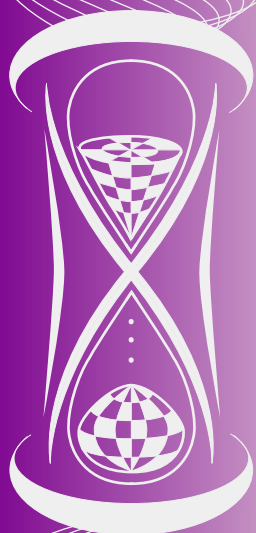


ISSN 2222-8772

МАТЕМАТИЧЕСКИЕ
СТРУКТУРЫ
И
МОДЕЛИРОВАНИЕ

№ 2 (74)
2025



**МИНИСТЕРСТВО НАУКИ
И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ**

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ОБРАЗОВАНИЯ
«ОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМ. Ф.М. ДОСТОЕВСКОГО»**

**МАТЕМАТИЧЕСКИЕ
СТРУКТУРЫ
И
МОДЕЛИРОВАНИЕ**

№ 2 (74)

**Омск
2025**

Редакционная коллегия

А. К. Гуц	главный редактор , председатель редакционной коллегии, доктор физ.-мат. наук, профессор, заслуженный работник высшей школы Российской Федерации, профессор кафедры информационных технологий и математики Сочинского государственного университета, г. Сочи
Н. Ф. Богаченко	зам. гл. редактора , технический редактор, канд. физ.-мат. наук, доцент, зав. кафедрой компьютерных технологий и сетей Омского государственного университета им. Ф. М. Достоевского, г. Омск
Д. Н. Лавров	зам. гл. редактора , выпускающий редактор, канд. техн. наук, доцент, проректор по образовательной деятельности и информатизации Нижневартковского государственного университета, г. Нижневартовск
В. В. Варламов	доктор физ.-мат. наук, профессор кафедры прикладной математики и информатики Сибирского государственного индустриального университета, г. Новокузнецк
В. П. Голубятников	доктор физ.-мат. наук, профессор, главный научный сотрудник Института математики СО РАН им. С. Л. Соболева, г. Новосибирск
С. И. Горлов	доктор физ.-мат. наук, профессор, ректор Нижневартковского государственного университета, г. Нижневартовск
А. Г. Гринь	доктор физ.-мат. наук, профессор, профессор кафедры информационной безопасности Омского государственного университета им. Ф. М. Достоевского, г. Омск
В. А. Еровенко	доктор физ.-мат. наук, профессор, г. Минск, Республика Беларусь
B. Zilber	Dr.Sc. (Phys.-Math.), Professor of Mathematical Logic, Mathematical Institute, University of Oxford, United Kingdom
А. Н. Кабанов	канд. физ.-мат. наук, доцент кафедры компьютерных технологий и сетей Омского государственного университета им. Ф. М. Достоевского, г. Омск
А. В. Копыльцов	доктор техн. наук, профессор, профессор кафедры информационных систем Санкт-Петербургского государственного электротехнического университета «ЛЭТИ» им. В. И. Ульянова (Ленина), г. Санкт-Петербург
А. Г. Коробейников	доктор техн. наук, профессор, зам. директора по науке Санкт-Петербургского филиала Института земного магнетизма, ионосферы и распространения радиоволн им. Н. В. Пушкина РАН, г. Санкт-Петербург
П. А. Корчагин	доктор техн. наук, профессор, ректор Омского государственного технического университета, г. Омск
V. Kreinovich	Ph.D. (Phys.-Math.), Professor, Computer Science Department, University of Texas at El Paso, Texas, USA
В. А. Плетюхов	доктор физ.-мат. наук, профессор, профессор кафедры общей и теоретической физики Брестского государственного университета им. А. С. Пушкина, г. Брест, Республика Беларусь
Л. Б. Соколинский	доктор физ.-мат. наук, профессор, зав. кафедрой системного программирования Южно-Уральского государственного университета (национального исследовательского университета), г. Челябинск
A. A. Fedorenko	Ph.D. (Phys.-Math.), Researcher (CR1) at the French National Centre of Scientific Research (CNRS) Laboratoire de Physique de l'ENS-Lyon, France
A. Jadczyk	Ph.D., Professor, Researcher, Laboratoire de Physique, Universite de Toulouse III et CNRS, France, and Ronin Institute, Montclair, New Jersey, USA

Учредитель

Федеральное государственное автономное образовательное учреждение высшего образования
«Омский государственный университет им. Ф. М. Достоевского»

Журнал зарегистрирован в Федеральной службе по надзору в сфере связи, информационных технологий и массовых коммуникаций.
Реестровая запись ПИ № ФС77-72200 от 15 января 2018 г.

Адрес редакции, издателя и типографии

644077, Омская обл., г. Омск,
пр-т Мира, д. 55а.

Дата выхода в свет: 10.06.2025.

Тираж 25 экз.

Свободная цена.

МАТЕМАТИЧЕСКИЕ СТРУКТУРЫ И МОДЕЛИРОВАНИЕ

Журнал основан в 1998 году. В журнале публикуются статьи, в которых излагаются результаты исследований по математике, теоретической физике, компьютерным наукам, философии и истории математики и информатики.

Все статьи журнала проходят обязательное рецензирование. Рефераты статей журнала опубликованы в «Реферативном журнале», «Mathematical Reviews» (США), «Zentralblatt für Mathematik» (Германия) и в РИНЦ (elibrary.ru). Журнал входит в Перечень рецензируемых научных изданий ВАК РФ, входящих в Международные реферативные базы данных (МРБД) и системы цитирования (Список от 31.12.2023, № 769), в которых должны быть опубликованы основные результаты диссертаций на соискание учёных степеней (Приказ Минобрнауки России от 25.07.2014 № 793). Имеет категорию K1.

Все статьи в журнале публикуются под лицензией Attribution 4.0 International (CC-BY).

Электронная версия журнала представлена в сети:

<http://msm.omsu.ru>

<http://msm.univer.omsk.su>

Подписной индекс по каталогу «Пресса России»: 94082

Электронная почта главного редактора:

aguts@mail.ru

Электронная почта выпускающего редактора:

dmtrl@ya.ru

СОДЕРЖАНИЕ

К 100-летию со дня рождения Ю.Ф. Борисова

А.К. Гуц. *Хроногеометрия профессора Ю.Ф. Борисова* 5

Фундаментальная математика и физика

И.П. Бесценный. *Об индуктивных исчислениях и вероятностной семантике* 12

В.В. Варламов. *Теоретико-групповое описание периодической системы элементов: VI. Восьмимерное обобщение* 18

А.К. Гуц. *Гравитация как квантовая запутанность массивных тел* 33

G.K. Sokolova. *On the Kirchhoff index for circulant graph with non-fixed jumps* 38

Прикладная математика и моделирование

С.А. Агалаков, А.А. Березин. *Повышение эффективности классификации фенотипов заболевания желудочно-кишечного тракта с помощью методов обработки данных* 45

А.А. Бугаенко, И.Ю. Крутова. *Численное моделирование решения полной системы уравнений Навье – Стокса при учёте действия силы Кориолиса* 55

А.К. Гуц. *Топологические полёты космических варп-кораблей Алькубьерре и потоки Риччи* 61

С.Л. Дерябин, А.П. Садов. *Задача о разрушении дамбы в модели мелкой воды* 72

А.А. Романова. *Сложность задачи календарного планирования со складываемыми ресурсами переменной стоимости* 83

Компьютерные науки

Т.Е. Болдовская, М.И. Ремизов, А.Е. Ветров. *Систематизация и сохранение исторических источников в условиях цифровой эпохи* 90

Продолжение на следующей странице

Наши публикации



С.В. Гусс. *Научно-техническое исследование программной составляющей беспилотных авиационных систем* 98

Т.В. Костеннов. *Использование нейронных сетей для создания системы принятия решений управления микроклиматом тепличной системы* 107

Т.М. Опарина. *Систематизация инструментов и методов для эффективной репликации базы данных* 117

С.Н. Чуканов, И.С. Чуканов, С.В. Лейхтер. *Обучение гранулярных когнитивных карт методом APSO* 123

Информационная безопасность

Т.В. Вахний, С.В. Вахний. *Теория игр в киберпространстве: рефлексивный подход к моделированию взаимодействия администратора и злоумышленника* 131

С.В. Усов. *О представлении некоторых ролевых моделей разграничения доступа объектно-ориентированной моделью HRU: наследование иерархически охватным листовым подходом* 138

Страница памяти

Ремесленников Владимир Никанорович (1938–2025)..... 146

ХРОНОГЕОМЕТРИЯ ПРОФЕССОРА Ю.Ф. БОРИСОВА

А.К. Гуц

д.ф.-м.н., профессор, e-mail: aguts@mail.ru

Сочинский государственный университет, Сочи, Россия

Аннотация. Замечательному новосибирскому геометру Ю.Ф. Борисову 15 июня 2025 г. исполнилось бы 100 лет. В статье излагается его подход к аксиоматизации специальной теории относительности, который он разрабатывал в течение 40 лет.

Ключевые слова: Ю.Ф. Борисов, Мир событий, хроногеометрия, аксиоматика, теория относительности.

К 100-летию со дня рождения
профессора Ю.Ф. Борисова

Когда пишут о научных исследованиях замечательного геометра Юрия Федоровича Борисова (1925–2007), а пишут о них геометры, как правило, не интересующиеся физикой или философией, то отмечают его основные геометрические работы, относящиеся к теории многообразий ограниченной кривизны и к проблеме регулярности гладких изометрических погружений таких многообразий в евклидовы пространства. Затем добавляют, что занимался профессор Борисов и основаниями теории относительности, или, как говорили в Новосибирске, *хроногеометрией* [1].

К хроногеометрии относятся следующие публикации Ю.Ф. Борисова:

1. Борисов Ю.Ф. О преобразованиях псевдоевклидова пространства, сохраняющих изотропность векторов // Известия высших учебных заведений. 1960. № 6. С. 31–39.
2. Борисов Ю.Ф. Об аксиоматическом определении групп Галилея и Лоренца // Сибирский математический журнал. 1978. Т. 19, № 6. С. 1237–1253.
3. Борисов Ю.Ф. Важное направление – хроногеометрия // Наука в Сибири. 12 января 1978 г.
4. Борисов Ю.Ф. К основаниям релятивистской кинематики // Сибирский математический журнал. 1986. Т. 27, № 3. С. 10–27.
5. Борисов Ю.Ф. Однозначная (с точностью до ориентации) определяемость времени пространственной структурой множества событий // Сибирский математический журнал. 1987. Т. 28, № 4. С. 57–64.
6. Борисов Ю.Ф. В четвёртом измерении // Наука в Сибири. 30 июня 1987 г.
7. Борисов Ю.Ф. Два вопроса о метрических основаниях римановой геометрии // Доклады РАН. 1994. Т. 336, № 2. С. 154–156.
8. Борисов Ю.Ф. Вклад А.Д. Александрова и его школы в прикладную математику // Сибирский журнал индустриальной математики. 2002. Т. 5, № 3. С. 3–4.
9. Борисов Ю.Ф. Вывод релятивистской кинематики из постулатов классической механики и ограниченной точности мыслимых экспериментов // Сибирский журнал индустриальной математики. 2008. Т. 11, № 1. С. 23–36.

10. Борисов Ю.Ф. Экспериментальное обоснование теории относительности, идеализм, позитивизм и материализм // Математические структуры и моделирование. 2012. № 25. С. 63–65.¹

Профессор Ю.Ф. Борисов, пережив блокаду Ленинграда, закончил Ленинградский университет в 1949 г., и дальнейшая его жизнь тесно связана с жизнью его учителя и друга Александра Даниловича Александрова. Геометрические исследования Ю.Ф. Борисова пересекаются во многом с идеями А.Д. Александрова. А эти идеи сформировали важнейшее направление геометрии XX в. – обобщённую риманову геометрию, известную сейчас как пространства Александрова.

Великий геометр А.Д. Александров – физик по образованию – проявил незаурядное предвидение реальности несиловой квантовой связи тел (запутанности) в квантовой механике [2], но начиная с середины 1960-х гг. погрузился в исследования по аксиоматической теории относительности, названные им хроногеометрией.

Ю.Ф. Борисов в 1960 г., как бы следуя за учителем², который в данное время больше интересовался основаниями теории относительности с акцентом на философскую сторону дела, обобщил теорему Александрова – Овчинниковой (анонсированную ещё в 1949 г. и полностью опубликованную в 1953 г.) на случай произвольного псевдоевклидова пространства.

Результат Ю.Ф. Борисова открывал путь аксиоматизации теории пространства-времени любой размерности и любой сигнатуры.

Схема аксиоматизации на тот момент была простой: рассматриваются два аффинных n -мерных пространства A^n и $\hat{A}^n$, в которых заданы семейства равных и параллельных конусов $\mathcal{C} = \{C_x : x \in A^n\}$ и $\hat{\mathcal{C}} = \{\hat{C}_{\hat{x}} : \hat{x} \in \hat{A}^n\}$, где $x, \hat{x}$ – вершины конусов C_x и $\hat{C}_{\hat{x}}$. Суть аксиоматизации состояла в том, чтобы вычислить группу $\mathcal{G}$ биективных отображений $f : A^n \rightarrow \hat{A}^n$, обладающих свойством $f(C_x) = \hat{C}_{f(x)}$.

В случае 4-мерных пространств $A^4, \hat{A}^4$ и эллиптических конусов группа $\mathcal{G}$, а это и есть теорема Александрова – Овчинниковой, совпадает с группой Пуанкаре, на которой построена специальная теория относительности.

На языке физиков аффинные пространства – это инерциальные системы отсчёта с координатами $\{x, y, z\}$ и со своим локальным временем t (со своими часами), а отображения f описывают пересчёты аффинных координат из одной системы отсчёта в другую. Смысл конусов прост: это либо уравнение распространения света

$$c^2 t^2 - x^2 - y^2 - z^2 = 0, \quad (1)$$

либо условие причинности

$$c^2 t^2 - x^2 - y^2 - z^2 = 0 \geq 0. \quad (2)$$

Наконец, равенство $f(C_x) = \hat{C}_{f(x)}$ означает либо закон постоянства скорости света, либо нерушимость причинно-следственных связей. Точнее, если в одной системе

¹Эта публикация появилась после смерти Ю.Ф. Борисова и представляет изложение его доклада на семинаре «Хроногеометрия» в Новосибирском университете в 197.. г. в той форме, как он запомнился А.К. Гуцу.

²За своим учителем Ю.Ф. Борисов следовал не только в науке, но и в жизни, переехав вместе с ним из Ленинграда в Сибирь, в новосибирский Академгородок. Они расстались в 1986 г., когда А.Д. Александров вернулся в Ленинград, а Ю.Ф. Борисов остался до конца жизни в Сибири.

отсчёта наблюдалось, что свет распространяется по закону (1), а после перехода в другую инерциальную систему отсчёта закон распространения света сохраняется, то пересчёт координат описывается преобразованием Лоренца.

Вне нашего внимания остаётся физическая сущность использованных аффинных пространств A^4 , $\hat{A}^4$. Они всего лишь удобная геометрическая конструкция. Так, во всяком случае, проповедуется в весьма популярном курсе теоретической физики Л.Д. Ландау и Е.М. Лифшица.

Если мы отождествляем пространства A^4 , $\hat{A}^4$, т. е. полагаем, что $A^4 = \hat{A}^4$, и, следовательно, имеем преобразование $f : A^4 \rightarrow A^4$ такое, что $f(C_x) = C_{f(x)}$, то мы имеем *движение* в *некотором* пространстве A^4 , которое сохраняет закон (1), т. е. соответствует принципу постоянства скорости света, либо сохраняет наблюдаемые причинно-следственные связи. Естественен вопрос: что это за пространство такое A^4 ? Ответ дал в 1908 г. Г. Минковский: это единое *реально существующее пространство-время*, и это есть *Мир событий*³. А. Эйнштейн с этим согласился лишь в 1921 г.

Пространство-время *абсолютно* [3], поскольку его метрика, т. е. способ вычисления расстояния между его *мировыми точками* (x, y, z, t) , как писал Г. Минковский, названными позже *событиями*,

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2 \quad (3)$$

не зависит от выбранной наблюдателем для его измерения (инерциальной) системы отсчёта, или, иначе, не меняется при преобразованиях, сохраняющих закон постоянства скорости света, т. е. при преобразованиях Лоренца, в то время как меняются расстояние

$$dl^2 = dx^2 + dy^2 + dz^2$$

между точками в пространстве и измеренные промежутки времени dt^2 между событиями. Расстояния и промежутки времени *относительны*, интервал (3) *абсолютен*! Значит, относительны и пространство, и время.

Обратим внимание, что мы представили формулировку А.Д. Александрова аксиоматики теории относительности. В ней исходным является 4-мерное аффинное пространство, т. е. пространство-время, объединившее в себе 3-мерное пространство и 1-мерное время. Для физиков система отсчёта – это не 4-мерное пространство, а 3-мерное пространство, в котором течёт своё локальное время. Иначе говоря, 4-мерное пространство-время у них исчезает, отходит на второй план, не абсолютно, реально пространство и время, тогда как у А.Д. Александрова, следовавшего постулату Г. Минковского, *реальным является 4-мерное пространство-время, а пространство и время – это всего лишь «тени»*.

Судьёй, решающим, кто прав, в физике всегда будет эксперимент: именно он сравнивает устройство реального внешнего мира и устройство этого же мира, как он видится физиками, иными словами, устройство внутреннего мира в сознании людей. На сегодня экспериментов, измеряющих абсолютность реальность Мира событий, т. е. единого пространства-времени, увы, крайне мало [1].

³Точнее, сам Г. Минковский говорил о *Мире*, состоящем из *мировых точек*.

Профессор Ю.Ф. Борисов строит свою хроногеометрию, в отличие от хроногеометрии А.Д. Александрова, не с божественных позиций, откуда виден лежащий внизу Мир событий, а с позиций живущих в этом Мире физиков, которым для проведения экспериментов, подтверждающих реальность их Мира, а значит и их самих, в силу способа их жития нужны системы отсчёта и часы, тикающие в этих системах отсчёта. Но основой для аксиоматики для Ю.Ф. Борисова, конечно же, является 4-мерный Мир событий.

Кратко Юрий Федорович так описал свой подход: «Если существует достаточно обширное множество инерциальных движений, евклидовости пространства, связанного с телом системы отсчёта, и верен принцип относительности Галилея для инерциальных движений, то имеет место альтернатива: либо все инерциальные системы отсчёта связаны с преобразованиями Галилея, либо все они связаны с преобразованиями Лоренца (тем самым – любое достоверно установленное противоречие с представлениями о пространстве и времени однозначно приводят к теории относительности, если не подвергать сомнению перечисленные выше предположения)» [3].

Свою хроногеометрию Ю.Ф. Борисов изложил в статье в 1978 г. Затем совершенствовал её в течение 40 лет (в трёх статьях, последняя [4] стала последней публикацией в его жизни). Его аксиоматика носит кинематический характер. Иначе говоря, он исходит из описания движения тел, точнее, он рассматривает инерциальные движения тел, которые используются как системы отсчёта, без выяснения причин, которые вызвали это движение. Как уже было сказано, основой для определения движений для него являются события, составляющие Мир событий $\mathcal{M} = \{x, y, \dots\}$. Пространства появляются как наборы *одноместных* событий, т. е. это те события, которые происходят, происходили или произойдут в *одном месте* пространства; а время появляется как наборы *одновременных* событий.

Пространство привязано к «допустимому телу» системы отсчёта, находящемуся в инерциальном движении. Множество «допустимых тел» обозначено через $\mathcal{T}$. Если $T_1, T_2 \in \mathcal{T}$ и $T_1 \neq T_2$, то отношения одноместности в T_1, T_2 различны, а именно хотя бы одна частица T_2 меняет с течением времени своё место в теле T_1 , т. е. T_2 движется относительно T_1 .

Вводятся два бинарных предиката⁴:

- 1) $\Pi_T(x, y)$ – события x и y одноместны в системе отсчёта T ;
- 2) $B_T(x, y)$ – события x и y одновременны в системе отсчёта T .

Тогда если $x \in \mathcal{M}$, то $x_{\Pi, T} = \{y \in \mathcal{M} | \Pi_T(x, y)\}$ и $x_{B, T} = \{y \in \mathcal{M} | B_T(x, y)\}$ являются классами событий, одноместных и соответственно одновременных с событием x .

Для всякого $T \in \mathcal{T}$ обозначим через

$$\Pi_T = \{x_{\Pi, T} | x \in \mathcal{M}\} \text{ и } B_T = \{x_{B, T} | x \in \mathcal{M}\}$$

– это множество, определяемых с помощью T мест и моментов местного времени всех событий, описанные на языке пространства событий $\mathcal{M}$.

⁴Мы меняем обозначения Ю.Ф. Борисова; нам они показались не совсем удачными.

Таким образом, у нас появились (относительные, привязанные к системе отсчёта, к телу T) пространство и время.

Если T^* – бесконечно простирающееся твёрдое тело, реализующее допустимое тело отсчёта T , то для наблюдателя, считающего тело T^* неподвижным, Π_T – множество точек окружающего пространства, а B_T – множество показаний часов, пребывающих в T^* .

Постулат Евклида. Множества Π_T и B_T мест и моментов местного времени событий, определяемые допустимыми телами отсчёта $T \in \mathcal{T}$, наделены согласованными структурами 3-мерного евклидова пространства E^3 и ориентированного 1-мерного евклидова пространства E_+^1 соответственно. Получаем пространства $\Pi_T(E^3)$ и $B_T(E_+^1)$.

Пусть $T \in \mathcal{T}$ и κ_Π – прямоугольная система координат с единичными масштабами на осях в $\Pi_T(E^3)$, а κ_B – декартова система координат в $B_T(E_+^1)$ с единичным масштабом и направлением оси, соответствующим ориентации пространства E_+^1 .

Для любого $x \in \mathcal{M}$ положим $x_{\Pi,T} = (x, y, z)$, $x_{B,T} = t$, $\kappa(x) = (x, y, z, t)$.

Отображение $\kappa : \mathcal{M} \rightarrow \mathbb{R}^4$, заданное формулой $\kappa(x) = (x, y, z, t)$ при любых x и фиксированных T , κ_Π и κ_B , как легко заметить, является биекцией.

Всякую биекцию $\kappa : \mathcal{M} \rightarrow \mathbb{R}^4$ указанного вида называем канонической системой координат в $\mathcal{M}$, соответствующей допустимому телу отсчёта T . Множество всех таких систем обозначаем через $K(T)$. Если $\kappa \in K(T)$, $x \in \mathcal{M}$ и $\kappa(x) = (x, y, z, t)$, то x, y, z называются пространственными координатами, а t – временной координатой события x в системе κ .

Множество $\bigcup_{T \in \mathcal{T}} K(T)$ всех канонических систем координат в множестве событий $\mathcal{M}$ обозначаем через $\mathcal{K}$.

Как известно, множество $\{\kappa' \circ \kappa^{-1} | \kappa', \kappa \in \mathcal{K}\}$ в классической механике совпадает с множеством преобразований Галилея, а в механике теории относительности – с множеством лоренцевых преобразований, содержащих некоторую константу $c > 0$. Однако в обоих случаях верна следующая теорема, которую мы принимаем за постулат.

Постулат Галилея. Для любых $T_1, T_2 \in \mathcal{T}$

$$\{\kappa_1 \circ \kappa_2^{-1} | \kappa_1 \in K(T_1), \kappa_2 \in K(T_2)\} = \{\kappa_2 \circ \kappa_1^{-1} | \kappa' \in K(T_1), \kappa \in K(T_2)\},$$

$$\{\kappa_1 \circ \kappa^{-1} | \kappa_1 \in K(T_1), \kappa \in \mathcal{K}\} = \{\kappa_2 \circ \kappa^{-1} | \kappa_2 \in K(T_2), \kappa \in \mathcal{K}\}.$$

Содержание этого постулата можно пересказать так: описание движений одного инерциального тела относительно другого не зависит от того, в каком из них находится наблюдатель, а описание движений всех инерциальных тел относительно одного из них не зависит от выбора последнего (центра вселенной).

Определение 1. Под кинематикой понимается пара $(\mathcal{M}, G^*(\mathcal{M}))$, где $G^*(\mathcal{M})$ – группа преобразований Мира событий $\mathcal{M}$, взаимно однозначно связанная с группой G всех преобразований $\kappa_2 \circ \kappa_1^{-1}$ канонических систем координат $\kappa_2, \kappa_1 \in \mathcal{K}$ условием

$$f \in G^*(\mathcal{M}) \Leftrightarrow (\kappa_0 \in \mathcal{K} \Rightarrow \kappa_0 \circ f \circ \kappa_0^{-1} \in G).$$

Поэтому кинематикой можно называть также пару $(\mathcal{M}, G)$.

Определение 2. Кинематику $(\mathcal{M}, G)$ называем классической, если G – группа преобразований Галилея, и релятивистской, если G – группа Лоренца.

Теорема 1. Кинематика, удовлетворяющая постулатам Евклида и Галилея, является либо классической кинематикой Галилея, либо релятивистской кинематикой Лоренца, и указанные возможности исключают друг друга.

Ю.Ф. Борисов в своей последней статье [4] интересуется, можно ли, проводя эксперименты, обосновать исторический выбор в пользу релятивистской кинематики. Всегда, говоря об этом, ссылаются на опыт Майкельсона.

Линейное преобразование $\mathbb{R}^4$ с матрицей (a_{ij}) , $i, j = 1, 2, 3, 4$, называем тривиальным, если матрица (a_{ij}) , где $i, j = 1, 2, 3$, ортогональная, $a_{i4} = a_{4i} = 0$ при $i = 1, 2, 3$ и $a_{44} = 1$.

Множество всех тривиальных преобразований обозначаем через $\mathbb{A}$.

Множество преобразований Галилея обозначим через Γ , а лоренцевых – через $L(c)$. Легко убедиться, что

$$\Gamma \cap L(c) = \mathcal{A}.$$

Теорема 2. Пусть $(\mathcal{M}, G)$ – кинематика, $\phi \in G \setminus \mathcal{A}$, (a_{ij}) – матрица ϕ , $P(\phi) = \sum_{j=1}^3 a_{4j}^2$. Если $P(\phi) = 0$, то кинематика $(\mathcal{M}, G)$ классическая, а если $P(\phi) > 0$, то релятивистская.

В силу ограниченной точности любых экспериментов возможность определения вида кинематики с помощью опыта накладывает на коэффициенты всякого линейного преобразования $\phi \in G \setminus \mathcal{A}$, где G – группа обнаруженного вида, следующее условие: никакое преобразование из $\tilde{G} \setminus \mathcal{A}$, где $\tilde{G}$ – группа другого вида, нельзя получить за счёт (достаточно) малых изменений коэффициентов рассматриваемого преобразования. Действительно, в противном случае не исключено, что более точный опыт даёт противоположный результат.

Согласно теореме 2 для всякого $f \in G \setminus \mathcal{A}$ число $P(f)$ служит индикатором кинематики $(\mathcal{M}, G)$. Так как $P(f)$ непрерывно зависит от коэффициентов преобразования f , то сформулированное выше необходимое условие экспериментальной распознаваемости вида кинематики с учётом ограниченной точности любых экспериментов равносильно принятию следующего постулата.

Постулат III. Пусть $(\mathcal{M}, G)$ – кинематика, удовлетворяющая постулатам Евклида и Галилея. Тогда для всякого преобразования $f \in G \setminus \mathcal{A}$ найдётся такое $\varepsilon > 0$, что для любого преобразования $\hat{f} \in G \setminus \mathcal{A}$, удовлетворяющего условию $|P(f) - P(\hat{f})| < \varepsilon$, либо $P(f) = P(\hat{f}) = 0$, либо $P(f) > 0$ и $P(\hat{f}) > 0$.

Теорема 3. Кинематика $(\mathcal{M}, G)$, удовлетворяющая постулатам Евклида и Галилея, тогда и только тогда является релятивистской, когда она удовлетворяет и постулату III. При этом постоянная $c > 0$, фигурирующая в определении лоренцевых преобразований из $G = L(c)$, однозначно определена выбором единиц длины и местного времени, используемых в определении канонических систем координат в $\mathcal{M}$.

Таким образом, отмечает Ю.Ф.Борисов, «согласно теореме 3 экспериментально распознаваемый вид кинематики предпрешён, такая кинематика является релятивистской».

Юрий Федорович сдал в редакцию журнала свою последнюю работу 5 июня 2007 г., посвятив её своей умершей (2004) жене: «Светлой памяти Людмилы Глебовны Борисовой». Он покинул наш Мир 19 октября 2007 г., но вечно продолжает существовать в Мире событий.

Закончить статью, не сказав о том, кем был Юрий Федорович, было бы неправильно. Но рассказать о нём непросто. С.С. Кутателадзе (добавление к статье [4]) сумел это сделать так, что неуместно было бы что-то к этому добавлять:

«Учёные бывают разные. Есть учёные в башнях из слоновой кости и есть учёные безбашенные. Есть учёные с большой дороги, от бога и от природы. Есть учёные члены, многочлены и малочлены. Есть учёные краснобаи, баи и не очень. Есть учёные по убеждениям и есть учёные время от времени.

Юрий Федорович Борисов был учёным на холме. Такой никому не мешает, никогда не чванится и не пыжится, не оскорбляет других ни ненавистью, ни небрежением. Учёный на холме не отражает чужой свет, а любит других, даря всем собственные откровения и излучая внутреннее тепло. Борисов внёс свою лепту и в математику, и в её преподавание, сохраняя достоинство независимости и подавая примеры для подражания».

Литература

1. Гуц А.К. Хроногеометрия. Аксиоматическая теория относительности. М.: УРСС, 2012.
2. Александров А.Д. О парадоксе Эйнштейна в квантовой механике // Доклады АН СССР. 1952. Т. 84, № 2. С. 253–256.
3. Борисов Ю.Ф. Важное направление – хроногеометрия // Наука в Сибири. 12 января 1978 г.
4. Борисов Ю.Ф. Вывод релятивистской кинематики из постулатов классической механики и ограниченной точности мыслимых экспериментов // Сибирский журнал индустриальной математики. 2008. Т. 11, № 1. С. 23–36.

CHRONOGEOMETRY OF PROFESSOR Y.F. BORISOV

A.K. Guts

Dr.Sc. (Phys.-Math.), Professor, e-mail: aguts@mail.ru

Sochi State University, Sochi, Russia

Abstract. The remarkable Novosibirsk geometer Yu.F. Borisov would have turned 100 on June 15, 2025. The article describes his approach to the axiomatization of special relativity, which he has been developing for 40 years.

Keywords: Yu.F. Borisov, the World of events, chronogeometry, axiomatics, Special theory of relativity.

Дата поступления в редакцию: 12.04.2025

ОБ ИНДУКТИВНЫХ ИСЧИСЛЕНИЯХ И ВЕРОЯТНОСТНОЙ СЕМАНТИКЕ

И.П. Бесценный

к.ф.-м.н., доцент, e-mail: ibests@mail.ru

Омский государственный университет им. Ф.М. Достоевского, Омск, Россия

Аннотация. Рассматриваются формальные эпистемические модальные исчисления, аксиомы которых отражают особенности определения достоверности индуктивных выводов. Так как семантика Крипке для них не подходит, то более естественной для индуктивных исчислений является вероятностная семантика.

Ключевые слова: модальная логика, эпистемические исчисления, вероятностная семантика.

Введение

В естественных науках на основании наблюдений и экспериментов выявляются закономерности, которые формулируются в виде эмпирических гипотез. Эти гипотезы впоследствии могут быть подтверждены или опровергнуты дальнейшими наблюдениями и экспериментами. Логический вывод из частных утверждений к общим называется индуктивным. Эпистемические модальные исчисления выражают степени доверия к эмпирическим гипотезам, их аксиомы основаны на осторожном отношении к достоверности индуктивных выводов.

Хорошо изучены свойства алетических и деонтических модальных исчислений, аксиоматика которых тесно связана со свойствами отношения достижимости соответствующей семантики Крипке [1]. Эпистемические модальные исчисления (в частности, рассматриваемые в данной статье индуктивные исчисления) не содержат аксиомы «сильной импликации», которая истинна в любой семантике Крипке. Так как семантика Крипке для них не подходит, то более естественной для индуктивных исчислений является вероятностная семантика.

1. Индуктивные исчисления

Рассмотрим формальные теории, основанные на исчислении высказываний генценовского типа – исчислении секвенций (ИС) [2, с. 11–17]. Будем использовать с сохранением нумерации из [2, с. 12–13] те же основные правила вывода, а также следующие допустимые правила (транзитивности и обращения):

$$\text{T. } \frac{\Gamma_1 \vdash \Phi; \Gamma_2, \Phi \vdash \Psi}{\Gamma_1, \Gamma_2 \vdash \Psi}.$$

$$9a. \frac{\Gamma, \Phi \vdash}{\Gamma \vdash \neg\Phi}.$$

$$9б. \frac{\Gamma \vdash \Phi}{\Gamma, \neg\Phi \vdash}.$$

Формулы индуктивных исчислений определяются как в ИС с добавлением фразы:

- Если Φ – формула, то $\Box\Phi, \Diamond\Phi$ – тоже формулы.

При этом полагаем по определению $\Diamond\Phi \equiv \neg\Box\neg\Phi$. Содержательный смысл формулы $\Box\Phi$ – «гипотеза Φ принимается». Тогда $\Box\neg\Phi$ означает «гипотеза Φ отвергается», а $\Diamond\Phi$ – «гипотеза Φ не отвергается».

Индуктивное исчисление Е получается добавлением в ИС схем аксиом:

$$E0. \vdash \Box(\Phi \rightarrow \Phi),$$

$$E1. \Box\Phi \vdash \Diamond\Phi,$$

$$E2. \Box(\Phi \rightarrow \Psi) \vdash \Box\Phi \rightarrow \Diamond\Psi,$$

и правила введения:

$$BB. \frac{\Phi \vdash \Psi}{\Box\Phi \vdash \Box\Psi}.$$

Аксиома E0 обязывает принимать как гипотезы все логические тавтологии. Из правила введения и аксиомы E0 следует допустимость правила Гёделя:

$$Г. \frac{\vdash \Phi}{\vdash \Box\Phi}.$$

Действительно,

$$\frac{\vdash \Box(\Phi \rightarrow \Phi); \frac{\frac{\vdash \Phi}{\Phi \rightarrow \Phi \vdash \Phi}^{(12)} \quad \Box(\Phi \rightarrow \Phi) \vdash \Box\Phi}{\Box(\Phi \rightarrow \Phi) \vdash \Box\Phi}^{(BB)}}{\vdash \Box\Phi}^{(T)}.$$

Используя правила 9ab и эквивалентность $\Diamond\Phi \equiv \neg\Box\neg\Phi$, нетрудно убедиться, что в Е допустимо также правило вывода:

$$BBa. \frac{\Phi \vdash \Psi}{\Diamond\Phi \vdash \Diamond\Psi}.$$

Аксиома E1 синтаксически совпадает с аксиомой D1 деонтического исчисления D ([3]). Осторожное отношение к достоверности эмпирических гипотез выражается в том, что секвенции вида $\Box\Phi \vdash \Phi$ и $\Phi \vdash \Diamond\Phi$ не доказуемы в исчислении Е. То есть принятие гипотезы не гарантирует её истинности, а истинность гипотезы не гарантирует её неотвергаемости. Однако E1 гарантирует внутреннюю непротиворечивость множества рассматриваемых гипотез: если гипотеза принимается, то она не отвергается.

Но самое существенное отличие индуктивных исчислений от алетических и деонтических содержится в схеме аксиом E2. Сравним её с аксиомой «сильной импликации» $\Box(\Phi \rightarrow \Psi) \vdash \Box\Phi \rightarrow \Box\Psi$, которая в E не доказуема.

Если мы принимаем сложную гипотезу, являющуюся импликацией, и принимаем её посылку, то, в отличие от классических дедуктивных рассуждений (правило *modus ponens*), нет достаточных оснований для принятия заключения, можно лишь утверждать, что заключение не отвергается.

Важным следствием схемы аксиом E2 является следующее утверждение:

Предложение 1. В исчислении E доказуема секвенция $\Box\Phi \wedge \Box\Psi \vdash \Diamond(\Phi \wedge \Psi)$.

Доказательство. В ИС доказуемы секвенции $\Phi, \Psi \vdash \Phi \wedge \Psi$; $\Box\Phi \wedge \Box\Psi \vdash \Box\Phi$; $\Box\Phi \wedge \Box\Psi \vdash \Box\Psi$. Соединяем с E2:

$$\frac{\frac{\Phi, \Psi \vdash \Phi \wedge \Psi}{\Phi \vdash \Psi \rightarrow \Phi \wedge \Psi} (7) \quad \frac{\Phi \vdash \Psi \rightarrow \Phi \wedge \Psi}{\Box\Phi \vdash \Box(\Psi \rightarrow \Phi \wedge \Psi)} (BB) \quad \Box(\Psi \rightarrow \Phi \wedge \Psi) \vdash \Box\Psi \rightarrow \Diamond(\Phi \wedge \Psi) (T)}{\Box\Phi \vdash \Box\Psi \rightarrow \Diamond(\Phi \wedge \Psi)} (T).$$

Теперь:

$$\frac{\Box\Phi \wedge \Box\Psi \vdash \Box\Phi; \Box\Phi \vdash \Box\Psi \rightarrow \Diamond(\Phi \wedge \Psi) (T)}{\Box\Phi \wedge \Box\Psi \vdash \Box\Psi \rightarrow \Diamond(\Phi \wedge \Psi)} (T) \quad \frac{\Box\Phi \wedge \Box\Psi \vdash \Box\Psi; \Box\Phi \wedge \Box\Psi \vdash \Box\Psi \rightarrow \Diamond(\Phi \wedge \Psi)}{\Box\Phi \wedge \Box\Psi \vdash \Diamond(\Phi \wedge \Psi)} (8).$$

■

Это означает, что, принимая каждую из двух гипотез по отдельности, мы не имеем права принять их конъюнкцию, их конъюнкция только лишь неотвергаема.

Отметим, что, используя правила ВВ и ВВа, легко доказываются следующие секвенции: $\Box\Phi \vee \Box\Psi \vdash \Box(\Phi \vee \Psi)$; $\Box(\Phi \wedge \Psi) \vdash \Box\Phi \wedge \Box\Psi$; $\Diamond\Phi \vee \Diamond\Psi \vdash \Diamond(\Phi \vee \Psi)$; $\Diamond(\Phi \wedge \Psi) \vdash \Diamond\Phi \wedge \Diamond\Psi$.

Теорема 1. В исчислении E доказуемы секвенции:

$$\Box(A \rightarrow B), \Box(\neg A \rightarrow B) \vdash \Diamond B.$$

$$\Box(A \rightarrow B), \Box(A \rightarrow \neg B) \vdash \neg\Box A.$$

Доказательство. Из предложения 1 следует доказуемость секвенции

$$\Box\Phi, \Box\Psi \vdash \Diamond(\Phi \wedge \Psi).$$

Подставляем: $\Phi = A \rightarrow B$, $\Psi = \neg A \rightarrow B$. Тогда

$$\Phi \wedge \Psi \equiv (\neg A \vee B) \wedge (A \vee B) \equiv (\neg A \wedge A) \vee B \equiv B.$$

Первая секвенция доказана.

Если взять $\Psi = A \rightarrow \neg B$, то

$$\Phi \wedge \Psi \equiv (\neg A \vee B) \wedge (\neg A \vee \neg B) \equiv \neg A \vee (B \wedge \neg B) \equiv \neg A.$$

Но $\Diamond\neg A \equiv \neg\Box A$. И вторая секвенция доказана.

