ вариантов размещения частиц со всеми возможными значениями их количеств $X = k$ среди первых m ячеек: $N_D = n^r = \sum_{k=0}^r N_D(k) = \sum_{k=0}^r C_r^k m^k (n - m)^{r-k}$, где все указанные в сумме значения $X = k$ допустимы, каждое слагаемое $N_D(k) = C_r^k m^k (n - m)^{r-k}$ – число исходов схемы D при $X = k$, а исходы схемы $S = D$ при каждом значении $X = k$ представляют собой группы ее исходов, не меняющих исходы схемы S^* , численностями $(n - m)^{r-k}$. Отсюда получаем для числа N_{S^*} таких групп формулу (7).

В схеме D разные исходы определяются разными составами номеров частиц в ячейках среди n ячеек в их порядке, а в схеме S^* – среди первых m ячеек. Тогда в схеме S^* вероятности ее исходов будут получаться сложением вероятностей исходов соответствующих групп размерами по $(n - m)^{r-k}$ равновероятных исходов схемы D , что приводит к результату (8).

Пучковая структура графа перечисления исходов схемы S^* в виде последовательностей размеров пучков i -й итерации ($i = \overline{0, 1}$) при $i = 0$ это $\bar{d}^{(0)} = (r + 1)$, при $i = 1$ это $\bar{d}^{(1)} = (C_r^0 m^0, \dots, C_r^r m^r)$.

Проверка на распределение вероятностей исходов схем D и S^* дает результаты – для схемы D : $n^r p_D = \sum_{k=0}^r C_r^k m^k (n - m)^{r-k}$, а для схемы S^* : $\sum_{k=0}^r C_r^k m^k p(k) = \sum_{k=0}^r C_r^k m^k (n - m)^{r-k} / n^r = 1$. $\square$

Пример 4. Пусть $n = 4$, $m = 2$, $r = 3$. Тогда $0 \leq k \leq 3$, $N_D = 4^3 = 64$ или через $N_D(k)$ при $N_D(k) = C_3^k 2^k 2^{3-k}$ получаем $N_D = C_3^0 2^0 2^3 + C_3^1 2^1 2^2 + C_3^2 2^2 2^1 + C_3^3 2^3 2^0 = 64$, $\bar{d}^{(0)} = (4)$, $\bar{d}^{(1)} = (C_3^0 2^0, C_3^1 2^1, C_3^2 2^2, C_3^3 2^3) = (1, 6, 12, 8)$, $N_{S^*} = C_3^0 2^0 + C_3^1 2^1 + C_3^2 2^2 + C_3^3 2^3 = 1 + 6 + 12 + 8 = 27$; распределение вероятностей исходов в схеме $D : \bar{p} = (1/64, \dots, 1/64)$; распределение вероятностей по (8) $\{p_j(k)\}$ в схеме S^* , с числами ее исходов совпадающих вероятностей по $C_r^k m^k$, т. е. соответственно равными $C_3^0 2^0 = 1$, $C_3^1 2^1 = 6$, $C_3^2 2^2 = 12$, $C_3^3 2^3 = 8$, и по (8):

$p_1(0) = 2^3/64 = 8/64$;
 $p_2(1) = \dots = p_7(1) = 2^2/64 = 4/64$;
 $p_8(2) = \dots = p_{19}(2) = 2^1/64 = 2/64$;
 $p_{20}(3) = \dots = p_{27}(3) = 2^0/64 = 1/64$, откуда проверка на распределение дает верный результат $(8/64) \cdot 1 + (4/64) \cdot 6 + (2/64) \cdot 12 + (1/64) \cdot 8 = 1$.

ЗАКЛЮЧЕНИЕ

1. Перечислительным методом (ПМ) по выбранным из указанных в аннотации направлениям предложены процедуры проведения доасимптотического анализа класса схем S^* в условиях схем $A-D$.

2. Построен метод (МГИ) пересчета результатов анализа схем S для схем S^* .

3. В результате доказательства теорем 1–4 (при равновероятности исходов схемы S) получены новые аналитические результаты численностей исходов схем S^* и их вероятностных распределений в условиях схем $B-D$ с просчетом численных примеров (условия схемы A изучались в [3, п. 2.23], а условия схем $B-D$ рассмотрены впервые).

ЛИТЕРАТУРА

1. Ивченко Г. И., Медведев Ю. И., Севастьянов Б. А. Размещение случайного числа частиц по ячейкам // Математические заметки. 1967. Т. 1, № 5. С. 549–554.
2. Колчин В. Ф., Севастьянов Б. А., Чистяков В. П. Случайные размещения. М.: Наука, 1976. 223 с.
3. Энатская Н. Ю. Доасимптотический анализ комбинаторных схем. М.: URSS, 2023. 536 с.

REFERENCES

1. Ivchenko G. I., Medvedev Yu. I., Sevastyanov B. A. Arrangement of a random number of particles in cells. *Matemeticheskie zametki = Mathematical notes*. 1967;1(5):549–554. (In Russ.)
2. Kolchin V. F., Sevastyanov B. A., Chistyakov V. P. Random allocations. Moscow: Nauka; 1976. 223 p. (In Russ.)
3. Enatskaya N. Yu. Pre-asymptotic analysis of combinatorial schemes. Moscow: URSS; 2023. 536 p. (In Russ.)

Поступила в редакцию / received: 14.12.2025; принята к публикации / accepted: 29.04.2026.
Автор заявляет об отсутствии конфликта интересов / The author declares no conflict of interest.

СВЕДЕНИЯ ОБ АВТОРЕ:

Энатская Наталия Юрьевна

канд. физ.-мат. наук, доцент Департамента
прикладной математики

e-mail: nat1943@mail.ru

CONTRIBUTOR:

Enatskaya, Natalia

Cand. Sci. (Phys.-Math.), Associate Professor

УДК 519.115:519.2

КОМБИНАТОРНЫЙ АНАЛИЗ СХЕМ РАЗМЕЩЕНИЯ ЧАСТИЦ С ЗАДАНЫМ МАКСИМАЛЬНЫМ УРОВНЕМ ЗАПОЛНЕНИЯ ЯЧЕЕК

Н. Ю. Энатская

*Московский институт электроники и математики,
Национальный исследовательский университет «Высшая школа экономики»
(ул. Таллинская, 34, Москва, Россия, 123458)*

Рассматриваемый класс схем размещения частиц по ячейкам характеризуется введением верхнего ограничения уровней заполнения ячеек с его обязательным достижением хотя бы в одной ячейке каждого исхода каждой схемы. Схемы различаются между собой парными качествами составляющих их элементов (ячеек и частиц) по их различимостям. Из направлений исследования схем выделяются представляющие наибольший интерес по нестандартным приемам доасимптотического анализа и получению новых результатов – это нумерованные неповторные перечисления исходов схем и нахождение их чисел.

Ключевые слова: размещение частиц; максимальный уровень заполнения ячеек; задача нумерации; моделирование

Для цитирования: Энатская Н. Ю. Комбинаторный анализ схем размещения частиц с заданным максимальным уровнем заполнения ячеек // Труды Карельского научного центра РАН. 2026. № 6. С. 139–147. doi: 10.17076/mat2272

N. Yu. Enatskaya. COMBINATORIAL ANALYSIS OF PARTICLE ALLOCATION SCHEMES WITH A GIVEN MAXIMUM CELL FILLING LEVEL

*National Research University Higher School of Economics, Moscow Institute of
Electronics and Mathematics (34 Tallinskaya St., 123458 Moscow, Russia)*

The class of particle cell arrangements under consideration is characterized by the introduction of an upper limit on the filling levels of the cells, which must be achieved in at least one cell of each outcome of each arrangement. The circuits are distinguished by the paired qualities of their constituent elements (cells and particles) according to their differences. Of the areas of study of schemes, those of greatest interest in non-standard methods of presymptotic analysis and obtaining new results are identified, these are numbered lists of the outcomes of schemes and finding their numbers.