■

Выводы будут сделаны после сравнения с теоремой 2 в следующем пункте.

Также как и для алетических и деонтических модальностей, с целью редукции идущих подряд модальных знаков, добавим к исчислению E две схемы аксиом:

$$A3. \Box\Phi \vdash \Box\Box\Phi.$$

$$A4. \neg\Box\Phi \vdash \Box\neg\Box\Phi.$$

Получим исчисление ES5, в котором (как и в S5) выполнено следующее правило редукции модальностей: любая цепочка идущих подряд модальных знаков эквивалентна последнему из них.

Предложение 2. В исчислении ES5 доказуемы секвенции:

$$\vdash \Box(\Box\Phi \rightarrow \Phi).$$

$$\vdash \Box(\Psi \rightarrow \Diamond\Psi).$$

Доказательство. В E доказуемы секвенции $\Box\neg\Box\Phi \vee \Box\Phi \vdash \Box(\neg\Box\Phi \vee \Phi)$; $\Box\neg\Psi \vee \Box\Diamond\Psi \vdash \Box(\neg\Psi \vee \Diamond\Psi)$. Они эквивалентны $\Diamond\Box\Phi \rightarrow \Box\Phi \vdash \Box(\Box\Phi \rightarrow \Phi)$ и $\Diamond\Psi \rightarrow \Box\Diamond\Psi \vdash \Box(\Psi \rightarrow \Diamond\Psi)$ соответственно.

Обращение A4 по правилам 9ab приводит к $\Diamond\Box\Phi \vdash \Box\Phi$, а подстановка в неё $\neg\Psi$ даёт $\Diamond\Psi \vdash \Box\Diamond\Psi$, т. е. доказуемы $\vdash \Diamond\Box\Phi \rightarrow \Box\Phi$ и $\vdash \Diamond\Psi \rightarrow \Box\Diamond\Psi$.

По правилу вывода T (транзитивность импликации) получаем требуемое. ■

Доказанные секвенции имеют следующие смыслы: «приемлемо считать, что если гипотеза принимается, то она истинна» и «приемлемо считать, что если гипотеза истинна, то она не отвергается» соответственно. Однако ни $\vdash \Box\Phi \rightarrow \Phi$, ни $\vdash \Psi \rightarrow \Diamond\Psi$ не доказуемы в индуктивных исчислениях.

2. Вероятностная семантика

Каждой формуле Φ сопоставляется некоторая вероятность $P(\Phi)$, которую можно интерпретировать как вероятность того, что данная гипотеза истинна. Способов сопоставления бесконечно много, но все они обязаны удовлетворять следующим свойствам:

$$B1. 0 \leq P(\Phi) \leq 1.$$

$$B2. P(\neg\Phi) = 1 - P(\Phi).$$

$$B3. \text{ если } \Phi \vdash \Psi, \text{ то } P(\Phi) \leq P(\Psi).$$

$$B4. P(\Phi \vee \Psi) = P(\Phi) + P(\Psi) - P(\Phi \wedge \Psi).$$

Из свойства B3 следует, что вероятность конъюнкции двух формул меньше, чем вероятность каждой из её компонент. Вероятность истинности тавтологий исчисления высказываний всегда равна 1. Если же $P(\Phi) = 1$, то этого недостаточно для истинности формулы Φ , нужно, чтобы $P(\Phi) = 1$ выполнялось для каждого вероятностного отображения. Кроме того, отметим, что вероятности истинности для эквивалентных формул равны между собой.

Вероятностная семантика – это пара отображений $\langle v, P \rangle$, заданных на множестве формул, где v может принимать только два значения (0 – «ложь», 1 – «истина») и согласовано с логическими связками (как в исчислении высказываний), а P – вероятностное отображение, удовлетворяющее свойствам В1–В4. Связь между этими отображениями для исчисления Е задаётся свойством

$$\text{ВЕ. } v(\Box\Phi) = 1 \Leftrightarrow P(\Phi) \geq 1 - \varepsilon,$$

где ε – некоторая заданная величина из интервала $[0, \frac{1}{3}]$. Тогда

$$P(\Phi) > \varepsilon \Leftrightarrow P(\neg\Phi) < 1 - \varepsilon \Leftrightarrow v(\Box\neg\Phi) = 0 \Leftrightarrow v(\neg\Box\neg\Phi) = v(\Diamond\Phi) = 1.$$

Получается, что утверждение «гипотеза принимается» истинно, если вероятность истинности гипотезы не менее $1 - \varepsilon \geq 66,67\%$, а утверждение «гипотеза отвергается» истинно, если вероятность истинности гипотезы не более $\varepsilon \leq 33,33\%$. Можно назвать ε неким «порогом недоверия».

Как обычно, формулы и секвенции будем называть тождественно истинными, если они истинны в любой вероятностной семантике. В [4] (метатеоремы 50 и 51) приведены доказательства непротиворечивости и полноты исчисления Е для такой семантики.

Теорема 2. В исчислении Е не доказуемы секвенции:

$$\Box(A \rightarrow B), \Box(\neg A \rightarrow B) \vdash \Box B;$$

$$\Box(A \rightarrow B), \Box(A \rightarrow \neg B) \vdash \Box\neg A.$$

Доказательство. Выберем $\varepsilon = 0,3 < 1/3$. Положим $P(A) = P(B) = 0,5$, тогда и $P(\neg A) = P(\neg B) = 1 - 0,5 = 0,5$. Если $P(A \wedge B) = 0,25$, то по симметрии $P(A \wedge \neg B) = P(\neg A \wedge B) = P(\neg A \wedge \neg B) = 0,25$. Следовательно,

$$\begin{aligned} P(A \rightarrow B) &= P(\neg A \vee B) = P(\neg A) + P(B) - P(\neg A \wedge B) = \\ &= 0,5 + 0,5 - 0,25 = 0,75 \geq 1 - \varepsilon = 0,7. \end{aligned}$$

Аналогично получаем, что $P(\neg A \rightarrow B) \geq 1 - \varepsilon$ и $P(A \rightarrow \neg B) \geq 1 - \varepsilon$. По свойству ВЕ посылки обеих секвенций истинны, а заключения ложны. Значит, они не доказуемы ввиду непротиворечивости исчисления Е. ■

Сравнивая результаты теоремы 1 и теоремы 2, ещё раз убеждаемся в осторожности индуктивных выводов: если мы принимаем две гипотезы о том, что В следует из обеих противоположных посылок, то В всего лишь не отвергаема. И наоборот, если из А следуют два противоречивых следствия, то А не отвергается, а только не принимается (до получения дополнительной информации).

В заключение заметим, что вероятностная семантика может применяться и для алетических модальных исчислений.

Например, вероятностная семантика $\langle v, P \rangle$ для исчисления Т (см. [3]) удовлетворяет свойствам В1–В4 и следующему условию

$$\text{ВТ. } v(\Box\Phi) = 1 \Leftrightarrow P(\Phi) = 1.$$

Из него, в частности, следует, что $v(\Diamond\Phi) = 1 \Leftrightarrow P(\Phi) \neq 0$. Получается, что «необходимость» означает равенство вероятности 1. Формальное сравнение с условием ВЕ говорит о том, что вероятностная семантика для Т является частным случаем индуктивной вероятностной семантики, когда «порог недоверия» $\varepsilon = 0$. Это согласуется с тем, что исчисление Т является расширением исчисления Е.

Литература

1. Goldblatt R. Logics of time and computation. Standford: CSLI Lecture Notes, Vol. 7. 1992. 131 p.
2. Судоплатов С.В., Овчинникова Е.В. Математическая логика и теория алгоритмов. М.: ИНФРА-М; Новосибирск: Изд-во НГТУ, 2004. 224 с.
3. Бесценный И.П. О редукции модальностей в деонтических исчислениях // Математические структуры и моделирование. 2017. № 2 (42). С. 5–10.
4. Костюк В.Н. Элементы модальной логики. Киев: Наукова думка, 1978. 180 с.

ON INDUCTIVE CALCULI AND PROBABILISTIC SEMANTICS

I.P. Bestsennyi

Ph.D. (Phys.-Math.), Associate Professor, e-mail: ibests@mail.ru

Dostoevsky Omsk State University, Omsk, Russia

Abstract. Formal epistemic modal calculi are considered, whose axioms reflect the features of determining the reliability of inductive inferences. Since Kripke semantics is not suitable for them, probabilistic semantics is more natural for inductive calculi.

Keywords: modal logic, epistemic calculi, probabilistic semantic.

Дата поступления в редакцию: 20.02.2025

ТЕОРЕТИКО-ГРУППОВОЕ ОПИСАНИЕ ПЕРИОДИЧЕСКОЙ СИСТЕМЫ ЭЛЕМЕНТОВ: VI. ВОСЬМИМЕРНОЕ ОБОБЩЕНИЕ

В.В. Варламов

д.ф.-м.н., e-mail: varlamov@subsiu.ru

Сибирский государственный индустриальный университет, Новокузнецк, Россия

Аннотация. Периодическая система химических элементов представляется в рамках весовой диаграммы алгебры Ли четвёртого ранга группы вращений 8-мерного псевдоевклидова пространства. Дается интерпретация спина как четвёртого генератора подалгебры Картана, двум собственным значениям которого соответствуют две 3-мерные проекции весовой диаграммы, содержащие элементы периодической системы от водорода до московия (первая проекция) и от гелия до оганессона (вторая проекция). Одним из главных преимуществ предложенной теоретико-групповой конструкции периодической системы является естественное включение антиматерии в общую схему.

Ключевые слова: периодическая система, алгебра Ли, весовая диаграмма, подалгебра Картана, спин, антиматерия.

Введение

Исторически сложилось так, что графическое представление периодического закона имеет очень важное значение. Первоначальные таблицы Менделеева и Мейера, помимо наглядного визуального образа, обозначают распределение химических элементов согласно периодам, группам и подгруппам, а также содержат связи между гомологичными элементами, что позволило Д.И. Менделееву предсказать новые на тот момент времени элементы (галлий и германий). Однако уже с первых лет после появления таблицы Менделеева (1869) стали предприниматься попытки выхода за пределы 2-мерного табличного представления. Так, в 1882 г. появляется пирамидальная модель Бэйли, далее усовершенствованная И. Томсеном (1895) и Н. Бором (1922). Более того, ранее таблиц Менделеева и Мейера появляются винтовые и спиральные модели («теллуrow винт» Шанкуртуа (1862) и спираль Хинрикса (1867)). Спираль – это 2-мерная фигура, в то время как винт имеет три измерения. Обе конструкции были использованы для того, чтобы дать чёткое представление о периодической системе элементов. Преимущество как спиральной, так и винтовой (спиралевидной) систем заключается в том, что они образуют единое целое. Проекцией винта на плоскость является спираль. Винтовая система обеспечивает дополнительное средство выражения аналогий в свойствах и, что также важно, различий. Поскольку это преимущество изначально недооценивалось, винт использовался ре-

же, чем спираль. Только после открытия изотопов некоторые исследователи пришли к выводу, что форма винта позволяет решить проблему включения изотопов в периодическую систему Менделеева. О различных графических представлениях периодической системы см. [1–3].

Главной особенностью перечисленных выше моделей (табличных, пирамидальных, спиральных и винтовых) является построение системы химических элементов в зависимости от возрастания атомного веса. Однако квантовая механика и первое построение квантовомеханической модели атома (модель Бора) привели к пониманию, что главной структурной характеристикой периодического закона является не линейное возрастание атомного веса, а структура, определяемая порядком квантовых чисел. Одним из первых, кто это осознал, был немецкий математик Г. Генцель. Во введении к статье [4] Г. Генцель пишет: «Прежние представления о периодической системе элементов как таблицы с горизонтальными и вертикальными рядами и столбцами (Мейер – Менделеев) или с соединительными линиями между родственными элементами¹ уже не соответствуют нынешнему пониманию систематики элементов, поскольку не согласуются с результатами волновой механики и их выводами, которые в достаточной мере справедливо отражают химию и физику периодической системы. Исходя из элементарной геометрической конфигурации концентрических правильных многоугольников, получается многослойная плоскость, на которой эти многоугольники с нечётным числом углов распределяются таким образом, что каждый многоугольник окружён контуром предыдущего. Эта плоскость, которая, таким образом, берет своё начало из одного из первых объектов геометрии, оказывается подходящим носителем системы химических элементов, поскольку её структура и топология полностью соответствуют теоремам, управляющим этой системой». В 1943 г. В. Финке в статье [5] представляет систему кругов и колец Генцеля в пространстве. Трёхмерная система Финке, подобно винту по отношению к спирали в системах, построенных согласно возрастанию атомного веса, обеспечивает дополнительное средство выражения аналогий в свойствах (линии Бэйли – Томсена – Бора). Однако общим недостатком систем Генцеля и Финке является искусственное представление четвёртого квантового числа s (спин) в виде двух точек на трансверсалах.

Настоящая статья является продолжением серии работ [6–11], в которых предпринимается попытка построения периодической системы химических элементов в рамках весовой диаграммы групповой алгебры некоторой группы симметрии. Так, в статье [11] построена весовая диаграмма алгебры Ли $\mathfrak{so}(4, 2)$ конформной группы $SO(4, 2)$. Поскольку $\mathfrak{so}(4, 2)$ является алгеброй третьего ранга, весовая диаграмма ($SO(4, 2)$ -башня) определена в 3-мерном весовом пространстве, в котором координатными осями выступают три генератора, образующие подалгебру Картана $\mathfrak{K} \subset \mathfrak{so}(4, 2)$. Однако для теоретико-группового описания структуры периодической системы, определяемой четвёркой квантовых чисел (n, l, m, s) , алгебры третьего ранга недостаточно. Необходим переход к алгебре Ли четвёртого ранга. В п. 5 настоящей статьи рассматривается группа вращений $SO(4, 4)$ 8-мерного псев-

¹Здесь, по всей видимости, Г. Генцель имеет в виду уже упоминавшиеся выше пирамидальные модели периодической системы Бэйли, Томсена и Бора, в которых родственные (гомологичные) элементы соединялись линиями.

доевклидова пространства $\mathbb{R}^{4,4}$. Алгебра Ли $\mathfrak{so}(4,4)$ группы $SO(4,4)$ имеет четвёртый ранг. Четвёртый генератор $\mathbf{L}_{78}$ (понимаемый как генератор спина) подалгебры Картана $\mathfrak{K} \subset \mathfrak{so}(4,4)$ коммутирует со всеми 15 генераторами подалгебры $\mathfrak{so}(4,2)$. Как следствие, базис алгебры двулистного накрытия $\mathbf{Spin}_+(4,4)$ расщепляется на два структурно идентичных базиса, каждый из которых изоморфен базису Яо [?] для групповой алгебры группы $\mathbf{Spin}_+(4,2) \simeq SU(2,2)$, накрывающей конформную группу. Это позволяет построить две 3-мерные проекции весовой диаграммы алгебры $\mathfrak{so}(4,4)$ с использованием базиса Маделунга. Первая проекция содержит все элементы периодической системы со значением спина $s = -1/2$, начиная с водорода **H** ($Z = 1$) до московия **Mc** ($Z = 115$), вторая проекция содержит все элементы со значением $s = +1/2$, начиная с гелия **He** ($Z = 2$) до оганессона **Og** ($Z = 118$).

1. Polygonfläche Генцеля

Главной идеей статьи [4] является расположение химических элементов периодической системы согласно квантовым числам n , l , m и s , которые Г. Генцель обозначает n_1 , n_2 , n_3 и n_4 ². Как известно, диапазоны изменения квантовых чисел следующие:

$$\begin{aligned} n &= 1, 2, 3, \dots, \nu, \dots, \\ l &= 0, 1, 2, \dots, n-1, \\ m &= 0, \pm 1, \pm 2, \dots, \pm(n-1), \\ s &= -1/2, +1/2. \end{aligned}$$

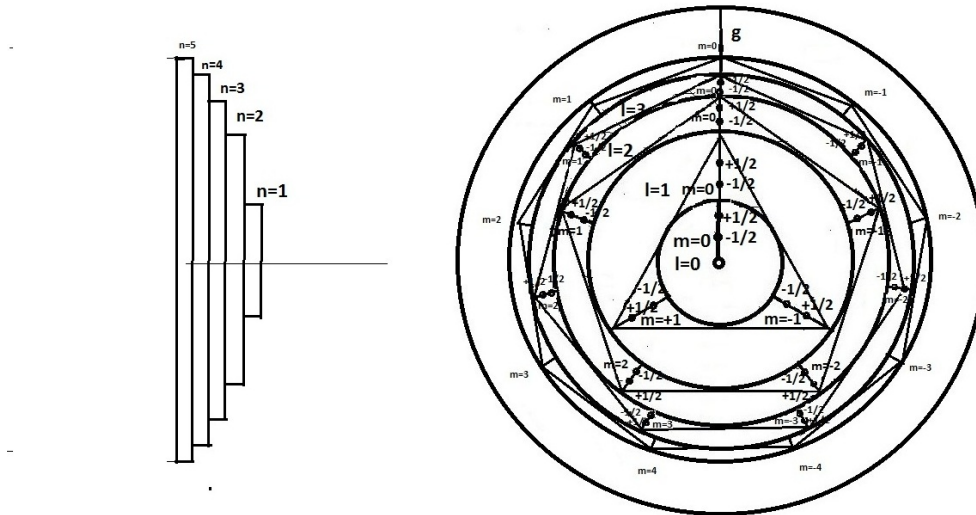


Рис. 1. Круги и кольца Генцеля

Общая схема строения листов и колец Генцеля представлена на рис. 1. Ряд наложенных друг на друга плоских листов (кругов) пронумерован снизу вверх: $n =$

²Далее мы будем придерживаться стандартных обозначений n , l , m , s .

$1, 2, 3, \dots, \nu, \dots$ Все листы содержат единичный круг, площадь которого обозначена $l = 0$, и полупрямую g с направлением $m = 0$. Единичный круг самого нижнего листа $n = 1$ разрезается по его краю и прикрепляется к листу $n = 2$, находящемуся над ним, так что из области $n = 1, l = 0$ можно выйти только забравшись на лист $n = 2$. На этом листе единичная окружность $n = 2, l = 0$ вписана в равносторонний треугольник (с углом, лежащем на g). Площадь кольца между единичной окружностью и окружностью, описанной вокруг треугольника, обозначается $l = 1$. В ней отмечены три трансверсали (поперечные линии) к углам: $m = 0, \pm 1$. Внешний край кольца $l = 1$ прикреплён к плоскости $n = 3$. Она содержит единичный круг $l = 0$, первое круговое кольцо $l = 1$ с треугольными трансверсалими $m = 0, \pm 1$ и второе круговое кольцо $l = 2$ с пятиугольными трансверсалими $m = 0, \pm 1, \pm 2$. Внешний край кольца $l = 2$ соединён с плоскостью $n = 4$ и т. д. В общем случае лист $n = \nu$ содержит круговые кольца $l = 0, 1, 2, \dots, \nu - 1$ и содержащиеся в них трансверсали $m = 0, \pm 1, \pm 2, \dots, \pm(\nu - 1)$.

Согласно рис. 1 предполагается, что на каждой трансверсали в отдельных круговых кольцах существует две симметрично расположенные точки, которые отличаются друг от друга четвёртым квантовым числом $s = \pm 1/2$ (спин). Все эти точки называются точками собственных значений Polygonfläche, а их четвёрки являются собственными значениями квантовой системы. Система собственных значений распределяется на поверхности следующим образом: на нижнем листе $n = 1$ имеется две точки собственных значений $n = 1, l = 0, m = 0, s = \pm 1/2$. На втором листе – 8 точек: $n = 2, l = 0, m = 0, s = \pm 1/2$; $l = 1, m = 0, \pm 1, s = \pm 1/2$. На третьем листе имеется 18 точек. В общем случае лист $n = \nu$ содержит $2\nu^2$ точек собственных значений. При этом данный лист n содержит кольца от $l = 0$ до $l = n - 1$, и в каждом кольце с номером l имеется $2l + 1$ трансверсалий так, что плоскость n содержит $\sum_{l=0}^{n-1} (2l + 1)$ трансверсалий каждая.

На рис. 2 показано применение Polygonfläche к периодической системе химических элементов.

В нижнем листе $n = 1$ лежат одно- и двухэлектронные системы (**H** и **He**), выше в $n = 2$ – восьмизрядная последовательность от лития к неону, на следующем листе – восемнадцатизрядная последовательность и т. д. На момент написания Г. Генцелем статьи [4] (1942) было известно только 94 элемента. С учётом современного знания периодической системы на рис. 2 представлено следующее расширение Polygonfläche: 1) добавлено 24 новых элемента, начиная от америция **Am** ($Z = 95$) до оганессона **Og** ($Z = 118$), что привело к расширению листа $n = 5$ до кольца $l = 3$ и листа $n = 7$ до кольца $l = 2$; 2) добавлен лист $n = 8, l = 0$, содержащий два гипотетических элемента **Uue** ($Z = 119$) и **Ubn** ($Z = 120$), поскольку эти элементы логически завершают f -оболочку.

2. Трёхмерная система Финке

В этом же году (1943) в «Zeitschrift für Physik» выходит статья В. Финке [5], посвящённая замечаниям по Polygonfläche Генцеля. Во-первых, В. Финке приводит Polygonfläche в соответствие с правилом Маделунга. Во-вторых, раздвигает в пространстве круги Генцеля, образуя тем самым 3-мерное представление периоди-

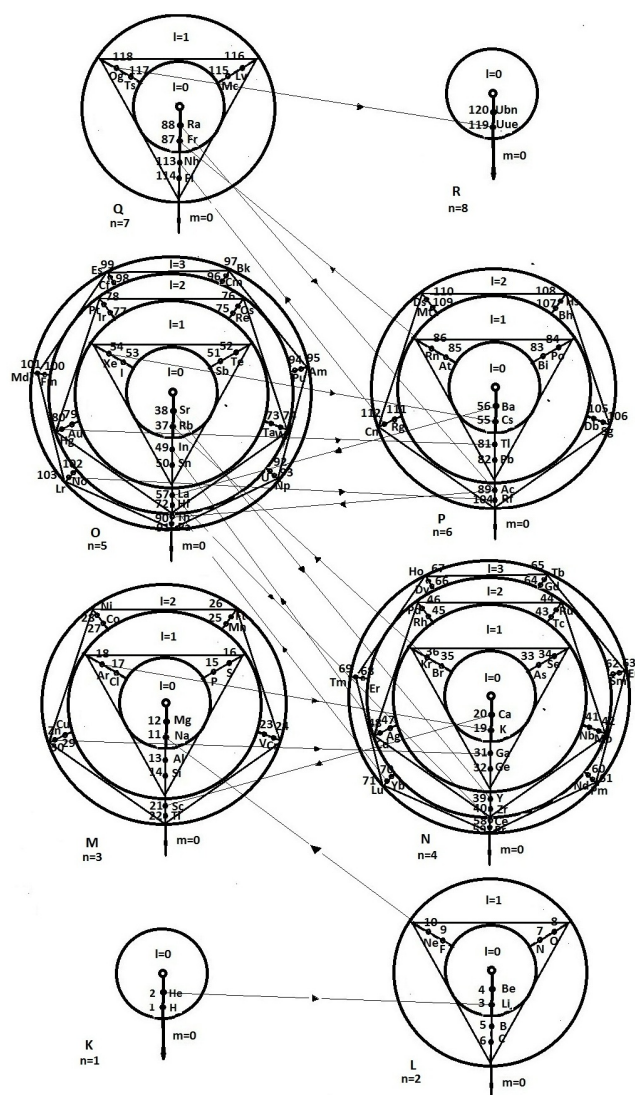


Рис. 2. Polygonfläche и периодическая система.

ческой системы³(см. рис. 3). Среди главных преимуществ 3-мерного представления следует отметить появление графической связи между гомологичными элементами (линии Бэйли – Томсена – Бора). Однако при этом теряется порядок заполнения квантовых уровней (кругов и колец Генцеля) периодической системы, который в Polygonfläche осуществляется с помощью направленных линий (см. рис. 2). В связи с этим, обе системы (Г. Генцеля и В. Финке) следует рассматривать как взаимодополняющие. Так же как и в случае с Polygonfläche, к системе Финке на рис. 3 добавлены неизвестные на момент 1943 г. элементы, т. е. расширены круги O и Q , а также добавлен круг R с гипотетическими элементами.

³Следует отметить, что 3-мерность Polygonfläche уже подразумевалась Г. Генцелем, исходя из строения многослойной поверхности, показанной на рис. 1.

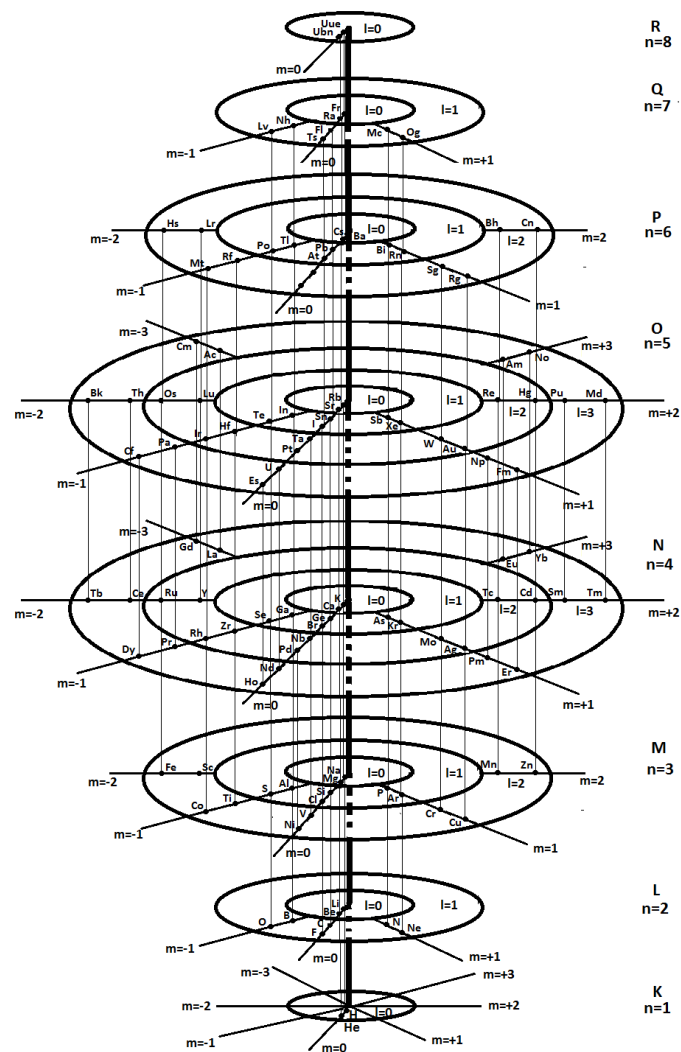


Рис. 3. Трёхмерная система Финке

3. Общее замечание к системам Генцеля и Финке

В обеих системах первые три квантовые числа n , l и m имеют ясное геометрическое представление, первоначально заимствованное Н. Бором в его планетарной модели атома из небесной механики (радиальные и азимутальные движения). В отличие от них четвертое квантовое число s (спин) на плоскости (Polygonfläche) или в пространстве (система Финке) имеет явно искусственное представление в виде двух точек на трансверсалиях. Всё это говорит о том, что понятие спина не вмещается в 3-мерное представление. Поэтому все попытки дать механическую интерпретацию спина в рамках 3-мерного пространства не имели успеха (например, первоначальное представление Уленбека – Гаудсмита о вращающемся волчке). *Спин – это объект 4-мерного мира*, или, может быть, более правильно, *эффект четвертого измерения*.

4. Группа $SO(4, 4)$

Специальная псевдоортогональная группа в восьми измерениях, $SO(4, 4)$, соответствует группе вращений 8-мерного псевдоевклидова пространства $\mathbb{R}^{4,4}$, или эквивалентно множеству 8×8 ортогональных матриц, оставляющих квадратичную форму

$$Q(\mathbf{r}) = x_1^2 + x_2^2 + x_3^2 + x_4^2 - x_5^2 - x_6^2 - x_7^2 - x_8^2 = \mathbf{r}^T \mathbf{r},$$

где $\mathbf{r} = [x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8]^T$, инвариантной.

Структура соответствующей алгебры Ли $\mathfrak{so}(4, 4)$ определяется коммутационными свойствами её генераторов $\mathbf{L}_{\alpha\beta}$. $\mathbf{L}_{\alpha\beta}$ образуют базис алгебры $\mathfrak{so}(4, 4)$. Число независимых генераторов легко найти: из 64 возможных комбинаций индексов α и β восемь комбинаций исчезают в силу $\mathbf{L}_{\alpha\alpha} = 0$, это уменьшает число генераторов до 56. Более того, в силу $\mathbf{L}_{\alpha\beta} = -\mathbf{L}_{\beta\alpha}$ остаётся только 28 независимых генераторов, число которых также может быть получено с помощью формулы $n(n-1)/2$, где $n = p + q$ – размерность пространства $\mathbb{R}^{p,q}$. Таким образом,

$$\mathbf{L} \Leftrightarrow \begin{bmatrix} 0 & \mathbf{L}_{12} & \mathbf{L}_{13} & \mathbf{L}_{14} & \mathbf{L}_{15} & \mathbf{L}_{16} & \mathbf{L}_{17} & \mathbf{L}_{18} \\ & 0 & \mathbf{L}_{23} & \mathbf{L}_{24} & \mathbf{L}_{25} & \mathbf{L}_{26} & \mathbf{L}_{27} & \mathbf{L}_{28} \\ & & 0 & \mathbf{L}_{34} & \mathbf{L}_{35} & \mathbf{L}_{36} & \mathbf{L}_{37} & \mathbf{L}_{38} \\ & & & 0 & \mathbf{L}_{45} & \mathbf{L}_{46} & \mathbf{L}_{47} & \mathbf{L}_{48} \\ & & & & 0 & \mathbf{L}_{56} & \mathbf{L}_{57} & \mathbf{L}_{58} \\ & & & & & 0 & \mathbf{L}_{67} & \mathbf{L}_{68} \\ & & & & & & 0 & \mathbf{L}_{78} \\ & & & & & & & 0 \end{bmatrix}. \quad (1)$$

Система 28 генераторов $\mathbf{L}_{\alpha\beta}$ алгебры $\mathfrak{so}(4, 4)$ удовлетворяет следующим перестановочным соотношениям:

$$[\mathbf{L}_{\alpha\beta}, \mathbf{L}_{\gamma\delta}] = i (g_{\alpha\delta} \mathbf{L}_{\beta\gamma} + g_{\beta\gamma} \mathbf{L}_{\alpha\delta} - g_{\alpha\gamma} \mathbf{L}_{\beta\delta} - g_{\beta\delta} \mathbf{L}_{\alpha\gamma}), \quad (2)$$

где $\alpha, \beta, \gamma, \delta = 1, \dots, 8$, $\alpha \neq \beta$, $\gamma \neq \delta$, при этом $g_{11} = g_{22} = g_{33} = g_{44} = 1$, $g_{55} = g_{66} = g_{77} = g_{88} = -1$. Таким образом, имеем 28-мерную алгебру Ли $\mathfrak{so}(4, 4)$.

Найдём максимальное подмножество коммутирующих генераторов алгебры $\mathfrak{so}(4, 4)$. Как известно, два генератора коммутируют, если они не имеют общих индексов. Легко видеть, что среди генераторов алгебры $\mathfrak{so}(4, 4)$ этому условию удовлетворяют четыре генератора:

$$\mathbf{L}_{12}, \mathbf{L}_{34}, \mathbf{L}_{56}, \mathbf{L}_{78}. \quad (3)$$

Четвёрка генераторов $\{\mathbf{L}_{12}, \mathbf{L}_{34}, \mathbf{L}_{56}, \mathbf{L}_{78}\}$ образует базис максимальной абелевой подалгебры $\mathfrak{K} \subset \mathfrak{so}(4, 4)$ (подалгебры Картана). Генераторы (3) называются генераторами Картана. Размерность подалгебры $\mathfrak{K}$ определяет ранг алгебры Ли, следовательно, $\mathfrak{so}(4, 4)$ является алгеброй Ли четвёртого ранга. Таким образом, все корневые и весовые диаграммы для $\mathfrak{so}(4, 2)$ будут 4-мерными⁴.

⁴Как следствие, это исключает явную визуализацию этих диаграмм для алгебры $\mathfrak{so}(4, 4)$.

Составим базис алгебры $\mathfrak{so}(4, 4)$ относительно максимальной компактной подгруппы

$$K = \mathrm{SU}(2) \otimes \mathrm{SU}(2) \otimes \mathrm{U}(1)$$

группы $\mathrm{SO}(4, 4)$ посредством следующих линейных комбинаций:

$${}^1\mathbf{K}_1 = 1/2 (\mathbf{L}_{23} + \mathbf{L}_{14}), \quad {}^1\mathbf{K}_2 = 1/2 (\mathbf{L}_{31} + \mathbf{L}_{24}), \quad {}^1\mathbf{K}_3 = 1/2 (\mathbf{L}_{12} + \mathbf{L}_{34}), \quad (4)$$

$${}^1\mathbf{J}_1 = 1/2 (\mathbf{L}_{23} - \mathbf{L}_{14}), \quad {}^1\mathbf{J}_2 = 1/2 (\mathbf{L}_{31} - \mathbf{L}_{24}), \quad {}^1\mathbf{J}_3 = 1/2 (\mathbf{L}_{12} - \mathbf{L}_{34}), \quad (5)$$

$${}^1\mathbf{T}_1 = 1/2 (-\mathbf{L}_{15} - \mathbf{L}_{26}), \quad {}^1\mathbf{T}_2 = 1/2 (\mathbf{L}_{25} - \mathbf{L}_{16}), \quad {}^1\mathbf{T}_0 = 1/2 (-\mathbf{L}_{12} - \mathbf{L}_{56}), \quad (6)$$

$${}^1\mathbf{S}_1 = 1/2 (-\mathbf{L}_{15} + \mathbf{L}_{26}), \quad {}^1\mathbf{S}_2 = 1/2 (-\mathbf{L}_{25} - \mathbf{L}_{16}), \quad {}^1\mathbf{S}_0 = 1/2 (\mathbf{L}_{12} - \mathbf{L}_{56}), \quad (7)$$

$${}^1\mathbf{P}_1 = 1/2 (-\mathbf{L}_{35} - \mathbf{L}_{46}), \quad {}^1\mathbf{P}_2 = 1/2 (\mathbf{L}_{45} - \mathbf{L}_{36}), \quad {}^1\mathbf{P}_0 = 1/2 (-\mathbf{L}_{34} - \mathbf{L}_{56}), \quad (8)$$

$${}^1\mathbf{Q}_1 = 1/2 (\mathbf{L}_{35} - \mathbf{L}_{46}), \quad {}^1\mathbf{Q}_2 = 1/2 (\mathbf{L}_{45} + \mathbf{L}_{36}), \quad {}^1\mathbf{Q}_0 = 1/2 (\mathbf{L}_{34} - \mathbf{L}_{56}), \quad (9)$$

$${}^2\mathbf{K}_1 = 1/2 (\mathbf{L}_{67} + \mathbf{L}_{58}), \quad {}^2\mathbf{K}_2 = 1/2 (-\mathbf{L}_{57} + \mathbf{L}_{68}), \quad {}^2\mathbf{K}_3 = 1/2 (\mathbf{L}_{56} + \mathbf{L}_{78}), \quad (10)$$

$${}^2\mathbf{J}_1 = 1/2 (\mathbf{L}_{67} - \mathbf{L}_{58}), \quad {}^2\mathbf{J}_2 = 1/2 (-\mathbf{L}_{57} - \mathbf{L}_{68}), \quad {}^2\mathbf{J}_3 = 1/2 (\mathbf{L}_{56} - \mathbf{L}_{78}), \quad (11)$$

$${}^2\mathbf{T}_1 = 1/2 (\mathbf{L}_{17} + \mathbf{L}_{28}), \quad {}^2\mathbf{T}_2 = 1/2 (-\mathbf{L}_{27} + \mathbf{L}_{18}), \quad {}^2\mathbf{T}_0 = 1/2 (\mathbf{L}_{12} + \mathbf{L}_{78}), \quad (12)$$

$${}^2\mathbf{S}_1 = 1/2 (\mathbf{L}_{17} - \mathbf{L}_{28}), \quad {}^2\mathbf{S}_2 = 1/2 (\mathbf{L}_{27} + \mathbf{L}_{18}), \quad {}^2\mathbf{S}_0 = 1/2 (-\mathbf{L}_{12} + \mathbf{L}_{78}), \quad (13)$$

$${}^2\mathbf{P}_1 = 1/2 (\mathbf{L}_{37} + \mathbf{L}_{48}), \quad {}^2\mathbf{P}_2 = 1/2 (-\mathbf{L}_{47} + \mathbf{L}_{38}), \quad {}^2\mathbf{P}_0 = 1/2 (\mathbf{L}_{34} + \mathbf{L}_{78}), \quad (14)$$

$${}^2\mathbf{Q}_1 = 1/2 (\mathbf{L}_{37} - \mathbf{L}_{48}), \quad {}^2\mathbf{Q}_2 = 1/2 (\mathbf{L}_{47} + \mathbf{L}_{38}), \quad {}^2\mathbf{Q}_0 = 1/2 (-\mathbf{L}_{34} + \mathbf{L}_{78}). \quad (15)$$

Базис (4)–(15) содержит 36 генераторов, что создаёт избыточную систему для алгебры $\mathfrak{so}(4, 4)$, поскольку последняя состоит из 28 независимых генераторов. Базис алгебры $\mathfrak{so}(4, 4)$ можно получить из (4)–(15), исключив восемь генераторов с помощью соотношений

$${}^1\mathbf{J}_3 + {}^1\mathbf{K}_3 = {}^1\mathbf{S}_0 - {}^1\mathbf{T}_0 = {}^2\mathbf{T}_0 - {}^2\mathbf{S}_0 = \mathbf{L}_{12}, \quad (16)$$

$${}^1\mathbf{J}_3 - {}^1\mathbf{K}_3 = {}^1\mathbf{P}_0 - {}^1\mathbf{Q}_0 = {}^2\mathbf{Q}_0 - {}^2\mathbf{P}_0 = -\mathbf{L}_{34}, \quad (17)$$

$${}^1\mathbf{P}_0 + {}^1\mathbf{Q}_0 = {}^1\mathbf{S}_0 + {}^1\mathbf{T}_0 = -{}^2\mathbf{K}_3 - {}^2\mathbf{J}_3 = -\mathbf{L}_{56}, \quad (18)$$

$${}^2\mathbf{K}_3 - {}^2\mathbf{J}_3 = {}^2\mathbf{T}_0 + {}^2\mathbf{S}_0 = {}^2\mathbf{P}_0 + {}^2\mathbf{Q}_0 = \mathbf{L}_{78}. \quad (19)$$

Это сводит число генераторов до 28, как и должно быть для алгебры $\mathfrak{so}(4, 4)$. Таким образом, множество из двенадцати генераторов ${}^1\mathbf{K}_3, {}^1\mathbf{J}_3, {}^1\mathbf{T}_0, {}^1\mathbf{S}_0, {}^1\mathbf{P}_0, {}^1\mathbf{Q}_0, {}^2\mathbf{K}_3, {}^2\mathbf{J}_3, {}^2\mathbf{T}_0, {}^2\mathbf{S}_0, {}^2\mathbf{P}_0, {}^2\mathbf{Q}_0$ в силу соотношений (16)–(19) эмулирует подалгебру Картана $\mathfrak{K} = \{\mathbf{L}_{12}, \mathbf{L}_{34}, \mathbf{L}_{56}, \mathbf{L}_{78}\}$ алгебры Ли $\mathfrak{so}(4, 4)$.