Key words: placement of particles; maximal filling level of cells; numbering task; modeling

For citation: Enatskaya N. Yu. Combinatorial analysis of particle allocation schemes with a given maximum cell filling level. *Trudy Karel'skogo nauchnogo tsentra RAN = Transactions of the Karelian Research Centre RAS*. 2026. No. 6. P. 139–147. doi: 10.17076/mat2272

ВВЕДЕНИЕ

В работах [1–4] проводился асимптотический анализ вероятностного поведения максимальных уровней заполнения ячеек в схемах равновероятного размещения различных частиц по различным ячейкам. Здесь представлено доасимптотическое исследование решения обратной экстремальной задачи исследования исходов в классе схем с заданным максимальным уровнем k заполнения ячеек при равновероятном размещении частиц по n ячейкам при всех парных качествах ячеек и частиц по их различимости. (Для сравнения: при прямой постановке задачи исследуются экстремальные значения характеристик размещения частиц по их возможным исходам.)

Анализ схем проводится авторским перечислительным методом (ПМ) [5]. В его основе – построение итерационного случайного процесса прямого бесповторного нумерованного перечисления его исходов в доасимптотической области изменения параметров.

Прием исследования схем состоит в непосредственном учете специфики их исходов в одной (начальной) схеме данного класса схем с дальнейшим пересчетом результатов ее анализа в остальных схемах с разными качествами по различимости составляющих их элементов. Выбор начальной схемы и порядков пересчета результатов схем определяется возможностью его проведения.

Рассматриваются 4 схемы размещения r частиц по n ячейкам во всех вариантах их качеств по различимости с заданным максимальным уровнем их заполнения k , причем $r \leq kn$:

схема А – частицы и ячейки различимы;

схема В – частицы различимы, ячейки неразличимы;

схема С – частицы неразличимы, ячейки различимы;

схема D – частицы неразличимы, ячейки неразличимы.

Специфика схем состоит в том, что число частиц в каждой ячейке должно быть не только не более k , но и с обязательным достижением этого уровня заполнения в каждом исходе каждой схемы, т. е. хотя бы в одной ячейке каждого ее исхода. Используя обсуждаемые здесь связи между исходами схем, будем строить процедуры их бесповторных перечислений и на этой основе проводить их комбинаторный анализ в наиболее рациональном (с нашей точки зрения) порядке, а именно, начиная со схе-

мы с наиболее легко интерпретируемыми для других схем исходами, как этапа их изучения.

Доасимптотический анализ схем А–D проводился в нескольких авторских работах [5, п. 2.1, 2.12] и [6]. Но в п. 2.1 рассматривалась другая задача нахождения чисел исходов схем при фиксации диапазонов допустимых уровней заполнения для каждой ячейки; в п. 2.12 изучался только случай размещения неразличимых частиц по различным ячейкам для разных диапазонов значений r – чисел частиц, для которых не было найдено единого решения для числа исходов схем, а в [6] для схем А–D решалась задача нахождения вероятностного распределения исходов схемы при фиксированном максимальном уровне заполнения ячеек.

Постановка задачи. Провести доасимптотический анализ схем А–D по направлениям ПМ: бесповторному перечислению их исходов и определению их численностей путем прямого вычисления в одной схеме и их пересчету для остальных схем.

Актуальность исследований комбинаторных схем размещения частиц по ячейкам с различными ограничениями связана как с их широкими приложениями, так и с перспективным конструированием анализа новых все усложняющихся комбинаторных схем, содержащих данные схемы в нейросетях. Примерами таких схем могут быть случайное размещение единиц продукции по контейнерам определенного объема для отправки до обязательного данного уровня заполнения хотя бы одного из них или распределение продукции (или кадров специалистов) по объектам без превышения данного количественного максимума при его достижении.

1. МЕТОДОЛОГИЯ И НОВИЗНА ПРИЕМА АНАЛИЗА СХЕМ

Основной используемый здесь метод доасимптотического анализа схем – перечислительный метод ПМ. Он состоит в перечислении и формировании их исходов [5].

Предлагается проводить исследование в одной из схем с учетом общей специфики их исходов, т. е. с уровнями заполнения ячеек, не превышающими заданный максимум с обязательным его достижением хотя бы в одной ячейке исхода. Выбор начальной схемы определяется возможностью ее непосредственного прямого анализа. А для остальных рассматриваемых схем строятся процедуры последовательных пересчетов их результатов из резуль-

татов ранее исследованных схем. Порядок таких пересчетов наиболее просто осуществляется для *близких схем* (с различиями в различимости для одного элемента), где их пары составляются по возможности таких пересчетов с сохранением учтенной в начальной схеме общей специфики их исходов.

В качестве начальной в классе изучаемых схем предлагается схема C , для которой решим задачу бесповторного перечисления и определения числа исходов сначала только с учетом первого условия – данного ограничения ($\leq k$) уровней заполнения всех ячеек. Для учета второго ограничения, проводя отбраковку исходов без достижения данного максимума уровня хотя бы в одной ячейке, получим все исходы схемы C . Бесповторное перечисление ее исходов с учетом всех ограничений в одной из данного класса схем дают основу для пересчета аналогичных результатов остальных схем A – D с уже учтенной спецификой исходов.

Дальнейшее исследование данного класса схем проводим без непосредственного прямого исследования каждой, реализуемой последовательными пересчетами результатов перечисления их исходов парами схем с учетом условий размещения в них от бесповторного перечисления их исходов до вычисления их численностей. Заключительные алгоритмические действия определения численностей исходов схем в результате этих пересчетов в разных вариантах начальных схем в парах приводим как утверждения теорем. Они сводятся к их *классификации, упорядочиванию и отбраковке повторов*. Под *классификацией* элементов совокупности понимается их объединение по группам совпадения определенного признака с определением размеров этих групп. При пересчете перечисления исходов схемы в других схемах результатом классификации является совокупность объединений элементов в группы, а для пересчета чисел исходов – размеры этих групп.

В наших исследованиях под элементами классифицируемой совокупности будут рассматриваться как исходы схем, так и ячейки их исходов.

Выделение исследуемых направлений анализа по ПМ вызвано возможностью получения их результатов пересчетом соответствующих результатов начальной схемы. Это связано с приемом использования групповых соответствий исходов в парах *близких схем* в схемах с большей дифференциацией исходов с исходами схем своих пар.

2. Виды исходов схем A , B , C , D

Исходы схем A – D записываются соответственно в терминах:

для схемы A : $\bar{m} = (m_1, \dots, m_n)$ – состав номеров частиц в n ячейках в порядке ячеек;

для схемы B : $\bar{m}^* = (m_1^*, \dots, m_n^*)$ – состав номеров частиц в n ячейках в порядке возрастания уровней заполнения ячеек, а при совпадающих уровнях заполнения ячеек – в порядке возрастания минимальных номеров в них;

для схемы C : $\bar{\eta} = (\eta_1, \dots, \eta_n)$ – набор уровней заполнения ячеек в порядке ячеек;

для схемы D : $\bar{\eta}^* = (\eta_1^*, \dots, \eta_n^*)$ – набор уровней заполнения ячеек в неубывающем порядке (или частоты разных полученных уровней заполнения ячеек).

3. Бесповторное перечисление исходов схем A , B , C , D и числа их исходов N_A , N_B , N_C , N_D

На исходы каждой схемы A – D наложено два ограничения:

- 1) в каждой ячейке $\leq k$ частиц;
- 2) заданный максимальный уровень заполнения должен достигаться в каждом исходе хотя бы в одной ячейке.

Выбор порядка проведения анализа схем начиная со схемы C обусловлен возможностью непосредственного учета данных ограничений в ней с сохранением их при последовательном попарном пересчете результатов для остальных схем. Здесь для примера рассматривается пересчет результатов шире, чем только парами близких схем, т. е. D из C , A из C или D и B из D или A с отличием качеств различимости в одном (близкие схемы) или двух из составляющих схемы элементов (ячеек и частиц) в этих парах.

Схема C

Схема C представляет собой схему сочетаний с повторением [5, п. 1.1.1.5] с данными выше ограничениями. В бесповторном перечислении исходов схемы C будем сначала учитывать только первое ограничение (учет второго ограничения произведем позже отбраковкой исходов с максимальным уровнем заполнения ячеек $< k$). Это значит, что исходы схемы C должны представлять собой перечисление в порядке ячеек уровней их заполнения $\leq k$ при $r \leq nk$. В процессе организации исходов схемы такого вида сначала будем их записывать в другом виде: последовательностей r нулей и $(n-1)$ единиц на $m = n+r-1$ местах, где подряд стоящими нулями (сериями нулей) обозначены частицы каждой ячейки, а единицами –

перегородки $\{t_i\}$, $i = \overline{1, n-1}$ между n ячейками. Здесь длины упорядоченных серий (подряд стоящих) нулей есть уровни заполнения соответствующих ячеек. Тогда среди m мест номера мест единиц определяют исход схемы, т. к. уровни заполнения ячеек в порядке ячеек $\{\bar{h}\} = (h_1, h_2, \dots, h_n)$ вычисляются из значений $\{t_i\}$, а места единиц должны быть выбраны так, чтобы все ячейки содержали не более k частиц.

Перечисление исходов схемы здесь состоит в последовательных действиях выбора допустимых мест расстановок внутренних перегородок между ячейками от t_1 до t_{n-1} . Доопределим $t_0 = 0$, $t_n = n + r$.

Число исходов схемы $N(k)$ вычисляется суммой размеров пучков предпоследней итерации, т. е. диапазонов варьирования значений мест для последней t_{n-1} -й внутренней перегородки.

Выбор мест этих перегородок $\bar{t} = (t_1, \dots, t_{n-1})$ по первому ограничению определен теоремой:

Теорема 1. *Диапазоны варьирования возможных вариантов последовательной установки мест внутренних перегородок t_i определяются при $r \leq nk$ для всех $i = \overline{1, n-1}$, начиная с $i = 1$ из рекуррентного соотношения*

$$l_i = \max(t_{i-1} + 1, r + i - k(n - i)) \leq t_i \leq \min(t_{i-1} + k + 1, r + i) = L_i. \quad (1)$$

Доказательство. Для нахождения мест внутренних перегородок $\bar{t} = (t_1, \dots, t_{n-1})$ должен выполняться ряд требований:

1. разности между местами границ должны быть $\leq (k + 1)$, откуда

$$t_i \leq t_{i-1} + k + 1; \quad (2)$$

2. правее значения t_i среди m мест поместятся остальные $n - 1 - i$ границ, откуда

$$t_i \leq m - (n - 1 - i) = r + i; \quad (3)$$

3. значение следующего элемента больше предыдущего, откуда

$$t_i \geq t_{i-1} + 1; \quad (4)$$

4. правее значения t_i среди m мест окажется мест не более, чем для $n - 1 - i$ остальных границ и $k(n - i)$ максимально допустимого в остальных ячейках схемы A числа частиц, откуда

$$t_i \geq m - (n - 1 - i) - k(n - i) = r + i - k(n - i). \quad (5)$$

Теперь, объединяя соотношения (2) и (3) в верхнюю границу для t_i , а (4) и (5) – в нижнюю границу для t_i , получаем при $i = \overline{1, n-1}$ формулу (1) диапазона изменения значений для t_i . Теорема 1 доказана.

Приведем пояснительный пример использования теоремы 1.

Пример 1. Пусть $n = 3$, $r = 5$, $k = 3$. Отсюда $m = 7$. Перечислим все исходы схемы, вычисляя пределы варьирования их элементов в указанных в (1) диапазонах и их размеры – размеры пучков по шагам (итерациям) перечисления исходов:

1-й шаг:

$l_1 = \max(1, 0) \leq t_1 \leq \min(4, 6)$, откуда $1 \leq t_1 \leq 4$, размер пучка равен 4;

2-й шаг:

$l_2 = \max(t_1 + 1, 4) \leq t_2 \leq \min(t_1 + 4, 6)$, откуда при

$t_1 = 1$ получаем $4 \leq t_2 \leq 5$, размер 1-го пучка равен 2;

$t_1 = 2$ получаем $4 \leq t_2 \leq 6$, размер 2-го пучка равен 3;

$t_1 = 3$ получаем $4 \leq t_2 \leq 7$, размер 3-го пучка равен 4;

$t_1 = 4$ получаем $5 \leq t_2 \leq 7$, размер 4-го пучка равен 3.

Отсюда число исходов схемы $N(k) = 2 + 3 + 4 + 3 = 12$.

Замечание 1. Это утверждение фактически доказано в [5, п. 2.13], но не сформулировано и не акцентировано там в виде теоремы, поэтому здесь оно оформлено и доказано полностью как теорема 1.

Число $N(k)$ исходов схемы теоремы 1 в графе их перечисления определяется итерационным процессом варьирования значений мест последовательной расстановки перегородок между ячейками среди лежащих в ряд частиц. Размеры пучков итераций задаются их полученными в теореме 1 диапазонами – например, для диапазона места перегородки $t_i \in [a, b]$ размер пучка в графе перечисления исходов схемы C равен $b - a + 1$.

Тогда число $N(k)$ вычисляется суммой размеров пучков предпоследней итерации, равной сумме размеров диапазонов варьирования мест t_{n-1} перегородки. $N(k) = \sum_j d_j$, где d_j – размер j -го пучка предпоследней $(n - 1)$ -й итерации.

В терминах наборов уровней заполнения ячеек по выбранным местам этих внутренних перегородок получим их через последовательные уровни заполнения ячеек $\bar{h} = (h_1, \dots, h_n)$

в их заданном в схеме порядке по нижеприведенной формуле

$$i = \overline{1, n}, \quad h_i = t_i - t_{i-1} - 1. \quad (6)$$

Например, если при $t_0 = 0$ получаем $t_1 = 4$, то это значит, что в первой ячейке уровень заполнения $h_1 = 4 - 0 - 1 = 3$ – в ней 3 частицы.

Замечание 2. По сути, итерационное варьирование места каждой перегородки от начала своего диапазона означает добавление частицы в предыдущую от этой перегородки ячейку, что следует из формулы (6).

Замечание 3. Число N_C исходов схемы C определяем из $N(k)$ подсчетом оставшихся исходов схемы из теоремы 1 после их отбраковки по второму ограничению, которое можно записать формулами двух видов, фильтрующих исходы схемы в теореме 1 от исходов с недостижением заданного максимального уровня заполнения хотя бы в одной ячейке: через а) индикаторные функции, б) разность исходов по теореме 1 с верхними ограничениями уровней заполнения ячеек k и $(k - 1)$.

Теорема 2. Число N_C исходов схемы C определяется формулой

а)

$$N_C = \sum_{\{\bar{h}\}} I_2 \left(\sum_{i=1}^n I_1(k - h_i) \right), \quad (7)$$

где первая сумма производится по перечислению исходов указанного под знаком суммы множества исходов $\bar{h}$ схемы из теоремы 1 и где индикаторные функции заданы следующим образом: $I_1(Z_1) = 0$ при $Z_1 > 0$ и $I_1(Z_1) = 1$ при $Z_1 = 0$; $I_2(Z_2) = 0$ при $Z_2 = 0$ и $I_2(Z_2) = 1$ при $Z_2 > 0$, или

б)

$$N_C = N(k) - N(k - 1). \quad (8)$$

Доказательство. а) Число исходов схемы $N(k)$ можно записать формулой через индикаторные функции $I_1(Z)$ и $I_2(Z)$ в виде суммы по перечислению исходов схемы $\bar{h}$ в теореме 1 в терминах уровней заполнения ячеек в доказываемой формуле (7), т. к. всегда $Z_1 = k - h_i \geq 0$ (по замечанию 3) как разность максимального уровня и уровня заполнения ячейки в исходах схемы теоремы 1, $I_1(Z_1) = I_1(k - h_i)$ – индикатор достижения в i -й ячейке заданного максимального уровня заполнения k , $I_2(Z_2) = I_2(\sum_{i=1}^n I_1(k - h_i))$ – индикатор достижения хотя бы по одной компоненте фиксированного исхода из $\{\bar{h}\}$ заданного максимального уровня заполнения k . Тогда $\sum_{\{\bar{h}\}} I_2(Z_2)$ – число

исходов схемы C , т. к. в них выполняются оба ограничения, что приводит к (7).

б) Доказательство формулы (8) следует из того, что $N(k)$ учитывает первое ограничение в числе исходов схемы C из теоремы 1, а вычитание из него $N(k - 1)$ (числа исходов с уровнями заполнения ячеек $\leq (k - 1)$) – и второе ограничение в них.