Из оставшихся 24 генераторов базиса (4)–(15) образуем генераторы Вейля посредством следующих линейных комбинаций:

$${}^1\mathbf{J}_\pm = {}^1\mathbf{J}_1 \pm i {}^1\mathbf{J}_2, \quad {}^1\mathbf{P}_\pm = {}^1\mathbf{P}_1 \pm i {}^1\mathbf{P}_2, \quad {}^1\mathbf{S}_\pm = {}^1\mathbf{S}_1 \pm i {}^1\mathbf{S}_2, \quad (20)$$

$${}^1\mathbf{K}_\pm = {}^1\mathbf{K}_1 \pm i {}^1\mathbf{K}_2, \quad {}^1\mathbf{Q}_\pm = {}^1\mathbf{Q}_1 \pm i {}^1\mathbf{Q}_2, \quad {}^1\mathbf{T}_\pm = {}^1\mathbf{T}_1 \pm i {}^1\mathbf{T}_2, \quad (21)$$

$${}^2\mathbf{J}_\pm = {}^2\mathbf{J}_1 \pm i {}^2\mathbf{J}_2, \quad {}^2\mathbf{P}_\pm = {}^2\mathbf{P}_1 \pm i {}^2\mathbf{P}_2, \quad {}^2\mathbf{S}_\pm = {}^2\mathbf{S}_1 \pm i {}^2\mathbf{S}_2, \quad (22)$$

$${}^2\mathbf{K}_{\pm} = {}^2\mathbf{K}_1 \pm i^2\mathbf{K}_2, \quad {}^2\mathbf{Q}_{\pm} = {}^2\mathbf{Q}_1 \pm i^2\mathbf{Q}_2, \quad {}^2\mathbf{T}_{\pm} = {}^2\mathbf{T}_1 \pm i^2\mathbf{T}_2. \quad (23)$$

12 генераторов (20)–(21) являются генераторами Вейля подалгебры $\mathfrak{so}(4, 2)$ (алгебры Ли конформной группы $SO(4, 2)$). Подробное исследование структуры генераторов Вейля (20)–(21) в зависимости от различных подалгебр алгебры $\mathfrak{so}(4, 2)$ содержится в [10].

Как уже отмечалось выше, четыре генератора (3) являются *генераторами Картана* подалгебры $\mathfrak{K} \subset \mathfrak{so}(4, 4)$. Они образуют базис 4-мерной ортогональной системы координат и располагаются в начале координат корневой диаграммы. Наряду с оставшимися 24 *генераторами Вейля* (20)–(23) они образуют *базис Картана – Вейля* для алгебры $\mathfrak{so}(4, 4)$:

$$\{ \mathbf{L}_{12}, \mathbf{L}_{34}, \mathbf{L}_{56}, \mathbf{L}_{78}, {}^1\mathbf{K}_+, {}^1\mathbf{K}_-, {}^1\mathbf{J}_+, {}^1\mathbf{J}_-, {}^1\mathbf{T}_+, {}^1\mathbf{T}_-, {}^1\mathbf{S}_+, {}^1\mathbf{S}_-, {}^1\mathbf{P}_+, {}^1\mathbf{P}_-, {}^1\mathbf{Q}_+, {}^1\mathbf{Q}_-, {}^2\mathbf{K}_+, {}^2\mathbf{K}_-, {}^2\mathbf{J}_+, {}^2\mathbf{J}_-, {}^2\mathbf{T}_+, {}^2\mathbf{T}_-, {}^2\mathbf{S}_+, {}^2\mathbf{S}_-, {}^2\mathbf{P}_+, {}^2\mathbf{P}_-, {}^2\mathbf{Q}_+, {}^2\mathbf{Q}_- \}. \quad (24)$$

Структура базиса (4)–(15) определяется следующими коммутационными соотношениями:

$$[{}^1\mathbf{K}_1, {}^1\mathbf{K}_2] = -i^1\mathbf{K}_2, \quad [{}^1\mathbf{K}_2, {}^1\mathbf{K}_3] = -i^1\mathbf{K}_1, \quad [{}^1\mathbf{K}_3, {}^1\mathbf{K}_1] = -i^1\mathbf{K}_2, \quad (25)$$

$$[{}^1\mathbf{J}_1, {}^1\mathbf{J}_2] = -i^1\mathbf{J}_2, \quad [{}^1\mathbf{J}_2, {}^1\mathbf{J}_3] = -i^1\mathbf{J}_1, \quad [{}^1\mathbf{J}_3, {}^1\mathbf{J}_1] = -i^1\mathbf{J}_2, \quad (26)$$

$$[{}^1\mathbf{K}_i, {}^1\mathbf{J}_j] = 0 \quad (i, j = 1, 2, 3), \quad (27)$$

$$[{}^1\mathbf{T}_1, {}^1\mathbf{T}_2] = i^1\mathbf{T}_2, \quad [{}^1\mathbf{T}_2, {}^1\mathbf{T}_0] = -i^1\mathbf{T}_1, \quad [{}^1\mathbf{T}_0, {}^1\mathbf{T}_1] = -i^1\mathbf{T}_2, \quad (28)$$

$$[{}^1\mathbf{S}_1, {}^1\mathbf{S}_2] = i^1\mathbf{S}_2, \quad [{}^1\mathbf{S}_2, {}^1\mathbf{S}_0] = -i^1\mathbf{S}_1, \quad [{}^1\mathbf{S}_0, {}^1\mathbf{S}_1] = -i^1\mathbf{S}_2, \quad (29)$$

$$[{}^1\mathbf{T}_i, {}^1\mathbf{S}_j] = 0 \quad (i, j = 0, 1, 2), \quad (30)$$

$$[{}^1\mathbf{P}_1, {}^1\mathbf{P}_2] = i^1\mathbf{P}_2, \quad [{}^1\mathbf{P}_2, {}^1\mathbf{P}_0] = -i^1\mathbf{P}_1, \quad [{}^1\mathbf{P}_0, {}^1\mathbf{P}_1] = -i^1\mathbf{P}_2, \quad (31)$$

$$[{}^1\mathbf{Q}_1, {}^1\mathbf{Q}_2] = i^1\mathbf{Q}_2, \quad [{}^1\mathbf{Q}_2, {}^1\mathbf{Q}_0] = -i^1\mathbf{Q}_1, \quad [{}^1\mathbf{Q}_0, {}^1\mathbf{Q}_1] = -i^1\mathbf{Q}_2, \quad (32)$$

$$[{}^1\mathbf{P}_i, {}^1\mathbf{Q}_j] = 0 \quad (i, j = 0, 1, 2), \quad (33)$$

$$[{}^2\mathbf{K}_1, {}^2\mathbf{K}_2] = i^2\mathbf{K}_2, \quad [{}^2\mathbf{K}_2, {}^2\mathbf{K}_3] = i^2\mathbf{K}_1, \quad [{}^2\mathbf{K}_3, {}^2\mathbf{K}_1] = i^2\mathbf{K}_2, \quad (34)$$

$$[{}^2\mathbf{J}_1, {}^2\mathbf{J}_2] = i^2\mathbf{J}_2, \quad [{}^2\mathbf{J}_2, {}^2\mathbf{J}_3] = i^2\mathbf{J}_1, \quad [{}^2\mathbf{J}_3, {}^2\mathbf{J}_1] = i^2\mathbf{J}_2, \quad (35)$$

$$[{}^2\mathbf{K}_i, {}^2\mathbf{J}_j] = 0 \quad (i, j = 1, 2, 3), \quad (36)$$

$$[{}^2\mathbf{T}_1, {}^2\mathbf{T}_2] = -i^2\mathbf{T}_2, \quad [{}^2\mathbf{T}_2, {}^2\mathbf{T}_0] = i^2\mathbf{T}_1, \quad [{}^2\mathbf{T}_0, {}^2\mathbf{T}_1] = i^2\mathbf{T}_2, \quad (37)$$

$$[{}^2\mathbf{S}_1, {}^2\mathbf{S}_2] = -i^2\mathbf{S}_2, \quad [{}^2\mathbf{S}_2, {}^2\mathbf{S}_0] = i^2\mathbf{S}_1, \quad [{}^2\mathbf{S}_0, {}^2\mathbf{S}_1] = i^2\mathbf{S}_2, \quad (38)$$

$$[{}^2\mathbf{T}_i, {}^2\mathbf{S}_j] = 0 \quad (i, j = 0, 1, 2), \quad (39)$$

$$[{}^2\mathbf{P}_1, {}^2\mathbf{P}_2] = -i^2\mathbf{P}_2, \quad [{}^2\mathbf{P}_2, {}^2\mathbf{P}_0] = i^2\mathbf{P}_1, \quad [{}^2\mathbf{P}_0, {}^2\mathbf{P}_1] = i^2\mathbf{P}_2, \quad (40)$$

$$[{}^2\mathbf{Q}_1, {}^2\mathbf{Q}_2] = -i^2\mathbf{Q}_2, \quad [{}^2\mathbf{Q}_2, {}^2\mathbf{Q}_0] = i^2\mathbf{Q}_1, \quad [{}^2\mathbf{Q}_0, {}^2\mathbf{Q}_1] = i^2\mathbf{Q}_2, \quad (41)$$

$$[{}^2\mathbf{P}_i, {}^2\mathbf{Q}_j] = 0 \quad (i, j = 0, 1, 2). \quad (42)$$

В силу соотношений (25)–(27) генераторы ${}^1\mathbf{K}$ и ${}^1\mathbf{J}$ образуют алгебру Ли $\mathfrak{so}(4)$ подгруппы $SO(4)$, при этом в силу (27) ${}^1\mathbf{K}$ и ${}^1\mathbf{J}$ образуют базисы двух независимых алгебр $\mathfrak{so}(3)$. Таким образом, алгебра Ли $\mathfrak{so}(4)$ подгруппы $SO(4)$ изоморфна прямой сумме

$$\mathfrak{so}(4) \simeq \mathfrak{su}(2) \oplus \mathfrak{su}(2). \quad (43)$$

Далее, в силу соотношений (28)–(30) и (31)–(33) генераторы ${}^1\mathbf{T}$, ${}^1\mathbf{S}$ и ${}^1\mathbf{P}$, ${}^1\mathbf{Q}$ образуют алгебры Ли $\mathfrak{so}(2, 2)$ подгрупп $SO(2, 2)$, при этом в силу (30) (соотв. 33) генераторы ${}^1\mathbf{T}$, ${}^1\mathbf{S}$ (соотв. ${}^1\mathbf{P}$, ${}^1\mathbf{Q}$) образуют базисы двух независимых алгебр $\mathfrak{so}(2, 1)$. Следовательно, по аналогии с алгеброй $\mathfrak{so}(4)$ получим алгебру Ли $\mathfrak{so}(2, 2)$, которая локально изоморфна прямой сумме двух алгебр $\mathfrak{so}(2, 1)$:

$$\mathfrak{so}(2, 2) \simeq \mathfrak{so}(2, 1) \oplus \mathfrak{so}(2, 1). \quad (44)$$

Совокупность генераторов (4)–(9) и соответствующие коммутационные соотношения (25)–(33) описывают базис Яо [?] алгебры Ли $\mathfrak{su}(2, 2) \simeq \mathfrak{so}(4, 2)$ группы псевдоунитарных унимодулярных матриц,

$$SU(2, 2) \simeq \mathbf{Spin}_+(2, 4) = \left\{ \begin{bmatrix} A & B \\ C & D \end{bmatrix} \in \mathbb{C}_4 : \det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = 1 \right\}$$

являющейся двулистным накрытием конформной группы. $SU(2, 2)$ является группой преобразований 4-мерного комплексного пространства $\mathbb{C}^4$ (твисторного пространства), оставляющей инвариантной квадратичную форму $|Z_1|^2 + |Z_2|^2 - |Z_3|^2 - |Z_4|^2$. При этом *твисторы* (векторы пространства $\mathbb{C}^4$) определяются как редуцированные спиноры для конформной группы (см. Приложение Б в [12]).

Легко видеть, что генераторы (10)–(15) второй половины базиса (4)–(15) в силу коммутационных соотношений (34)–(42) описывают базис, структурно идентичный (с обратным знаком) базису Яо. Таким образом, *базис (4)–(15) алгебры Ли спинорной группы*

$$\mathbf{Spin}_+(4, 4) = \{s \in \text{Mat}_8(\mathbb{R}) \oplus \text{Mat}_8(\mathbb{R}) \mid N(s) = 1\}.$$

имеет структуру, расщеплённую на две половины.

5. Спин как четвёртый генератор подалгебры Картана

Как уже отмечалось выше в п. 4, в системах Генцеля и Финке четвёртое квантовое число s , обозначаемое в обеих системах двумя точками на трансверсалиях, не имеет ясного геометрического представления. Как следствие, никакая 3-мерная система не является полным представлением периодического закона. Спин s есть четвёртая степень свободы, поэтому для его адекватного представления требуется 4-мерная система. Ниже мы покажем, что такой системой является весовая диаграмма алгебры $\mathfrak{so}(4, 4)$.

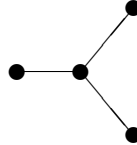
Четыре генератора (3) образуют базис подалгебры Картана $\mathfrak{K}$ алгебры $\mathfrak{so}(4, 4)$, среди которых первые три генератора составляют базис подалгебры Картана алгебры $\mathfrak{so}(4, 2)$ (алгебра Ли конформной группы $SO(4, 2)$). Согласно *водородной реализации* алгебры $\mathfrak{so}(4, 2)$ (представление Барута [13, 14], см. также [10]) первый генератор $\mathbf{L}_{12}$ соответствует третьей компоненте углового момента, $\mathbf{L}_{12} = \mathbf{L}_3$, второй

генератор $\mathbf{L}_{34}$ соответствует третьей компоненте вектора Лапласа – Рунге – Ленца, $\mathbf{L}_{34} = \mathbf{A}_3$. Наконец, третий генератор $\mathbf{L}_{56}$ задаёт третью компоненту радиальной части волновой функции водородоподобной системы, $\mathbf{L}_{56} = \Delta_3$. В свою очередь, генераторы $\mathbf{L}_3$ и $\mathbf{A}_3$ образуют базис подалгебры Картана алгебры Ли $\mathfrak{so}(4)$ группы $SO(4)$, в рамках которой В. Паули [15] объяснил полное вырождение спектра водорода. Три генератора $\mathbf{L}_{12}$, $\mathbf{L}_{34}$ и $\mathbf{L}_{56}$ образуют базис 3-мерной системы координат, в которой строится весовая диаграмма алгебры $\mathfrak{so}(4, 2)$ в виде двойной опрокинутой пирамиды ($SO(4, 2)$ -башня) (см. рис. 6 в [11]). Собственные значения генераторов $\mathbf{L}_{56}$, $\mathbf{L}_{12}$, $\mathbf{L}_{34}$ соответствуют квантовым числам n, l, m . Включение четвёртого генератора $\mathbf{L}_{78}$ приводит к 8-мерному обобщению $\mathfrak{so}(4, 4)$, что позволяет отождествить собственные значения генератора $\mathbf{L}_{78}$ с четвёртым квантовым числом s . Таким образом,

$$\begin{aligned}\mathbf{L}_{56} &\longrightarrow n, \\ \mathbf{L}_{12} &\longrightarrow l, \\ \mathbf{L}_{34} &\longrightarrow m, \\ \mathbf{L}_{78} &\longrightarrow s.\end{aligned}$$

6. Весовая диаграмма алгебры $\mathfrak{so}(4, 4)$

В отличие от весовой диаграммы алгебры Ли $\mathfrak{so}(4, 2)$, допускающей визуализацию в 3-мерном пространстве [11], весовая диаграмма алгебры $\mathfrak{so}(4, 4)$ определена в 4-мерном пространстве, что, естественно, существенным образом затрудняет её непосредственную визуализацию. Диаграмма Дынкина алгебры $\mathfrak{so}(4, 4) \simeq D_4$ имеет вид



Однако, используя свойства базиса (4)–(15) алгебры $\mathfrak{so}(4, 4)$, заключающееся в расщеплении на два структурно идентичных (изоморфных) базиса Яо подалгебры $\mathfrak{so}(4, 2)$, мы можем частично визуализировать весовую диаграмму алгебры $\mathfrak{so}(4, 4)$ посредством *двух 3-мерных проекций*.

Весовое пространство определяется четырьмя генераторами Картана $\mathbf{L}_{56} = \Delta_3$, $\mathbf{L}_{12} = \mathbf{L}_3$, $\mathbf{L}_{34} = \mathbf{A}_3$ и $\mathbf{L}_{78}$, которые служат базисом 4-мерной ортогональной системы координат. Определим кет-вектор

$$\left| \nu, \dot{\nu}; \lambda, \dot{\lambda}; \mu, \dot{\mu}; \sigma, \dot{\sigma} \right\rangle, \quad (45)$$

где $\nu, \dot{\nu} \in \{0, 1/2, 1, 3/2, \dots\}$, $\lambda \in \{-\nu, -\nu + 1, \dots, \nu - 1, \nu\}$, $\dot{\lambda} \in \{-\dot{\nu}, -\dot{\nu} + 1, \dots, \dot{\nu} - 1, \dot{\nu}\}$, $\mu \in \{-\lambda, -\lambda + 1, \dots, \lambda - 1, \lambda\}$, $\dot{\mu} \in \{-\dot{\lambda}, -\dot{\lambda} + 1, \dots, \dot{\lambda} - 1, \dot{\lambda}\}$, $\sigma, \dot{\sigma} \in \{-1/2, +1/2\}$.

Связь чисел, входящих в кет-вектор (45), с квантовыми числами n, l, m и s кет-вектора

$$|n, l, m, s\rangle \quad (46)$$

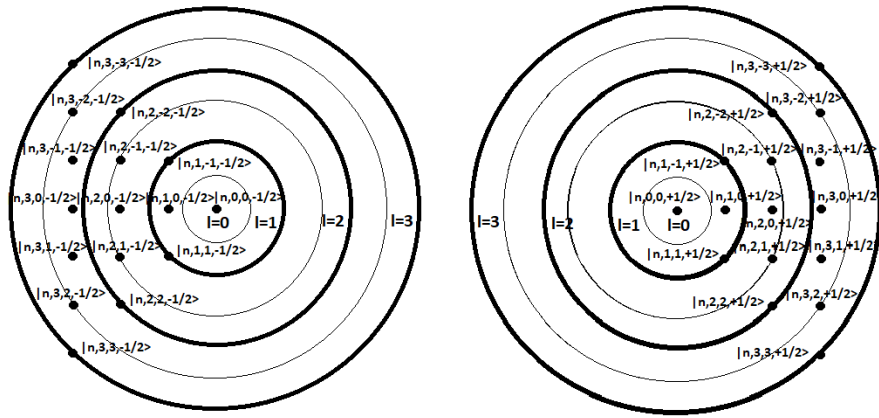


Рис. 4. Двумерные сечения 3-мерных проекций весовой диаграммы алгебры $\mathfrak{so}(4, 4)$ с кет-векторами $|n, l, m, -1/2\rangle$ и $|n, l, m, +1/2\rangle$

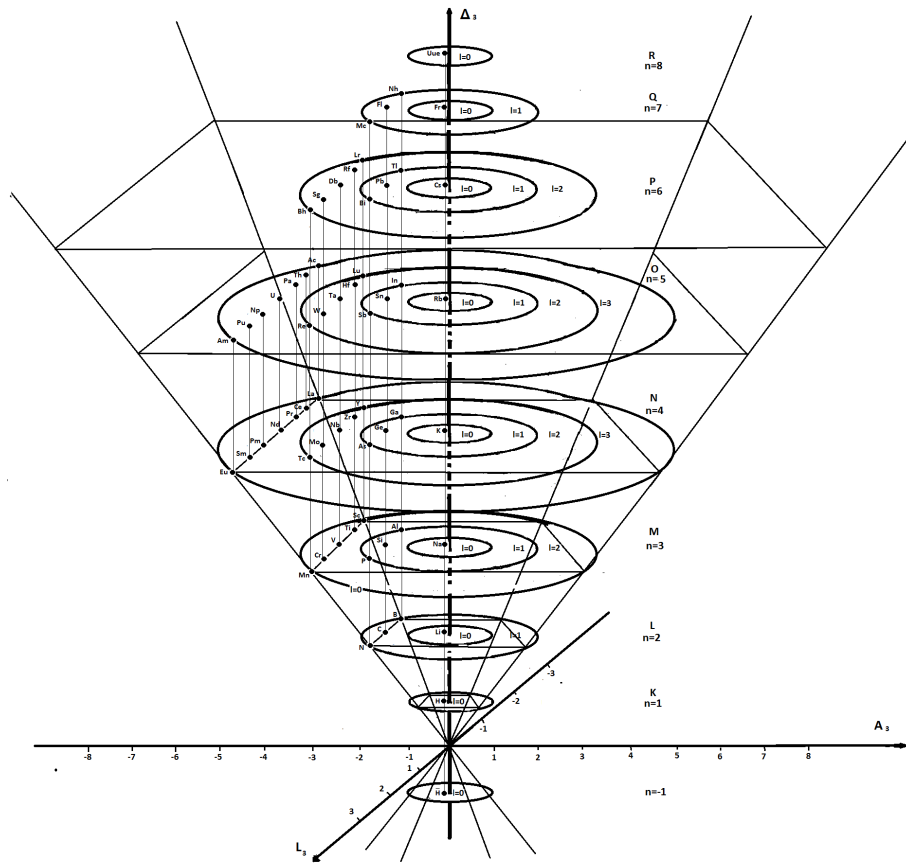


Рис. 5. $SO(4, 2)$ -башня для $s = -1/2$ с элементами $\mathbf{H} = |1, 0, 0, -1/2\rangle, \dots, \mathbf{Uue} = |8, 0, 0, -1/2\rangle$

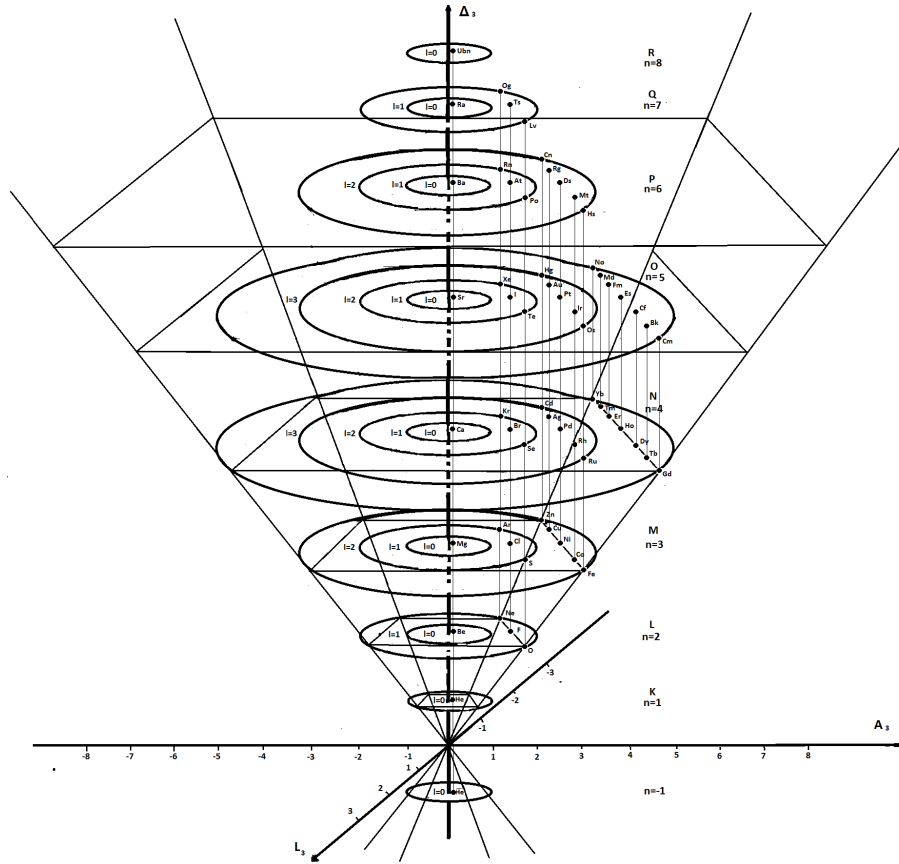


Рис. 6. $SO(4, 2)$ -башня для $s = +1/2$ с элементами $\mathbf{He} = |1, 0, 0, +1/2\rangle, \dots, \mathbf{Ubn} = |8, 0, 0, +1/2\rangle$

базиса Маделунга задаётся следующими соотношениями:

$$n = |\nu - \dot{\nu}|, \quad l = |\lambda - \dot{\lambda}|, \quad m = |\mu - \dot{\mu}|, \quad s = |\sigma - \dot{\sigma}|.$$

При этом главное квантовое число n изменяется в пределах

$$-|\nu - \dot{\nu}|, -|\nu - \dot{\nu}| + 1, -|\nu - \dot{\nu}| + 2, \dots, |\nu - \dot{\nu}|. \quad (47)$$

На рис. 4 представлены 2-мерные сечения двух 3-мерных проекций весовой диаграммы алгебры $\mathfrak{so}(4, 4)$.

Первая 3-мерная проекция, соответствующая элементам с $s = -1/2$, показана на рис. 5. Верхняя пирамида $SO(4, 2)$ -башни, начиная с водорода $\mathbf{H} = |1, 0, 0, -1/2\rangle$ до гипотетического элемента $\mathbf{Uue} = |8, 0, 0, -1/2\rangle$, образует *пирамиду материи*. Отражённая от плоскости (L_3, A_3) нижняя пирамида начинается с антиводорода $\bar{\mathbf{H}} = |-1, 0, 0, -1/2\rangle$ и образует *пирамиду антиматерии*⁵.

Вторая 3-мерная проекция весовой диаграммы алгебры $\mathfrak{so}(4, 4)$, соответствующая элементам с $s = +1/2$, показана на рис. 6. В этом случае пирамида материи

⁵Первые свидетельства существования антиводорода были получены на ускорителях в 1990-х гг. Впервые захват атомов антиводорода в ловушку продемонстрировала группа Antihydrogen Laser Physics Apparatus (ALPHA) в ЦЕРН в 2010 г.

содержит элементы от гелия $\mathbf{He} = |1, 0, 0, +1/2\rangle$ до оганессона $\mathbf{Og} = |7, 1, 1, +1/2\rangle$, включая гипотетический элемент $\mathbf{Ubn} = |8, 0, 0, +1/2\rangle$. Пирамида антиматерии начинается с антигелия $\overline{\mathbf{He}} = |-1, 0, 0, +1/2\rangle$.

Гомологичные элементы на диаграммах (рис. 5 и 6) соединяются вертикальными линиями (линии Бэйли – Томсена – Бора). На диаграммах видно, что, начиная с листа $n = 5$ (5-го этажа $SO(4, 2)$ -башни), остаются незаполненными следующие круговые кольца (кольца Генцеля): кольцо $l = 4$ для листа $n = 5$, кольца $l = 3, l = 4, l = 5$ для листа $n = 6$, а также $l = 2, l = 3, l = 4, l = 5$ для листа $n = 7$. На этих уровнях располагаются гипотетические элементы таблицы Сиборга и таблицы 10-периодического расширения (см. [7, 8]), которые на диаграммах рис. 5 и 6 опущены с целью не загромождать построение. Так, для элементов **Ac, Th, Pa, U, Np, Pu, Am** уровня ($n = 5, l = 3$) гомологичными элементами являются **Ute, Uqn, Uqu, Uqb, Uqt, Uqq, Uqp** уровня ($n = 6, l = 3$) таблицы Сиборга и т. д. Из построения видно, что полностью «заселёнными» (экспериментально открытыми элементами) являются только первые четыре этажа $SO(4, 2)$ -башни. Кроме того, начиная с четвёртого этажа (круга Генцеля) $n = 4$, стабильность элементов возрастает по мере приближения к радиальной оси Δ_3 .

Литература

1. Spronsen J.W. van. The Periodic System of Chemical Elements. Amsterdam. New York: Elsevier, 1969.
2. Mazurs E.G. Graphic Representations of the Periodic System During One Hundred Years. Tuscaloosa: University of Alabama Press, 1974.
3. Scerri E. Periodic Table: its story and its significance. New York: Oxford University Press, 2019.
4. Haenzel G. Die Polygonfläche und das Periodische System der Elemente // Zeitschrift für Physik. 1943. Vol. 120. P. 283–300.
5. Finke W. Bemerkungen zu G. Haenzel „Die Polygonfläche und das Periodische System der Elemente“ // Zeitschrift für Physik. 1943. Vol. 121. P. 586–587.
6. Варламов В.В. Теоретико-групповое описание периодической системы элементов // Математические структуры и моделирование. 2018. № 2 (46). С. 5–23.
7. Варламов В.В. Теоретико-групповое описание периодической системы элементов II.: Таблица Сиборга // Математические структуры и моделирование. 2019. № 1 (49). С. 5–21.
8. Варламов В.В. Теоретико-групповое описание периодической системы элементов III.: 10-периодическое расширение // Математические структуры и моделирование. 2019. № 3 (51). С. 5–20.
9. Varlamov V.V., Pavlova L.D, Babushkina O.S. Group Theoretical Description of the Periodic System // Symmetry. 2022. Vol. 14. P. 137.
10. Варламов В.В. Теоретико-групповое описание периодической системы элементов IV.: Групповая алгебра // Математические структуры и моделирование. 2024. № 1 (69). С. 18–31.
11. Варламов В.В. Теоретико-групповое описание периодической системы элементов V.: Весовая диаграмма // Математические структуры и моделирование. 2024. № 3 (71). С. 4–18.

12. Варламов В.В. О системе аксиом нелокальной квантовой теории // Математические структуры и моделирование. 2017. № 4 (44). С. 5–25.
13. Barut A.O. Group Structure of the Periodic System / The Structure of Matter: Rutherford Centennial Symposium. Christchurch, New Zeland: University of Canterbury Press, 1972. P. 126–136.
14. Adams B.G., Čížek, Paldus J. Representation Theory of $SO(4,2)$ for the Perturbation Treatment of Hydrogenic-Type Hamiltonians by Algebraic Methods // Int. J. Quant. Chem. 1982. Vol. 21. P. 153–171.
15. Pauli W. Über das Wasserstoffspektrum vom Standpunkt der neuen Quantenmechanik // Zeitschrift für Physik. 1926. Vol. 36. P. 336–363.

GROUP THEORETICAL DESCRIPTION OF PERIODIC SYSTEM OF ELEMENTS: VI. AN EIGHT-DIMENSIONAL GENERALIZATION

V.V. Varlamov

Dr.Sc. (Phys.-Math.), e-mail: varlamov@sibsiu.ru

Siberian State Industrial University, Novokuznetsk, Russia

Abstract. The periodic system of chemical elements is represented within the framework of the weight diagram of the Lie algebra of the fourth rank of the rotation group of an eight-dimensional pseudo-Euclidean space. Spin is interpreted as the fourth generator of the Cartan subalgebra, whose two eigenvalues correspond to two three-dimensional projections of the weight diagram containing elements of the periodic system from hydrogen to moscovium (the first projection) and from helium to oganesson (the second projection). One of the main advantages of the proposed group-theoretic construction of the periodic system is the natural inclusion of antimatter in the general scheme.

Keywords: periodic system, Lie algebra, weight diagram, Cartan subalgebra, spin, antimatter.

Дата поступления в редакцию: 01.12.2024

ГРАВИТАЦИЯ КАК КВАНТОВАЯ ЗАПУТАННОСТЬ МАССИВНЫХ ТЕЛ

А.К. Гуц

д.ф.-м.н., профессор, e-mail: aguts@mail.ru

Сочинский государственный университет, Сочи, Россия

Аннотация. Показано, что формулу Ньютона для силы всемирного тяготения можно представить как силу, отвечающую квантовой запутанности тяготеющих тел. Обсуждается природа гравитации.

Ключевые слова: гравитация, формула Ньютона, запутанность.

Введение

Дав выражения для *силы* взаимодействия тел с массами M и m , воспринимаемое людьми как притяжение тел, и объяснив посредством понятия *притяжение* падение тел на Землю, И. Ньютон ничего не сказал о природе притяжения. Природа гравитации оставалась загадкой для физиков вплоть до создания А. Эйнштейном общей теории относительности (ОТО). Притяжение тел по А. Эйнштейну – это результат искривления 4-мерного пространства-времени.

Заметим, что основополагающие уравнения Эйнштейна

$$R_{ik} - \frac{1}{2}g_{ik}R = \kappa T_{\text{мат.},ik}$$

были выведены Д. Гильбертом из вариационного принципа

$$\delta \int R \sqrt{-g} d^4x - \kappa \delta \int L_{\text{мат.}} \sqrt{-g} d^4x = 0.$$

Однако константа $\kappa = 8\pi G/c^4$ была найдена только при выписывании уравнений Эйнштейна для слабого метрического поля g_{ik} и при сравнении полученного уравнения с уравнением Пуассона

$$\Delta U = 4\pi G \rho_{\text{мат.}},$$

которому удовлетворяет ньютоновский потенциал гравитационного поля $U = -GM/r$, где r – расстояние между телами.

Именно величина r , а точнее, её квадрат, и фигурирует во всех высказываниях И. Ньютона в его гениальном труде «Математические начала натуральной философии». Эти высказывания на языке математики и дают следующее знаменитое выражение для силы притяжения двух тел:

$$F = G \frac{Mm}{r^2}. \quad (1)$$

Почему именно r^2 ? Объясняя этот факт, можно сослаться на эксперименты, можно сослаться на то, что именно в таком виде формула хорошо объясняет как движение планет, так и падение тел на Земле. Кроме того, существуют работы, показывающие, что если взять другую степень величины r , то они подходят лишь для пространства иной размерности, чем 3.

Но хотелось бы чисто математически получить r^2 , опираясь на вторую по значимости физическую теорию, открытую в XX в., т. е. на квантовую механику. Тем самым была бы установлена квантовая природа ньютоновской гравитации.

1. Дальнодействие гравитации

Сила притяжения Ньютона передаётся от тела к телу мгновенно, т. е. с бесконечной скоростью. Теория относительности, как известно, не допускает движения материальных тел быстрее скорости света. Гравитационные волны и квантовые носители гравитации, гравитоны, перемещаются со скоростью света. То есть гравитация Эйнштейна, базирующаяся на постулате псевдоримановости метрики g_{ik} пространства-времени, казалось бы, отвергает дальнодействие ньютоновской гравитации. Но это только на первый взгляд. Дело в том, что псевдориманова метрика появляется на базовом гладком многообразии M^4 априори сразу и везде, т. е. искривление даётся изначально, и как оно появилось там или здесь ОТО никогда не поясняла. Грубо говоря, искривление даётся везде на M^4 *мгновением мысли*. А это явно форма дальнодействия. Поэтому фактически дальнодействие в ОТО постулируется при произнесении фразы: «Рассмотрим лоренцево многообразие M^4 с метрикой g_{ik} , удовлетворяющее уравнениям Эйнштейна с правой частью, заданной тензором $T_{\text{мат},ik}$ ». Другой вопрос, как это дальнодействие проявляется в ОТО.

К сказанному остаётся только добавить, что аналог формулы (1) появляется в пространстве-времени Шварцильда, описывающем гравитационное поле Солнца (или планет) в виде, предполагающем всё то же дальнодействие притяжения [1, с. 26]:

$$F = G \frac{Mm}{r^2 \sqrt{1 - \frac{r_g}{r}}}. \quad (2)$$

Однако квантовая механика предсказала существование мгновенной передачи воздействия, названного *запутанностью* (сцепленностью, квантовой корреляцией, несиловым воздействием – А.Д. Александров), подтверждённой экспериментально для частиц. Таким образом, появилась возможность объяснить ньютоновское притяжение (а также кулоновское взаимодействие между точечными зарядами) как квантовую запутанность.

2. Новая формула для силы притяжения

Оставляя в стороне ОТО, где сила притяжения может иметь форму, отличную от формулы Ньютона, и возвращаясь к ньютоновской гравитации, постараемся объяснить природу величины r^2 посредством постулирования конкретной формы кван-

тового дальнего действия, которое вошло в квантовую механику посредством понятия *запутанности* квантовых объектов.

Пусть дано евклидово 3-мерное пространство V^3 , которое рассматриваем как 3-мерное векторное пространство. Вместо формулы (1) вводим формулу

$$F = G \frac{Mm}{\left\| \frac{1}{\sqrt{2}}(|r_M\rangle \otimes |r_m\rangle - |r_m\rangle \otimes |r_M\rangle) \right\|^2}. \quad (3)$$

Здесь $\vec{r}_M, \vec{r}_m$ – радиус-векторы центров двух тяготеющих масс M и m в пространстве V^3 , в обозначениях квантовой механики записываемые как $|r_M\rangle \otimes |r_m\rangle$ и $|r_m\rangle \otimes |r_M\rangle$ соответственно, а тензоры $\vec{r}_M \otimes \vec{r}_m, \vec{r}_m \otimes \vec{r}_M$, или $|r_M\rangle \otimes |r_m\rangle$ и $|r_m\rangle \otimes |r_M\rangle$, – элементы пространства $V^3 \otimes V^3$.

В знаменателе формулы (3) стоит квадрат нормы пространства $V^3 \otimes V^3$ неразложимого (запутанного) элемента этого тензорного произведения векторных пространств.

Фактически мы считаем запутанными физические объекты с массами M и m . Неразложимость, или запутанность, тензора

$$\frac{1}{\sqrt{2}}(|r_M\rangle \otimes |r_m\rangle - |r_m\rangle \otimes |r_M\rangle) \quad (4)$$

показана ниже в п. 3.

Следовательно, формула (3) сразу постулирует мгновенный способ передачи ньютоновского притяжения, указывая тем самым на квантовую природу гравитации.

Формула (3) – это не что иное, как формула (1). Покажем это.

Возьмём векторы $\vec{r}_M, \vec{r}_m$ следующим образом:

$$\vec{r}_M = (0, 0, 1), |\vec{r}_M| = ||r_M\rangle| = 1 \text{ (эталонный вектор)}, \vec{r}_M \perp \vec{r}, \vec{r} = \vec{r}_m - \vec{r}_M.$$

Тогда

$$\sin \angle(\vec{r}_M, \vec{r}_m) = \frac{r}{||r_m\rangle|}.$$

Вычисляя знаменатель в (3), получаем

$$\begin{aligned} & \left\| \frac{1}{\sqrt{2}}(|r_M\rangle \otimes |r_m\rangle - |r_m\rangle \otimes |r_M\rangle) \right\|^2 = \\ &= \frac{1}{2}[(|r_M\rangle \otimes |r_m\rangle - |r_m\rangle \otimes |r_M\rangle, |r_M\rangle \otimes |r_m\rangle - |r_m\rangle \otimes |r_M\rangle)] = \\ &= \frac{1}{2}[(|r_M\rangle \otimes |r_m\rangle, |r_M\rangle \otimes |r_m\rangle) + (|r_m\rangle \otimes |r_M\rangle, |r_m\rangle \otimes |r_M\rangle) - \\ & - (|r_M\rangle \otimes |r_m\rangle, |r_m\rangle \otimes |r_M\rangle) - (|r_m\rangle \otimes |r_M\rangle, |r_M\rangle \otimes |r_m\rangle)] = \\ &= \frac{1}{2}[(|r_M\rangle, |r_M\rangle)(|r_m\rangle, |r_m\rangle) + (|r_m\rangle, |r_m\rangle)(|r_M\rangle, |r_M\rangle) - \\ & - (|r_M\rangle, |r_m\rangle)(|r_m\rangle, |r_M\rangle) - (|r_m\rangle, |r_M\rangle)(|r_M\rangle, |r_m\rangle)] = \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2}[2||r_m\rangle|^2||r_M\rangle|^2 - 2|(|r_m\rangle, |r_M\rangle)|^2] = ||r_m\rangle|^2 - ||r_m\rangle|^2 \cos^2 \angle(\vec{r}_M, \vec{r}_m) = \\
&= ||r_m\rangle|^2 \sin^2 \angle(\vec{r}_M, \vec{r}_m) = ||r_m\rangle|^2 \frac{r^2}{||r_m\rangle|^2} = r^2.
\end{aligned}$$

Как видим, знаменатели в (1) и (3) совпадают, что мы и хотели показать.

Если в формулу (2) вместо r подставить выражение для r , вычисляемое по формуле (4), то и в ОТО гравитацию можно объяснять запутанностью двух массивных объектов.

3. Проверка запутанности элемента (4)

Раскладываем векторы $\vec{r}_M, \vec{r}_m$ по базисным $\vec{e}_1, \vec{e}_2, \vec{e}_3$. Имеем

$$\vec{r}_m = \sum_{i=1}^3 a_i \vec{e}_i, \quad \vec{r}_M = \sum_{j=1}^3 b_j \vec{e}_j.$$

Раскладываем по базисным векторам $\vec{e}_i \otimes \vec{e}_j$ пространства $V^3 \otimes V^3$ тензор

$$\vec{r}_m \otimes \vec{r}_M - \vec{r}_M \otimes \vec{r}_m = \sum_{i,j=1}^3 \alpha_{ij} \vec{e}_i \otimes \vec{e}_j.$$

Получаем матрицу, помня, что $\vec{r}_M = (0, 0, 1)$:

$$A = (\alpha_{ij}) = \begin{pmatrix} 0 & 0 & a_1 b_3 \\ 0 & 0 & a_2 b_3 \\ -b_3 a_1 & -b_3 a_2 & a_3 b_3 - b_3 a_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & a_1 \\ 0 & 0 & a_2 \\ -a_1 & -a_2 & 0 \end{pmatrix}.$$

Выписываем миноры 2-го порядка:

$$M_{22} = \begin{vmatrix} 0 & a_1 \\ -a_1 & 0 \end{vmatrix} = a_1^2, \quad M_{11} = \begin{vmatrix} 0 & a_2 \\ -a_2 & 0 \end{vmatrix} = a_2^2, \quad M_{12} = \begin{vmatrix} 0 & a_2 \\ -a_1 & 0 \end{vmatrix} = a_1 a_2.$$

Все эти миноры равны нулю, если $a_1 = a_2 = 0$. Но тогда $r = 0$, т. е. массы имеют один центр. Мы же рассматриваем разнесённые в пространстве тела (точечные частицы). Значит хотя бы $a_1 \neq 0$ или $a_2 \neq 0$. Получаем, что один из миноров отличен от нуля. Значит $\text{rang}(A) > 1$. И в силу критерия запутанности, приведённого в [2, с. 31], наш тензор (4) запутан.

4. Носитель гравитации

Если это гравитон, то его траектория, как и фотона, изотропна в пространстве-времени, т. е. для неё интервал $ds^2 = 0$ и его собственное время равно нулю, т. е. его перемещение для него мгновенно. Но есть ещё одна сущность, которая перемещается таким образом. Это *время*, как утверждал Н.А. Козырев [3], и что он наблюдал, ставя эксперименты, подтверждённые сотрудниками Института математики СО АН СССР академиками М.М. Лаврентьевым, И.А. Егановой, М.К. Луцетом, С.Ф. Фоминых [4].

Заключение

Природа гравитации, как известно, самим И. Ньютоном никак не была описана – он предпочитал обсчитываемые факты, а не фантазии на тему их сущности. Запутанность, во всяком случае, экспериментально зафиксирована, и экспериментаторы вознаграждены Нобелевской премией. Роль Н.А. Козырева в этих изысканиях на научных форумах не обсуждается. Если обратиться к статье [5]¹, то запутанность частиц – это порождение кротовой норы (и наоборот), соединяющей тяготеющие частицы, именно тяготеющие, иных мы не знаем. Иначе говоря, природа гравитации носит топологический характер – тяготение идёт не по кратчайшей в *неизменном* пространстве, а посредством нарушения односвязности пространства.

Литература

1. Зельдович Я.Б., Новиков И.Д. Общая теория относительности и астрофизика // Эйнштейновский сборник. М.: Наука, 1966. С. 18–120.
2. Кронберг Д.А., Ожигов Ю.И., Чернявский А.Ю. Алгебраический аппарат квантовой информатики. М.: МГУ, Факультет ВМК, 2016.
3. Козырев Н.А. Астрономическое доказательство реальности четырехмерной геометрии Минковского // Проявление космических факторов на Земле и звёздах: сб. ст. М.-Л.: Всесоюз. астроном.-геодез. о-во, 1980. С. 85–93.
4. Лаврентьев М.М., Еганова И.А., Луцет М.К., Фоминых С.Ф. О дистанционном воздействии звезд на резистор // Докл. АН СССР. 1990. Т. 314, № 2. С. 352–355.
5. Maldacena J., Susskind L. Cool horizons for entangled black holes. ArXiv:1306.0533v2.

GRAVITY AS QUANTUM ENTANGLEMENT OF MASSIVE BODIES

A.K. Guts

Dr.Sc. (Phys.-Math.), Professor, e-mail: aguts@mail.ru

Sochi State University, Sochi, Russia

Abstract. It is shown that Newton's formula for the force of universal gravitation can be represented as a force corresponding to the quantum entanglement of gravitating bodies. The nature of gravity is discussed.

Keywords: gravity, Newton's formula, entanglement.

Дата поступления в редакцию: 23.02.2025

¹Цитата: «Очень заманчиво думать, что любая система, коррелированная с EPR, связана каким-то ER-мостом, хотя в целом этот мост может быть высококвантовым объектом, который еще предстоит независимо определить» [5]. Здесь EPR – это запутанность, ER-мост – кротовая нора.

ON THE KIRCHHOFF INDEX FOR CIRCULANT GRAPH WITH NON-FIXED JUMPS

G.K. Sokolova^{1,2}

Ph.D. student, Assistant, e-mail: g.sokolova@g.nsu.ru

¹Novosibirsk State University, Novosibirsk, Russia

²Novosibirsk State Technical University, Novosibirsk, Russia

Abstract. This article deals with a graph invariant called the Kirchhoff index. An explicit analytical formula for the Kirchhoff index of circulant graphs with non-fixed jumps is defined.

Keywords: circulant graph with non-fixed jumps, Laplacian eigenvalues, Kirchhoff index of a graph.

Introduction

The Kirchhoff index $Kf(G)$ of a finite connected graph G was originally defined by D.J. Klein and M. Randić [1] as the mean resistance distance between its vertices, in other words,

$$Kf(G) = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n r_{ij},$$

where r_{ij} is the resistance distance between the vertices v_i and v_j , i.e. r_{ij} is equal to the resistance between equivalent points on an associated electrical network obtained by replacing each edge of G by a unit resistor. Later the simple formula

$$Kf(G) = n \sum_{j=2}^n \frac{1}{\lambda_j},$$

relating the Kirchhoff index to the spectrum of the Laplace matrix was independently found in [2] by I. Gutman, B. Mohar and [3] by H.Y. Zhu, D.J. Klein and I. Lukovits. Kirchhoff indices for various graph families have been studied, for example, in [4–8]. In particular, the analytical formula for the Kirchhoff index for a circulant graph with fixed jumps was found in paper [8], where the asymptotics of this formula was also investigated.

The main goal of this paper is to find explicit analytical formulas for the Kirchhoff indices of circulant graphs with non-fixed jumps. We will present these formulas as sums of finitely many terms whose number is independent of n , and each of these terms amounts to a rational function evaluated at the roots of some fixed polynomial.

1. Preliminaries and preliminary results

Consider the finite connected graph G_n , that is, a graph containing one component of connectivity with finite sets of vertices $V(G)$ and edges $E(G)$. Suppose a graph G_n allows multiple edges, but not loops. The following class of circulant graphs is considered throughout this paper.

Definition 1. A graph $G_n = C_{\beta n}(s_1, \dots, s_k, \alpha_1 n, \dots, \alpha_\ell n)$ is called a *circulant graph with non-fixed jumps* $1 \leq s_1 < \dots < s_k < \lfloor \frac{\beta n}{2} \rfloor$ and $1 \leq \alpha_1 < \dots < \alpha_\ell \leq \lfloor \frac{\beta}{2} \rfloor$ on βn vertices if any i -th vertex is adjacent to vertices $i \pm s_1, i \pm s_2, \dots, i \pm s_k$ and $i \pm \alpha_1 n, i \pm \alpha_2 n, \dots, i \pm \alpha_\ell n$ modulo βn . Here β and ℓ are positive integers and n is assumed to be sufficiently large.

Note that if $\alpha_\ell < \lfloor \frac{\beta}{2} \rfloor$, then the graph G_n does not contain multiple edges. If $\alpha_\ell = \lfloor \frac{\beta}{2} \rfloor$, then any i -th vertex is connected to the vertex $i \pm \frac{\beta n}{2}$ modulo βn by two parallel edges.

Let $A = \{a_{uv}\}_{u,v \in V(G)}$ be the *adjacency matrix* of G , where a_{uv} is the number of edges between vertices u and v of G . Let us introduce a *valency matrix* $D = \{d_{vv}\}_{v \in V(G)}$, where d_{vv} is a degree of the vertex $v \in V(G)$ which may be determined by $d_{vv} = \sum_{u \in V(G)} a_{uv}$.

Then the matrix $\mathcal{L} = D - A$ is called the *Laplace matrix* or *Laplacian* of the graph G .

Associate with each graph $G_n = C_{\beta n}(s_1, \dots, s_k, \alpha_1 n, \dots, \alpha_\ell n)$ the associated Laurent polynomial

$$L(z) = 2(k + \ell) - \sum_{i=1}^k (z^{s_i} + z^{-s_i}) - \sum_{m=1}^{\ell} (z^{\alpha_m n} + z^{-\alpha_m n}),$$

describing the structure of the Laplace matrix $\mathcal{L}$ of a given graph. Note that the numbering of the vertices of a circulant graph G_n may be chosen in such a way that the adjacency matrix A and the Laplace matrix $\mathcal{L}$ of G_n are circulant. Recall that a matrix of order n is called *circulant* if it has the form

$$\text{circ}(x_1, x_2, \dots, x_n) = \begin{pmatrix} x_1 & x_2 & x_3 & \dots & x_n \\ x_n & x_1 & x_2 & \dots & x_{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_2 & x_3 & x_4 & \dots & x_0 \end{pmatrix}.$$

This means the Laplacian of the graph G_n may be defined as the matrix

$$\mathcal{L} = L(T) = 2(k + \ell)\mathbb{E} - \sum_{i=1}^k (T^{s_i} + T^{-s_i}) - \sum_{m=1}^{\ell} (T^{\alpha_m n} + T^{-\alpha_m n}),$$

where $T = \text{circ}(0, 1, 0, \dots, 0)$ is circulant matrix of order βn , which represents the cyclic shift operator $T : (x_1, x_2, \dots, x_{\beta n-1}, x_{\beta n}) \rightarrow (x_2, x_3, \dots, x_{\beta n}, x_1)$, and $\mathbb{E}$ is the identity matrix of the respective order. Fined the spectrum of the Laplace matrix. Suppose that λ be

an eigenvalue of $\mathcal{L}$ and v be the respective eigenvector. As is known, then $\det(\mathcal{L} - \lambda \mathbb{E}) = 0$ and the following system of linear equations holds

$$\left[(2(k + \ell) - \lambda) \mathbb{E} - \sum_{i=1}^k (T^{s_i} + T^{-s_i}) - \sum_{m=1}^{\ell} (T^{\alpha_m n} + T^{-\alpha_m n}) \right] v = 0. \quad (1)$$

Note that [9] the powers of the primitive root of unity $\zeta_{\beta n}^j$ are the eigenvalues of T , here $j = 0, 1, \dots, \beta n - 1$ and $\zeta_{\ell} = e^{i \frac{2\pi}{\ell}}$. This means that the matrix T is similar to the diagonal matrix $\mathbb{T} = \text{diag}(1, \zeta_{\beta n}, \dots, \zeta_{\beta n}^{\beta n-1})$ and the unit vectors $e_{j+1} = (0, \dots, 0, \underbrace{1}_{j+1\text{-th}}, 0, \dots, 0)$ of length βn are Laplacian eigenvectors. The matrix of system (1) is written in diagonal form

$$\left[(2(k + \ell) - \lambda) \mathbb{E} - \sum_{i=1}^k (\mathbb{T}^{s_i} + \mathbb{T}^{-s_i}) - \sum_{m=1}^{\ell} (\mathbb{T}^{\alpha_m n} + \mathbb{T}^{-\alpha_m n}) \right] e_j = 0.$$

From this relation it follows the eigenvalues λ_j of the Laplacian $\mathcal{L}$ are given by the formula

$$\lambda_j = 2(k + \ell) - \sum_{i=1}^k (\zeta_{\beta n}^{js_i} + \zeta_{\beta n}^{-js_i}) - \sum_{m=1}^{\ell} (\zeta_{\beta n}^{j\alpha_m n} + \zeta_{\beta n}^{-j\alpha_m n}) = L(\zeta_{\beta n}^j).$$

Recall the considered graph G_n is assumed to be connected. This means that $\lambda_0 = 0$ and $\lambda_j > 0$ for $j = 1, 2, \dots, \beta n - 1$.

In conclusion of this section, we present Theorem and Lemma from the article [8] which are necessary to prove the main result of this paper.

Theorem 1. *The Kirchhoff index of the circulant graph $G = C_n(s_1, s_2, \dots, s_k)$ can be calculated as*

$$Kf_G = \frac{n}{12 \sum_{i=1}^k s_i^2} \left(n^2 - \frac{\sum_{i=1}^k s_i^4}{\sum_{i=1}^k s_i^2} \right) + \sum_{p=2}^{s_k} \frac{n^2 \mathcal{U}_{n-1}(w_p)}{Q'(w_p)(1 - \mathcal{T}_n(w_p))},$$

where w_p is a root of the polynomial $Q(w) = \sum_{j=1}^k (2 - 2\mathcal{T}_{s_j}(w))$ distinct from 1, where $\mathcal{T}_n(w)$ and $\mathcal{U}_n(w)$ are the Chebyshev polynomials of the first and the second kind respectively.

Lemma 1. *Consider two nonconstant polynomials $P(w)$ and $R(w)$ of degrees n and m , respectively. Denote the roots of $P(w)$ by $\alpha_1, \alpha_2, \dots, \alpha_n$ and the roots of $R(w)$ by $\beta_1, \beta_2, \dots, \beta_m$. Suppose that $R(w)$ lacks multiple roots and that $P(w)$ and $R(w)$ lack common roots. Then*

$$\sum_{j=1}^n \frac{1}{R(\alpha_j)} = - \sum_{j=1}^m \frac{1}{R'(\beta_j)} \frac{P'(\beta_j)}{P(\beta_j)}.$$

2. Kirchhoff index for circulant graph

In this section, an explicit analytical formula for the Kirchhoff index of a circulant graph with non-fixed jumps is given. The formula contains a sum whose terms amount to analytical expressions evaluated at the roots of a prescribed polynomial of degree s_k .

Consider the Laurent polynomial $L(z)$ for $G_n = C_{\beta n}(s_1, \dots, s_k, \alpha_1 n, \dots, \alpha_\ell n)$ and represent it as the sum $L(z) = P(z) + p(z^n)$ of polynomials

$$P(z) = 2k - \sum_{i=1}^k (z^{s_i} + z^{-s_i}), \quad p(z) = 2\ell - \sum_{m=1}^{\ell} (z^{\alpha_m} + z^{-\alpha_m}).$$

Introduce the following set of polynomials

$$P_u(z) = P(z) + p(\zeta_{\beta n}^{un}), \quad u = 0, 1, \dots, \beta n - 1, \quad (2)$$

Where it is easy to see that

$$p(\zeta_{\beta n}^{un}) = 4 \sum_{m=1}^{\ell} \sin^2 \left(\frac{u \alpha_m \pi}{\beta} \right).$$

Let $\mathcal{T}_n(w) = \cos n\theta$ be the Chebyshev polynomial of the first kind [10], here $\theta = \arccos w$. Since the equality $\mathcal{T}_n \left(\frac{z+z^{-1}}{2} \right) = \frac{z^n + z^{-n}}{2}$ satisfies for the Chebyshev polynomial $\mathcal{T}_n(w)$, the polynomial $P(z)$ may be written as

$$P(z) = Q(w) = \sum_{i=1}^k (2 - 2\mathcal{T}_{s_i}(w)),$$

where $w = \frac{z+z^{-1}}{2}$. Thus, for polynomials (2) the following representation is valid

$$P_u(z) = Q_u(w) = \sum_{i=1}^k (2 - 2\mathcal{T}_{s_i}(w)) + 4 \sum_{m=1}^{\ell} \sin^2 \left(\frac{u \alpha_m \pi}{\beta} \right).$$

Note that the roots of the polynomials $P_u(z)$ and $Q_u(w)$ are related by the following fact.

Remark 1. If the quantities $z_k, \frac{1}{z_k}$, for $k = 1, 2, \dots, s$, are the roots of the polynomial $P_u(z)$, then the numbers $w_k = \frac{z_k + z_k^{-1}}{2}$ are the roots of the polynomial $Q_u(w)$.

Suppose that the polynomials $L(z)$ and $Q(w)$ lack multiple zeros. Since eigenvalues of the Laplacian $\lambda_j = L(e^{i \frac{2\pi j}{\beta n}})$ of the graph G_n , then

$$\lambda_j = P_j(e^{i \frac{2\pi j}{\beta n}}) = Q \left(\cos \frac{2\pi j}{\beta n} \right) + 4 \sum_{m=1}^{\ell} \sin^2 \frac{j \alpha_m \pi}{\beta}.$$

Let us formulate the main Theorem of this article.

Theorem 2. The Kirchhoff index of a graph $G_n = C_{\beta n}(s_1, \dots, s_k, \alpha_1 n, \dots, \alpha_\ell n)$ with jumps $1 \leq s_1 < \dots < s_k < \lceil \frac{\beta n}{2} \rceil$ and $1 \leq \alpha_1 < \dots < \alpha_\ell \leq \lceil \frac{\beta}{2} \rceil$ can be calculated using the formula

$$Kf_{G_n} = \frac{\beta n}{12 \sum_{i=1}^k s_i^2} \left(n^2 - \frac{\sum_{i=1}^k s_i^4}{\sum_{i=1}^k s_i^2} \right) + \sum_{\substack{\omega: Q(\omega)=0 \\ \omega \neq 1}} \frac{\beta^2 n^2 \mathcal{U}_{\beta n-1}(\omega)}{Q'(\omega)(1 - \mathcal{T}_{\beta n}(\omega))},$$

where $\mathcal{T}_{\beta n}(w)$ and $\mathcal{U}_{\beta n-1}(w)$ are the Chebyshev polynomials of the first and the second kind respectively.

Proof. Since Laplacian eigenvalues of the graph G_n are $\lambda_j = L(e^{i \frac{2\pi j}{\beta n}})$, then the Kirchhoff index is written as

$$Kf_{G_n} = \beta n \sum_{j=1}^{\beta n-1} \frac{1}{\lambda_j} = \beta n \sum_{j=1}^{\beta n-1} \frac{1}{Q(\cos \frac{2\pi j}{\beta n}) + 4 \sum_{m=1}^{\ell} \sin^2 \frac{j \alpha_m \pi}{\beta}}.$$

Let us introduce a substitution in the index of the series $j = \beta t + u$, then $0 \leq t \leq n-1$ and $0 \leq u \leq \beta-1$. The Kirchhoff index may be rewritten by formula

$$Kf_{G_n} = \sum_{t=1}^{n-1} \frac{\beta n}{Q(\cos \frac{2\pi t}{n})} + \sum_{u=1}^{\beta-1} \sum_{t=0}^{n-1} \frac{\beta n}{Q(\cos \frac{2\pi(\beta t+u)}{\beta n}) + 4 \sum_{m=1}^{\ell} \sin^2 \frac{u \alpha_m \pi}{\beta}},$$

where the terms of type

$$\sum_{t=1}^{n-1} \frac{n}{Q(\cos \frac{2\pi t}{n})} = Kf_G$$

is Kirchhoff index $Kf(G)$ for a circulant graph with fixed jumps G . This means that it is necessary to calculate only the sum

$$\sum_{u=1}^{\beta-1} \sum_{t=0}^{n-1} \frac{\beta n}{Q(\cos \frac{2\pi(\beta t+u)}{\beta n}) + 4 \sum_{m=1}^{\ell} \sin^2 \frac{u \alpha_m \pi}{\beta}}.$$

Note that the numbers $\alpha_t^0 = \cos \frac{2\pi t}{n}$ are all roots of the polynomial $\mathcal{T}_n(w) - 1$, where $\beta = 1$ and $t = 1, 2, \dots, n$. Then the numbers $\alpha_t^u = \cos \frac{2\pi(\beta t+u)}{\beta n}$ are roots for the polynomial $\frac{\mathcal{T}_{\beta n}(w)-1}{\mathcal{T}_n(w)-1}$. Introduce the following notations

$$P(w) = \mathcal{T}_{\beta n}(w) - 1, \quad R(w) = Q(w) + 4 \sum_{m=1}^{\ell} \sin^2 \frac{u \alpha_m \pi}{\beta}.$$

Denote the roots of $P(z)$ by α_t^u , here $t = 1, 2, \dots, n$, and $u = 1, 2, \dots, \beta - 1$, the roots of $R(w)$ by β_j^u where $j = 1, 2, \dots, q$ and $u = 1, 2, \dots, \beta - 1$. Note that

$$R(\alpha_t^u) = Q(\alpha_t^u) + 4 \sum_{m=1}^{\ell} \sin^2 \frac{u\alpha_m\pi}{\beta} > 0,$$

in other words, the polynomial $R(w)$ does not vanish at the roots of polynomial $P(w)$. It follows from this that the polynomials $P(w)$ and $R(w)$ lack common roots. Observe $P'(w) = \beta n \mathcal{U}_{\beta n-1}(w)$ and $R'(w) = Q'(w)$. According to Lemma 1, the equality holds

$$\sum_{u=1}^{\beta-1} \sum_{t=0}^{n-1} \frac{1}{R(\alpha_t^u)} = \sum_{u=1}^{\beta-1} \sum_{j=1}^q \frac{\beta n \mathcal{U}_{\beta n-1}(\beta_j^u)}{Q'(\beta_j^u)(1 - \mathcal{T}_{\beta n}(\beta_j^u))},$$

Thus, we obtain

$$K f_{G_n} = \beta K f_G + \sum_{u=1}^{\beta-1} \sum_{j=1}^q \frac{\beta^2 n^2 \mathcal{U}_{\beta n-1}(\beta_j^u)}{Q'(\beta_j^u)(1 - \mathcal{T}_{\beta n}(\beta_j^u))}.$$

Combining formula (1) for $K f_G$ with the obtained formula we get the required result. ■

Acknowledgments

The work is supported by the Mathematical Center in Akademgorodok (Novosibirsk).

REFERENCES

1. Klein D.J., Randić M. Resistance distance // Journal of Mathematical Chemistry. 1993. Vol. 12, no. 1-4. P. 81–95.
2. Gutman I., Mohar B. The quasi-Wiener and the Kirchhoff indices coincide // Journal of Chemical Information & Computer Sciences. 1996. Vol. 36. P. 982–985.
3. Zhu H.Y., Klein D.J., Lukovits I. Extensions of the Wiener number // Journal of Chemical Information and Modeling. 1996. Vol. 36, no. 3. P. 420–428.
4. Zhang H., Yang Y. Resistance distance and Kirchhoff index in circulant graphs // International Journal of Quantum Chemistry. 2007. Vol. 107, no. 2. P. 330–339.
5. Luzhen Y. On the Kirchhoff index of some toroidal lattices // Linear and Multilinear Algebra. 2011. Vol. 59, no. 6. P. 645–650.
6. Cinkir Z. Effective resistances and Kirchhoff index of ladder graphs // Journal of Mathematical Chemistry. 2016. Vol. 54, no. 4. P. 955–966.
7. Baigonakova G.A., Mednykh A.D. Elementary formulas for Kirchhoff index of Möbius ladder and prism graphs // Siberian Electronic Mathematical Reports. 2019. Vol. 16. P. 1654–1661.
8. Mednykh A.D., Mednykh I.A. Kirchhoff index for circulant graphs and its asymptotics // Siberian Mathematical Journal. 2024. Vol. 65, no. 6, P. 1359–1372.
9. Davis P.J. Circulant Matrices. New York: AMS Chelsea Publishing, 1994.
10. Mason J.C., Handscomb D.C. Chebyshev Polynomials. New York: CRC Press. Boca Raton, 2003.

**ОБ ИНДЕКСЕ КИРХГОФФА ДЛЯ ЦИРКУЛЯНТНОГО ГРАФА
С НЕФИКСИРОВАННЫМИ СКАЧКАМИ**

Г.К. Соколова^{1,2}

аспирант, ассистент, e-mail: g.sokolova@g.nsu.ru

¹Новосибирский государственный университет, Новосибирск, Россия

²Новосибирский государственный технический университет, Новосибирск, Россия

Аннотация. Рассматривается инвариант графа, называемый индексом Кирхгоффа. Определяется явная аналитическая формула индекса Кирхгоффа для циркулянтных графов с нефиксированными скачками.

Ключевые слова: циркулянтный граф с нефиксированными скачками, собственные значения Лапласиана графа, индекс Кирхгоффа для графа.

Дата поступления в редакцию: 11.03.2025

ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ КЛАССИФИКАЦИИ ФЕНОТИПОВ ЗАБОЛЕВАНИЯ ЖЕЛУДОЧНО-КИШЕЧНОГО ТРАКТА С ПОМОЩЬЮ МЕТОДОВ ОБРАБОТКИ ДАННЫХ

С.А. Агалаков¹

к.ф.-м.н., доцент, e-mail: agalakovsa@gsuite.omsu.ru

А.А. Березин²

аспирант, e-mail: andreyberezin55@gmail.com

¹Омский государственный университет им. Ф.М. Достоевского, Омск, Россия

²Омский государственный технический университет, Омск, Россия

Аннотация. Статья посвящена проблеме повышения точности классификации фенотипов заболевания желудочно-кишечного тракта. В ходе предварительных исследований была выявлена модель CatBoost, которая демонстрирует наилучшие результаты в данной задаче, достигая точности 92,85 % на тренировочной выборке и 79,31 % на тестовой выборке. Главной целью данной работы является улучшение точности предсказаний модели, решающей задачу классификации заболеваний, за счёт исследования особенностей исходного набора данных. Предлагаются методы обработки данных, направленные на увеличение качества предсказаний модели, включая поиск и удаление аномальных данных в выборке. Также рассматривается вопрос выбора критерия «аномальности» для специфичных данных. Кроме того, в статье обсуждаются методы работы с проблемой дисбаланса классов, что является важным аспектом для повышения общей эффективности классификации.

Ключевые слова: машинное обучение, задача классификации, логистическая регрессия, нейронные сети, поиск аномальных данных, дисбаланс классов, аугментация данных.

Введение

Нейронные сети и модели машинного обучения набирают особую популярность в наши дни. Эти технологии способны решать множество задач, демонстрируя широкий спектр применения: от голосовых ассистентов до консультантов в промышленности и других профессиональных сферах.

Зачастую на начальном этапе освоения машинного обучения создание моделей кажется весьма простым: модель быстро строится, показывает высокие результаты, и, на первый взгляд, всё работает идеально. Например, в статье авторов [1] рассматривается очень популярная задача распознавания рукописных цифр из базы данных MNIST, где авторы определили модель свёрточной нейросети с наилучшими гиперпараметрами, показав выдающуюся точность, равную 99,41 %. Но тонкость в том,

что для этой задачи был подготовлен и создан датасет из 60 тысяч изображений, каждое из которых было обработано непосредственно для практического обучения начинающих специалистов в области распознавания изображений, что обеспечивало лёгкость в работе с моделью и её высокие показатели точности.

Однако когда дело касается прикладных задач с реальными данными в областях экономики, производства, техники и медицины, то тогда точность стандартных моделей машинного обучения получается невысокой. Это связано с тем, что входные данные зачастую нестандартные, они могут содержать различные искажения или выбросы, что и приводит к менее впечатляющим результатам. В частности, медицинские данные могут содержать аномалии по нескольким причинам, таким как ненормальное состояние пациента, ошибки приборов или записи наблюдений.

1. Постановка задачи исследования

В данной работе рассматривается задача улучшения точности классификации фенотипов заболеваний желудочно-кишечного тракта (ЖКТ) за счёт очистки и нормализации исходного набора данных.

Основная цель исследования – увеличение точности предсказаний через анализ особенностей набора данных, устранение возможных шумов и выбросов, а также корректное представление признаков, значимых для точной классификации.

В процессе предварительных экспериментов наилучшая модель продемонстрировала наилучшие результаты, достигнув точности 92,85 % на тренировочной выборке и 79,31 % на тестовой. Однако наблюдаемое расхождение между этими показателями указывает на потенциальные ограничения качества данных. В этой связи главное внимание в задаче уделено обработке данных, так как её улучшение должно способствовать более точному выявлению закономерностей в рамках задачи классификации, что, в свою очередь, увеличит обобщающую способность модели.

2. Поиск оптимальной модели задачи классификации

Исследование направлено на распределение пациентов по одному из шести классов фенотипов на основании данных анкетирования и медицинских показателей. В исходном наборе данных содержится информация о 281 пациенте, представленная в таблице, для каждого из которых определено 112 признаков. Эти признаки охватывают различные аспекты состояния здоровья человека и его образа жизни.

Цель классификации – определить принадлежность каждого пациента к одной из следующих групп фенотипов заболеваний ЖКТ: здоровые, пациенты с постинфекционным синдромом раздражённого кишечника (ПИ-СРК), пациенты с ожирением, лица с коморбидными состояниями, пациенты с эссенциальным фенотипом и те, у кого проявляется смешанный фенотип.

В процессе решения поставленной задачи изначально были проанализированы модели логистической регрессии и кластерного анализа [2]. Однако, учитывая их низкую эффективность для рассматриваемого набора данных, была проведена оценка более сложных моделей, включая деревья решений, ансамбли деревьев и искусственные нейронные сети [3]. В итоге наилучший результат был достигнут с исполь-

зованием модели градиентного бустинга CatBoost [4], который продемонстрировал точность 92,85 % на тренировочной выборке и 79,31 % на тестовой выборке (табл. 1).

Также следует рассмотреть дополнительные метрики: точность и полноту для каждого класса модели CatBoost на тренировочной и тестовой выборках (рис. 1).

Таблица 1. Результаты классификации на исходных данных

Модель	Точность на тренировочной выборке, %	Точность на тестовой выборке, %
Логистическая регрессия	75,00	68,97
Кластеризация методом k -средних	—	23,49
MLPClassifier	81,74	79,31
Нейросеть Keras	73,00	68,97
Решающее дерево	82,53	62,06
Случайный лес	77,78	72,41
XGBoost	67,86	62,07
CatBoost	92,85 %	79,31 %

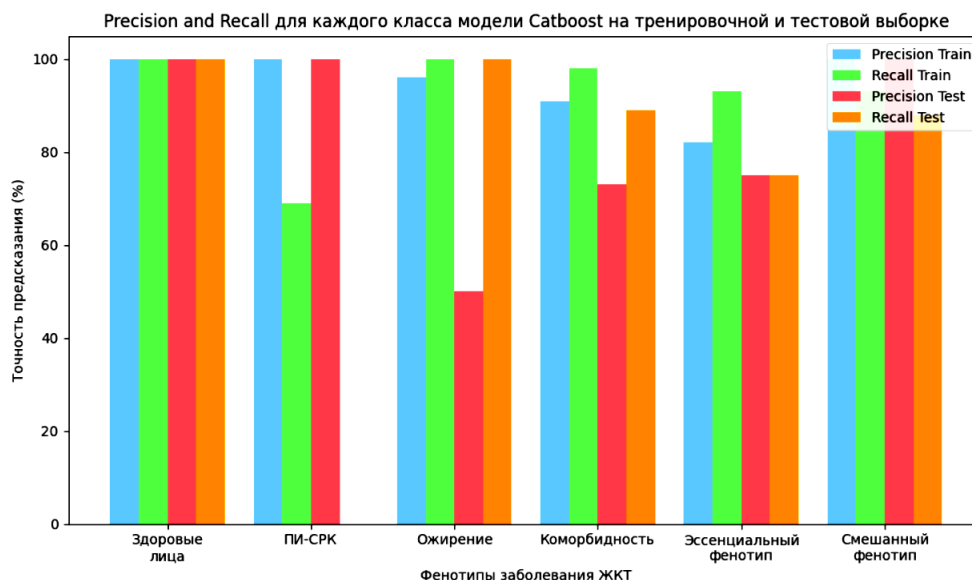


Рис. 1. Диаграмма точности и полноты CatBoost по классам

Например, при анализе полноты (recall) для каждого класса на тестовой выборке наблюдается следующее: для двух классов модель корректно определяет все наблюдения, для двух других классов полнота составляет около 89 %.

Однако для класса «1» (ПИ-СРК) полнота равна нулю. Более детальный анализ предсказанных значений показал, что в тестовой выборке представлены два наблюдения данного класса, и оба были классифицированы неверно. Это приводит к зна-

чению полноты, равному нулю, что может свидетельствовать о наличии проблемы с количеством или качеством данных для этого класса в выборке.

3. Поиск аномальных данных

В задаче классификации, когда различные модели показывают низкую точность, имеет смысл проанализировать датасет на наличие аномальных данных. Возникает вопрос о надёжности собранных наблюдений, так как данные могут содержать аномалии по разным причинам: нестабильное состояние пациентов, ошибки оборудования или погрешности при записи данных. Для улучшения качества модели будет проведена очистка данных от выбросов с использованием различных методов.

3.1. Критерий аномальности на основе графика расстояний между метками

Следующая идея основана на методах удаления выбросов, аналогичных подходу для линейной регрессии, где вычисляются остатки – разница между предсказанными и фактическими значениями (рис. 2).

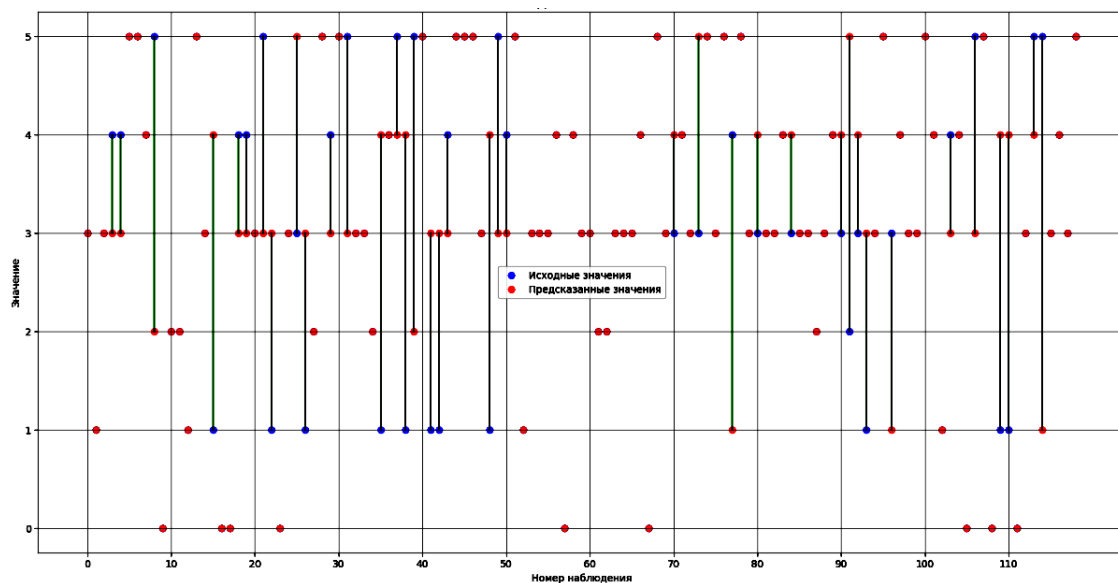


Рис. 2. График исходных значений и предсказанных

Проблемой является то, что для логистической регрессии этот метод сложнее применить, поскольку она выводит не непрерывные значения, а классы. В нашем случае результатом классификации являются метки $\{0, 1, 2, 3, 4, 5\}$. Если предположить, что эти метки имеют смысловое различие, например, 0 соответствует «здоровым», а 5 – «смешанному фенотипу», то сильное расхождение между исходной меткой и предсказанным значением можно рассматривать как выброс. На основании этого было принято решение удалить данные, где разница между исходной меткой и предсказанным классом превышает 2.

После удаления таких наблюдений точность моделей логистической регрессии и нейросетей снизилась по сравнению с исходными результатами. Это свидетельствует о том, что данный подход оказался неэффективным для улучшения качества классификации в рамках поставленной задачи. Скорее всего, это объясняется тем, что удаление наблюдений с расхождением меток больше чем на 2 могло привести к потере значимой информации, которая могла способствовать более точной дифференциации классов. Таким образом, данный метод обработки данных не подходит для решения задачи классификации фенотипов заболеваний ЖКТ, и необходимы другие подходы для выявления и устранения аномальных данных.

3.2. Критерий аномальности на основе выходных значений вероятностей логистической регрессии

С учётом того, что удаление выбросов на основе расстояний между метками, как это делается в линейной регрессии, не привело к улучшению результатов, была предложена идея: анализировать не предикторы, а вероятности принадлежности наблюдений к классам, которые также предоставляются моделью логистической регрессии [5]. В частности, мы сосредоточим внимание на ошибочных вероятностях: если предсказание оказывается неверным с вероятностью более 50 %, мы будем считать такое наблюдение выбросом.

Применение этого подхода к поиску и удалению выбросов позволило добиться увеличения точности моделей. В частности, для MLPClassifier точность на тестовой выборке повысилась до 82,60 %, а на тренировочной – 100 % (табл. 2).

Таблица 2. Модели с наилучшей точностью после удаления выбросов

Модель	Параметры	Точность на тренировочной выборке, %	Точность на тестовой выборке, %
MLPClassifier	{solver = "sgd", hidden_layer_sizes: (42, 50, 6), alpha = 0.1, max_iter = 100}	100	82,60
XGBoost	{learning_rate = 0.1, max_depth = 2, n_estimators = 50}	95,65	78,26

Важно отметить, что в ходе этого процесса из исходного датасета было исключено 46 наблюдений из 281, что составляет приблизительно 16 % от общего объема данных. Таким образом, данный метод представляется более эффективным в сравнении с предыдущим подходом.

4. Проблема несбалансированности классов

После применения множества моделей, подбора параметров и анализа аномалий необходимо провести детальный анализ распределения классов в данных. На предыдущих этапах работы не уделялось внимания количеству наблюдений и их распределению по классам, однако если рассчитать общее число наблюдений, то можно увидеть следующее распределение по классам (табл. 3).

Таблица 3. Количество наблюдений по классам

Фенотип	Количество пациентов
Здоровые лица	18
ПИ-СРК	31
Ожирение	24
Коморбидность	91
Эссенциальный	51
Смешанный	66

Явно наблюдается неравномерное распределение пациентов по классам. Это несоответствие может быть причиной низкой точности классификации для всех моделей машинного обучения. В частности, некоторые классы могут быть недостаточно представлены в обучающем наборе, что приводит к тому, что нейросеть может не выявлять закономерности в данных, а просто определять результат в пользу мажоритарного класса.