Замечание 4. В связи с тем, что перечисление и анализ исходов остальных схем будет производиться на основе перечисления исходов схемы C (т. е. при полученном наборе уровней заполнения ячеек схемы C), то их исходы будут удовлетворять первому и второму ограничениям.

Схема D

В исходах схемы D , получаемых из схемы C , нужно учесть неразличимость ячеек. Для этого упорядочиваем уровни заполнения ячеек в исходах схемы C , например, в неубывающем порядке с последующей отбраковкой повторов исходов. Так получаем перечисление исходов схемы D , а их число N_D определяем по теореме 3 на основании классификации исходов схемы C по совпадению составов уровней заполнения ячеек.

Теорема 3. Число N_D исходов схемы определяется формулой из исходов схемы C

$$N_D = \sum_{\{\bar{\mu}\}} I_3(\mu_j), \quad (9)$$

где суммирование производится по перечислению исходов указанного под знаком суммы множества при $j = \overline{1, s}$, индикаторная функция $I_3(Z) = 1$ при $Z > 0$ и $I_3(Z) = 0$ при $Z = 0$, а $\bar{\mu} = (\mu_1, \dots, \mu_j, \dots, \mu_s)$ – размеры групп классификации исходов $\{\bar{\eta}\}$ по совпадениям их составов уровней заполнения ячеек, $\mu_j - j$ -й из размеров групп разных составов уровней заполнения ячеек схемы C , а s – число их разных составов.

Доказательство. Исходы схемы D получаем из исходов $\{\bar{\eta}\}$ схемы C (после их упорядочивания) в виде $\{\bar{\eta}^*\}$ в результате отбраковки повторов в $\{\bar{\eta}^*\}$. Тогда числом N_D является размерность совокупности $\bar{\mu} = (\mu_1, \dots, \mu_j, \dots, \mu_s)$ размеров групп классификации по совпадениям их составов уровней заполнения ячеек. Получение этой размерности реализуется по (9), т. к. $I_3(\mu_j) = 1$ при $\mu_j > 0$, а $\mu_j > 0$ по всем разным исходам из $\{\bar{\eta}^*\}$, что и дает размерность $\bar{\mu}$.

Замечание 5. Формула (9) приведена для единообразия представления чисел исходов схем в виде вычисляемых формул.

Приведем пояснительный пример применения теоремы 3.

Пример 2. Пусть в условиях примера 1 $n = 3$, $r = 5$, $k = 3$. Тогда в терминах уровней заполнения ячеек по (6) из результатов примера 1 после отбраковки по второму ограничению исходов $(1, 2, 2)$, $(2, 1, 2)$, $(2, 2, 1)$ в порядке их получения выпишем все 9 исходов схемы C : $(0, 2, 3)$, $(0, 3, 2)$, $(1, 1, 3)$, $(1, 3, 1)$, $(2, 0, 3)$, $(2, 3, 0)$, $(3, 0, 2)$, $(3, 1, 1)$, $(3, 2, 0)$;

упорядочим уровни заполнения ячеек в неубывающем порядке, получим $(0, 2, 3)$, $(0, 2, 3)$, $(1, 1, 3)$, $(1, 1, 3)$, $(0, 2, 3)$, $(0, 2, 3)$, $(0, 2, 3)$, $(1, 1, 3)$, $(0, 2, 3)$;

классифицируем исходы по их совпадению (это значит, что после j -го следующего исхода вычеркиваем ее повторы, считая их общее число со сравниваемым μ_j), получим $\bar{\mu} = (\mu_1, \mu_2) = (6, 3)$ с размерностью, равной 2; отсюда по (9) $N_D = 2$.

Схема A

В исходах схемы A , получаемых а) из схемы C , нужно учесть различимость частиц, или б) из схемы D – различимость и частиц, и ячеек. Поэтому для перечисления исходов схемы A совершаем последовательность действий по разным схемам: а) бесповторно перечисляем все исходы схемы C и при каждом производим перечисление исходов соответствующей схемы перестановок с повторением [5, п. 1.7]), учитывающей все возможные составы частиц в ячейках, или б) бесповторно перечисляем все исходы схемы D и при каждом производим перечисление исходов, соответствующих двум последовательным действиям – перестановок с повторением, учитывающим все возможные порядки наборов уровней заполнения ячеек, и при каждом из них перечисляем все соответствующие им составы частиц с числами разных исходов N_A , определяемых соответственно формулами по следующему утверждению.

Теорема 4. Число N_A исходов схемы определяется формулой

а) из исходов схемы C

$$N_A = \sum_{\{\bar{\eta}\}} \frac{r!}{\prod_{i=1}^n \eta_i!} \quad (10)$$

или

б) из исходов схемы D

$$N_A = \sum_{\{\bar{\eta}^*\}} \frac{r!}{\prod_{i=1}^n \eta_i^*!} \frac{n!}{\prod_{j=1}^s \mu_j!}. \quad (11)$$

Суммы в (10) и (11) производятся по перечислению исходов указанных под знаками сумм

множеств, где μ_j – размеры групп классификации среди n ячеек исхода по совпадению их уровней заполнения, откуда $s \leq n$.

Доказательство. Все исходы схемы A должны учитывать все составы частиц в ячейках с уровнями заполнения схемы C во всех их порядках.

Формула (10) представляет собой сумму по перечислению исходов схемы C со слагаемыми всех разных составов частиц в ячейках, определяемых по схеме перестановок с повторением [5, п. 1.7].

Формула (11) представляет собой сумму по перечислению исходов схемы D (разных составов уровней заполнения ячеек) со слагаемыми – всех разных составов частиц в ячейках во всех порядках составов уровней заполнения ячеек, определяемых по двум последовательным схемам перестановок с повторением [5, п. 1.7 и 1.12].

Схема B

В исходах схемы B , получаемых а) из схемы D , нужно учесть различимость частиц, или б) из схемы A – неразличимость ячеек. Для этого будем приводить исходы этих схем к виду исходов схемы B .

В схеме D будем по известным составам возрастающих уровней заполнения ячеек $\{\bar{\eta}^*\}$ размещать по ним различимые частицы, а в схеме A – упорядочивать ячейки (например, в порядке роста в них уровней заполнения, а ячейки с совпадающими уровнями заполнений в порядке роста номеров частиц ячеек с минимальным номером) с отбраковкой повторов. Поэтому для перечисления исходов схемы B совершаем действия:

а) перечисляем все исходы схемы D с заменой уровней заполнения ячеек на составы частиц в ячейках по схеме перестановок с повторением [5, п. 1.7] с их перечислением в ячейках с совпадающими уровнями заполнений в заранее определенном порядке (например, роста номеров частиц ячеек с минимальным номером) или

б) перечисляем все исходы схемы A , например, в порядке неубывания уровней заполнения ячеек, а среди ячеек с совпадающими уровнями – в порядке роста номеров частиц ячеек с минимальным номером с отбраковкой повторов.

Тогда по этим алгоритмическим действиям число N_B определяется соответственно формулам по следующему утверждению.

Теорема 5. Число N_B исходов схемы определяется формулой

а) из исходов схемы D

$$N_B = \sum_{(\bar{\eta}^*)} \frac{r!}{\prod_{i=1}^n \eta_i^*! \prod_{j=1}^s \mu_j!}, \quad (12)$$

где $j = \overline{1, s} \leq n$, μ_j – j -й размер группы классификации n ячеек каждого исхода схемы D – η^* из $\{\eta^*\}$ по совпадениям их уровней заполнения, а суммирование проводится по указанному под ней множеству,

или

б) из исходов схемы A

$$N_B = \sum_{\{\bar{M}\}} I_3(M_j), \quad (13)$$

где $\bar{M} = (M_1, \dots, M_j, \dots, M_S)$ – размеры групп классификации исходов схемы A из $\{\bar{m}\}$ по совпадениям составов их частиц по ячейкам, S – число разных групп – порядков ячеек каждого состава частиц схемы A , суммирование проводится по указанному под знаком суммы множеству, а функция $I_3(Z)$ определена в теореме 3.

Доказательство. Все исходы схемы B должны учитывать все составы частиц в ячейках с уровнями заполнения схемы C без учета их порядка, т. е. в одном заранее установленном порядке, что соответствует неразличимости ячеек в схеме B .

Формула (12) представляет собой сумму данного вида слагаемых по перечислению исходов схемы D со слагаемыми всех разных составов частиц в ячейках.

Формула (13) представляет собой сумму данного вида слагаемых по перечислению размеров групп классификации исходов схемы по совпадениям их составов частиц в ячейках.

В соответствии с предложенными выше процедурами бесповторного перечисления исходов схемы B число ее исходов N_B определяется:

а) из перечисления исходов схемы D по формуле (12) с суммой, учитывающей все разные составы уровней заполнения ячеек в указанном в схеме D порядке ячеек в количествах исходов схемы перестановок с повторением [5, п. 1.7], деленных на числа всех порядков ячеек с совпадающими уровнями их заполнений (с учетом неразличимых порядков ячеек в схеме B) в количествах произведений факториалов количеств совпадения размеров уровней заполнения ячеек; $s \leq n$, т. к. это размерность упорядоченных чисел исходов совпадений составов n ячеек после отбраковки возможных повторов;

или

б) из перечисления исходов схемы A – по формуле (13), получаемой как сумма чисел групп классификации исходов схемы A с совпадающими составами частиц в ячейках; тогда N_B получаем как размерность вектора $\bar{M}$ классификации исходов схемы A с совпадающими составами частиц в любом или в одном заранее заданном порядке, например, с упорядочением всех составов частиц в ячейках в схеме A (в порядке роста уровней заполнения, а для ячеек с совпадающими размерами уровней заполнения – в порядке роста содержащихся в них минимальных номеров частиц) как число таких разных исходов, т. е. число разных количеств по совпадениям исходов $\{m^*\}$.

Тогда по этим перечислительным действиям число N_B определяется соответственно формулами по данному утверждению.

Замечание 6. Число исходов схемы B из исходов схемы A есть число разных составов частиц в ячейках схемы A и определяется

1) числом упорядоченных составов частиц в ячейках исходов схемы A с отбраковкой повторов или

2) размерностью количеств исходов схемы A по совпадению составов частиц в ячейках, что соответствует вычислению N_D по (13) и приведено для единообразия представления чисел исходов схем в виде вычисляемых формул.

Теоремы 2–5 являются новыми и ранее не публиковались.

Вычисления и пересчеты чисел исходов схем A , B , C , D на числовом примере

Схема C

Пример 3. Пусть, как в примерах 1 и 2, $n = 3$, $r = 5$, $k = 3$. Отсюда $m = 7$. В связи с этим частично исходы примеров 2 и 3 совпадают, но для наглядности разбора примера 3 будут здесь повторены. Вычислим число исходов нашей схемы разными способами: непосредственно с использованием теоремы 1 и по формулам (7) и (8) теоремы 2.

По результату примера 1 из всех $N(k) = 12$ исходов перечислим исходы схемы сочетаний с повторением с максимальным уровнем $k = 3$ заполнения ячеек в терминах мест внутренних перегородок $\{(t_1, t_2)\} = (1, 4), (1, 5), (2, 4), (2, 5), (2, 6), (3, 4), (3, 5), (3, 6), (3, 7), (4, 5), (4, 6), (4, 7)$. По (6) переведем эти исходы в принятый для схемы C вид наборов уровней заполнения ячеек в порядке ячеек – соответственно получим исходы: $(0, 2, 3), (0, 3, 2), (1, 1, 3), (1, 2, 2), (1, 3, 1), (2, 0, 3), (2, 1, 2),$

(2,2,1), (2,3,0), (3,0,2), (3,1,1), (3,2,0), из которых отбраковываются 3 исхода: (1,2,2), (2,1,2), (2,2,1), где заданный максимальный уровень $k = 3$ не достигнут; остальные исходы – все $N_C = 9$ исходов схемы C .

По удалении из перечисления исходов схемы теоремы 1 $\{\bar{h}\}$ недопустимых исходов для схемы C по второму ограничению в терминах мест внутренних перегородок ячеек (2,5), (3,5), (3,6) получаем скорректированную пучковую структуру второй итерации итогового графа перечисления исходов схемы C : 2,2,2,3, откуда число исходов как их сумма $N_C = 2 + 2 + 2 + 3 = 9$, что совпадает с ранее полученным результатом.

По (7) $N_C = 6I_2(1) + 3I_2(1) = 6 + 3 = 9$, где первое слагаемое соответствует количеству исходов с уровнями заполнения ячеек (0,2,3) во всех порядках, а второе – с уровнями заполнения ячеек (1,1,3) во всех порядках.

Для вычисления N_C по (8) аналогично расчету $N(3)$ вычисляем $N(2)$, где получаем $2 \leq t_1 \leq 3$, размер пучка равен 2; при

$t_1 = 2$ получаем $5 \leq t_2 \leq 5$, размер 1-го пучка равен 1,

$t_1 = 3$ получаем $5 \leq t_2 \leq 6$, размер 2-го пучка равен 2.

Отсюда число исходов $N(k) = N(2) = 1 + 2 = 3$. Теперь по (8) находим $N_C = 12 - 3 = 9$.

Схема D

Для пересчета исходов схемы D из исходов схемы C достаточно их упорядочить по возрастанию уровней заполнения с отбраковкой повторов. В условиях примера 3 это даст $N_D = 2$ исхода (0,2,3) и (1,1,3).

По (9) $N_D = I_3(6) + I_3(3) = 1 + 1 = 2$.

Схема A

В условиях примера 3 получаем по (10)

$$N_A = 6 \frac{5!}{0!2!3!} + 3 \frac{5!}{1!1!3!} = 120,$$

где среди исходов схемы C уровни заполнений (0,2,3) встречаются в $3!/1!1!1! = 6$ порядках, а уровни заполнения (1,1,3) – в $3!/2!1! = 3$ порядках. Это значит, что в сумме формулы (10) количества слагаемых видов $5!/0!2!3!$ и $5!/1!1!3!$ равны соответственно 6 и 3.

В условиях примера 3 получаем по (11)

$$N_A = \frac{5!}{0!2!3!} \frac{3!}{1!1!1!} + \frac{5!}{1!1!3!} \frac{3!}{2!1!} = 60 + 60 = 120.$$

Схема B

В условиях примера 3 в случае а) пересчета N_B из схемы D получаем по (12)

$$N_B = \frac{5!}{0!2!3!1!1!1!} + \frac{5!}{1!1!3!2!1!} = 20.$$

В условиях примера 3 в случае б) пересчета N_B из схемы A получаем по (13) число всех исходов схемы B после описанного упорядочивания исходов схемы A . А это есть число всех разных групп классификации упорядоченных исходов схемы A по совпадению в них составов частиц в ячейках, откуда и из определения функции I_3 следует формула (13), записывающая отбраковку повторов после упорядочивания в исходах схемы A .

Для наглядности перечислим эти 20 исходов с допустимыми уровнями заполнения (0,2,3) и (1,1,3), заключая в скобки составы ячеек с более чем одной частицей:

(0, (1, 2), (3, 4, 5)), (0, (1, 3), (2, 4, 5)),
 (0, (1, 4), (2, 3, 5)), (0, (1, 5), (2, 3, 4)),
 (0, (2, 3), (1, 4, 5)), (0, (2, 4), (1, 3, 5)),
 (0, (2, 5), (1, 4, 5)), (0, (3, 4), (1, 2, 5)),
 (0, (3, 5), (1, 2, 4)), (0, (4, 5), (1, 2, 3)),
 (1, 2, (3, 4, 5)), (1, 3, (2, 4, 5)),
 (1, 4, (2, 3, 5)), (1, 5, (2, 3, 4)),
 (2, 3, (1, 4, 5)), (2, 4, (1, 3, 5)),
 (2, 5, (1, 3, 4)), (3, 4, (1, 2, 5)),
 (3, 5, (1, 2, 4)), (4, 5, (1, 2, 3)).

Полученный результат для числа исходов схемы B совпадает с непосредственным пересчетом из числа исходов схемы A .

Приведение к виду исходов схемы B исходов схемы A в результате данного их упорядочения с отбраковкой повторов означает, что в результате этого в исходах схемы B каждая группа совпадающих в любом порядке составов уровней заполнения ячеек схемы A уменьшается в число разных порядков этих составов и в произведение факториалов количеств совпадающих уровней заполнения в группе.