4.1. Методы работы с несбалансированными классами

В настоящее время тема дисбаланса классов является актуальной и важной в машинном обучении. Существует несколько способов, которые могут повысить точность классификации при наличии несбалансированных классов [6]:

1. **Сбор дополнительных данных.** К сожалению, в рамках данной работы нет возможности собрать новые наблюдения.
2. **Сокращение количества наблюдений для выравнивания классов по миноритарному представленному классу.** Не подходит для данной задачи, так как выборка уже небольшая. Уменьшение количества наблюдений с 281 до 108 приведёт к значительной потере данных, что негативно скажется на способности моделей извлекать уникальные признаки, важные для дальнейшей диагностики новых пациентов.
3. **Добавление копий.** Данный метод также не подходит, поскольку он подразумевает генерацию множества одинаковых данных, что не гарантирует оптимальность модели. Это может привести к переобучению, так как модель будет обучаться на одних и тех же примерах.
4. **Взвешивание классов при обучении модели и применение стратификации при разделении данных на обучающий и тестовый наборы.** Эти методы были реализованы в данной работе, однако не привели к улучшению результатов.
5. **Аугментация данных.** Данный подход представляет собой интересное решение, заключающееся в генерации новых образцов на основе имеющихся данных с помощью изменения некоторых признаков определёнными алгоритма-

ми. Аугментация способствует увеличению точности, и поэтому её следует рассмотреть более подробно.

4.2. Аугментация данных

Алгоритм SMOTE включает в себя следующие шаги [7]:

1. **Определение наименее представленного класса.**
2. **Определение количества синтетических образцов.**
3. **Определение соседей:** для каждого образца в миноритарном классе выбираются ближайшие соседи – схожие образцы из того же класса, используя метод k -ближайших соседей.
4. **Генерация синтетических образцов:** выбирается один из ближайших соседей, и на основе разности между ним и текущим образцом генерируется новый синтетический образец. Эта разность умножается на случайное число от 0 до 1, и результат добавляется к исходному образцу (рис. 3).

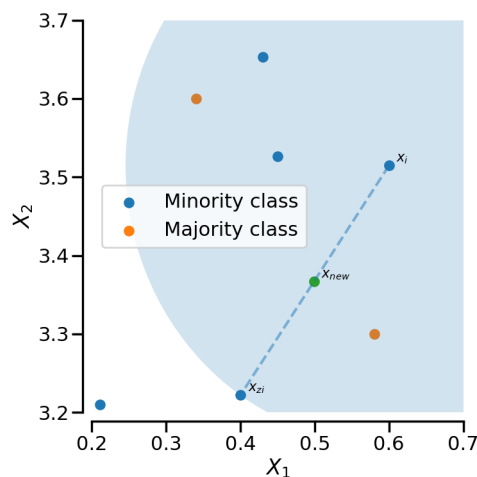


Рис. 3. Визуализация алгоритма SMOTE

Шаги 3 и 4 повторяются до тех пор, пока количество образцов в классах не станет равным.

Существует несколько модификаций данного алгоритма, среди которых SMOTE, ADASYN, Border-SMOTE.

- **ADASYN** работает аналогично обычному SMOTE, однако количество сгенерированных образцов для каждого наблюдения пропорционально количеству образцов, которые не принадлежат тому же классу, что и данное наблюдение, в заданном окружении. Это приводит к большему количеству сгенерированных образцов в областях, где не соблюдается правило ближайшего соседа [6].

- **Border-SMOTE** также генерирует синтетические образцы между ближайшими соседями, но сосредоточен на наблюдениях, находящихся на границе классов. Генерация синтетических образцов выполняется только для таких граничных наблюдений, что помогает минимизировать ошибки классификации.

В результате применения алгоритма SMOTE на новой выборке с синтезированными образцами была достигнута лучшая точность классификации для всех моделей (табл. 4).

Таблица 4. Результаты на выборке с синтезированными данными

Модель	Train, %	Test, %	Описание
Логистическая регрессия	86,88	79,63	SMOTE после удаления выбросов
Решающее дерево	92,91	79,63	Border-SMOTE после удаления выбросов
MLPClassifier	100,00	90,74	SMOTE после удаления выбросов
Случайный лес	88,33	83,33	SMOTE после удаления выбросов
XGBoost	100,00	87,27	SMOTE без нормирования и удаления выбросов
Нейросеть Keras	97,50	87,03	SMOTE после удаления выбросов

В ходе этого этапа исследования было установлено, что метод ADASYN не подходит для данного набора данных, поскольку в нем существуют области, где ни один из ближайших соседей не принадлежит к классу большинства. Это приводит к возникновению ситуации, когда происходит деление на ноль и возникают значения NaN.

В результате простой реализации SMOTE продемонстрировала лучшие результаты. Однако, учитывая, что часть данных является искусственно сгенерированной, был проведён дополнительный эксперимент. В этом эксперименте выборка была разделена на тренировочную и тестовую так, чтобы тестовая выборка вовсе не содержала синтезированных образцов. Таким образом, модели обучались на данных с искусственными образцами, а затем тестировались на части исходных данных, которые не участвовали в аугментации. Результаты этого эксперимента представлены в табл. 5.

Как видно из результатов, точность на тестовой выборке значительно ниже, чем на тренировочной, причём разница почти вдвое. Это свидетельствует о том, что метод SMOTE не применим для нашей задачи, так как сгенерированные данные не обеспечивают адекватной обобщающей способности моделей.

Заключение

В данной работе была рассмотрена задача повышения качества классификации фенотипов заболеваний ЖКТ с использованием различных методов обработки дан-

Таблица 5. Результаты на выборке без искусственных данных

Модель	Точность на тренировочной выборке, %	Точность на тестовой выборке, %
Логистическая регрессия	83,94	55,17
Решающее дерево	94,51	44,82
MLPClassifier	100,00	58,62
Случайный лес	88,61	68,96
XGBoost	100,00	55,17

ных. В результате тщательного поиска были выявлены ключевые проблемы в датасете, такие как наличие выбросов и дисбаланс классов, которые отрицательно сказались на точности классификации. Использование методов, таких как логистическая регрессия и градиентный бустинг CatBoost, позволило достичь значительных результатов, особенно после очистки данных, что подтверждает важность обработки данных для повышения качества предсказаний.

Одним из основных достижений работы стало успешное применение подхода к удалению аномальных данных на основе вероятностей логистической регрессии, что позволило улучшить точность классификации до 82,60 % для модели MLPClassifier. Несмотря на то, что метод аугментации данных с использованием SMOTE не дал ожидаемых результатов, он всё же привлёк внимание к возможности улучшения классификации с помощью синтетических образцов. Этот опыт подчёркивает, что не все подходы могут быть универсальными, и важно находить методы, подходящие конкретному набору данных.

Полученные результаты подтверждают, что более глубокий анализ данных и правильный выбор методов их обработки могут существенно влиять на качество предсказаний. Для дальнейшего исследования в данной области необходимо собрать дополнительные данные и привлечь медицинских специалистов для более точного определения классов и выбора признаков, имеющих наибольшее значение для диагностики. Это позволит значительно повысить эффективность моделей и улучшить качество диагностики заболеваний ЖКТ.

Литература

1. Shevlyakov A.N., Berezin A.A. Recognition of the MNIST dataset with defective rows // Journal of Physics: Conference Series: 15, Virtual, Online, 09–11 November 2021. Virtual, Online, 2022. P. 012031.
2. Агалаков С.А. Статистические методы анализа данных : учеб. пособие. Омск: Изд-во Ом. гос. ун-та, 2017. 92 с.
3. Николенко С., Кадури А., Архангельская Е. Глубокое обучение. СПб.: Питер, 2018. 480 с.
4. Официальная документация CatBoost. URL: <https://catboost.ai/en/docs/> (дата обращения: 25.02.2024).

5. Anomaly Detection. URL: https://vs.inf.ethz.ch/edu/HS2011/CPS/papers/chandola09_anomaly-detection-survey.pdf (дата обращения: 10.11.2024).
6. Eight tactics to Combat Imbalanced Classes in Your Machine Learning Dataset. URL: <https://machinelearningdev.blogspot.com/2016/08/8-tactics-to-combat-imbalanced-classes.html> (дата обращения: 10.11.2024).
7. Официальная документация imbalanced-learn (SMOTE). URL: https://imbalanced-learn.org/stable/over_sampling.html\#smote-variants (дата обращения: 10.11.2024).

IMPROVING THE EFFICIENCY OF CLASSIFICATION OF GASTROINTESTINAL DISEASE PHENOTYPES USING DATA PROCESSING METHODS

S.A. Agalakov¹

Ph.D. (Phys.-Math.), Associate Professor, e-mail: agalakovsa@gsuite.omsu.ru

A.A. Berezin²

Ph.D. Student, e-mail: andreyberezin55@gmail.com

¹Dostoevsky Omsk State University, Omsk, Russia

²Omsk State Technical University, Omsk, Russia

Abstract. The article is devoted to the problem of improving the accuracy of classification of phenotypes of diseases of the gastrointestinal tract. In the course of preliminary research, the CatBoost model was identified, which demonstrates the best results in this task, achieving an accuracy of 92.85 % in the training sample and 79.31 % in the test sample. The main goal of this work is to improve the accuracy of predictions of the model solving the problem of classification of diseases by studying the characteristics of the original data set. Data processing methods aimed at increasing the quality of model predictions, including the search and removal of abnormal data in the sample, are proposed. The issue of choosing the criterion of "abnormality" for specific data is also considered. In addition, the article discusses methods of dealing with the problem of class imbalance, which is an important aspect for improving the overall efficiency of classification.

Keywords: machine learning, classification problem, logistic regression, neural networks, search for abnormal data, class imbalance, data augmentation.

Дата поступления в редакцию: 11.11.2024

ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ РЕШЕНИЯ ПОЛНОЙ СИСТЕМЫ УРАВНЕНИЙ НАВЬЕ – СТОКСА ПРИ УЧЁТЕ ДЕЙСТВИЯ СИЛЫ КОРИОЛИСА

А.А. Бугаенко^{1,2}

ведущий инженер, e-mail: bugaenkoaa@vniitf.ru

И.Ю. Крутова¹

к.ф.-м.н., доцент, e-mail: IYKrutova@mephi.ru

¹Снежинский физико-технический институт Национальный исследовательский ядерный университет «МИФИ», Снежинск, Россия

²РФЯЦ-ВНИИТФ имени академика Е.И. Забабахина, Снежинск, Россия

Аннотация. Рассматривается полная система уравнений Навье – Стокса, в которой учтены сжимаемость газа и диссипативные эффекты вязкости и теплопроводности так, что общие законы сохранения массы, импульса и энергии выполняются. В силу нелинейности данной системы построение её решений является достаточно трудоёмким. Как следствие, применяется процедура линеаризации на точном решении (однородном покое) полной системы уравнений Навье – Стокса при учёте сил тяжести и Кориолиса.

Ключевые слова: полная система уравнений Навье – Стокса (ПСУНС), сила Кориолиса, линеаризация.

Введение

Природные восходящие закрученные потоки (ВЗП): торнадо, тропические циклоны, огненные вихри – представляют собой сложные и ещё достаточно мало изученные явления с точки зрения их возникновения и продолжительного функционирования. Надёжное теоретическое изучение этих потоков возможно только с использованием системы уравнений газовой динамики при учёте действия сил тяжести и Кориолиса.

В монографиях [1–5] (более подробную библиографию см. в [3, 4]) с использованием этой математической модели – система уравнений газовой динамики при учёте действия сил тяжести и Кориолиса – и с применением методологии характеристической задачи Коши [6–8] проведены аналитические и численные исследования течений воздуха в природных ВЗП.

В силу нелинейности системы уравнений газовой динамики построение её решений является достаточно трудоёмким. Это и послужило причиной линеаризации системы уравнений газовой динамики на её точных решениях [9].

Численное моделирование на основе линеаризованной системы уравнений газовой динамики ВЗП с учётом силы Кориолиса [9] показало, что процессы изменения газодинамических параметров, включая скоростные характеристики 3-мерных

нестационарных ВЗП для среднего по интенсивности циклона, вызванных вертикальным продувом, при увеличении времени счёта не выходят на стационарный режим.

При численном построении сложных неоднородных и нестационарных течений газа для увеличения устойчивости счёта часто применяют различные алгоритмы, сглаживающие решения. Для того чтобы подобная процедура имела обоснованный физический смысл, в работе рассматривается полная система уравнений Навье – Стокса (ПСУНС), в которой учтены сжимаемость газа и диссипативные эффекты вязкости и теплопроводности так, что общие законы сохранения массы, импульса и энергии выполняются. А также выполняются законы термодинамики благодаря учёту уравнений состояния.

Далее в работе приведена линеаризация PSUНС на точном решении, описывающем однородный покой.

Линеаризация PSUНС

PSUНС в изэнтропическом случае для идеального политропного газа с уравнением состояния $p = \rho^\gamma / \gamma$ при учёте действия силы Кориолиса и силы тяжести имеет следующий вид [2, 3, 10]:

$$\left\{ \begin{array}{l} \rho_t + u\rho_x + v\rho_y + w\rho_z + \rho(u_x + v_y + w_z) = 0, \\ u_t + uu_x + vu_y + wu_z + \frac{T}{\gamma\rho}\rho_x + \frac{1}{\gamma}T_x = av - bw + \\ + \frac{\mu_0}{\rho} \left(u_{xx} + \frac{3}{4}u_{yy} + \frac{3}{4}u_{zz} + \frac{1}{4}v_{xy} + \frac{1}{4}w_{xz} \right), \\ v_t + uv_x + vv_y + wv_z + \frac{T}{\gamma\rho}\rho_y + \frac{1}{\gamma}T_y = -au + \\ + \frac{\mu_0}{\rho} \left(\frac{3}{4}v_{xx} + v_{yy} + \frac{3}{4}v_{zz} + \frac{1}{4}u_{xy} + \frac{1}{4}w_{yz} \right), \\ w_t + uw_x + vw_y + ww_z + \frac{T}{\gamma\rho}\rho_z + \frac{1}{\gamma}T_z = bu - g + \\ + \frac{\mu_0}{\rho} \left(\frac{3}{4}w_{xx} + \frac{3}{4}w_{yy} + w_{zz} + \frac{1}{4}u_{xz} + \frac{1}{4}v_{yz} \right), \\ T_t + uT_x + vT_y + wT_z + (\gamma - 1)T(u_x + v_y + w_z) = \frac{\kappa_0}{\rho} (T_{xx} + T_{yy} + T_{zz}) + \\ + \frac{\mu_0\gamma(\gamma - 1)}{2\rho} [(u_x - v_y)^2 + (u_x - w_z)^2 + (v_y - w_z)^2 + \\ + \frac{3}{2}((u_y + v_x)^2 + (u_z + w_x)^2 + (v_z + w_y)^2)] \end{array} \right. \quad (1)$$

Здесь ρ – плотность газа; $\gamma = \text{const} > 1$ – показатель политропы идеального газа, $\mathbf{V} = (u, v, w)$ – вектор скорости газа с его проекциями на декартовы оси Ox , Oy , Oz ; $\mathbf{\Omega} = (0; \Omega_2; \Omega_3)$ – вектор угловой скорости вращения Земли вокруг своей оси, где $\Omega_2 = \Omega \cos \psi$, $\Omega_3 = \Omega \sin \psi$, ψ – широта точки, в которой находится начало декартовой системы координат (x, y, z) , вращающейся вместе с Землёй; $a = 2\Omega_3$;

$b = 2\Omega_2$; $\mathbf{g} = (0, 0, -g)$, $g = \text{const} > 0$ – ускорение свободного падения; T – температура газа; μ_0 , κ_0 – постоянные коэффициенты вязкости и теплопроводности. Индексами t, x, y, z обозначаются соответствующие частные производные.

Благодаря наличию в правых частях уравнений движения – второго, третьего и четвёртого уравнений системы (1) – слагаемых, отвечающих за присутствие ускорения Кориолиса, в поток газа вносится дополнительный внешний импульс. Именно он и приводит к закрутке вязкого теплопроводного газа в придонной части вертикального теплового восходящего потока.

В системе (1) при введении безразмерных переменных масштабные значения связаны следующими соотношениями [1–5]:

$$u_{00} = \sqrt{R\gamma T_{00}} = c_{00}; \quad t_{00} = \frac{r_{00}}{u_{00}}; \quad p_{00} = R\rho_{00}T_{00}; \quad e_{00} = c_{v0}T_{00},$$

где $\gamma = 1 + R/c_{v0}$, R – универсальная газовая постоянная, c_{v0} – удельная теплоёмкость. Следовательно, за масштабные значения скорости, давления и внутренней энергии выбраны соответственно скорость звука, давление и внутренняя энергия газа с параметрами ρ_{00} , T_{00} . В качестве масштабного значения скорости берётся значение $u_{00} = c_{00} = \frac{1}{3}10^3$ м/с, близкое к значению скорости распространения звука в воздухе при стандартных условиях, т. е. при $\rho_{00} = 1,29 \text{ кг/м}^3$, $T_{00} = 288 \text{ К}$.

Безразмерные значения констант Ω и g задаются следующим образом:

$$\Omega = \frac{2\pi}{24 * 3600c} * t_{00}, \quad g = g_* \frac{T_{00}}{c_{00}}.$$

Постоянные безразмерные значения коэффициентов вязкости μ_0 и теплопроводности κ_0 определяются следующими соотношениями:

$$\mu_0 = \frac{4\mu_*}{3\rho_{00}u_{00}r_{00}}, \quad \kappa_0 = \frac{\kappa_*}{c_{v0}\rho_{00}u_{00}r_{00}},$$

где μ_* , κ_* – размерные значения постоянных коэффициентов вязкости и теплопроводности. Для воздуха обычно полагают $\gamma = 1,4$, и тогда $\kappa_0 \approx 1,46\mu_0$.

Система (1) имеет смешанный тип: первое уравнение – уравнение неразрывности – образует гиперболическую часть системы, так как определяет в течениях сжимаемого теплопроводного вязкого газа наличие слабого разрыва на контактной поверхности [1]; второе и третье уравнения – уравнения движения и энергии – составляют параболическую часть системы, так как содержат вторые производные компонент скорости и температуры по пространственной переменной. Заметим, что $\mu_0 = \kappa_0 = 0$, т. е. при отсутствии вязкости и теплопроводности ПСУНС (1) переходит в систему уравнений газовой динамики, рассмотренной в статье [11].

У системы (1) в случае $g = 0$ имеется точное решение:

$$\rho = 1; \quad T = 1; \quad u = v = w = 0. \quad (2)$$

Линеаризация ПСУНС (1) на точном решении (2) состоит в том, что решение этой системы представляется в виде

$$\rho = 1 + \tilde{\rho}; \quad T = 1 + \tilde{T}; \quad \mathbf{V} = \tilde{\mathbf{V}},$$

и эти выражения подставляются в систему (1), предварительно умножив все уравнения, кроме первого, на ρ :

$$\left\{ \begin{array}{l} (1 + \tilde{\rho})(1 + \tilde{\rho})_t + (1 + \tilde{\rho})\tilde{u}(1 + \tilde{\rho})_x + (1 + \tilde{\rho})\tilde{v}(1 + \tilde{\rho})_y + (1 + \tilde{\rho})\tilde{w}(1 + \tilde{\rho})_z + \\ + (1 + \tilde{\rho})(1 + \tilde{\rho})(\tilde{u}_x + \tilde{v}_y + \tilde{w}_z) = 0, \\ (1 + \tilde{\rho})\tilde{u}_t + (1 + \tilde{\rho})\tilde{u}\tilde{u}_x + (1 + \tilde{\rho})\tilde{v}\tilde{u}_y + (1 + \tilde{\rho})\tilde{w}\tilde{u}_z + \frac{(1 + \tilde{T})}{\gamma}(1 + \tilde{\rho})_x + \\ + (1 + \tilde{\rho})\frac{1}{\gamma}(1 + \tilde{T})_x = (1 + \tilde{\rho})a\tilde{v} - (1 + \tilde{\rho})b\tilde{w} + \\ + \mu_0 \left(\tilde{u}_{xx} + \frac{3}{4}\tilde{u}_{yy} + \frac{3}{4}\tilde{u}_{zz} + \frac{1}{4}\tilde{v}_{xy} + \frac{1}{4}\tilde{w}_{xz} \right), \\ (1 + \tilde{\rho})\tilde{v}_t + (1 + \tilde{\rho})\tilde{u}\tilde{v}_x + (1 + \tilde{\rho})\tilde{v}\tilde{v}_y + (1 + \tilde{\rho})\tilde{w}\tilde{v}_z + \frac{(1 + \tilde{T})}{\gamma}(1 + \tilde{\rho})_y + \\ + (1 + \tilde{\rho})\frac{1}{\gamma}(1 + \tilde{T})_y = -(1 + \tilde{\rho})a\tilde{u} + \mu_0 \left(\frac{3}{4}\tilde{v}_{xx} + \tilde{v}_{yy} + \frac{3}{4}\tilde{v}_{zz} + \frac{1}{4}\tilde{u}_{xy} + \frac{1}{4}\tilde{w}_{yz} \right), \\ (1 + \tilde{\rho})\tilde{w}_t + (1 + \tilde{\rho})\tilde{u}\tilde{w}_x + (1 + \tilde{\rho})\tilde{v}\tilde{w}_y + (1 + \tilde{\rho})\tilde{w}\tilde{w}_z + \frac{(1 + \tilde{T})}{\gamma}(1 + \tilde{\rho})_z + \\ + (1 + \tilde{\rho})\frac{1}{\gamma}(1 + \tilde{T})_z = (1 + \tilde{\rho})b\tilde{u} + \mu_0 \left(\frac{3}{4}\tilde{w}_{xx} + \frac{3}{4}\tilde{w}_{yy} + \tilde{w}_{zz} + \frac{1}{4}\tilde{u}_{xz} + \frac{1}{4}\tilde{v}_{yz} \right), \\ (1 + \tilde{\rho})(1 + \tilde{T})_t + (1 + \tilde{\rho})\tilde{u}(1 + \tilde{T})_x + (1 + \tilde{\rho})\tilde{v}(1 + \tilde{T})_y + (1 + \tilde{\rho})\tilde{w}(1 + \tilde{T})_z + \\ + (1 + \tilde{\rho})(\gamma - 1)(1 + \tilde{T})(\tilde{u}_x + \tilde{v}_y + \tilde{w}_z) = \\ = \kappa_0 \left((1 + \tilde{T})_{xx} + (1 + \tilde{T})_{yy} + (1 + \tilde{T})_{zz} \right) + \\ + \frac{\mu_0 \gamma (\gamma - 1)}{2} [(\tilde{u}_x - \tilde{v}_y)^2 + (\tilde{u}_x - \tilde{w}_z)^2 + (\tilde{v}_y - \tilde{w}_z)^2 + \\ + \frac{3}{2} ((\tilde{u}_y + \tilde{v}_x)^2 + (\tilde{u}_z + \tilde{w}_x)^2 + (\tilde{v}_z + \tilde{w}_y)^2)]. \end{array} \right. \quad (3)$$

Слагаемые, не содержащие функций и производных с тильдой, взаимно уничтожаются, поскольку выражения (2) задают точное решение системы (1).

Затем нелинейные выражения с функциями и производными с тильдой отбрасываются. В результате получается следующая линейная система уравнения с частными производными, где для простоты знак тильды опущен [11]:

$$\left\{ \begin{array}{l} \rho_t + (\tilde{u}_x + \tilde{v}_y + \tilde{w}_z) = 0, \\ \tilde{u}_t + \frac{\tilde{\rho}_x}{\gamma} + \frac{\tilde{T}_x}{\gamma} = a\tilde{v} - b\tilde{w} + \mu_0 \left(\tilde{u}_{xx} + \frac{3}{4}\tilde{u}_{yy} + \frac{3}{4}\tilde{u}_{zz} + \frac{1}{4}\tilde{v}_{xy} + \frac{1}{4}\tilde{w}_{xz} \right), \\ \tilde{v}_t + \frac{\tilde{\rho}_y}{\gamma} + \frac{\tilde{T}_y}{\gamma} = -a\tilde{u} + \mu_0 \left(\frac{3}{4}\tilde{v}_{xx} + \tilde{v}_{yy} + \frac{3}{4}\tilde{v}_{zz} + \frac{1}{4}\tilde{u}_{xy} + \frac{1}{4}\tilde{w}_{yz} \right), \\ \tilde{w}_t + \frac{\tilde{\rho}_z}{\gamma} + \frac{\tilde{T}_z}{\gamma} = b\tilde{u} + \mu_0 \left(\frac{3}{4}\tilde{w}_{xx} + \frac{3}{4}\tilde{w}_{yy} + \tilde{w}_{zz} + \frac{1}{4}\tilde{u}_{xz} + \frac{1}{4}\tilde{v}_{yz} \right), \\ \tilde{T}_t + (\gamma - 1)(\tilde{u}_x + \tilde{v}_y + \tilde{w}_z) = \kappa_0 \left(\tilde{T}_{xx} + \tilde{T}_{yy} + \tilde{T}_{zz} \right). \end{array} \right. \quad (4)$$

Заключение

В работе проведена линеаризация линейной системы уравнений с частными производными на точном решении, описывающем однородный покой. Присутствие силы тяжести в ПСУНС приводит к достаточно затруднительной процедуре линеаризации. Поэтому полагается нулём. И потом уже явно учитывается в линеаризованной ПСУНС.

В монографии [6] показано, что процедура линеаризации квазилинейного уравнения с частными производными на его точном решении и построение решения полученного линейного уравнения фактически определяют первое слагаемое конкретного бесконечного ряда по степеням формального малого параметра ε . Этот ряд решает специальным образом поставленную характеристическую задачу Коши стандартного вида [6] и при условии аналитичности входных данных задачи сходится в окрестности рассматриваемой точки. Следовательно, решение линеаризованной задачи в сумме с точным решением, на котором проведена линеаризация, дают первые два слагаемых бесконечного сходящегося ряда, задающего новое решение исходного нелинейного уравнения с частными производными.

Литература

1. Баутин С.П. Торнадо и сила Кориолиса. Новосибирск: Наука, 2008. 96 с.
2. Разрушительные атмосферные вихри: теоремы, расчёты, эксперименты / С.П. Баутин [и др.]. Новосибирск: Наука ; Екатеринбург: УрГУПС. 2013. 216 с.
3. Разрушительные атмосферные вихри и вращение Земли вокруг оси / С.П. Баутин [и др.]. Екатеринбург: УрГУПС. 2017. 216 с.
4. Баутин С.П., Крутова И.Ю. Аналитическое и численное моделирование течений газа при учёте действия силы Кориолиса. Екатеринбург: УрГУПС, 2019. 181 с.
5. Баутин С.П., Обухов А.Г. Математическое моделирование разрушительных атмосферных вихрей. Новосибирск: Наука, 2012. 152 с.
6. Баутин С.П. Характеристическая задача Коши и её приложения в газовой динамике. Новосибирск: Наука, 2009. 368 с.
7. Баутин С.П. Аналитические решения задачи о движении поршня // Численные методы механики сплошной среды. Новосибирск: ВЦ СО АН СССР, 1973. Т. 4, № 1. С. 3–15.
8. Баутин С.П. Характеристическая задача Коши для квазилинейной аналитической системы // Дифференциальные уравнения. 1976. Т 12, № 11. С. 2052–2063.
9. Баутин С.П., Крутова И.Ю. Линеаризованная система уравнений газовой динамики при учёте действия силы Кориолиса: препринт. Снежинск: СФТИ НИЯУ МИФИ, 2019. 52 с.
10. Баутин С.П., Обухов А.Г. Численное моделирование трёхмерных нестационарных течений сжимаемого вязкого теплопроводного газа. Екатеринбург: УрГУПС, 2020. 287 с.
11. Крутова И.Ю., Бугаенко А.А. Численное моделирование на основе линеаризованной системы уравнений газовой динамики восходящих закрученных потоков с учётом силы Кориолиса // ВАНТ. Серия Математическое моделирование физических процессов. 2024. Вып. 1.

NUMERICAL MODELING OF THE SOLUTION OF PSUNS TAKING INTO ACCOUNT THE EFFECT OF CORIOLIS FORCE

A.A. Bugaenko^{1,2}

Lead Engineer, e-mail: bugaenkoaa@vniitf.ru

I.Yu. Krutova¹

Ph.D. (Eng.), Associate Professor, e-mail: IYKrutova@mephi.ru

¹Snezhinsk Institute of Physics and Technology National Research Nuclear University "MEPhI",
Snezhinsk, Russia

²Federal State Unitary Enterprise "Russian Federal Nuclear Center – All-Russian Research
Institute of Technical Physics named after Academician E.I. Zababakhin", Snezhinsk, Russia

Abstract. The paper considers the complete system of Navier-Stokes equations, which takes into account the compressibility of the gas and the dissipative effects of viscosity and thermal conductivity so that the general laws of conservation of mass, momentum and energy are satisfied. Due to the nonlinearity of this system, constructing its solutions is quite labor-intensive. As a consequence, a linearization procedure is used on the exact solution (homogeneous rest) of the complete system of Navier-Stokes equations, taking into account the forces of gravity and Coriolis.

Keywords: full system of Navier-Stokes equations (FSNE), Coriolis force, linearization.

Дата поступления в редакцию: 19.01.2025

ТОПОЛОГИЧЕСКИЕ ПОЛЁТЫ КОСМИЧЕСКИХ ВАРП-КОРАБЛЕЙ АЛЬКУБЬЕРРЕ И ПОТОКИ РИЧЧИ

А.К. Гуц

д.ф.-м.н., профессор, e-mail: aguts@mail.ru

Сочинский государственный университет, Сочи, Россия

Аннотация. Показана возможность описания полётов сверхбыстрых космических варп-кораблей Алькубьерре, меняющих топологию и геометрию 3-мерного физического пространства посредством потоков Риччи.

Ключевые слова: варп-корабль, Алькубьерре, изменения топологии и геометрии, потоки Риччи, кротовая нора, знак энергии.

Введение

Реализация проекта Алькубьерре [1] космического сверхбыстрого корабля с радиусом варп-пузыря 100 м, который наблюдателем на Земле воспринимается как сверхсветовой, как показали расчёты, требует энергии порядка $(1/4)M_{\odot}$ [2, р. 73]. Поскольку это значительно больше, чем энергия $(1/100)M_{\odot}$, необходимая для создания 4-мерной кротовой норы [3], то образуется кротовая нора и варп-пузырь, содержащий варп-корабль, уходит в неё [4]. Пространство при этом теряет связность – образуются две компоненты связности. Последняя оценка, полученная автором летом 1981 г. и связанная с образованием отрыва части 3-пространства, т. е. с изменением связности 3-многообразия, была найдена при допущении изменения гладкости римановой метрики пространства в процессе её изменения со временем. Фактически это перекликается с образованием сингулярностей в римановом пространстве, которые появляются при исследовании потоков Риччи в конечное время [5].

1. Теорема Гамильтона

Обращение к исследованиям по потокам Риччи, первые публикации Р. Гамильтона, которые появились позже, лишь в 1982 г., позволяет сейчас более обстоятельно изучить целый ряд проблем, возникающих при полётах варп-кораблей Алькубьерре.

Эволюция римановой 3-метрики $g(t)$, которая приводит к некоторой выделенной римановой метрике на данном замкнутом ориентируемом дифференцируемом 3-мерном многообразии M^3 , при использовании потоков Риччи, характеризуется дифференциальным уравнением

$$\frac{\partial g}{\partial t} = -2Ric_g(t), \quad (1)$$

определённым на $[0; T)$, и относится к бесконечномерному пространству римановых метрик на многообразии M^3 . Обычно уравнение (1) записывается в некоторой системе локальных координат – локальных карт из атласа карт многообразия M^3 :

$$\frac{\partial g_{ik}}{\partial t} = -2R_{ik}.$$

Фундаментальный результат Р. Гамильтона гарантирует существование и единственность потока Риччи в течение короткого промежутка времени. Для заданной гладкой начальной метрики g_0 на компактном многообразии M^3 существует единственное гладкое решение $g(t)$ уравнения потока Риччи в течение короткого времени. Это гарантирует, что поток может быть запущен из любой гладкой начальной точки. Иначе говоря, варп-корабль может выйти на сверхсветовой режим, **изменяющий топологию и ведущий к сингулярности**, в любой точке пространства.

Справедлива следующая теорема.

Теорема 1. (Гамильтон, ([5], 1982). Если g_0 – гладкая метрика на компактном многообразии M^3 , то поток Риччи с $g(0) = g_0$ имеет единственное решение $g(t)$ на максимальном временном интервале $t \in [0, T)$, где $T \leq \infty$. Причём если $T < \infty$, то

$$\lim_{t \rightarrow T} \sup_{x \in M^3} |Riem(x, t)| = \infty,$$

$$\text{где } |Riem| = \sqrt{R_{ijkl}R^{ijkl}}.$$

Как видим, по мере развития потока Риччи могут образовываться сингулярности, в которых кривизна становится неограниченной. Одна из возможных сингулярностей изображена на рис. 1 (слева) [6].

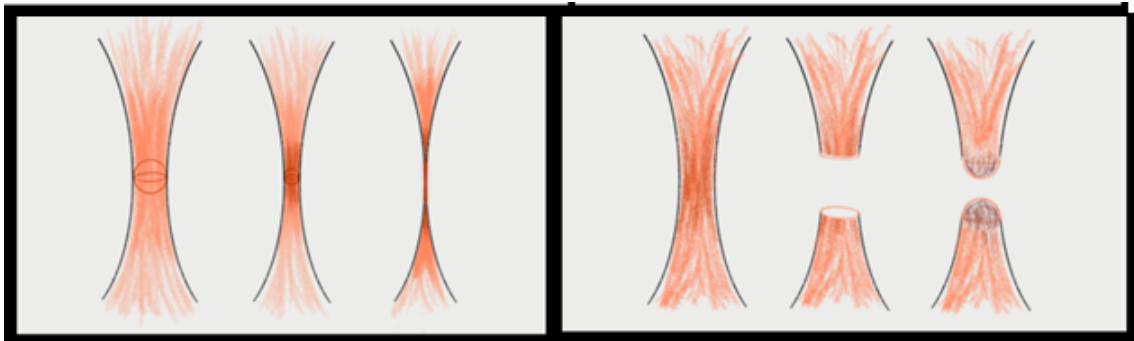


Рис. 1. Поток Риччи приближается к сингулярности (слева). Использование хирургии для укрощения сингулярностей (справа) [6]

Заметим, что, вообще говоря, кривизна неограниченно возрастает лишь на некоторой части многообразия; при этом говорят, что поток имеет сингулярность.

Фундаментальная проблема теории потоков Риччи – это описать все возможные геометрии сингулярностей. С точки зрения теории полётов варп-кораблей решение этой проблемы даёт представление, что может ожидать варп-корабль при переходе на сверхсветовой режим.

2. Древние потоки Риччи – это модели сингулярностей

Форму сингулярности можно «разглядеть» благодаря процедуре масштабирования – параболического растяжения – в точке x_0 , где $R(x_0, t) \rightarrow \infty$ при $t \rightarrow T$.

Масштабирование заключается в растяжении метрики и замедлении течения времени, которые позволяют получить новое решение уравнения потока.

Как это делается?

Если даны поток Риччи $g(t)$ на $M \times [0; T)$, точка x_0 и момент времени t_0 , то определяется растяжение как поток, заданный формулой

$$g_0(t) = Q_0 \cdot g\left(t_0 + \frac{t}{Q_0}\right), \quad Q_0 = R(x_0, t_0).$$

Оно определено на промежутке $[-t_0 Q_0; (T - t_0) Q_0)$ и нормируется условием $R_{g_0}(x_0, 0) = 1$ (рис. 2) [7].

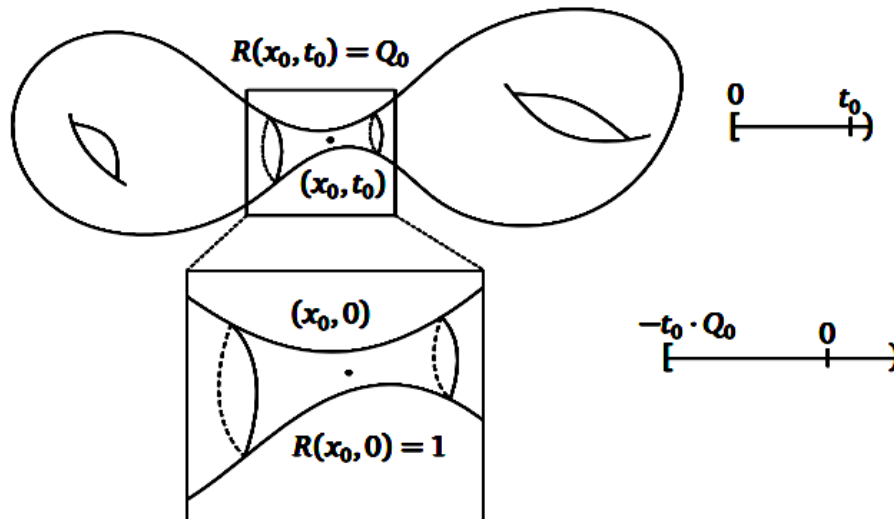


Рис. 2. Масштабирование в точке x_0 [7]

При определённых предположениях увеличенный, т. е. параболически растянутый, поток стремится к предельному потоку Риччи, называемому *моделью сингулярности*. Модели сингулярности – это *древние потоки Риччи*, т. е. их можно бесконечно распространять в прошлое. Понимание возможных моделей сингулярности в потоке Риччи является важнейшей исследовательской задачей.

Рассмотрим процесс получения предельного потока подробнее. Рассматривается последовательность масштабирований в точках (x_k, t_k) , для которых $Q_k = R(x_k, t_k) \rightarrow +\infty$ и скалярная кривизна на $M \times [0, t_k]$ в которых максимальна. Тогда последовательность параболических растяжений метрик $g_k(t)$ в (x_k, t_k) имеет ограниченную кривизну на отрезках $[-t_k Q_k, 0]$, сходящихся к $(-\infty, 0]$ [7].

Г. Перельман показал, что при определённых условиях гарантируется плавная сходимость последовательности $\{(M, g_k(t), x_k)\}$ (или некоторой её подпоследовательности) к предельному древнему потоку Риччи $(M_\infty, g_\infty(t), x_\infty)$ [7]. Заметим, что

топология многообразия M_∞ не обязана совпадать с топологией многообразия M [8, р. 2].

Полученный древний поток Риччи является потоком на $(-\infty, 0]$ ограниченной и при этом ненулевой кривизны. Более того, теорема Гамильтона – Айви о сжатии позволяет показать, что кривизна сечений предельного потока положительна или равна нулю [7].

3. Сингулярности в потоке Риччи

По Р. Гамильтону, Г. Перельману предельные потоки (blow-up limit flow) в точках максимальной кривизны приводят к одной из следующих двух моделей сингулярности [9, р. 86] и [10]:

1. Сжимающаяся круглая сферическая форма S^3 или её фактор S^3/Γ (который сжимается в точку за конечное время) (см. рис. 3, а [6, р. 31]).
2. Круглый бесконечный цилиндр $S^2 \times \mathbb{R}$ (шейка), либо его Z_2 -фактор, содержащий одностороннюю проективную плоскость (см. рис. 3, b [9]).