Отсюда получаем для составов уровней заполнения ячеек (0,2,3) и (1,1,3) соответственно из 60 и 60 исходов схемы A в схеме B – $60/(6 \cdot 1!1!1!) = 10$ и $60/((3 \cdot 2!1!)) = 10$, т. е. всего $N_B = 20$.

Выводы

1. Предложен подход к исследованию данного класса схем, состоящий в последовательном пересчете непосредственно полученных результатов начальной схемы с учетом данных в классе ограничений с их сохранением для остальных схем.

2. При выборе начальной схемы по возможности проведения ее наиболее простого прямого анализа предложены парные порядки пересчета результатов в остальных схемах с отличием различимости в составляющих их элементах.

ЛИТЕРАТУРА

1. Викторова И. И. Об асимптотическом поведении максимума в равновероятной полиномиальной схеме // Математические заметки. 1969. Т. 3, № 3. С. 305–316.
2. Колчин В. Ф. О предельном поведении крайних членов вариационного ряда в полиномиальной схеме // Теория вероятностей и ее применения. 1969. Т. 14, № 3. С. 476–487.
3. Колчин В. Ф., Севастьянов Б. А., Чистяков В. П. Случайные размещения. М.: Наука, 1976. 223 с.
4. Хакимуллин Е. Р. О предельном поведении максимального заполнения в равновероятной схеме размещения частиц комплектами // Математические заметки. 1981. Т. 30, № 2. С. 277–289.
5. Энатская Н. Ю. Доасимптотический анализ комбинаторных схем. М.: URSS, 2023. 536 с.
6. Энатская Н. Ю. Вероятностный анализ максимальных заполнений в моделях размещения частиц по ячейкам // Труды Карельского научного центра РАН. 2024. № 4. С. 49–54. doi: 10.17076/mat1856

Поступила в редакцию / received: 14.12.2025; принята к публикации / accepted: 25.05.2026.
Автор заявляет об отсутствии конфликта интересов / The author declares no conflict of interest.

СВЕДЕНИЯ ОБ АВТОРЕ:

Энатская Наталия Юрьевна

канд. физ.-мат. наук, доцент Департамента
прикладной математики

e-mail: nat1943@mail.ru

REFERENCES

1. Viktorova I. I. Asymptotic behavior of maximum of an equiprobable polynomial scheme. *Mathematical Notes*. 1969;5(3):184–191. doi: 10.1007/BF01388624
2. Kolchin V. F. On the limiting behavior of extreme order statistics in a polynomial scheme. *Theory of Probability and its Applications*. 1969;14(3):458–469. doi: 10.1137/1114058
3. Kolchin V. F., Sevastyanov B. A., Chistyakov V. P. Random allocations. Moscow: Nauka; 1976. 223 p. (In Russ.)
4. Khakimullin E. R. Asymptotic behavior of the maximum occupancy in an equiprobable scheme of allocation of particles by complexes. *Mathematical Notes*. 1981;30(2):626–633. doi: 10.1007/BF01708846
5. Enatskaya N. Yu. Pre-asymptotic analysis of combinatorial schemes. Moscow: URSS; 2023. 536 p. (In Russ.)
6. Enatskaya N. Yu. Probability analysis of maximum occupation in models of particle placement to cells. *Trudy Karelskogo nauchnogo tsentra RAN = Transactions of the Karelian Research Centre RAS*. 2024;4:49–54. (In Russ.). doi: 10.17076/mat1856

CONTRIBUTOR:

Enatskaya, Natalia

Cand. Sci. (Phys.-Math.), Associate Professor

ПАМЯТИ АНАТОЛИЯ ДМИТРИЕВИЧА СОРОКИНА (1936–2025)



14 ноября 2025 года после тяжелой болезни на 90-м году ушел из жизни наш коллега Анатолий Дмитриевич Сорокин, замечательный человек, много сделавший для сохранения в Карелии математического направления, руководивший Отделом математики и анализа данных в трудные годы перестройки.

Анатолий Дмитриевич Сорокин родился 12 октября 1936 г. в Петрозаводске в семье слесаря Кировской железной дороги. После окончания железнодорожной средней школы № 9 в 1954 г. будущая специальность была выбрана без колебаний: Ленинградский политехнический институт, электромеханический факультет. Окончив успешно вуз в 1960 г. и проработав два года в конструкторском бюро в Выборге, поступил в очную аспирантуру Карельского филиала АН СССР. После окончания аспирантуры в 1965 г. был распределен на работу в Карельский НИИ лесной промышленности, где проработал до 1975 г. сначала в должности старшего научного сотрудника, затем заведу-

ющего лабораторией автоматизации учета вычислительных работ.

В 1971 г. защитил кандидатскую диссертацию по специальности «Техническая кибернетика» в Ленинградском политехническом институте. В январе 1975 г. прошел по конкурсу на должность заведующего лабораторией автоматизации научных исследований Отдела математических методов и вычислительной техники КФ АН СССР, и с этого времени его научная и научно-организационная деятельность свыше 40 лет была неразрывно связана с Карельским филиалом АН СССР.

Обладея организаторскими способностями, коммуникабельностью в отношениях с сотрудниками, являясь исключительно добросовестным и инициативным работником, А. Д. Сорокин привлек внимание руководства Карельского филиала АН СССР и был выдвинут на должность заместителя председателя Президиума КФ АН СССР по научной работе. В этой должности он работал с 1977 по 1992 г., совмещая административную деятельность с руководством лабораторией.

Под его научным руководством и при непосредственном участии в филиале было создано новое научное направление — автоматизация научных исследований, объединившее усилия сотрудников различных научных специальностей в создании систем автоматизации экспериментов и обработки данных. Под руководством А. Д. Сорокина формировались центры коллективного пользования по хроматографии и радиоизотопным методам исследований с автоматизацией сбора и обработки данных. Он проводил активную работу по координации и операции научно-исследовательских работ в области информатизации и автоматизации научных исследований на всесоюзном уровне,

участвовал в работе международных конференций и совещаний.

В 1989 г. решением Президиума АН СССР А. Д. Сорокину присвоено ученое звание старшего научного сотрудника по специальности «применение вычислительной техники и математических методов в научных исследованиях». В 1992 г., после избрания по конкурсу на должность заведующего Отделом ММАНИП Карельского научного центра РАН, А. Д. Сорокин переходит на работу в отдел, уделяя большое внимание подготовке и росту научных кадров, оснащенности отдела техническими средствами, новому направлению исследований – информационно-телекоммуникационным системам. Исследования поддерживались грантами РФФИ, РГНФ и международного фонда ИНТАС.

При поддержке руководства КарНЦ и Отделения математики РАН А. Д. Сорокин ведет большую подготовительную работу по преобразованию отдела в институт. Отдел получает разрешение на подготовку аспирантов и соискателей ученой степени кандидата наук, с октября 1993 г. переводится на самостоятельный баланс с правами юридического лица, имеет свой Устав, гербовую печать, проходит аккредитацию. Отдел осуществляет научно-методическое руководство работами по организации компьютерной сети КарНЦ и созданию инфраструктуры доступа к сети Internet, поддержке WWW-сервера центра, активно сотрудничает с вузовской наукой, проводит международные конференции и семинары, исследования поддерживаются РФФИ, РГНФ, научными фондами Правительства РФ и Института «Открытое общество».

В июне 1999 г. Постановлением Президиума РАН отдел преобразован в Институт прикладных математических исследований КарНЦ РАН. А. Д. Сорокин назначен заместителем директора по научной работе, одновременно руководит научной темой по созданию и развитию электронных информаци-

онных ресурсов в КарНЦ. Разработанный подход к созданию и использованию электронных научных информационных ресурсов в виде комплексной распределенной информационно-аналитической системы поддержки и сопровождения научных исследований представляется перспективным и соответствует современному отечественному и зарубежному уровню исследований и разработок в этой области. Запущен в эксплуатацию вычислительный кластер, создан центр коллективного пользования КарНЦ РАН «Центр высокопроизводительной обработки данных». Организационно-технические мероприятия позволили добиться надежного функционирования компьютерной сети КарНЦ РАН и обеспечили высокую пропускную способность каналов связи.