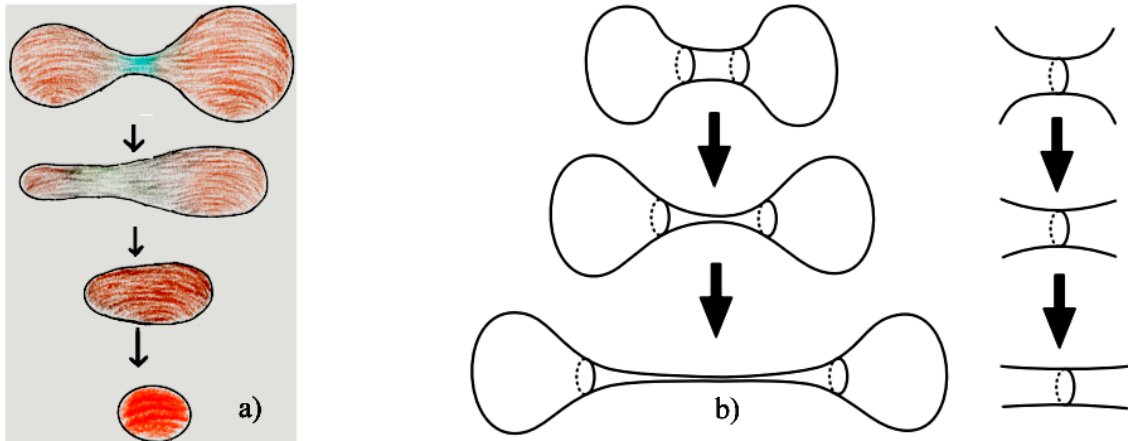


Рис. 3. Две модели сингулярности: а – круглая сфера S^3/Γ [6] и б – цилиндр (шейка) [9]

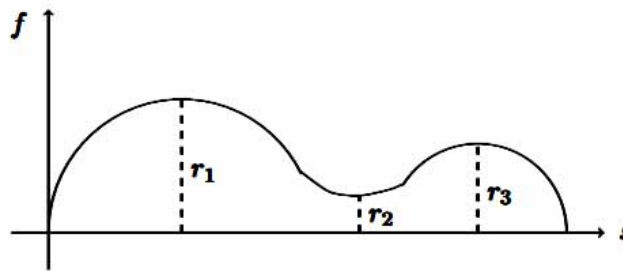
4. Метрика, описывающая образование сингулярности с тонким цилиндром (шеей)

Пусть дано многообразие $M^3 = S^1 \cup (S^2 \times I) \cup S_2$ с исходной метрикой

$$g_0 = \{1 + [f'(s)]^2\}ds^2 + [f(s)]^2g_{S^2},$$

где вид функции $f(s)$ дан на рис. 4 [11, р. 25].

Тогда поток Риччи эволюционирует за конечное время до сингулярности, представленной на рис. 5, если $r_1 \ll r_2, r_3$ [11, р. 25] и [12], которая затем отщипывается.

Рис. 4. График функции $f(s)$ [11]Рис. 5. Пример $r_2 \ll r_1, r_3$ с разрушающейся шеей (отщипывается) [11]

5. Потоки Риччи с хирургией

Как показал Г. Перельман, если сингулярности появляются за конечное время, то они выглядят только как сжимающиеся сферы или цилиндры. Далее он разрезает многообразие вдоль сингулярностей, разделяя многообразие на несколько частей, а затем продолжает поток Риччи на каждой из этих частей. Эта процедура известна как поток Риччи с хирургическим вмешательством [8].

Понимание природы образования сингулярностей важно, поскольку они результат старта варп-корабля и, следовательно, должны рассматриваться как нанесение «ущерба» пространству. Устранение же этих сингулярностей имеет решающее значение как для понимания долгосрочного поведения потока, так и для обнаружения наличия естественного (природного, самоустраниющегося, самовосстанавливающегося) или искусственного механизмов устранения нанесённого «ущерба».

Для изучения этих явлений математиками используются такие методы, как анализ раздувания (масштабирование), который предполагает изменение масштаба метрики вблизи сингулярностей и использование специальных математических операций — потоков Риччи с хирургией (Г. Перельман) — для изменения многообразия в сингулярных точках. На рис. 1 (справа) показано, как ликвидируется один из типов сингулярности, ведущий к потере связности многообразия, т. е. к отрыву варп-пузыря от пространства.

Для 3-мерного риманова многообразия Г. Перельман показал, как **продолжить движение за сингулярности**, используя операции хирургии на многообразии¹.

Необходимую работу в этом направлении, отвечая на утверждение Г. Перельмана, проделали авторы статьи [14].

Они «рассмотрели поток Риччи на S^{n+1} , начинающийся с вращательно-

¹Точнее, Г. Перельман писал: «Вероятно, что, перейдя к пределу в этой конструкции, можно было бы получить канонически определённый поток Риччи через сингулярности, но на данный момент у меня нет доказательства этого» [13, р. 37].

симметричной метрики g_0 . В п. 4 показали, что если g_0 достаточно сжата, то поток Риччи за конечное время T сформирует сингулярность «шейка» и вычислили точную асимптотику профиля решения вблизи сингулярной гиперповерхности при $t \nearrow T$. Позже, в работе [14], авторы доказали, что можно определить плавную прямую эволюцию потока Риччи через сингулярность «шейка». Они добились этого, по сути, взяв предел потоков Риччи с хирургией, и, следовательно, показали, что гипотеза Г. Перельмана на самом деле верна в случае «шейки» сферы S^3 . Поскольку плавная прямая эволюция выполняет хирургию в сингулярное время $T = 0$ и на масштабе 0, то во все последующие моменты времени поток будет состоять из двух непересекающихся гладких потоков Риччи на *паре* многообразий [15].

Искусственная, проводимая математиком процедура отрезания и заклеивания «дыры» половиной шара, ведущая к отрыву варп-пузыря от пространства, – это то, что делает варп-двигатель с пространством, будучи искусственной инженерной конструкцией.

Уход варп-корабля в 4-мерную кротовую нору – это порождение сингулярности типа, показанного на рис. 3 и 5. На рис. 6 показан дальнейший возможный процесс с разрывом многообразия на два «куска». Но детали того, как на самом деле происходит этот разрыв и как следует проводить данную «хирургию», очень сложны [9, р. 28].



Рис. 6. «Пережим» и «отщипывание» шейки в 3-пространстве посредством потока Риччи [9]

6. Управляемые потоки Риччи с хирургией

Особый интерес для нас представляют управляемые 3-мерные потоки [16, гл. 15], т. е. в 4-мерном пространстве-времени, чьи сингулярности строго контролируются как топологически, так и геометрически. При определённых гипотезах, которым должны удовлетворять эти эволюции доказывається, что эти управляемые 3-мерные потоки Риччи с хирургией всегда существуют для любого компактного 3-многообразия в качестве начальной метрики. Иначе говоря, существуют потоки Риччи с хирургией, определённые для любого времени с любыми нормализованными начальными условиями.

7. Потоки Риччи и уравнения Эйнштейна

Как видим, потоки Риччи с хирургией дают инструмент для исследования деталей, сопровождающих полёт варп-корабля.

Действительно, при включении варп-двигателя на сверхсветовой режим, а это означает выделение энергии, описываемой тензором энергии-импульса ${}^{(4)}T_{\alpha\beta}$, про-

исходит изменение геометрии пространства-времени в соответствии с уравнениями Эйнштейна:

$${}^{(4)}R_{\alpha\beta} = k \left({}^{(4)}T_{\alpha\beta} - \frac{1}{2} {}^{(4)}g_{\alpha\beta} {}^{(4)}T \right), \quad (2)$$

$$\alpha, \beta = 0, 1, 2, 3.$$

Рассматриваем метрику ${}^{(4)}g_{\alpha\beta}$ пространства-времени в синхронной системе отсчёта:

$$ds^2 = {}^{(4)}g_{\alpha\beta} dx^\alpha dx^\beta = dx^{0^2} - g_{ik} dx^i dx^k,$$

$$i, k = 1, 2, 3.$$

Пространственная часть уравнений (2) имеет вид [17, с. 372]:

$$R_{ik} + \frac{1}{2} \frac{\partial^2 g_{ik}}{\partial t^2} + \frac{1}{4} \left[\frac{\partial \ln \det ||g_{mn}||}{\partial t} \cdot \frac{\partial g_{ik}}{\partial t} - 2g^{ls} \frac{\partial g_{il}}{\partial t} \frac{\partial g_{ks}}{\partial t} \right] =$$

$$= k \left({}^{(4)}T_{ik} + \frac{1}{2} g_{ik} {}^{(4)}T \right). \quad (3)$$

С учётом того, что

$${}^{(4)}g_{ik} = -g_{ik},$$

или, более кратко,

$$R_{ik} + \frac{g_{kj}}{2\sqrt{\det ||g_{mn}||}} \frac{\partial}{\partial t} \left(\sqrt{\det ||g_{mn}||} g^{js} \frac{\partial g_{is}}{\partial t} \right) = k \left({}^{(4)}T_{ik} - \frac{1}{2} g_{ik} {}^{(4)}T \right), \quad (4)$$

задавая тем самым тензор Риччи R_{ik} .

8. Оценка энергии формирования кротовой норы

Поток Риччи сопровождается «потоком» энергии ${}^{(4)}T_{\alpha\beta}$. И это та энергия, которая обеспечивает успешную работу варп-двигателя на сверхсветовом режиме. Следовательно, для оценки требуемой энергии для полёта варп-корабля, с учётом наличия потока Риччи $g(t)$, и принимая

$${}^{(4)}T_{\alpha\beta} = \varepsilon u_\alpha u_\beta, \quad u_\alpha = (1, 0, 0, 0),$$

имеем ${}^{(4)}T = \varepsilon$ и благодаря формуле (3)

$$k\varepsilon g_{ik} = 2R_{ik} + \frac{\partial^2 g_{ik}}{\partial t^2} + \frac{1}{2} \left[\frac{\partial \ln \det ||g_{mn}||}{\partial t} \cdot \frac{\partial g_{ik}}{\partial t} - 2g^{ls} \frac{\partial g_{il}}{\partial t} \frac{\partial g_{ks}}{\partial t} \right]. \quad (5)$$

Откуда получаем для плотности требуемой энергии для варп-полёта следующую оценку:

$$\varepsilon = \frac{2}{3k} R + \frac{1}{3k} g^{ik} \frac{\partial^2 g_{ik}}{\partial t^2} + \frac{1}{6k} g^{ik} \left[\frac{\partial \ln \det ||g_{mn}||}{\partial t} \cdot \frac{\partial g_{ik}}{\partial t} - 2g^{ls} \frac{\partial g_{il}}{\partial t} \frac{\partial g_{ks}}{\partial t} \right], \quad (6)$$

или, более кратко, опираясь на (4),

$$\varepsilon = \frac{2}{3k}R + \frac{1}{3k\sqrt{\det ||g_{mn}||}} \frac{\partial}{\partial t} \left(\sqrt{\det ||g_{mn}||} g^{is} \frac{\partial g_{is}}{\partial t} \right). \quad (7)$$

Манипуляции, совершаемые варп-двигателем, соответствуют хирургии удаляющей сингулярности. Затем повторный поток Риччи на каждом из образовавшихся кусков многообразия, возможность которых доказал Г. Перельман, приводит их соответственно к состояниям оставленного пространства и пространства варп-пузыря, перемещающегося в 4-мерном пространстве-времени.

Обратим внимание еще на то, что кроме чисто пространственной составляющей уравнений Эйнштейна имеют место еще два уравнения [17, с. 372]:

для ${}^{(4)}R_0^0 -$

$$\varepsilon = \frac{2}{k} {}^{(4)}R_0^0 = -\frac{1}{k} \frac{\partial^2 \ln \det ||g_{mn}||}{\partial t^2} - \frac{1}{2k} g^{is} g^{jk} \frac{\partial g_{ik}}{\partial t} \frac{\partial g_{js}}{\partial t}$$

и для ${}^{(4)}R_i^0 -$

$${}^{(4)}R_i^0 = \frac{1}{2} \left[\nabla_s g^{sk} \frac{\partial g_{ik}}{\partial t} - \frac{\partial^2}{\partial x^i \partial t} \ln \det ||g_{mn}|| \right] = k {}^{(4)}T_i^0 = 0.$$

Хотелось бы думать, что все эти уравнения согласуются и не ведут к противоречивым выводам, касающимся величины энергии, сопровождающей полёт варп-корабля.

9. Знак энергии

Вопрос о знаке плотности энергии ε имеет принципиальное значение. В случае $\varepsilon > 0$, мы имеем дело с традиционной энергией, а в случае $\varepsilon < 0$ – с экзотической энергией.

Поток Риччи сохраняет положительную скалярную кривизну $R \geq 0$ [18, р. 16]. Это следует из уравнения

$$\frac{\partial R}{\partial t} = \Delta R + 2R_{ik}R^{ik}$$

и наблюдения, что $|Ric|^2 = R_{ik}R^{ik} \geq 0$. Напротив, отрицательность скалярной кривизны не сохраняется. Поэтому формула (7), скорее всего, говорит нам, что энергия, потребляемая варп-кораблём, положительна. В пользу этого говорит теорема о сжатии Гамильтона – Айви, показывающая, что отрицательная часть кривизны становится пренебрежимо малой по сравнению со скалярной кривизной $R(x, t)$.

В какой мере энергия, рассчитываемая по формуле (6) или (7), связана с энергией, необходимой для формирования кротовой норы? Если образуется 4-мерная кротовая нора, т. е. появляется в результате потока Риччи с хирургией при разделении пространства на два куска, энергия, как показано в работах [4, 19], использовавших другие методы, должна быть положительной.

10. В каком месте пространства успешен быстрый старт варп-корабля

Ответ, казалось бы, мы находим у Г. Перельмана: поток Риччи **не может быстро превратить почти евклидову область** в очень искривлённую независимо от того, что происходит далеко [13, р. 4]. Другими словами, для топологической метаморфозы, производимой варп-кораблём, предполагающей сильное искривление пространства, влекущее образование сингулярности, геометрия пространства, окружающего варп-корабль, не должна быть почти евклидовой, т. е. имеющей метрику

$$dl^2 = \delta_{ik} dx^i dx^k = dr^2 + r^2(d\theta^2 + \sin^2 \theta d\phi^2).$$

Окружающее пространство планет или Солнца описывается метрикой Шварцшильда

$$dl_{\text{Шв.}}^2 = \frac{dr^2}{1 - \frac{r_g}{r}} + r^2(d\theta^2 + \sin^2 \theta d\phi^2),$$

которая, как видим, описывает геометрию, отличную от евклидовой. Напрашивается вывод: успех быстрого запуска варп-корабля обеспечивается близостью места запуска к тяготеющей звезде.

Однако ответ Г. Перельмана должно уточнить. Варп-двигатель, выходящий на сверхсветовой режим, порождает плотность энергии в варп-пузыре гораздо большую, чем 10^{30} эрг/см^3 . А это означает существенное искривление окружающего пространства.

Заключение

Обращение к теории потоков Риччи, меняющих геометрию и топологию 3-пространства, естественно, когда мы начинаем исследовать детали ухода варп-корабля в кротовую нору. Более того, геометро-топологическая метаморфоза изменения 3-пространства при образовании кротовой норы (потеря связности пространством) в статье [3] была описана весьма абстрактно – была дана скорее модель этой метаморфозы на одном базовом множестве, позволившая провести оценку необходимой энергии, но теперь появилась возможность уточнять детали полёта варп-корабля. Хотя и в настоящее время сама хирургия сингулярностей с точки зрения физики нуждается в подробном изучении.

Литература

1. Alcubierre M. The Warp drive: Hyper – fast travel within general relativity // Classical and Quantum Gravity. 1994. Vol. 11. P. 73–77. URL: <http://arXiv:gr-qc/0009013> (дата обращения: 10.11.2024).
2. Pfenning M.J., Ford L.H. Quantum Inequality Restrictions on Negative Energy Densities in Curved Spacetimes. URL: <https://arxiv.org/pdf/gr-qc/9805037> (дата обращения: 10.11.2024).
3. Гуц А.К. Изменение топологии физического пространства в замкнутой вселенной // Известия вузов. Физика. 1982. № 5. С. 23–26.

4. Гуц А.К. Топологический характер работы варп-двигателя Алькубьерре при выходе на сверхсветовую скорость // Эффективное обеспечение научно-технического прогресса: исследование задач и поиск решений: сб. ст. Междунар. науч.-практ. конф. (г. Магнитогорск, РФ, 25 августа 2024г.). Уфа: Аэтерна, 2024. С. 6–9.
5. Hamilton R.S. Three-manifolds with positive Ricci curvature // Journal of Differential geometry. 1982. Vol. 17 (2). P. 255–306.
6. Khan G. An Illustrated Introduction to the Ricci Flow. URL: <http://arXiv:2201.04923v1> (дата обращения: 19.12.2024).
7. Бессьер Л., Бессон Ж., Буало М. Доказательство гипотезы Пуанкаре (по работам Г. Перельмана) // Математическое просвещение. 2019. Вып. 24. С. 53–69.
8. Perelman G. Ricci flow with surgery on three-manifolds. URL: <http://arXiv:math/0303109v1> (дата обращения: 19.12.2024).
9. Sheridan N. Hamilton's Ricci Flow. Honours Thesis. Melbourne, 2006.
10. Риччи поток. URL: https://ru.wikibrief.org/wiki/Ricci_flow#Blow-up_limits_of_singularities (дата обращения: 19.12.2024).
11. Bamler R. Ricci Flow. Lecture Notes by Otis Chodosh and Christos Mantoulidis. URL: <https://web.stanford.edu/~ochodosh/Bamler-RFnotes.pdf> (дата обращения: 19.12.2024).
12. Angenent S., Knopf D. An example of neckpinching for Ricci flow on S^{n+1} // Mathematical Research Letters. 2004. Vol. 11, No. 4. P. 493–518.
13. Perelman G. The entropy formula for the Ricci flow and its geometric applications. URL: <http://arXiv:math.DG/0211159v1> (дата обращения: 19.12.2024).
14. Angenent S.B., Caputo M.C., Knopf D. Minimally invasive surgery for Ricci flow singularities. URL: <https://arxiv.org/pdf/0907.0232> (дата обращения: 10.02.2025).
15. Lakzian S. Intrinsic flat continuity of Ricci flow through neckpinch singularities // Geom Dedicata. 2015. Vol. 179. P. 69–89.
16. Morgan J.W., Gang Tian. Ricci Flow and the Poincare Conjecture. URL: <https://arxiv.org/pdf/math/0607607v2> (дата обращения: 19.12.2024).
17. Ландау Л.Д., Лифшиц Е.М. Теоретическая физика: в 10 т. М.: ФМ, 1968. Т. 2. Теория поля.
18. Hamilton R. The formations of singularities in the Ricci Flow // Surveys in Differential Geometry. 1993. Vol. 2, No. 1. P. 7–136. URL: <https://dx.doi.org/10.4310/SDG.1993.v2.n1.a2> (дата обращения: 10.02.2025).
19. Гуц А.К. Энергия, необходимая для порождения варп-двигателем кротовых нор // 18-я Российская гравитационная конференция – Международная конференция по гравитации, космологии и астрофизике «RusGrav-18». Тезисы докладов. Казань: Изд-во КФУ, 2024. С. 73–75.

ALCUBIERRE SPACE WARP SHIP TOPOLOGICAL FLIGHTS AND RICCI FLOWS

A.K. Guts

Dr.Sc. (Phys.-Math.), Professor, e-mail: aguts@mail.ru

Sochi State University, Sochi, Russia

Abstract. In the paper the possibility of describing the flights of ultra-fast Alcubierre warp spaceships that change the topology and geometry of 3-dimensional physical space using Ricci flows is demonstrated.

Keywords: warp ship, Alcubierre, changes in topology and geometry, Ricci flows, worm hole, sign of energy.

Дата поступления в редакцию: 10.02.25

ЗАДАЧА О РАЗРУШЕНИИ ДАМБЫ В МОДЕЛИ МЕЛКОЙ ВОДЫ

С.Л. Дерябин

д.ф.-м.н., профессор, e-mail: SDeryabin@usurt.ru

А.П. Садов

к.ф.-м.н., доцент, e-mail: alsadov@yandex.ru

Уральский государственный университет путей сообщения, Екатеринбург, Россия

Аннотация. Рассматриваются одномерные течения воды, возникающие после разрушения дамбы. Предполагается, что в начальный момент времени справа от дамбы – покоящаяся жидкость, слева – сухой берег. В качестве математической модели используется система уравнений мелкой воды. В системе уравнений мелкой воды в переменную x вводится автомодельная особенность, и для полученной системы ставится задача Коши с данными на звуковой характеристике. Решение начально-краевой задачи строится в виде степенного ряда. Коэффициенты ряда находятся при интегрировании обыкновенных дифференциальных уравнений. Далее для доказательства сходимости этого ряда ставится начально-краевая задача в пространстве других независимых переменных. Решение этой начально-краевой задачи строится в виде своего сходящегося степенного ряда, и доказываем эквивалентность решений первой и второй начально-краевых задач.

Ключевые слова: распад специального разрыва, система уравнений мелкой воды, начально-краевая задача, нестационарные автомодельные переменные, степенные ряды.

Введение

Для описания распространения длинных волн используются классические уравнения теории мелкой воды [1, 2]. Среди всех рассматриваемых задач для численного моделирования наиболее сложной оказалась задача о выходе волны на берег, когда на некоторой линии (линии уреза) глубина жидкости обращается в нуль. Заметим, что в модели мелкой воды система уравнений фактически совпадает с системой уравнений газовой динамики для изэнтропических течений политропного газа с показателем политропы $\gamma = 2$. Разница только в том, что вместо плотности ρ в системе уравнений теории мелкой воды неизвестной функцией является глубина H . В газовой динамике свободная поверхность, выделяемая с помощью соотношения $\rho = 0$, соответствует границе «газ-вакуум», а в модели мелкой воды равенство $H = 0$ задаёт границу уреза. Поэтому для решения поставленной задачи используется методология исследования задач об истечении газа в вакуум [3]. В случае гладкой границы уреза с использованием этой методики была решена задача о выходе волны на берег

в 1-мерном [4] и 2-мерном [5] случаях в пространстве специальным образом введённых новых независимых переменных. Вместе с тем восстановление этих решений в физическом пространстве вызывает большие трудности. В данной работе введение автомодельной особенности в независимую переменную позволило решить краевую задачу с данными на звуковой характеристике в физическом пространстве.

1. Постановка задачи

В момент $t = 0$ дамба Γ с уравнением $x = x_{00}$ отделяет покоящуюся жидкость от сухого берега. В задаче предполагается, что жидкость находится справа от дамбы, а сухой берег – слева от неё, и на жидкость действует сила тяжести. Будет предполагаться, что в начальный момент времени $t = 0$ высота дамбы $H(x_{00}) = H_{00}$, скорость жидкости $u(x_{00}) = u_{00}$ (см. рис. 1).

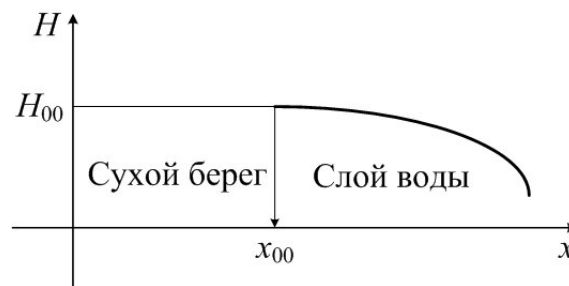


Рис. 1. Конфигурация течения в начальный момент времени

В рассматриваемой задаче в момент $t = 0$ начинается движение жидкости, определяемое заданными аналитическими функциями $u_0(x)$, $H_0(x)$, которое в дальнейшем называется невозмущённой волной.

Также в момент $t = 0$ происходит разрушение дамбы, и возникает течение, граничащее с невозмущённой волной и называемое далее возмущённой волной. Это другое течение – возмущённая волна – отделено от невозмущённой волны линией Γ_1 , являющейся звуковой характеристикой, на которой имеет место слабый разрыв. С другой стороны, возмущённая волна ограничена линией Γ_0 :

$$H(t, x)|_{\Gamma_0} = 0.$$

Такая задача называется задачей о распаде специального разрыва. Требуется построить возмущённую волну, а также найти закон движения Γ_0 .

Схематично конфигурация течения, возникшего после момента времени $t = 0$, изображена на рис. 2.

В предположении, что дно горизонтально, система уравнений мелкой воды для одномерных течений записывается в виде [4]:

$$H_t + H_x u + H u_x = 0, \tag{1}$$

$$u_t + u_x u + g H_x = 0.$$

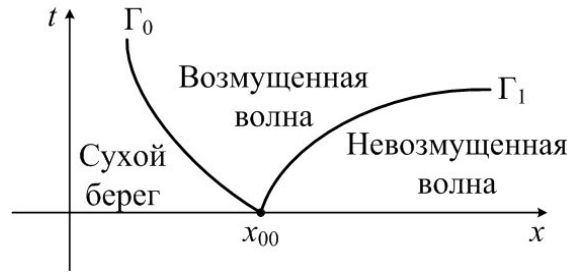


Рис. 2. Конфигурация течения после распада специального разрыва

Здесь H – полная глубина жидкости от непроницаемого дна до свободной поверхности, u – скорость жидкости, g – ускорение свободного падения.

Линия $x = x_0(t)$, для которой

$$H(t, x_0(t)) = 0, \quad (2)$$

называется границей уреза. В каждый момент времени она отделяет слой жидкости от сухого берега. Будет предполагаться, что волна набегаёт на берег справа налево.

Поскольку рассматриваемая система (1) является системой типа Ковалевской, а начальные данные – это аналитические функции, то задача Коши для системы (1) имеет [6] аналитическое решение:

$$u = u^0(t, x), \quad H = H^0(t, x).$$

В системе (1) сделаем замену переменных: вместо функции $H(t, x)$ введём функцию $C(t, x)$ по следующей формуле: $H = C^2$. После преобразований получим

$$C_t + uC_x + \frac{1}{2}Cu_x = 0, \quad (3)$$

$$u_t + uu_x + 2gCC_x = 0.$$

В этих переменных фоновое течение имеет вид

$$u = u^0(t, x), \quad C = C^0(t, x) = \sqrt{H^0(t, x)}, \quad C_{00} = \sqrt{H_{00}}. \quad (4)$$

Закон движения характеристики $\Gamma_1 : x = x_1(t)$ определяется из решения дифференциальной задачи [7]

$$x_{1t} = u^0(t, x_1) + \sqrt{g}C^0(t, x_1), \quad x_1(0) = x_{00}. \quad (5)$$

По теореме Ковалевской задача (5) имеет единственное аналитическое решение. Представим его в виде ряда по степеням t :

$$x_1(t) = \sum_{k=0}^{\infty} x_{1k} \frac{t^k}{k!}. \quad (6)$$

Из уравнения (5) имеем $x_{10} = x_{00}$, $x_{11} = u_{00} + \sqrt{g}C_{00}$.

Продифференцируем по t уравнение (5), получим

$$x_{1tt} = u_t^0 + \sqrt{g}C_t^0.$$

Отсюда

$$x_{12} = u_t^0(0, x_{00}) + \sqrt{g}C_t^0(0, x_{00}).$$

При помощи решения (6) однозначно строится закон движения Γ_1 :

$$x = x_1(t) = x_{00} + tx_2(t),$$

где $x_2(t)$ – аналитическая функция.

Условия на характеристике Γ_1 будут иметь вид

$$u(t, x)|_{x=x_1(t)} = u^0(t, x_1(t)), \quad C(t, x)|_{x=x_1(t)} = C^0(t, x_1(t)). \quad (7)$$

Для построения возмущённой волны в системе (1) введём нестационарные автомодельные переменные по следующим формулам:

$$t = t', \quad y = \frac{x - x_{00}}{t}.$$

В результате этой замены получим систему:

$$tC_t + (u - y)C_y + \frac{1}{2}Cu_y = 0, \quad (8)$$

$$tu_t + (u - y)u_y + 2gCC_y = 0$$

с условиями на характеристике Γ_1

$$u(t, y)|_{y=x_2(t)} = u^0(t, x_1(t)), \quad C(t, y)|_{y=x_2(t)} = C^0(t, x_1(t)). \quad (9)$$

2. Построение возмущённой волны в пространстве независимых переменных t, y

Построим формальное решение задачи (8)–(9) в виде ряда по степеням t :

$$f(t, y) = \sum_{n=0}^{\infty} f_n(y) \frac{t^n}{n!}, \quad f = \{C, u\}. \quad (10)$$

В системе (8) положим $t = 0$. Получим уравнения для определения нулевых коэффициентов ряда (10)

$$(u_0 - y)C_{0y} + \frac{1}{2}C_0u_{0y} = 0, \quad (11)$$

$$(u_0 - y)u_{0y} + 2gC_0C_{0y} = 0.$$

Ненулевое решение системы (11) получается, если определитель системы равен нулю, т. е.

$$(u_0 - y)^2 = gC_0^2$$

или

$$u_0 - y = \pm\sqrt{g}C_0.$$

Поскольку при $t = 0$ на характеристике Γ_1 $y = u_{00} + \sqrt{g}C_{00}$, то окончательно имеем

$$u_0 - y = -\sqrt{g}C_0. \quad (12)$$

Подставляя $u_0 - y$ во второе уравнение системы (11), получим

$$u_{0y} = 2\sqrt{g}C_{0y}.$$

Интегрируя, имеем

$$u_0 = 2\sqrt{g}C_0 + D.$$

Из условий (9) определяем произвольную постоянную D и получаем

$$u_0 = 2\sqrt{g}C_0 + u_{00} - 2\sqrt{g}C_{00}.$$

Подставляя u_0 в уравнение (12), после преобразований будем иметь

$$C_0 = \frac{1}{3\sqrt{g}}(y - u_{00} + 2\sqrt{g}C_{00}), \quad u_0 = \frac{2}{3}y - \frac{2}{3}\sqrt{g}C_{00} + \frac{1}{3}u_{00}. \quad (13)$$

Также справедливы соотношения

$$C_{0y} = \frac{1}{3\sqrt{g}}, \quad u_{0y} = \frac{2}{3}y. \quad (14)$$

Систему (8) продифференцируем по t , положим $t = 0$ и с учётом (12), (14) имеем

$$\begin{aligned} C_0 u_{1y} - 2\sqrt{g}C_0 C_{1y} + \frac{2}{3\sqrt{g}}u_1 + \frac{8}{3}C_1 &= 0, \\ -C_0 u_{1y} + 2\sqrt{g}C_0 C_{1y} + \frac{5}{3\sqrt{g}}u_1 + \frac{2}{3}C_1 &= 0. \end{aligned} \quad (15)$$

При сложении уравнений системы (15) получаем

$$\frac{7}{3\sqrt{g}}u_1 + \frac{10}{3}C_1 = 0$$

или

$$u_1 = -\frac{10}{7}\sqrt{g}C_1, \quad u_{1y} = -\frac{10}{7}\sqrt{g}C_{1y}.$$

Подставляя u_1 и u_{1y} во второе уравнение (15), после преобразований имеем

$$\sqrt{g}C_0 C_{1y} - \frac{1}{2}C_1 = 0.$$

Подставляя C_0 в уравнение, после преобразований получим

$$(y - u_{00} + 2\sqrt{g}C_{00}) C_{1y} - \frac{3}{2}C_1 = 0.$$

Интегрируя уравнение, имеем

$$C_1 = C_{10} (y - u_{00} + 2\sqrt{g}C_{00})^{\frac{3}{2}},$$

$$u_1 = -\frac{10}{7}\sqrt{g}C_{10} (y - u_{00} + 2\sqrt{g}C_{00})^{\frac{3}{2}}.$$

(16)

Произвольная постоянная C_{10} определится из условий (9).

Замечание 1. В частном случае, когда в начальный момент времени слой жидкости имеет постоянную глубину и покоится $H = H_0(x) = H_{00}$, $u = u_0(x) = 0$, произвольная постоянная $C_{10} = 0$. В этом случае ряд (10) обрывается, коэффициенты $C_n = u_n = 0$, $n \geq 1$, и задача (8)–(9) имеет точное решение.

$$H = C^2 = \frac{1}{9g} \left(\frac{x - x_{00}}{t} + 2\sqrt{gH_{00}} \right)^2,$$

$$u = \frac{2}{3} \left(\frac{x - x_{00}}{t} - \sqrt{gH_{00}} \right).$$

Систему (8) продифференцируем n раз по t , положим $t = 0$ и с учётом (12), (14) будем иметь

$$C_0 u_{ny} - 2\sqrt{g}C_0 C_{ny} + \frac{2}{3\sqrt{g}}u_n + \left(2n + \frac{2}{3}\right) C_n = F_{1n}(y),$$

(17)

$$-C_0 u_{ny} + 2\sqrt{g}C_0 C_{ny} + \frac{1}{\sqrt{g}} \left(n + \frac{2}{3}\right) u_n + \frac{2}{3}C_n = F_{2n}(y).$$

Здесь $F_{1n}(y)$, $F_{2n}(y)$ – функции, известным образом зависящие от уже найденных коэффициентов ряда (10).

При сложении уравнений системы (17) получаем

$$\frac{1}{\sqrt{g}} \left(n + \frac{4}{3}\right) u_n + \left(2n + \frac{4}{3}\right) C_n = F_{1n}(y) + F_{2n}(y) = F_n^+(y).$$

Или

$$u_n = -\frac{6n + 4}{3n + 4}\sqrt{g}C_n + \frac{3}{3n + 4}\sqrt{g}F_n^+(y),$$

$$u_{ny} = -\frac{6n + 4}{3n + 4}\sqrt{g}C_{ny} + \frac{3}{3n + 4}\sqrt{g}F_{ny}^+(y).$$

Подставляя u_n, u_{ny} во второе уравнение (17), после преобразований имеем

$$\sqrt{g} \left(2 + \frac{6n+4}{3n+4} \right) C_0 C_{ny} + \frac{2}{3} \left(1 - \frac{(3n+2)^2}{3n+4} \right) C_n = F_n(y).$$

Здесь функция $F_n(y) = F_{2n}(y) + \frac{3\sqrt{g}}{3n+4} C_0 F_{ny}^+(y) - \frac{3n+2}{3n+4} F_n^+(y)$.

Подставляя C_0 в уравнение, после преобразований получим

$$(y - u_{00} + 2\sqrt{g}C_{00}) C_{ny} - \frac{3}{2} n C_n = \frac{3n+4}{4n+4} F_n(y). \quad (18)$$

Интегрируя уравнение (18), имеем

$$\begin{aligned} C_n &= (y - u_{00} + 2\sqrt{g}C_{00})^{\frac{3}{2}n} (C_{n0} + \\ &+ \frac{3n+4}{4n+4} \int_{x_2(0)}^y F_n(y) (y - u_{00} + 2\sqrt{g}C_{00})^{-\frac{3}{2}n-1} dy), \\ u_n &= -\frac{6n+4}{3n+4} \sqrt{g} (y - u_{00} + 2\sqrt{g}C_{00})^{\frac{3}{2}n} (C_{n0} + \\ &+ \frac{3n+4}{4n+4} \int_{x_2(0)}^y F_n(y) (y - u_{00} + 2\sqrt{g}C_{00})^{-\frac{3}{2}n-1} dy) + \\ &+ \frac{3}{3n+4} \sqrt{g} F_n^+(y). \end{aligned} \quad (19)$$

Произвольные постоянные C_{n0} определяются из условий (9).

Таким образом, построенное в виде ряда (10) формальное решение задачи (8)–(9) является единственным.

Теорема 1. *Задача (8)–(9) имеет единственное аналитическое решение, которое представляется в виде сходящегося ряда (10).*

Доказательство теоремы проводится методом мажорант, применение которого изложено в [8].

С помощью простейших преобразований решение (10) восстанавливается в физическом пространстве переменных t, x .

$$H = C^2 \left(t, \frac{x - x_{00}}{t} \right), \quad u = u \left(t, \frac{x - x_{00}}{t} \right).$$

3. Построение решения в пространстве независимых переменных t, C

Для завершения исследования необходимо показать, что решение, построенное в [4], совпадает с (10). Для этого решение задачи о разрушении дамбы построим в пространстве, где в системе (3) за независимые переменные берутся t, C , а за неизвестные функции – x, u . То есть переменные x и C меняются ролями. Якобиан преобразования $J = x_C$.

В результате такой замены вместо системы (3) получается система

$$x_t = u + \frac{1}{2}Cu_C, \quad (20)$$

$$x_C u_t + (u - x_t)u_C + 2gC = 0.$$

Для системы (20) начальные данные (7) на характеристике Γ_1 перепишутся в виде:

$$u(t, x)|_{\Gamma_1} = u(t, x_1(t)), \quad x|_{\Gamma_1} = x_1(t). \quad (21)$$

Течение в области между Γ_1 и Γ_0 будем строить как решение системы (20) с данными на характеристике (21). Поскольку Γ_1 – характеристика кратности один, то для получения единственного локально-аналитического решения необходимо задать одно дополнительное условие [8]. Этим условием в пространстве переменных t, C служит [3] граничное условие вертикали

$$x(0, C) = x_{00}. \quad (22)$$

Разложим решение задачи (20)–(21) в ряд по степеням t :

$$\mathbf{q}(t, C) = \sum_{k=0}^{\infty} \mathbf{q}_k(C) \frac{t^k}{k!}, \quad \mathbf{q} = \{x, u\}. \quad (23)$$

В силу условия (22) $x_0 = x_{00}$.

В системе (20) положим $t = 0$ и с учётом (21), (22) получим

$$x_1 = 3\sqrt{g}C - 2\sqrt{g}C_{00} + u_{00}, \quad u_0 = 2\sqrt{g}C - 2\sqrt{g}C_{00} + u_{00}.$$

Систему (20) продифференцируем k раз по t , положим $t = 0$ и с учётом (22) получим уравнения для определения коэффициентов ряда (23):

$$x_{k+1} = u_k + \frac{1}{2}Cu_{kC} + G_{1k}(C), \quad (24)$$

$$Cu_{kC} - \frac{3}{2}ku_k = G_{2k}(C).$$

Здесь $G_{1k}(C)$, $G_{2k}(C)$ – функции, известным образом зависящие от уже найденных коэффициентов ряда (23).

Интегрируя второе уравнение системы (24), имеем

$$u_k = C^{\frac{3}{2}k} \left(C_{k0} + \int_{C_{00}}^C G_{2k}(C) C^{-\frac{3}{2}k-1} dy \right).$$

Теорема 2. Область сходимости рядов (23) покрывает всю зону течения от Γ_1 до Γ_0 включительно. При этом закон движения границы уреза определяется по следующей формуле:

$$x_t = x_{00} + tu_{00}.$$

Доказательство теоремы аналогично доказательству из [3] и проводится по методике [8], позволяющей установить неограниченность области сходимости рядов по соответствующей переменной.

Теорема 3. Краевая задача (8)–(9) эквивалентна начально-краевой задаче (20)–(22).

Доказательство. На основании теоремы 2 можно утверждать, что в виде сходящегося ряда (23) единственным образом построено решение начально-краевой задачи (20)–(22). Полученное решение имеет вид

$$x = x_{00} + (3\sqrt{g}C - 2\sqrt{g}C_{00} + u_{00})t + \dots + x_k(C)\frac{t^k}{k!} + \dots, \quad (25)$$

$$u = (2\sqrt{g}C - 2\sqrt{g}C_{00} + u_{00}) + u_1(C)t + \dots + u_k(C)\frac{t^k}{k!} + \dots$$

В первом соотношении системы (25) вместо x введём $y = \frac{x - x_{00}}{t}$. Получим

$$y = (3\sqrt{g}C - 2\sqrt{g}C_{00} + u_{00}) + \dots + x_k(C)\frac{t^{k-1}}{k!} + \dots \quad (26)$$

Продифференцируем соотношение (26) по C , положим $t = 0$ и будем иметь

$$y_C = 3\sqrt{g} \neq 0.$$

Следовательно, по теореме о существовании обратной функции соотношение (26) задаёт локально-аналитическую функцию $C = C(t, y)$. Тогда решение задачи (20)–(22) можно переписать в виде

$$C = C(t, y), \quad (27)$$

$$u = (2\sqrt{g}C(t, y) - 2\sqrt{g}C_{00} + u_{00}) + \dots + u_k(C(t, y))\frac{t^k}{k!} + \dots$$

Так как при $t > 0$ якобиан преобразования $J = x_c \neq 0$, то решение (27) удовлетворяет системе (8). Также оно удовлетворяет условиям (9).