А. Д. Сорокин автор более 100 научных работ, из них 80 печатных, соавтор четырех электронных коллекций научных информационных ресурсов, зарегистрированных в депозитарии электронных изданий НТИЦ «Информрегистр». Он являлся членом Совета по информатизации при Главе РК, заместителем председателя научно-технического Совета по телекоммуникациям при Президиуме КарНЦ РАН.

За успехи в научной и научно-организационной деятельности А. Д. Сорокин награжден Почетной грамотой Президиума РАН и профсоюза работников РАН, Почетной грамотой Республики Карелия, юбилейной медалью ФНПР «100 лет профсоюзам России», медалью ордена «За заслуги перед Отечеством 2 степени», ему присвоено звание «Заслуженный деятель науки РК».

Анатолия Дмитриевича отличали высокая ответственность, трудолюбие, доброжелательность, скромность и порядочность, и он всегда поддерживал молодежь. Светлая память о нем навсегда останется в памяти и сердцах знавших его людей.

Коллеги из ИПМИ КарНЦ РАН

ПРАВИЛА ДЛЯ АВТОРОВ

Серия «Математическое моделирование и информационные технологии»

(требования к работам, представляемым к публикации
в «Трудах Карельского научного центра Российской академии наук»)

«Труды Карельского научного центра Российской академии наук» (далее – Труды КарНЦ РАН) публикуют результаты завершённых оригинальных исследований в различных областях современной науки: теоретические и обзорные статьи, сообщения, материалы о научных мероприятиях (симпозиумах, конференциях и др.), персоналии (юбилеи и даты, утраты науки), статьи по истории науки. Представляемые работы должны содержать новые, ранее не публиковавшиеся данные.

Статьи проходят обязательное рецензирование. Решение о публикации принимается редакционной коллегией серии или тематического выпуска Трудов КарНЦ РАН после рецензирования, с учётом научной значимости и актуальности представленных материалов. Редколлегия серий и отдельных выпусков Трудов КарНЦ РАН оставляет за собой право возвращать без регистрации рукописи, не отвечающие настоящим правилам.

При получении редакцией рукопись регистрируется (в случае выполнения авторами основных правил ее оформления) и направляется на отзыв рецензентам. Отзыв состоит из ответов на типовые вопросы анкеты и может содержать дополнительные расширенные комментарии. Кроме того, рецензент может внести замечания и правки в текст рукописи. Авторам высылаются электронная версия анкеты и комментарии рецензентов. Доработанный экземпляр автор должен вернуть в редакцию вместе с первоначальным экземпляром и ответом на все вопросы рецензента не позднее чем через месяц после получения рецензии. Перед опубликованием авторам высылаются электронная версия статьи, которую авторы вычитывают и заверяют.

Журнал имеет систему электронной редакции на базе Open Journal System (OJS), позволяющую вести представление и редактирование рукописи, общение автора с редколлегиями серий и рецензентами в электронном формате и обеспечивающую прозрачность процесса рецензирования при сохранении анонимности рецензентов (<http://journals.krc.karelia.ru/>).

Содержание выпусков Трудов КарНЦ РАН, аннотации и полнотекстовые электронные версии статей, а также другая полезная информация, включая настоящие Правила, доступны на сайтах – <http://transactions.krc.karelia.ru/>; <http://journals.krc.karelia.ru/>

Почтовый адрес редакции: 185910, г. Петрозаводск, ул. Пушкинская, 11, КарНЦ РАН, редакция Трудов КарНЦ РАН. Телефон: (8142) 762018

ПРАВИЛА ОФОРМЛЕНИЯ РУКОПИСИ

Статьи публикуются на русском или английском языке. Рукописи должны быть тщательно выверены и отредактированы авторами.

Объём рукописи (включая таблицы, список литературы, подписи к рисункам, рисунки) не должен превышать: для обзорных статей – 30 страниц, для оригинальных – 25, для сообщений – 15, для хроники и рецензий – 5–6. Объём рисунков не должен превышать 1/4 объёма статьи. Рукописи большего объёма (в исключительных случаях) принимаются при достаточном обосновании по согласованию с ответственным редактором.

При оформлении рукописи применяется полуторный межстрочный интервал, шрифт Times New Roman, кегль 12, выравнивание по обоим краям. Размер полей страницы – 2,5 см со всех сторон. Все страницы, включая список литературы и подписи к рисункам, должны иметь сплошную нумерацию в правом нижнем углу. Страницы с рисунками не нумеруются.

Рукописи подаются в электронном виде в систему электронной редакции на сайте <http://journals.krc.karelia.ru> либо высылаются на e-mail: trudy@krc.karelia.ru, или же представляются в редакцию лично (г. Петрозаводск, ул. Пушкинская, 11, каб. 502).

Для публикации в выпусках серии «Математическое моделирование и информационные технологии» рукописи принимаются в формате .tex (LaTeX 2ε) с использованием стилевого файла, который находится по адресу: <http://transactions.krc.karelia.ru/section.php?id=755>.

Статья в файле с расширением .pdf загружается на сайт журнала <http://journals.krc.karelia.ru>. Исходный файл с расширением .tex и необходимые рисунки загружаются на 4-м шаге «Загрузка дополнительных файлов».

Обязательные элементы рукописи располагаются в следующем порядке:

УДК курсивом в левом верхнем углу первой страницы; заглавие статьи на русском языке полужирным шрифтом; инициалы и фамилии авторов на русском языке полужирным шрифтом; полное название и полный почтовый адрес организации – места работы каждого автора в именительном падеже на русском языке курсивом (если авторов несколько и работают они в разных учреждениях, следует отметить арабскими цифрами соответствие фамилий авторов аффилированным организациям; автора, ответственного за переписку, следует отметить звездочкой и указать в аффилиации его электронный адрес); аннотация на русском языке; ключевые слова на русском языке; указание источников финансирования выполненных исследований на русском языке.

Далее располагаются все вышеуказанные элементы на английском языке.

Текст статьи (статьи экспериментального характера, как правило, должны иметь разделы: **Введение. Материалы и методы. Результаты и обсуждение. Выводы** либо **Заключение**); благодарности; списки литературы на языке оригинала (**Литература**) и на английском языке (**References**); таблицы на русском и английском языках (на отдельных листах); рисунки (на отдельных листах); подписи к рисункам на русском и английском языках (на отдельном листе).

На отдельном листе дополнительные сведения об авторах: фамилии, имена, отчества всех авторов полностью на русском и английском языке; должности, ученые звания, ученые степени авторов; адрес электронной почты каждого автора; можно указать телефон для контакта редакции с авторами статьи.

ЗАГЛАВИЕ СТАТЬИ должно точно отражать ее содержание и состоять из 8–10 значащих слов.

АННОТАЦИЯ должна быть лишена вводных фраз, создавать возможно полное представление о содержании статьи и иметь объем не менее 200 слов. Рукопись с недостаточно раскрывающей содержание аннотацией может быть отклонена.

Отдельной строкой приводится перечень КЛЮЧЕВЫХ СЛОВ (как правило, не менее пяти). Ключевые слова или словосочетания отделяются друг от друга точкой с запятой, в конце точка не ставится.

Раздел «Материалы и методы» должен содержать сведения об объекте исследования с обязательным указанием латинских названий и сводок, по которым они приводятся, авторов классификаций и пр. Транскрипция географических названий должна соответствовать атласу последнего года издания. Единицы физических величин приводятся по Международной системе СИ. Желательна статистическая обработка всех количественных данных. Необходимо возможно точнее обозначать местонахождения (в идеале – с точным указанием географических координат).

Изложение результатов должно заключаться не в пересказе содержания таблиц и графиков, а в выявлении следующих из них закономерностей. Автор должен сравнить полученную им информацию с имеющейся в литературе и показать, в чем заключается ее новизна. На табличный и иллюстративный материал следует ссылаться так: на рисунки, фотографии и таблицы в тексте (рис. 1, рис. 2, табл. 1, табл. 2 и т.д.), фотографии, помещаемые на вклейках (рис. I, рис. II). Обсуждение завершается формулировкой в разделе «Заключение» основного вывода, которая должна содержать конкретный ответ на вопрос, поставленный во «Введении». Ссылки на литературу в работах **серии «Математическое моделирование и информационные технологии»** даются цифрами.