Покажем, что при $t = 0$ нулевые коэффициенты рядов (10) и (23) совпадают.

В решении (27) положим $t = 0$ и, учитывая (26), получим

$$y = 3\sqrt{g}C|_{t=0} - 2\sqrt{g}C_{00} + u_{00}, \quad u|_{t=0} = 2\sqrt{g}C|_{t=0} - 2\sqrt{g}C_{00} + u_{00}. \quad (28)$$

Из первого соотношения (28) найдём $C|_{t=0}$ и, подставляя во второе соотношение (28), будем иметь

$$C|_{t=0} = \frac{1}{3\sqrt{g}}(y + 2\sqrt{g}C_{00} - u_{00}), \quad (29)$$

$$u|_{t=0} = \frac{2}{3}y - \frac{2}{3}\sqrt{g}C_{00} + \frac{1}{3}u_{00}.$$

Условия (29) совпадают с условиями (13). В силу единственности построенных решений задачи (8)–(9) и задачи (20)–(22) решение (10) совпадает с решением (23). Теорема доказана. ■

Заключение

1. Таким образом, с помощью системы уравнений мелкой воды построены течения жидкости, возникающие после разрушения дамбы.
2. Введение в систему автомодельной особенности в переменную x позволило решить краевую задачу с данными на звуковой характеристике в физическом пространстве.
3. Решение краевой задачи было построено в виде сходящегося степенного ряда.
4. В частном случае покоящейся жидкости решение задачи получено в виде конечных формул.
5. Исследование задачи о разрушении дамбы в специальном функциональном пространстве позволило получить закон движения границы уреза в виде конечной формулы.

Литература

1. Хакимзянов Г.С., Шокин Ю.И., Барахнин В.Б., Шокина Н.Ю. Численное моделирование течений жидкости с поверхностными волнами. Новосибирск: Изд-во СО РАН, 2001. 393 с.
2. Баутин С.П., Дерябин С.Л., Мезенцев А.В., Чуев Н.П. Начально-краевые задачи для моделирования движения сплошной среды с особенностями на свободной границе. Новосибирск: Наука, 2015. 191 с.
3. Баутин С.П., Дерябин С.Л. Математическое моделирование истечения идеального газа в вакуум. Новосибирск: Наука, 2005. 390 с.
4. Bautin S.P., Deryabin S.L., Sommer A.F., Khakimzyanov G.S., Shokina N. Yu. Use of analytic solutions in the statement of difference boundary conditions on movable shoreline // Russian Journal of Numerical Analysis and Mathematical Modeling. 2011. Vol. 26, No. 4. P. 353–377.
5. Bautin S.P., Deryabin S.L. Two-dimensional solutions of the equations shallow-water theory in the neighbourhood of a shore line boundary // Journal of Applied Mathematics and Mechanics. 2015. Vol. 79, Is. 4. P. 358–366.
6. Курант Р. Уравнения с частными производными. М.: Мир, 1964. 830 с.
7. Овсянников Л.В. Лекции по основам газовой динамики. М.; Ижевск: Ин-т компьютер. исслед., 2003. 336 с.
8. Баутин С.П. Характеристическая задача Коши и её приложения в газовой динамике. Новосибирск: Наука, 2009. 368 с.

DAM BREAK PROBLEM IN SHALLOW WATER MODEL**S.L. Deryabin**

Dr.Sc. (Phys.-Math.), Professor, e-mail: SDeryabin@usurt.ru

A.P. Sadov

Ph.D. (Phys.-Math.), Associate Professor, e-mail: alsadov@yandex.ru.ru

Ural State University of Railway Engineering, Yekaterinburg, Russia

Abstract. The paper considers one-dimensional water flows that occur after the destruction of the dam. It is assumed that at the initial moment of time there is a resting liquid to the right of the dam, and a dry shore to the left. A system of shallow water equations is used as a mathematical model. In the system of shallow water equations, a self-similar feature is introduced into the variable x and for the resulting system, the Cauchy problem is posed with data on the sound characteristic. The solution of the initial boundary value problem is constructed in the form of a power series. The coefficients of the series are found by integrating ordinary differential equations. Next, to prove the convergence of this series, an initial boundary value problem is posed in the space of other independent variables. The solution to this initial boundary value problem is constructed in the form of its own convergent power series, and the equivalence of the solutions to the first and second initial boundary value problems is proven.

Keywords: decay of a special gap, a system of shallow water equations, an initial boundary value problem, nonstationary self-similar variables, power series.

Дата поступления в редакцию: 01.11.2024

СЛОЖНОСТЬ ЗАДАЧИ КАЛЕНДАРНОГО ПЛАНИРОВАНИЯ СО СКЛАДИРУЕМЫМИ РЕСУРСАМИ ПЕРЕМЕННОЙ СТОИМОСТИ

А.А. Романова

к.ф.-м.н., доцент, e-mail: RomanovaAA@omsu.ru

Омский государственный университет им. Ф.М. Достоевского, Омск, Россия

Аннотация. Исследуется сложность задачи календарного планирования со складировемыми ресурсами и критерием минимизации затрат на их приобретение при условии переменной стоимости ресурса и ограниченного склада. Показано, что в общем случае задача NP-трудна в сильном смысле. Разработан алгоритм динамического программирования, выделены псевдополиномиально разрешимые случаи.

Ключевые слова: расписание, складировемые ресурсы, динамическое программирование, полиномиальная сводимость.

Введение

В последние годы в теории и на практике возникает множество разнообразных задач календарного планирования проектов с учётом ограничений на ресурсы. В рамках проекта нужно выполнить комплекс взаимосвязанных работ. При этом каждая работа характеризуется длительностью и определёнными ресурсными потребностями и связанными с ними ограничениями. Необходимо, с учётом ограничений на ресурсы, определить расписание выполнения работ проекта, при котором значение целевой функции оптимально.

Задачи календарного планирования исследуются многими авторами. В случае складировемых ресурсов большая часть работ посвящена классическим постановкам, в которых в каждый период времени поступает определённое количество ресурса. В [1] показана полиномиальная разрешимость задачи с критерием минимизации срока завершения проекта. В [2] доказано, что для критерия максимизации чистой приведённой прибыли задача является NP-трудной в сильном смысле.

В данной работе рассматривается задача календарного планирования со складировемыми ресурсами и критерием минимизации затрат на приобретение ресурсов. Особенностью постановки является то, что стоимость ресурса меняется в течение горизонта планирования. Целью работы является анализ сложности задачи.

1. Постановка задачи

В работе рассматривается задача календарного планирования в следующей постановке. Имеется проект, состоящий из множества $V = \{1, \dots, n\}$ взаимосвязан-

ных работ. Для каждой работы $j \in V$ известна длительность p_j её выполнения. Прерывания выполнения работ не допускаются. Взаимосвязь между работами определяется технологией выполнения проекта и задаётся с помощью ациклического ориентированного графа $G = (V, E)$, в котором вершины соответствуют работам, а дуги – отношениям предшествования $i \rightarrow j$, что означает, что работа j не может начаться до завершения работы i . Задан директивный срок T выполнения проекта, к которому все работы должны быть завершены.

Для выполнения работ требуются складываемые ресурсы m видов. Известна потребность q_{rj} работы j в ресурсе вида r в каждый период времени её выполнения, $r = 1, \dots, m; j = 1, \dots, n$. Ресурсы для выполнения работ проекта необходимо приобретать. В силу целочисленности длительностей работ будем считать, что покупка ресурсов производится в целочисленные моменты времени. Также предполагаем, что приобретение ресурсов происходит в начале периода $t, t = 1, \dots, T$. Складываемые ресурсы характеризуются тем, что суммарное количество ресурса, закупленного к периоду времени t , не должно быть меньше суммарного требуемого ресурса к периоду t для всех $t = 1, \dots, T$.

Можно рассматривать классический случай неограниченного склада, когда излишек ресурса можно хранить в любом количестве. В случае ограниченного склада излишек ресурса r , т. е. ресурс, который предполагается потратить позже, в каждый период времени не должен превышать размера склада для ресурса r – величины $V_r, r = 1, \dots, m$.

В данной работе рассматривается случай изменяющейся во времени стоимости ресурса, что является актуальным в настоящее время и часто встречается на практике. Таким образом, известна цена c_{rt} единицы ресурса r в периоде времени $t, r = 1, \dots, m, t = 1, \dots, T$.

Решение определяется набором $\{s_j\}_{j=1}^n$ времён начала выполнения работ проекта и объёмом y_{rt} закупаемого ресурса r в период времени $t, t = 1, \dots, T$. Расписание и план закупок ресурсов определяют допустимое решение, если выполняются ограничения предшествования, время завершения всего проекта не превышает T , суммарный объём потребления каждого ресурса к периоду времени t не превосходит объёма закупленного ресурса, и если рассматривается случай ограниченного склада, то объём ресурса r , хранящегося на складе, не превышает V_r в каждый период времени $t, r = 1, \dots, m, t = 1, \dots, T$. Задача заключается в нахождении допустимого решения, при котором суммарные затраты на приобретение ресурсов минимальны.

Предполагаем, что все исходные числовые данные, кроме, может быть, стоимостей ресурсов, являются целочисленными. В связи с этим можем ограничиться рассмотрением расписаний, в которых работы начинаются в целочисленные моменты времени, а объём закупаемого ресурса в каждый период времени также является целочисленной величиной. Если допустимое решение не удовлетворяет этим условиям, очевидно, можно его преобразовать в решение нужной структуры, не увеличив значение целевой функции.

Рассматриваемая задача с возобновимыми ресурсами и постоянной стоимостью ресурсов, известная как RACP (Resource Availability Cost Problem), NP-трудна [3]. Ранее автором в совместных работах [4, 5] при стоимости ресурса, зависящей от

объёма его покупки, изучена сложность задачи, предложены алгоритмы решения.

В данной работе изучается задача в случае, когда цена за единицу ресурса меняется в течение периода планирования. Для неограниченного склада в [6] показана полиномиальная разрешимость задачи. Настоящая работа посвящена исследованию случая ограниченного склада. В п. 2 доказана NP-трудность в сильном смысле задачи; в п. 3 предлагается алгоритм динамического программирования и проводится анализ его трудоёмкости.

2. Сложность задачи в общем случае

В данном параграфе показана NP-трудность в сильном смысле задачи в общем случае. Доказательство проводится по аналогии с доказательством NP-трудности задачи календарного планирования с критерием максимизации чистой приведённой прибыли [2].

Теорема 1. *Задача календарного планирования со складываемыми ресурсами и критерием минимизации суммарных затрат на их приобретение в случае ограниченного склада и цен ресурсов, зависящих от времени, NP-трудна в сильном смысле даже при единичных длительностях работ.*

Доказательство. Проведём полиномиальное сведение задачи о максимальной клике к рассматриваемой. Сформулируем задачу распознавания, соответствующую задаче о максимальной клике. Даны граф $\Gamma = (V^\Gamma, E^\Gamma)$ и натуральное число $y_0 \leq |V^\Gamma|$. Существует ли клика в графе размера не меньше y_0 ? Будем считать без ограничения общности, что изолированных вершин граф не имеет, так как заведомо они не могут содержаться в клике.

Построим соответствующий класс примеров задачи календарного планирования следующим образом. Пусть число работ равно $n = |V^\Gamma| + |E^\Gamma|$, при этом будем выделять работы-вершины, соответствующие вершинам графа Γ , и работы-рёбра, соответствующие рёбрам графа Γ . Граф предшествования G двудольный: первая доля содержит работы-вершины, вторая – работы-рёбра; при этом работа-вершина предшествует работе-ребру, если и только если в графе Γ соответствующие вершина и ребро инцидентны. Кроме того, положим $m = 1$, $T = 3$. Длительности всех работ и потребности в ресурсе единичные: $p_j = 1$, $q_{1j} = 1$, $j = 1, \dots, n$. Стоимости за единицу ресурса равны $c_{11} = 2$, $c_{12} = 1$, $c_{13} = |V^\Gamma| + 1$; объём склада $V_1 = |E^\Gamma| - \frac{y_0(y_0-1)}{2}$.

Докажем, что граф Γ содержит клику размера не меньше y_0 тогда и только тогда, когда в соответствующей задаче календарного планирования существует допустимое решение с целевой функцией, не превышающей $y_0 + |V^\Gamma| + |E^\Gamma|$.

Пусть в графе Γ имеется клика размера не меньше y_0 , построим требуемое допустимое расписание в соответствующей задаче календарного планирования. Выделим в графе Γ клику размера y_0 и соответствующие вершины-работы выполним в первом периоде времени. Во втором периоде времени выполним оставшиеся вершины-работы и работы-рёбра, соответствующие рёбрам клики; наконец, в третьем периоде времени выполним работы-рёбра, которым предшествуют вершины-работы, не входящие в клику. Для выполнения работ приобретём ресурс в количе-

стве y_0 в начале первого периода. Во втором периоде купим остальные требуемые ресурсы в количестве $|V^\Gamma| + |E^\Gamma| - y_0$, при этом в этом же периоде потратим на выполнение работ $|V^\Gamma| - y_0 + \frac{y_0(y_0-1)}{2}$ единиц ресурса, а оставшиеся $|E^\Gamma| - \frac{y_0(y_0-1)}{2}$ отправим на склад, объёма которого хватает. Ресурсы для выполнения работ в третьем периоде времени берём со склада. Имеем допустимое решение со значением целевой функции $2y_0 + |V^\Gamma| + |E^\Gamma| - y_0 = y_0 + |V^\Gamma| + |E^\Gamma|$. Значит, требуемое ограничение по целевой функции выполнено.

Докажем в обратную сторону. Пусть имеется допустимое решение задачи из построенного класса примеров задачи календарного планирования. Пусть y_t , как и прежде, объём ресурса, закупленного в периоде t , $t = 1, \dots, 3$. Пусть в первый период времени выполняется k работ. Исходя из структуры графа предшествования, это работы-вершины. Так как изолированных вершин по предположению в графе предшествования нет, то во второй период времени выполняются оставшиеся $|V^\Gamma| - k$ работ-вершин (иначе все работы не успеют выполняться за три периода времени), а также e работ-рёбер, связанных с работами из первого периода времени. В третий период выполняются оставшиеся $|E^\Gamma| - e$ работ-рёбер. Докажем, что работы-вершины, выполняющиеся в первый период времени, соответствуют вершинам клики в исходном графе Γ , и их количество k равно y_0 . В силу допустимости решения должны быть выполнены неравенства (1)–(7):

$$2y_1 + y_2 + (|V^\Gamma| + 1)y_3 \leq y_0 + |V^\Gamma| + |E^\Gamma|, \quad (1)$$

$$k \leq y_1, \quad (2)$$

$$|V^\Gamma| + e \leq y_1 + y_2, \quad (3)$$

$$|V^\Gamma| + |E^\Gamma| \leq y_1 + y_2 + y_3, \quad (4)$$

$$y_1 - k \leq |E^\Gamma| - \frac{y_0(y_0 - 1)}{2}, \quad (5)$$

$$y_1 + y_2 - |V^\Gamma| - e \leq |E^\Gamma| - \frac{y_0(y_0 - 1)}{2}, \quad (6)$$

$$y_1 + y_2 + y_3 - |V^\Gamma| - |E^\Gamma| \leq |E^\Gamma| - \frac{y_0(y_0 - 1)}{2}. \quad (7)$$

Неравенство (1) задаёт ограничение по целевой функции, ограничения (2)–(4) означают, что в каждый период времени общее количество истраченного ресурса не превышает закупленного к этому периоду ресурса; неравенства (5)–(7) задают ограничения по вместимости склада.

При сложении неравенств (1), (2) и (4) получим неравенство $k + |V^\Gamma|y_3 \leq y_0$, которое в силу условия $y_0 \leq |V^\Gamma|$ и целочисленности y_3 может быть верно только при $y_3 = 0$. Также из получившегося неравенства следует, что $k \leq y_0$.

При сложении неравенств (4) и (6) с учётом $y_3 = 0$ получим неравенство $e \geq \frac{y_0(y_0 - 1)}{2}$. С другой стороны, из структуры примеров рассматриваемого класса следует, что $e \leq \frac{k(k - 1)}{2}$. Поэтому имеем неравенство

$$\frac{y_0(y_0 - 1)}{2} \leq \frac{k(k - 1)}{2},$$

откуда после преобразований получаем $(y_0 - k)(y_0 + k - 1) \leq 0$. Заметим, что y_0 – натуральное число, а k – целое, неотрицательное число. Если $k = 0$, то и $e = 0$, в этом случае при сложении неравенств (4) и (6) получим неверное неравенство $\frac{y_0(y_0 - 1)}{2} \leq 0$. Значит, $k \geq 1$. Таким образом, $y_0 + k - 1 \geq 0$, значит, $y_0 - k \leq 0$. Вспоминая, что $k \leq y_0$, заключаем, что $k = y_0$. Ранее было получено, что $\frac{y_0(y_0 - 1)}{2} \leq e \leq \frac{k(k - 1)}{2}$, значит, $e = \frac{y_0(y_0 - 1)}{2}$. Так как e – это количество работ-рёбер, связанных с y_0 работами-вершинами из первого периода времени, то вершины исходного графа Γ , соответствующие работам, выполняющимся в первом периоде времени, вместе с инцидентными рёбрами образуют клику размера y_0 . Теорема доказана. ■

3. Алгоритм динамического программирования

Рассмотрим предлагаемый в работе алгоритм динамического программирования. Для снижения громоздкости описания приведём алгоритм для случая одного ресурса (при этом опустим в обозначениях индекс $r = 1$). Используем идею алгоритма динамического программирования для классической задачи календарного планирования с возобновимыми ресурсами [7]. Для этого разобьём частичный порядок на минимальное количество цепей. В работе [7] эта задача сводится к построению максимального паросочетания в двудольном ориентированном графе (V_1, V_2, E) , в котором множества вершин V_1 и V_2 каждой доли содержат по n вершин, а дуга (i, j) ($i \in V_1, j \in V_2$) принадлежит множеству E , если работа i предшествует работе j . После нахождения в этом графе максимального паросочетания, что является полиномиально разрешимой задачей [8], совмещаем соответствующие вершины двух долей. В получившемся графе с n вершинами имеется дуга (i, j) , если соответствующая дуга входит в паросочетание двудольного графа. Построенный граф состоит из $k = n - U$ цепей, где U – число дуг в паросочетании.

Пусть получилось k цепей. Обозначим через S_l суммарную длительность работ в цепи l . Введём состояния системы $(t, \mathbf{x}, W)$. Здесь вектор $\mathbf{x} = (x_1, x_2, \dots, x_k)$, $x_l \in \{0, 1, \dots, S_l\}$, задаёт состояние выполнения работ проекта, достигнутое в момент t , $t = 0, 1, \dots, T$. При этом $\mathbf{0} = (0, 0, \dots, 0)$ – начальное состояние, когда ни одна работа не начала выполняться, а $\mathbf{S} = (S_1, S_2, \dots, S_k)$ – конечное состояние, когда все работы завершены. Величина $W \in \{0, 1, \dots, V\}$ задаёт остаток ресурса на складе в момент t . Работы каждой цепи выполняются последовательно. Отметим, что в отношении предшествования могут находиться работы разных цепей. Пусть, например, работа j_1 цепи S_1 предшествует работе j_2 цепи S_2 . Если состояние $\mathbf{x}$ определяет, что первая из этих работ ещё не завершена, а вторая уже начала своё выполнение, то в силу заданного на множестве работ частичного порядка состояние $\mathbf{x}$ недопустимо. Чтобы проверить соблюдение таких условий, нужно перебрать все пары работ, находящихся в стадии выполнения, например, используя матрицу смежности графа предшествования. Обозначим через $F(t, \mathbf{x}, W)$ наименьшие суммарные затраты на приобретение ресурсов от начала проекта до состояния $(t, \mathbf{x}, W)$.

Пара (δ, y) задаёт управление. Булевы векторы $\delta = (\delta_1, \delta_2, \dots, \delta_k)$ определяют переход из состояния $\mathbf{x} - \delta$ в состояние $\mathbf{x}$; y – объем ресурса, приобретённого в периоде t . Процесс перехода соответствует одновременному выполнению в периоде t работ тех цепей l , для которых $\delta_l = 1$, а также приобретению ресурса в количестве y . Вектор δ является допустимым, если соблюдаются условия непрерывности выполнения работ и ограничения предшествования между работами разных цепей. Пусть N_δ – множество работ всех цепей, выполняющихся при управлении δ , тогда управление y является допустимым в состоянии $(t, \mathbf{x}, W)$, если

$$\max\{0, W + \sum_{j \in N_\delta} q_j - V\} \leq y \leq W + \sum_{j \in N_\delta} q_j. \quad (8)$$

Пусть $\Delta_{(\mathbf{x}, W)}$ – множество допустимых управлений для состояния $(t, \mathbf{x}, W)$. Выпишем рекуррентное соотношение:

$$F(t, \mathbf{x}, W) = \min_{(\delta, y) \in \Delta_{(\mathbf{x}, W)}} \left(c_t y + F(t-1, \mathbf{x} - \delta, W - y + \sum_{j \in N_\delta} q_j) \right).$$

Полагаем $F(0, \mathbf{0}, 0) = 0$ для начального состояния. В лексикографическом порядке перебирая все состояния, находим оптимальное значение целевой функции $F(T, \mathbf{S}, 0)$. При этом для недопустимых состояний $(0, \mathbf{x}, W)$, где $\mathbf{x}$ отлично от нулевого вектора и/или W не равно нулю, и других недопустимых состояний, вызванных ограничениями предшествования и вместимостью склада, полагаем значение функционала равным $+\infty$. Обратный ход по восстановлению расписания и плана закупок ресурса производится стандартным образом.

Алгоритм обобщается на случай m ресурсов, в этом случае W будет задавать вектор остатков ресурсов на складе по видам. Оценим трудоёмкость алгоритма. Трудоёмкость формирования цепей равна $O(n^{2,5})$ [8]. Количество состояний не превышает $O(T(S_1+1)(S_2+1) \dots (S_k+1)(V_1+1)(V_2+1) \dots (V_m+1))$. Для каждого из этих состояний происходит перебор не более чем 2^k булевых векторов δ . Для каждого из этих булевых векторов необходимо проверить согласованность с частичным порядком, выполнение работ без прерываний, что требует не более $O(k^2)$ операций. Кроме того, для каждого из оставшихся после предыдущих проверок δ нужно перебрать все y , удовлетворяющие неравенству (8). Это с учётом m ресурсов требует $(V_1+1)(V_2+1) \dots (V_m+1)$ операций. Итого, трудоёмкость алгоритма составляет $O(n^{2,5} + 2^k k^2 T(S_1+1)(S_2+1) \dots (S_k+1)(V_1+1)^2(V_2+1)^2 \dots (V_m+1)^2)$.

Так как $\sum_{l=1}^k S_l = P$, где P – сумма длительностей всех работ, то произведение

$(S_1+1)(S_2+1) \dots (S_k+1)$ можно оценить сверху величиной $\left(\frac{P+k}{k}\right)^k$. Таким обра-

зом, предложенный алгоритм динамического программирования для задачи календарного планирования при фиксированных k и m является псевдополиномиальным. Представляется целесообразным его практическое использование для небольших k , т. е. в случае так называемых «узких» графов предшествования. Для решения задач большой размерности перспективной является разработка алгоритмов приближенного решения.

Литература

1. Гимади Э.Х., Залюбовский В.В., Севастьянов С.В. Полиномиальная разрешимость задач календарного планирования со складываемыми ресурсами и директивными сроками // Дискретный анализ и исследование операций. Междунар. конф. Новосибирск: Изд-во Ин-та математики, 2000. Сер. 2. Т. 7, № 1. С. 9–34.
2. Сервах В.В., Щербинина Т.А. О сложности одной задачи календарного планирования со складываемыми ресурсами // Вестник НГУ. Серия: Математика. Механика. Информатика. 2008. Т. 8, вып. 3. С. 105–112.
3. Möhring R.H. Minimizing Costs of Resource Requirements in Project Networks Subject to a Fixed Completion Time // Operations Research. 1984. Vol. 32. P. 89–120.
4. Пирогов А.Ю., Романова А.А. О сложности одной задачи построения расписания с минимальными затратами на приобретение используемых ресурсов // Россия молодая: передовые технологии в промышленность. Омск: Изд-во ОмГТУ, 2015. С. 69–73.
5. Romanova A.A. Minimizing resource cost in project scheduling problem with accumulative resources // Journal of Physics: Conf. Series. 2019. Vol. 1260. P. 082005.
6. Полшкова В.В. Минимизация затрат в задаче календарного планирования с переменной стоимостью ресурсов // Молодёжь третьего тысячелетия: сб. науч. ст. XLVIII регион. студенч. науч.-практ. конф.: в 2 ч. Омск : Издательство ОмГУ, 2024. Ч. 2. С. 401–404.
7. Сервах В.В. Эффективно разрешимый случай задачи календарного планирования с возобновляемыми ресурсами // Дискретный анализ и исследование операций. 2000. Серия 2. Т. 7, № 1. С. 75–82.
8. Papadimitriou C.H., Kenneth S. Combinatorial Optimization: Algorithms and Complexity. New Jersey: Englewood Cliffs, 1982.

THE COMPLEXITY OF PROJECT SCHEDULING PROBLEM WITH ACCUMULATIVE RESOURCES OF TIME-DEPENDENT COST

A.A. Romanova

Ph.D. (Phys.-Math.), Associate Professor, e-mail: RomanovaAA@omsu.ru

Abstract. We consider the Project Scheduling Problem with accumulative resources of time-dependent cost and the case of limited storage. We prove its NP-hardness in a strong sense. Also a dynamic programming algorithm has been developed and pseudopolynomially solvable cases have been identified.

Keywords: schedule, accumulative resource, dynamic programming, polynomial-time reduction.

Дата поступления в редакцию: 01.10.2024

СИСТЕМАТИЗАЦИЯ И СОХРАНЕНИЕ ИСТОРИЧЕСКИХ ИСТОЧНИКОВ В УСЛОВИЯХ ЦИФРОВОЙ ЭПОХИ

Т.Е. Болдовская^{1,2}

к.т.н., доцент, e-mail: boldovskaya73@gmail.com

М.И. Ремизов^{1,3}

младший научный сотрудник, e-mail: remizov.maximus.maxim6@gmail.com

А.Е. Ветров^{1,3}

младший научный сотрудник, e-mail: aleksandrvetrv838@gmail.com

¹Омский государственный технический университет, Омск, Россия

²Омский государственный университет им. Ф.М. Достоевского, Омск, Россия

³Томский государственный университет, Томск, Россия

Аннотация. Статья посвящена проблеме сохранения исторического наследия через цифровизацию письменных источников. Рассматриваются существующие электронные хранилища и платформы, а также процесс разработки информационно-аналитического онлайн-ресурса «Православный ландшафт таёжной Сибири». Основное внимание уделяется унификации данных, созданию многоуровневой системы хранения и проектированию баз данных.

Ключевые слова: исторические источники, база данных, унификация данных, электронные архивы, цифровая гуманитаристика, цифровое наследие, ГИС-технологии, межконфессиональные связи.

Введение

Существует множество источников исторической информации, одним из них являются письменные материалы, содержащие информацию о культурном наследии мировой истории, благодаря им можно восстановить наиболее полную историческую картину прошлого. К сожалению, подлинники исторических документов в процессе естественного старения, ненадлежащего хранения и обращения могут быть утрачены, что несёт огромный урон для культурного наследия. Для борьбы с этой проблемой разрабатываются различные электронные хранилища данных, но их главным недостатком является нестабильность. Программы и платформы сменяются так быстро, что невозможно организовать надёжную сохранность данных, ведь уже через несколько лет платформа будет попросту недоступна из-за несовместимости с новой системой [1]. В свою очередь, оцифрованные экземпляры исторических документов не имеют единого стандарта, возникает риск несоответствия копии оригиналу, что заметно усложняет работу с историческими данными [2]. Информационно-аналитический онлайн-ресурс «Православный ландшафт таёжной Сибири» разрабатывается для решения этих проблем. Текущая цель проекта – отработать применение цифровых технологий на ограниченной территории, чтобы в

дальнейшем их можно было масштабировать и использовать для решения более глобальных задач в области сохранения и исследования исторического наследия. Все исторические данные, загруженные в нашу систему, будут оцифрованы и приведены к единому стандарту, что ускоряет исторические исследования и повышает удобство при работе с данными, ведь пользователь сможет просматривать данные в удобном для него формате, например, в текстовой или в графической форме. Это также позволит сократить использование непосредственно самих письменных экземпляров, ведь теперь появится возможность работать одновременно с множеством источников в одном месте, это способствует более длительному сохранению физических оригиналов в хорошем состоянии. Кроме того, благодаря многоуровневой системе хранения цифровых копий ценных исторических данных обеспечивается их долговременная сохранность и доступность, минимизируя вероятность необратимой потери информации.

1. Примеры других платформ

На сегодняшний день существует множество веб-платформ, которые нацелены на создание цифровых архивов. Например: «Europeana» – проект, основной целью которого является создание цифровой библиотеки, музея и архива Европы. Он был запущен в 2008 г. Европейской комиссией и является коллективным усилием более 3500 учреждений, включая музеи, библиотеки, архивы и галереи из разных стран Европы. Цель проекта «Europeana» – собрать, сохранить и сделать доступными цифровые копии культурных объектов и артефактов, связанных с историей, культурой и наследием Европы [3].

«Mapping the Republic of Letters» – масштабная платформа, которая стремится воссоздать и визуализировать сети общения между интеллектуальными деятелями Европы в период с XVI по XVIII в. Целью проекта является представление исторического контекста, в котором развивались и распространялись идеи, знания и культурные ценности этого времени. Проект основан на обширном анализе письменной переписки, архивных документов и других исторических источников. Исследователи проекта изучают и систематизируют сотни тысяч писем, которые были обменены между известными мыслителями, учёными, писателями и философами того периода. Они анализируют содержание писем, контексты общения и сети связей, чтобы понять, как эти интеллектуальные деятели взаимодействовали друг с другом и как их идеи распространялись по всей Европе. Визуализация этих сетей общения и обмена знаниями является ключевым аспектом проекта «Mapping the Republic of Letters». С использованием географических и социальных данных исследователи создают интерактивные карты и графические представления, которые позволяют исследователям и публике лучше понять масштаб и значимость обмена идеями в этом периоде истории. Эти визуализации помогают выявить основные центры интеллектуальной активности, показать связи между различными деятелями и организациями, а также иллюстрировать потоки идей и знаний [4].

«PRO Сибирь» – онлайн-сервис, который в настоящее время предлагает доступ к более 120 тысячам источников, включая периодические издания, рукописи, книжные памятники, материалы исследовательских экспедиций, архивы ученых и дру-

гие ресурсы, отражающие важность Сибири и Дальнего Востока с точки зрения их территориального, экономического и культурно-исторического значения. Главной особенностью платформы «PRO Сибирь» является её интерактивный характер, который позволяет пользователям принимать активное участие в её развитии благодаря интеграции поисковых, аналитических и коммуникационных сервисов. На платформе пользователи могут выступать в роли исследователей и волонтеров, а также использовать различные функциональные возможности. Они могут распознавать тексты, размещать свои собственные оцифрованные материалы, описывать и тегировать информацию, создавать тематические коллекции, публиковать статьи в блогах, находить единомышленников и предлагать новые идеи. Таким образом, платформа предоставляет возможность активного взаимодействия пользователей, содействуя развитию и обогащению её содержимого [5].

Анализируя работу данных платформ, было принято решение разработать веб-сервис, который представлял бы собой некие исторические данные, экстрагированные из подлинников исторических документов и представленные на современной карте. При этом карта будет иметь поддержку графов, фильтров и слоев для более детального анализа.

2. Проектирование базы данных

Разработка базы данных для исторических исследований представляет собой сложный и многогранный процесс, требующий как технического проектирования, так и глубокого понимания предметной области. Обработка исторических данных связана с определёнными специфическими трудностями. Высокая степень взаимосвязанности объектов усложняет процесс их структурирования, поскольку один и тот же объект может присутствовать в разных исторических контекстах, различающихся во времени, пространстве и тематической направленности. Особое внимание было уделено точной репрезентации межобъектных связей, а также стандартизации терминологического аппарата, что позволило исключить дублирование и гарантировать высокую точность обработки данных. Важнейшей задачей стало создание системы, способной обрабатывать большие объёмы информации и гибко настраиваться под разнообразные исследовательские запросы. В этой связи критически важным стал выбор оптимальной методологии проектирования базы данных, обеспечивающей надёжное хранение информации наряду с удобным доступом для аналитических операций и визуализационных процедур.

Основываясь на проведённом анализе, было принято решение применить нисходящую методологию разработки базы данных ввиду её большей применимости к сложным информационным системам с множеством атрибутов, где идентификация всех функциональных зависимостей является сложной задачей. В противоположность восходящей методике, удобной для создания менее сложных баз данных с ограниченными наборами атрибутов, нисходящая стратегия позволяет глубже проработать семантику предметной области. Основные достоинства нисходящего метода заключаются в минимизации рисков возникновения ошибок в процессе эксплуатации информационной системы, высокой степени формализованности проекторочных процедур и снижении трудовых затрат на нормализацию даталогической

модели до необходимой нормальной формы. Процесс нисходящего проектирования состоит из пяти последовательных стадий [6]:

1. Анализ предметной области. На данном этапе осуществляется всесторонний анализ предметной области с последующей декомпозицией на отдельные элементы. В результате получается описание внешнего уровня базы данных, которая создаётся для пользователей и абстрагируется от особенностей реализации.
2. Разработка информационно-логической модели предметной области. Информационно-логическая модель отражает предметную область в виде совокупности информационных объектов и их структурных связей. Её разработка начинается с построения инфологической модели, затем на её основе строится концептуальная модель данных.
3. Формирование даталогической модели базы данных. Даталогическая модель базы данных отражает логические взаимосвязи между элементами данных, независимо от их содержания и физической организации. Она разрабатывается на основе информационно-логической модели предметной области с учётом конкретной реализации системы управления базой данных и специфики конкретной предметной области.
4. Этап нормализации. Здесь осуществляется процесс нормализации, который предполагает последовательное создание ряда моделей и применение нисходящего подхода к проектированию с минимальными корректировками. В качестве основы для нормализации используется даталогическая модель данных, разработанная на предшествующем этапе.
5. Формирование физической модели базы данных. На этом этапе осуществляется разработка физической модели базы данных, которая включает определение местоположения данных, методов доступа к ним, а также выбор техники индексирования. Данная модель также известна как внутренняя модель системы.

Концептуальное проектирование представляет из себя основу всего процесса проектирования баз данных. Для того чтобы база данных адекватно отражала предметную область, нужно было хорошо представлять себе все особенности, присущие данной предметной области, и уметь отобразить их в базе данных. Поэтому перед началом проектирования базы данных необходимо было детально изучить предоставленную нам историческую информацию, территориально-административное деление российской империи, территориально-административную иерархию Русской православной церкви, чтобы хорошо представлять себе все особенности, присущие данной предметной области. На данном этапе значительное содействие было оказано коллегами из Томского государственного университета. В рамках серии организованных онлайн-конференций были проведены детальные обсуждения и представлены исчерпывающие ответы на возникающие исследовательские и технические вопросы. Инфологическая модель предметной области может описываться как аналитическими, так и графическими средствами. Графическое представление является наиболее наглядным и простым для восприятия и анализа. Кроме того, графическое представление поддерживается системами автоматизации проектирования БД. Поэтому для описания объектов и их связей мы воспользовались графиче-

ским способом отображения, позволяющим построить модель, условно называемую «объект – свойство – связь». При исследовании и анализе предметной области были выделены следующие классы объектов: населённый пункт, церковь, образование, причт, население, социальные конфликты, религиозные девиации. Были отражены связи между объектом и характеризующими его свойствами. Представленное изображение на рис. 1 иллюстрирует результат установления взаимосвязи «объект – свойство», сформированной в ходе проведённого исследования. Связь сущности и единичного свойства обозначается одинарной стрелкой, а связь между сущностью и множественным свойством обозначается двойной стрелкой. Также постоянные свойства обозначены непрерывной линией. Если свойство присутствует у всех объектов данного класса, оно обозначается буквой «S».



Рис. 1. Пример изображения получившейся в результате работы связи «объект – свойство»

Помимо связей между объектами и их свойствами, были зафиксированы связи между объектами разных классов (рис. 2). Двойная стрелка говорит о том, что в населённом пункте может проживать несколько объектов населения.

В ходе работы над даталогической моделью выяснилось, что Postgres не предоставляет функциональности для работы с ГИС-данными, поэтому необходимо было установить расширение PostGIS, которое является мощным инструментом для хранения, обработки и анализа пространственных данных, без которого было не обойтись в данном случае. В результате даталогического проектирования было выделено несколько сущностей:

1. Образование. Содержит сведения об учебных заведениях, включая координаты, название, принадлежность к приходу, вид учебного заведения, численность учащихся по полу, количество неправославных учеников, а также подробную информацию об учителях (ФИО, дата рождения, пол, сословие, образование, место получения образования, годовое содержание).

2. Церковь. Детализирует храмы и церкви, включая их названия, приходы, даты просьбы о постройке, закладки, освящения, типы проектов, материальные аспекты строительства и содержания (средства прихожан, частные пожертвования, государственные субсидии), страховые данные (сумму и срок страховки), материалы сооружений, библиографические ссылки, координаты, наличие приделов (наименование, материал, сумма ремонта, характер ремонта), библиотек (количество томов в библиотеке, источники пополнения библиотеки), культовых сооружений, обществ и организаций, корреспондентов «Томских епархиальных ведомостей», а также информацию о наличии попечительства и расходах за год и суммах на конец отчётного периода.

3. Населённый пункт. Хранит данные о населённых пунктах, такие как название, дата основания, причина основания, географические координаты, административная принадлежность (уезд, волость, приход, благочиние), роль населённого пункта

(волостной центр, приходский центр, центр благочиния), тип населённого пункта, количество дворов, площадь земельных участков, удалённость от ключевых объектов инфраструктуры (уездный центр, благочинный, транспортные пути, реки), инфраструктура и окружающая среда.

4. Причт. Содержит информацию о священнослужителях, включающую приход, сан, ФИО, дату рождения, вид и размер содержания, этапы карьеры, награды и назначения.

5. Население. Предоставляет демографические данные населения, включая распределение по полу, этнической принадлежности, вероисповеданиям и направлениям.

6. Религиозные девиации. Собирает сведения о религиозных отклонениях, включая вовлечённые конфессии, виды действий, местоположения событий и библиографические ссылки.

7. Социальные конфликты. Документирует социальные конфликты, их предметы, внутренние и внешние конфликтующие стороны, координаты событий и библиографические ссылки.

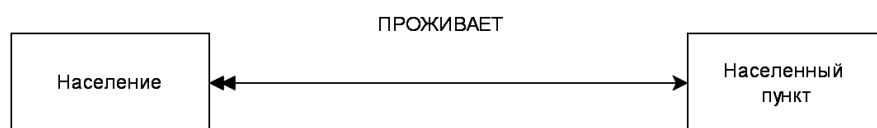


Рис. 2. Пример изображения получившейся в результате работы связи «объект – связь»

В результате проведённых исследований была разработана концептуальная схема базы данных для информационно-аналитической веб-платформы «Православный ландшафт таёжной Сибири», включающая более 40 взаимосвязанных таблиц. Полученная схема демонстрирует сложную систему иерархических связей между разнородными элементами различных тематических областей через центральную таблицу «Населённые пункты» и предоставляет возможность углублённого исследования множества сфер жизни конкретного региона. Такой подход обусловлен тесной корреляцией исторических данных с конкретными географическими локациями, включая городские поселения, сельские общины и прочие типы населённых пунктов.