ТАБЛИЦЫ нумеруются в порядке упоминания их в тексте, каждая таблица имеет свой заголовок. Заголовки таблиц, заголовки и содержание столбцов, строк, а также примечания приводятся на русском и английском языках. Диаграммы и графики не должны дублировать таблицы. Материал таблиц должен быть понятен без дополнительного обращения к тексту. Все сокращения, использованные в таблице, поясняются в Примечании, расположенном под ней. При повторении цифр в столбцах нужно их повторять, при повторении слов – в столбцах ставить кавычки. Таблицы могут быть книжной или альбомной ориентации.

РИСУНКИ представляются отдельными файлами с расширением TIFF (*.TIF) или JPG. При первичной подаче материала в редакцию рисунки вставляются в общий текстовый файл. При сдаче материала, принятого в печать, все рисунки должны быть представлены в виде отдельных файлов в вышеуказанном формате. Графические материалы могут быть снабжены указанием желательного размера рисунка, пожеланиями и требованиями к конкретным иллюстрациям. На каждый рисунок должна быть как минимум одна ссылка в тексте.

ПОДПИСИ К РИСУНКАМ приводятся на русском и английском языках, должны содержать достаточную информацию для того, чтобы приводимые данные могли быть понятны без обращения к тексту (если эта информация уже не дана в другой иллюстрации). Аббревиации расшифровываются в подрисунковых подписях, детали на рисунках следует обозначать цифрами или буквами, значение которых также приводится в подписях.

СОКРАЩЕНИЯ. Разрешаются лишь общепринятые сокращения – названия мер, физических, химических и математических величин и терминов и т. п. Все прочие сокращения должны быть расшифрованы, за исключением небольшого числа общеупотребительных.

БЛАГОДАРНОСТИ. Располагаются после основного текста статьи отдельным абзацем, в котором авторы выражают признательность частным лицам, сотрудникам учреждений и организациям, оказавшим содействие в проведении исследований и подготовке статьи.

ИНФОРМАЦИЯ О КОНФЛИКТЕ ИНТЕРЕСОВ. При подаче статьи авторы должны раскрыть потенциальные конфликты интересов, которые могут быть восприняты как оказавшие влияние на результаты или выводы, представленные в работе. Если конфликт интересов отсутствует, следует об этом сообщить в отдельной формулировке.

СПИСОК ЛИТЕРАТУРЫ оформляется по ГОСТ Р 7.0.5-2008. Источники располагаются в алфавитном порядке. Все ссылки даются на языке оригинала (названия на японском, китайском и других языках, использующих нелатинский шрифт, пишутся в русской транскрипции). Сначала приводится список работ на русском языке и на языках с близким алфавитом (украинский, болгарский и др.), а затем – работы на языках с латинским алфавитом. В списке литературы между инициалами авторов ставится пробел.

REFERENCES. Приводится отдельным списком, повторяя все позиции основного списка литературы. Библиографические записи источников оформляются согласно стилю Vancouver (см. примеры в ГОСТ Р 7.0.7-2021 и образцы ниже). Заголовки русскоязычных работ приводятся на английском языке; для журналов и сборников, в которых размещены цитируемые работы, указывается параллельное английское наименование (при его наличии) либо русскоязычное наименование приводится в латинской транслитерации (вариант BSI) с переводом на английский язык. Прочие элементы библиографической записи приводятся на английском языке (русскоязычное название издательства транслитерируется). При наличии переводной версии источника в References желательно указать ее. Библиографические описания прочих работ приводятся на языке оригинала.

Для каждого источника обязательно указание DOI при его наличии; если приводится адрес интернет-страницы источника (URL), нужно указать дату обращения к ней.

Transactions of the Karelian Research Centre of the Russian Academy of Sciences
No. 6, 2026
MATHEMATICAL MODELING AND INFORMATION TECHNOLOGIES

TABLE OF CONTENTS

A. V. Borodina. A GRAPH OF SOCIAL CONNECTIONS OF RUSSIAN WRITERS: A NETWORK ANALYSIS OF THE LITERARY COMMUNITY BASED ON THE UNIFIED STATE EXAM (USE) CODIFIER . . .	5
K. V. Gubareva, E. Yu. Prosviryakov, A. V. Eremin. UNSTEADY ISOBARIC FLOW OF A VISCOUS INCOMPRESSIBLE FLUID IN A THIN LAYER WITH PERMEABLE BOUNDARIES	20
Yu. V. Zaika, E. K. Kostikova. PARAMETRIC IDENTIFICATION OF A HYDROGEN THERMAL DESORPTION MODEL BY THE CONJUGATE EQUATIONS METHOD	32
A. V. Ivanov. ON THE TOPOLOGICAL DIMENSION OF MAXIMAL LINKED SYSTEMS	47
A. N. Kirillov, A. M. Sazonov. CONTROLLED DYNAMICS OF THE BIOLOGICAL WASTEWATER TREATMENT PROCESS UNDER VARIABLE POLLUTION INPUT FLOW	54
N. N. Kukushkina. POTENTIAL IN A NON-COOPERATIVE TASK SCHEDULING GAME WITH LINEAR EXTERNALITIES	61
Yu. L. Pavlov. MORE ON THE SIZES OF TREES IN THE GALTON–WATSON FOREST	73
Yu. L. Pavlov. ON THE RATE OF NORMAL APPROXIMATION FOR THE DISTRIBUTION OF THE NUMBER OF TREES OF A GIVEN SIZE IN A GALTON–WATSON FOREST	82
A. A. Pechnikov. ANALYSIS OF THE WORLD TRADE NETWORK USING THE FRIENDSHIP INDEX AND DISPARITY FILTER	89
A. A. Rogov, K. A. Kulakov, N. D. Moskin. DECISION TREES WITH A LINEAR COMBINATION OF ATTRIBUTES	100
I. A. Cheplyukova. ON THE GLOBAL CLUSTERING COEFFICIENT OF A CONDITIONAL CONFIGURATION GRAPH	106
I. A. Cheplyukova. LIMIT THEOREMS FOR THE NUMBER OF VERTICES WITH A GIVEN DEGREE IN A CONDITIONAL CONFIGURATION GRAPH	114
A. N. Chuprunov, K. N. Iakovlev. ON THE CONVERGENCE OF GIVEN VALUE CELLS NUMBERS TO AN EMPIRICAL PROCESS	127
N. Yu. Enatskaya. PROBABILISTIC RANDOM NUMBER OF PARTICLES CELL PLACEMENT MODELS	132
N. Yu. Enatskaya. COMBINATORIAL ANALYSIS OF PARTICLE ALLOCATION SCHEMES WITH A GIVEN MAXIMUM CELL FILLING LEVEL	139
BEREAVEMENTS	
In memory of Anatoli D. Sorokin (1936–2025)	148
INSTRUCTIONS FOR AUTHORS	150

Научный журнал

**Труды Карельского научного центра
Российской академии наук**
№ 6, 2026

**МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ
И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ**

*Печатается по решению Ученого совета
Федерального исследовательского центра
«Карельский научный центр Российской академии наук»*

Выходит 12 раз в год

Издание зарегистрировано Федеральной службой по надзору в сфере связи,
информационных технологий и массовых коммуникаций
Регистрационная запись ПИ № ФС 77-72429 от 28.02.2018 г.

Редактор А. И. Мокеева
Компьютерная верстка Л. Э. Бюркланд

Подписано в печать 23.06.2026. Дата выхода 30.06.2026. Формат 60x84^{1/8}.
Печать офсетная. Уч.-изд. л. 13,3. Усл. печ. л. 17,9.
Тираж 100 экз. Заказ 901. Цена свободная

Учредитель и издатель: Федеральное государственное бюджетное учреждение науки
Федеральный исследовательский центр «Карельский научный центр Российской академии наук»
185910, г. Петрозаводск, ул. Пушкинская, 11

Оригинал-макет: Редакция научного издания «Труды КарНЦ РАН»

Типография: Редакционно-издательский отдел КарНЦ РАН
185030, г. Петрозаводск, пр. А. Невского, 50