Отдельно на рис. 3 выделен фрагмент базы данных, описывающий населённый пункт и его отношение к волости, уезду, приходу и благочинию. Дополнительно представлено отношение, обеспечивающее детализацию характеристик населённого пункта в различные периоды его существования. В этом контексте также включены таблицы, отражающие проведённые инспекционные мероприятия в населённых пунктах, включая отдельную таблицу для идентификации лица, возглавляющего проверку, с указанием его должности, причины проведения проверки и типа принятого воздействия по её результатам. В результате все атрибуты отношений являются простыми, а используемые домены содержат только скалярные величины. Каждый ключевой атрибут демонстрирует функциональную зависимость от первичного ключа, исключая возможность декомпозиции. Все неключевые атрибуты

характеризуются отсутствием транзитивной зависимости от первичного ключа. Исходя из вышесказанного, можно заключить, что структура базы данных соответствует требованиям третьей нормальной формы.

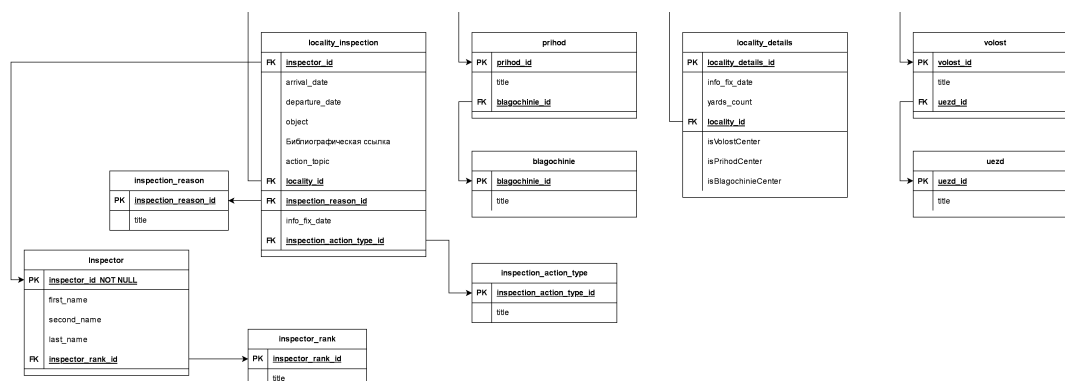


Рис. 3. Фрагмент базы данных информационно-аналитического онлайн-ресурса «Православный ландшафт таёжной Сибири», описывающей населённый пункт

Таким образом, на данном этапе проекта была создана нормализованная база данных, отвечающая требованиям современных информационных хранилищ. Особое внимание было уделено учёту сложных межобъектных связей внутри исторической инфраструктуры. Были идентифицированы основные параметры структурирования данных, что обусловило необходимость модификации классических методик проектирования.

Заключение

В результате проведённых исследований была разработана реляционная база данных, представляющая собой фундаментальную компоненту разрабатываемой веб-платформы, ориентированной на систематизацию, хранение и доступ к историческим данным. Следует подчеркнуть, что процесс проектирования осуществлялся в тесном сотрудничестве с научным коллективом Томского государственного университета, что обеспечило учёт сложных взаимосвязей, характерных для данной предметной области. Наличие большого числа количественных и качественных показателей позволит проводить статистический анализ, а обширные библиографические ссылки делают возможным углублённое исследование исторических источников.

На текущий момент проект находится в фазе активной реализации, включая создание пользовательского графического интерфейса и расширение функционала платформы для проведения аналитических операций. Промежуточные результаты формируют основу для последующего развития проекта и демонстрируют потенциал использования передовых информационных технологий в сохранении и исследовании исторического наследия. Достигнутые результаты служат фундаментом для дальнейшего совершенствования платформы и её применения в историко-культурных исследованиях.

Благодарности

Данная работа была выполнена при поддержке гранта № 23-78-10119, выделенного Российским научным фондом.

Литература

1. Антопольский А.Б. Информационные ресурсы России. Москва: Либерия, 2004. 423 с.
2. Юмашева Ю.Ю. Научное издание исторических документов в электронной среде: проблемы источниковедения и археографии // Письменное наследие и информационные технологии (El'Manuscript-2016) / материалы Междунар. конф. Вильнюс, 2016. С. 334–339.
3. Europeana: исторический информационный ресурс. URL: <https://www.europeana.eu/en> (дата обращения: 20.02.2025).
4. Mapping the Republic of Letters. URL: <http://republicofletters.stanford.edu/> (дата обращения: 20.02.2025).
5. PRO Сибирь: исторический информационный ресурс. URL: <https://prosiberia.tsu.ru/> (дата обращения: 20.02.2025).
6. Илюшечкин В.М. Основы использования и проектирования баз данных: учеб. для вузов. М.: Юрайт, 2024. 213 с.

SYSTEMATIZATION AND PRESERVATION OF HISTORICAL SOURCES IN THE DIGITAL ERA

T.E. Boldovskaya^{1,2}

Ph.D. (Techn.), Associate Professor, e-mail: boldovskaya73@gmail.com

M.I. Remizov^{1,3}

Junior Researcher, e-mail: remizov.maximus.maxim6@gmail.com

A.E. Vetrov^{1,3}

Junior Researcher, e-mail: aleksandrvetrv838@gmail.com

¹Omsk State Technical University, Omsk, Russia

²Dostoevsky Omsk State University, Omsk, Russia

³Tomsk State University, Tomsk, Russia

Abstract. The article is devoted to the problem of preserving historical heritage through the digitalization of written sources. It examines existing electronic repositories and platforms, as well as the process of developing the information-analytical online resource "Orthodox Landscape of Taiga Siberia." The focus is on data unification, the creation of a multi-level storage system, and database design.

Keywords: historical sources, database, data unification, electronic archives, digital humanities, digital heritage, GIS technologies, interfaith relations.

Дата поступления в редакцию: 24.02.2025

НАУЧНО-ТЕХНИЧЕСКОЕ ИССЛЕДОВАНИЕ ПРОГРАММНОЙ СОСТАВЛЯЮЩЕЙ БЕСПИЛОТНЫХ АВИАЦИОННЫХ СИСТЕМ

С.В. Гусс

старший преподаватель, e-mail: sviat@v-guss.ru

Омский государственный университет им. Ф.М. Достоевского, Омск, Россия

Аннотация. Рассматривается идея создания лаборатории научно-технического анализа программной составляющей беспилотных авиационных систем. Представлены особенности приобщения молодёжи к инженерным и научным исследованиям. Даются схемы, обобщающие деятельность в рамках исследовательских работ.

Ключевые слова: летающая робототехника, БПЛА, БАС, программотехника.

Введение

Одна из основных повесток на ближайшие годы в области цифровых технологий – повышение спроса на инженерные кадры. Это вызвано необходимостью разработки и внедрения новых технологий для ускорения цифровой трансформации. Главные темы: интернет вещей, искусственный интеллект, машинное обучение. Особое место занимает робототехника и её специализированная область – летающая робототехника, которая имеет дело с беспилотными авиационными системами (БАС) и включает в себя не только беспилотные летательные аппараты (БПЛА), но и наземные станции управления, системы связи и программное обеспечение (ПО) для управления полётами и координации действий БПЛА. К разработке сложного ПО необходимо подходить комплексно [1], используя инженерную методологию [2]. В работе будет представлен проект исследовательской лаборатории научно-технического анализа программной составляющей БАС. Это, в свою очередь, можно рассматривать как шаг на пути к открытию реальной, физической лаборатории.

1. Предпосылки

В конце ноября автор статьи защитил итоговый проект в Университете Иннополис в рамках курса «Разработка программного обеспечения для полётного контроллера БАС», проходящего в рамках федерального проекта «Кадры для БАС», что ещё больше мотивировало автора на создание и развитие научно-исследовательской лаборатории, связанной с техническими исследованиями (анализ, разработка, проектирование, тестирование, обобщение, систематизация, дистилляция в контексте

DDD и т. д.) программного обеспечения для сферы БАС. Был представлен проект «БАС для тестирования и отладки алгоритмов в лаборатории. Создание программного обеспечения для полётного контроллера БАС в сфере образования». В проекте была описана деятельность лаборатории с позиции научного руководителя. Основная задача лаборатории – задать основу для систематизации и повторного использования наработок в виде знаний, конструкций, алгоритмов, исходных кодов и программ. На повестке стоял вопрос об актуальности, особенностях и необходимости такой лаборатории. Представлена общая характеристика работ, проводимых в лаборатории, связь научного исследования с программной разработкой, создание инженерных артефактов, переход от проектных к производственным решениям. К концу был представлен план дальнейшего развития проекта, речь шла о разработке и тестировании алгоритмов управления и коммуникации устройств, сборе и накоплении предметных знаний, полезных данных статистики и о расширении области применения программных решений.

На рис. 1 представлен фрагмент презентации проекта, где отражены главные (но не единственные) технологии и методы разработки.

Главные технологии и методы разработки

- ❑ **Имитационное и полунатурное моделирование:** использование стендов для изучения поведения летательного аппарата и системы управления в различных режимах.
- ❑ **Математическое моделирование:** применение методов теории управления и вычислительной математики для анализа и оптимизации алгоритмов управления.
- ❑ **Создание программно-аппаратной платформы моделирования:** разработка системы для тестирования алгоритмов пилотирования и визуализации полёта.
- ❑ **Обеспечение безопасности и предсказуемости** поведения систем автоматического управления: соответствие требованиям к управлению на всех этапах полёта.



Рис. 1. Технологии и методы разработки

В целях информационной поддержки лаборатории, освещения её деятельности и отслеживания актуального состояния сферы БАС был создан онлайн-курс «Научно-техническое исследование программной составляющей БАС» [3]. В основе курса лежит важная идея, навеянная университетской инициативой приобщения молодежи к науке (в рамках ОмГУ им. Ф.М. Достоевского), и желание потрудиться на этом поле, сделать что-то полезное, поделиться знанием и опытом.

2. Лаборатория научно-технического анализа программной составляющей БАС

Приобщение к научно-технической деятельности в рамках лаборатории происходит на примере такого объекта исследования, как архитектура и программирование вычислительной системы полётного контроллера БАС. Рассматриваются методология и принципы научного исследования, этапы, подготовка к эксперименту, составление и оформление технических отчётов, распространение и обсуждение результатов научного труда.

В процессе знакомства с методологией научно-технического исследования разбираются научные статьи, накапливаются полезные материалы о предметной области, развивается своё, индивидуальное видение проблем, возможных путей их решения, готовится почва для собственного исследования студентов.

Цель – достичь понимания того, куда и зачем двигаться в своих исследованиях в процессе работы над курсовой или дипломной работой (или проектом), магистерской или кандидатской диссертацией.

Объект исследования – программная составляющая БАС. А вот предмет исследования у каждого свой. Также, как цель, задачи, актуальность и планируемые результаты исследования, которые вполне могут стать научным открытием. Все эти моменты и «обкатываются» в рамках лаборатории.

Для кого это может быть интересно

Для школьников (возможно), студентов (наверное), магистрантов (скорее всего), аспирантов (желательно) и всех, кто желает приобщиться к научным изысканиям в области программотехники (программирование, разработка программного обеспечения, программная инженерия) и системотехники (конструкция, аппаратные особенности, микропроцессорная начинка) БАС.

Начальные требования

Интерес, живой и неподдельный, к прикладной, технической, инженерной науке, направленной на расширение знаний в сфере беспилотной авиации (управление полётом; программирование автономного следования по заданному пути, сбора и обработки информации, преодоления препятствий и т. д.).

3. Описание деятельности в рамках лаборатории

Направления (согласно ПИШ – Передовые Инженерные школы):

- Авиационная техника.
- Цифровые технологии.
- Программная инженерия.

Далее в общем виде представлены ключевые моменты деятельности лаборатории, которые в дальнейшем, в процессе работы лаборатории, необходимо будет детализировать.

Исследование предметной области

Конечный результат научного исследования, имеющего место в процессе разработки программного обеспечения (в рамках инженерной дисциплины), – предметно-ориентированные проектные решения для конкретной предметной области (т. е. «области знаний или деятельности, в которой пользователь использует программное обеспечение»). К ним относят различные технические и инженерные артефакты (т. е. промежуточные продукты проекта разработки), такие как описания концепций, модели (систематизированные описания), шаблоны (документов, интерфейсов (пользователя, модулей системы), анализа, дизайна архитектуры и др.), компоненты (подсистемы, классы, функции) и т. д., которые можно повторно использовать в разных программных проектах на разных этапах жизненного цикла продукта (уровнях абстракции) (см. рис. 2).

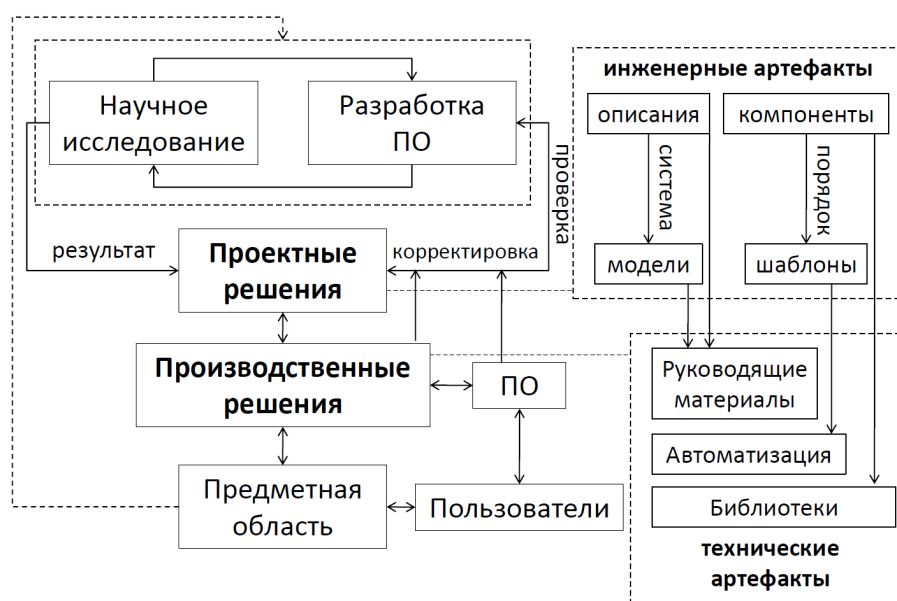


Рис. 2. Исследование предметной области

Подготовка проектных решений

Центральная идея предметно-ориентированного проектирования – накопление опыта создания программных средств для конкретной предметной области в виде повторно применимых элементов ПО (модулей, библиотек компонентов, каркасов, инструментов, специальных языков, профилей настройки инструментов разработки, сценариев, схем и документации для различных этапов процесса разработки) (см. рис. 3).

Архитектурное моделирование

В предметно-ориентированном проектировании под архитектурной средой (framework) понимают одну из форм программной инфраструктуры, реализующей

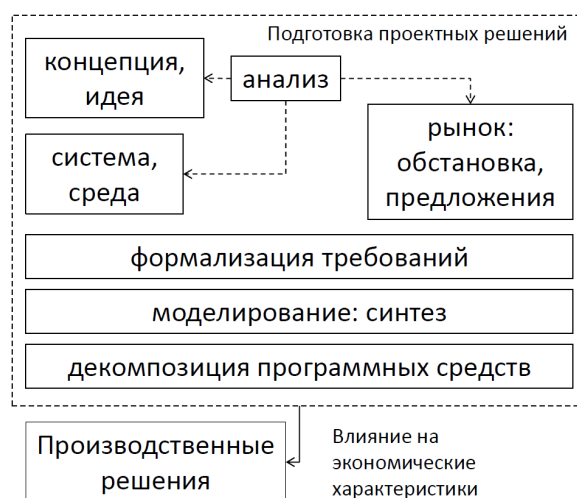


Рис. 3. Подготовка проектных решений

нужные функции. Она навязывает другим уровням архитектуры приложения определённые способы реализации, решает технические задачи и позволяет разработчику уровня предметной области (собственно ключевых функциональных компонентов системы) сосредоточиться на решении проблемы из предметной области, не отвлекаясь на второстепенные детали. Архитектурная среда – это уже сделанный «выбор проектных решений в предметной области» (см. рис. 4).

В рамках объектно-ориентированного проектирования тоже используется такое понятие, как «архитектурная среда», однако исторически сложилось, что в русскоязычной литературе она обозначена таким термином, как «каркас» (в английском это всё тот же framework). «Каркас – набор взаимодействующих классов, составляющих повторно используемый дизайн для конкретного класса программ» [4].

В программной инженерии можно встретить термин «стиль» (тоже framework в оригинале). В стандарте IEEE 1320.2 «Standard for Conceptual Modeling Language» отмечается, что «framework представляет собой проект (модель или код), предназначенный для многократного повторного использования, который может быть усовершенствован (специализирован) и расширен и на основе которого может быть реализована некоторая часть полных функциональных возможностей для многих приложений» [5].

Также используется словосочетание «архитектурная модель», под этим будет пониматься нечто, соответствующее двум следующим классическим (устоявшимся) определениям: «архитектура – концепция взаимосвязи элементов сложной структуры, включающая компоненты логической, физической и программной структур» [6]; «модель – материальный объект, система зависимостей или программа, адекватная объекту» [6]. Архитектурная модель будет выражаться логическими схемами, программной реализацией, физическими характеристиками и границами среды выполнения, а также поясняться соответствующими диаграммами.

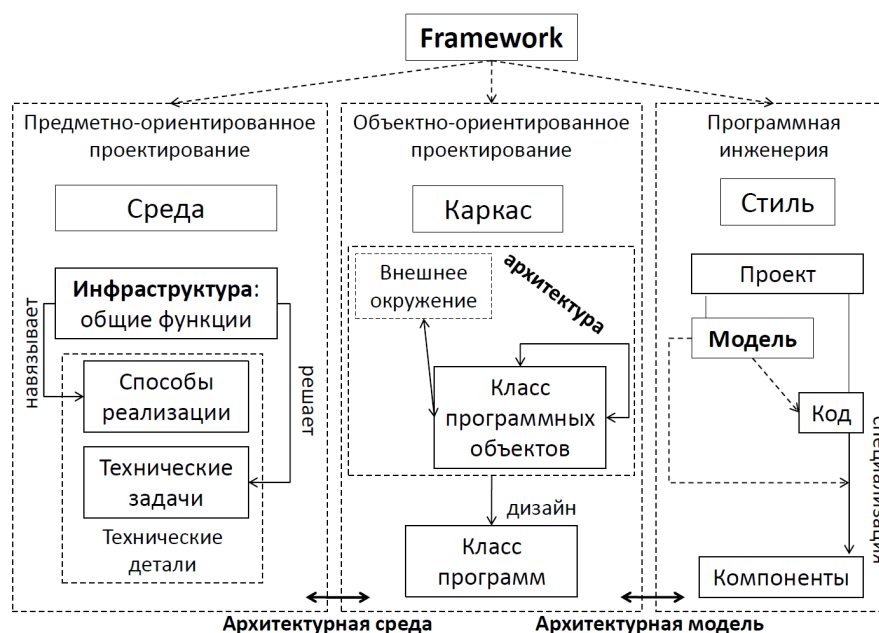


Рис. 4. Архитектурное моделирование

Уточнение знаний

Каждая предметная область имеет свои характерные особенности, которые определяют свойства проектов и форму соответствующих проектных решений. Нельзя сказать заранее, какие средства нужны для решения той или иной задачи из предметной области. Необходимо провести исследовательские работы (с постепенной переработкой знаний и уточнением), разработать ряд экспериментальных проектов, проанализировать и обобщить полученные артефакты, найти смысловое ядро, суть, объединяющую приложения из предметной области (см. рис. 5).

Совершение открытий

Работа в рамках научного исследования предполагает целенаправленное размышление над объектом исследования. Гипотеза Сепира – Уорфа, также известная как гипотеза лингвистической относительности, предполагает, что структура языка определяет мышление и способ познания реальности. Согласно этой гипотезе, люди, говорящие на разных языках, по-разному воспринимают мир и мыслят. Тот факт, что мы говорим на одном и том же государственном языке, не означает что мы используем этот язык одинаково. Да, у нас может быть один общий синтаксис и одна грамматика, но семантика далеко не всегда одна. Из-за того, что мы чувствуем мир немного по-разному, по своему его переживаем, смысл, который мы вкладываем в слова и их сочетания, может различаться. От субъективности в таком случае уйти сложно. Но наука, насколько мы знаем, должна стремиться к объективности (см. рис. 6).

Методология научного исследования предлагает схемы, которые пусть и не гарантируют, но способствуют выработке объективного знания.

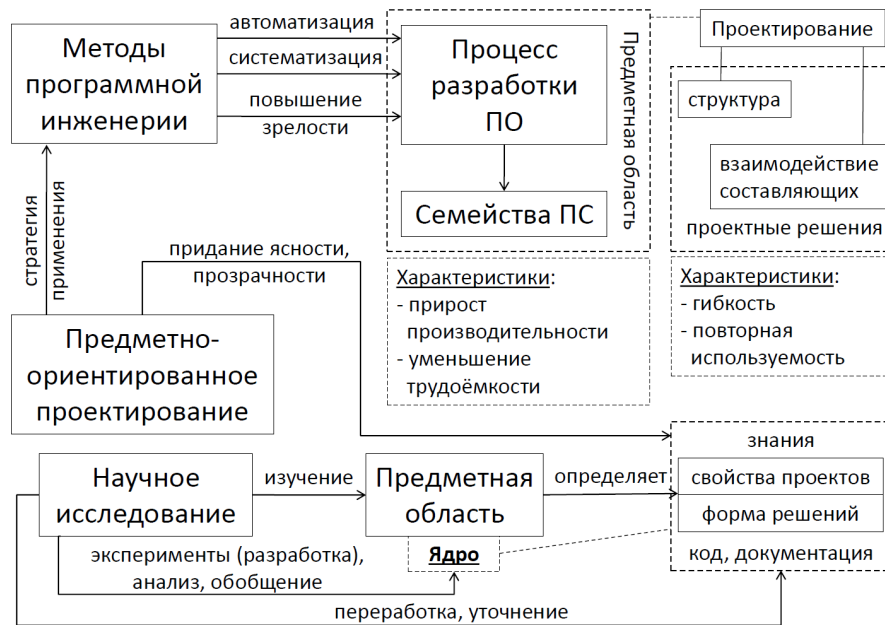


Рис. 5. Уточнение знаний



Рис. 6. Совершение открытий

Нам понадобится программа научного исследования:

- Ограничение предметной области.
- Фокусировка познавательной деятельности.
- Определение проблемы, гипотезы, объекта, предмета.

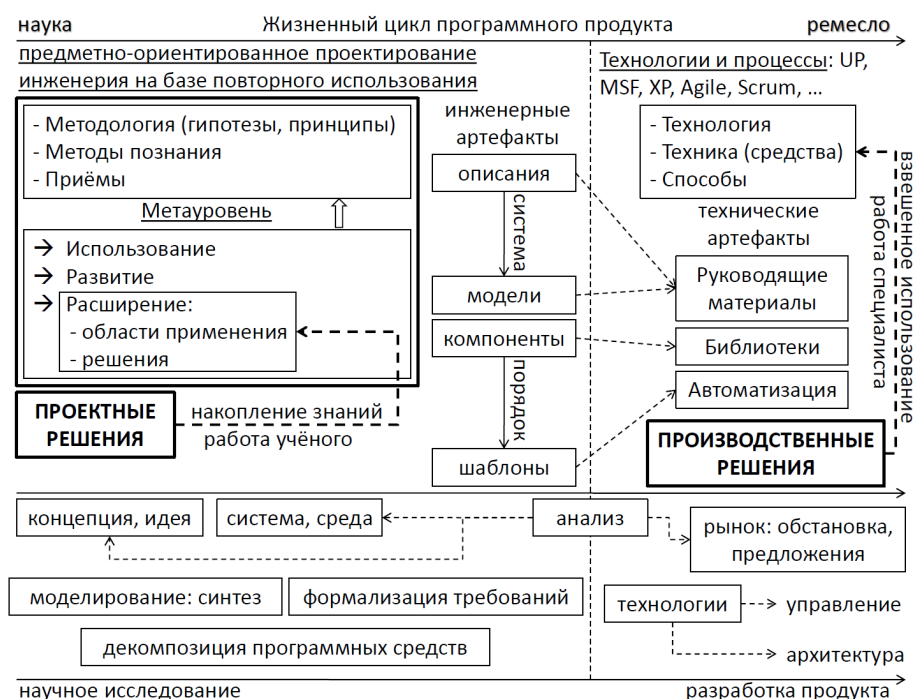


Рис. 7. Переход к производственным решениям

- Проекция объекта. Предположения о сущности объекта.
- Выделение характеристик, качеств, свойств объекта.
- Анализ динамики объекта, как системы, её развития и преобразований.
- Детализация объекта и его содержательно-смыслового представления.
- Типизация объекта. Описание принципов и закономерностей.
- Концептуальное и логическое моделирование.
- Дистилляция знаний. Отсеивание второстепенных деталей. Определение предмета. Выход к смысловому ядру исследования.

Программа исследования – это вопросы, которые стоит внимательно проработать. Помогает определиться с собственной, авторской программой исследования. Это отразится на характере исследования и поможет окончательно сформулировать цель и задачи исследования, выделить предмет и выбрать методы исследования. Работая в рамках программы, мы должны прийти к значимым результатам исследования, обладающим научной новизной.

Переход к производственным решениям

Следующая схема (рис. 7) условно намекает нам на то, что наука – это то, что левее, а производство – это то, что правее. А программотехника – это вся та концептуальная среда, тот контекст, в котором происходит всё это. «Концептуальная среда – это место существования изучаемого многомерного объекта, в котором он зарождается, получает развитие и наполняется новым фактическим материалом» [7]. Переход от абстрактного к конкретному, от теории к практике и обратная связь, повторение, корректировка и обобщение.

Заключение

В статье был рассмотрен проект лаборатории научно-технического анализа программной составляющей БАС. Предложены схемы, позволяющие составить некоторое представление о деятельности в рамках лаборатории. Впереди ещё много работы по детализации деятельности, внедрению идеи и привлечению молодых исследователей.

Литература

1. Гусс С.В. Разработка семейства программных систем в специфической предметной области // Математические структуры и моделирование. 2011. Вып. 22. С. 55–68.
2. Гусс С.В. Пакет вспомогательных средств для построения семейства программных систем в определённой предметной области // Программная инженерия. 2011. № 1. С. 25–33.
3. Научно-техническое исследование программной составляющей БАС. Stepik. URL: <https://stepik.org/course/216056> (дата обращения: 24.11.2024).
4. Приёмы объектно-ориентированного проектирования. Паттерны проектирования / Э. Гамма [и др.]. СПб.: Питер, 2008. 366 с.
5. Батоврин В.К. Системная и программная инженерия. Словарь-справочник. М.: ДМК Пресс, 2010. 280 с.
6. Першиков В.И., Савинков В.М. Толковый словарь по информатике. М.: Финансы и статистика, 1991. 543 с.
7. Синельникова Л.Н. Концептуальная среда фронтального дискурса в гуманитарных науках // Russian Journal of Linguistics. 2020. Т. 24, № 2. С. 467–492.

SCIENTIFIC AND TECHNICAL RESEARCH OF THE SOFTWARE COMPONENT OF UNMANNED AIRCRAFT SYSTEMS

S.V. Guss

Assistant Professor, e-mail: sviat@v-guss.ru

Dostoevsky Omsk State University, Omsk, Russia

Abstract. The idea of creating a laboratory for scientific and technical analysis of the software component of unmanned aircraft systems is considered. The features of involving young people in engineering and scientific research are presented. Schemes are given that summarize the activities within the framework of research work.

Keywords: flying robotics, UAV, UAS, software engineering.

Дата поступления в редакцию: 24.11.2024

ИСПОЛЬЗОВАНИЕ НЕЙРОННЫХ СЕТЕЙ ДЛЯ СОЗДАНИЯ СИСТЕМЫ ПРИНЯТИЯ РЕШЕНИЙ УПРАВЛЕНИЯ МИКРОКЛИМАТОМ ТЕПЛИЧНОЙ СИСТЕМЫ

Т.В. Костеннов

аспирант, e-mail: timofey.kostenov@gmail.com

Омский государственный университет им. Ф.М. Достоевского, Омск, Россия

Аннотация. Описывается механизм использования нейронных сетей для прогнозирования показателей микроклимата в системе принятия решений тепличной системы. Рассматривается использование нескольких рекуррентных нейронных сетей внутри системы принятия решений. Внимание уделяется правильному определению места системы принятия решений в функциональной схеме тепличной системы.

Ключевые слова: тепличная система, системный анализ, прогнозирование, микроклимат, нейронная сеть, система принятия решений.

Введение

Продовольственная безопасность является одной из актуальных проблем государства, большие усилия агропромышленного комплекса направлены на её обеспечение [1]. В рамках решения данной проблемы большую значимость получают методы выращивания агрокультур в регионах с холодным климатом. Подобные методы, кроме очевидных результатов в виде повышения совокупного урожая, позволяют, например, уменьшить затраты на транспортировку продукции, что, в свою очередь, сделает её более доступной для потребителя. Одним из таких методов является метод выращивания растений в закрытых грунтах, широко применяемый тепличными хозяйствами по всему миру. Применение теплиц в промышленности позволяет максимально использовать биологический потенциал растений благодаря возможности управления многими факторами роста. Доля теплиц в сельском хозяйстве растёт, и по прогнозам специалистов данная тенденция будет сохраняться в ближайшие годы [2].

Общий курс на автоматизацию и уменьшение участия человека в базовых операциях находит всё больше поддержки. Одним из применений автоматизации в области тепличных хозяйств являются управляемые алгоритмами теплицы. Для таких тепличных систем, в которых получилось добиться высокого уровня автоматизации, участие человека сводится к минимуму, но не исчезает полностью. Человек принимает основные решения по функционированию тепличной системы. Для того чтобы минимизировать человеческий фактор в принятии таких решений, можно использовать системы принятия решений.

В рамках данной статьи рассматривается теоретическая модель системы принятия решений на основе нескольких рекуррентных нейронных сетей (RNN). Используя различные данные о функционировании тепличной системы и её микроклимате система принятия решений формирует рекомендации по основным решениям.

Для создания подобной системы принятия решений можно рассмотреть тепличную систему с точки зрения системного анализа. Это поможет унифицировать и структурировать моделируемую систему.

1. Сбор информации о среде

Рассмотрим простейшую тепличную систему, использующую следующие датчики:

- датчик освещённости внутри тепличной системы;
- датчик температуры воздуха внутри тепличной системы;
- датчик влажности воздуха внутри тепличной системы;
- датчик температуры воздуха во внешней среде;
- датчик влажности воздуха во внешней среде.

И следующие устройства регуляции среды:

- нагреватель воздуха;
- охладитель воздуха;
- увлажнитель воздуха;
- осушитель воздуха;
- устройства полива;
- устройства освещения.

На основе информации о температуре и влажности внутри тепличной системы можно вычислить меру насыщенности воздуха влагой или дефицит давления водяного пара (VPD, vapour pressure deficit). Эта мера характеризует разницу между давлением насыщенного пара (максимальным количеством влаги в воздухе) и фактическим содержанием пара в воздухе (текущего количества влаги в воздухе). Этот параметр находит всё большее применение в современном растениеводстве и экологических исследованиях [3].

С помощью аналитических методов, таких как моделирование данных, можно предсказать значения температуры и влажности. Для этого необходимо подготовить данные для реализации и обучения модели.

2. Принцип работы системы принятия решений

Для реализации моделей будет использована RNN – вид нейронных сетей, где связи между элементами образуют направленную последовательность. Использование RNN находит всё большее применение в системах интернета вещей из-за высоких темпов роста объёмов информации [4]. Однако при использовании обычных RNN можно столкнуться с проблемой исчезающих градиентов – результаты вычислений могут не учитывать долгосрочные зависимости в данных. Это происходит из-за того, что градиенты при обратном распространении экспоненциально затухают или растут при обновлении весов. Для решения проблемы долгосрочной памяти

можно использовать различные архитектуры RNN. Рассмотрим более подробно две из них: Long Short-Term Memory (LSTM) и Gated Recurrent Units (GRU).

2.1. LSTM

LSTM (долгая краткосрочная память) – архитектура RNN, предложенная в 1997 г. Ж. Хохрайтером и Ю. Шмидхубером (см. рис. 1). Основным нововведением архитектуры являлась невосприимчивость к длительности временных разрывов, что давало LSTM преимущество по отношению к обычным RNN. Достигалось вышеописанное с помощью ячейки состояния (cell state) и трёх «ворот», управляющих информацией.

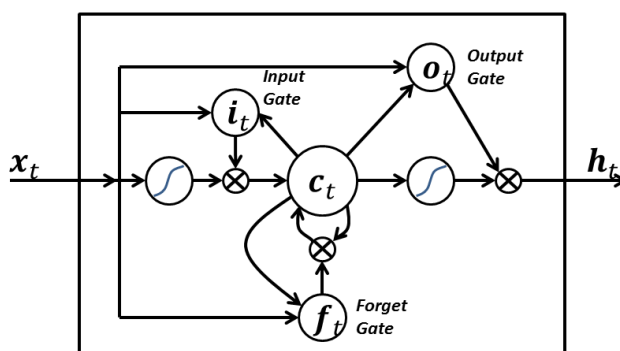


Рис. 1. Архитектура LSTM

Ключевые компоненты LSTM можно выделить следующим образом:

- Cell State – передаёт информацию через время почти без изменений, позволяет сохранять долгосрочные зависимости: C_t .
- Forget Gate – решает, какую информацию удалить из ячейки, вектор забвения: f_t .
- Input Gate – определяет, какую новую информацию добавить в ячейку, вектор входного вентиля: i_t .
- Output Gate – решает, какую информацию передать на следующий шаг, вектор выходного вентиля, кандидат на выход: o_t .

Благодаря своей архитектуре LSTM эффективно работают с длинными последовательностями, например, текстом, временными рядами, распознаванием речи, генерации музыкальных композиций, распознавании рукописного ввода.

2.2. GRU

GRU (управляемые рекуррентные блоки), представленные в 2014 г., – это упрощённая версия LSTM. Они объединяют некоторые компоненты для уменьшения вычислительной сложности. Ключевые компоненты GRU можно выделить следующим образом:

- Update Gate – определяет, какая информация сохранится из предыдущего состояния: z_t .

- Reset Gate – решает, какую часть прошлой информации игнорировать: r_t .
- Скрытое состояние: h_t .

Основным отличием от LSTM стало отсутствие ячейки состояния (C_t), осталось только скрытое состояние (h_t). Input Gate и Forget Gate объединены в Update Gate. Отсутствует Output Gate, выход формируется напрямую из обновлённого состояния. К основным преимуществам GRU относится меньшее число входных параметров, а значит более быстрое обучение.

Рассмотренные архитектуры решают проблему долгосрочных зависимостей, но выбор зависит от задачи и вычислительных ресурсов. LSTM мощнее для сложных последовательностей, в то время как GRU – эффективный компромисс между скоростью и точностью. В дальнейшем предлагается использовать GRU как более быстрый алгоритм. Для упрощения понимания внутренняя реализация GRU не будет рассматриваться совместно с входными параметрами. Обе архитектуры могут быть использованы в дальнейших предложенных схемах.

2.3. Входные и выходные данные

Рассмотрим параметры, участвующие в моделировании нейронной сети № 1:

- дата и время;
- температура воздуха внутри тепличной системы;
- влажность воздуха внутри тепличной системы;
- температура воздуха снаружи тепличной системы;
- влажность воздуха снаружи тепличной системы.

В результате набор данных содержит пять характеристик, содержит данные за две недели, с шагом в 10 минут. Таким образом, задача сводится к предсказанию временных рядов, для чего RNN часто используются и хорошо себя зарекомендовали [5].

Для обучения моделей данные следует разделить на несколько частей: 70 % – обучающие, 20 % – проверочные, 10 % – тестовые.

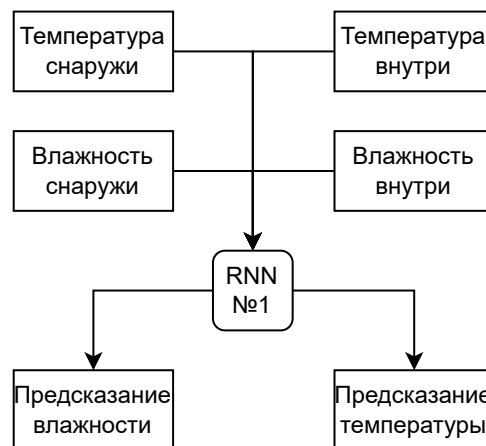


Рис. 2. Схема параметров нейронной сети № 1

Схема входных и выходных параметров нейронной сети № 1 представлена на рис. 2.

После обучения модель будет способна предсказывать влажность и температуру внутри тепличной системы. Эти данные предполагается использовать для других моделей с целью предсказания других параметров.

Рассмотрим параметры, участвующие в моделировании нейронной сети № 2:

- дата и время;
- температура воздуха внутри тепличной системы;
- влажность воздуха внутри тепличной системы;
- предсказание температуры воздуха внутри тепличной системы;
- предсказание влажности воздуха внутри тепличной системы.

В результате набор данных содержит пять характеристик, содержит данные за две недели, с шагом в 10 минут. Для обучения моделей данные следует разделить на несколько частей: 60 % – обучающие, 30 % – проверочные, 10 % – тестовые. Большой процент данных, отведённых на проверочную выборку, обусловлен использованием результатов работы другой нейронной сети.

Схема входных и выходных параметров нейронной сети № 2 представлена на рис. 3.

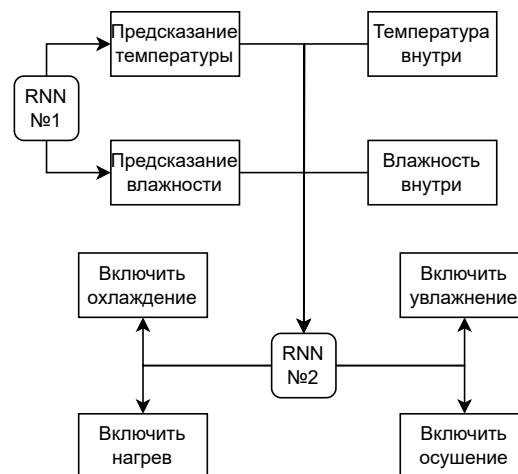


Рис. 3. Схема параметров нейронной сети № 2

После обучения модель будет способна выдавать рекомендации о включении средств регуляции среды. Стоит отметить, что при включении одних средств регуляции среды противоположные им отключаются. Например, при включении охлаждения будет автоматически отключён нагрев. Эти данные предполагается использовать для модели более высокого уровня с целью помощи в принятии решений.

Рассмотрим параметры, участвующие в моделировании нейронной сети № 3:

- дата и время;
- температура воздуха внутри тепличной системы;
- влажность воздуха внутри тепличной системы;
- предсказание температуры воздуха внутри тепличной системы;

- предсказание влажности воздуха внутри тепличной системы;
- дефицит давления пара;
- мощность, подаваемая на устройство освещения.

В результате набор данных содержит семь характеристик, содержит данные за две недели, с шагом в 10 минут. Для обучения моделей данные следует разделить на несколько частей: 60 % – обучающие, 30 % – проверочные, 10 % – тестовые. Большой процент данных, отведённых на проверочную выборку, обусловлен использованием результатов работы другой нейросети.

Схема входных и выходных параметров нейронной сети № 3 в представлена на рис. 4.

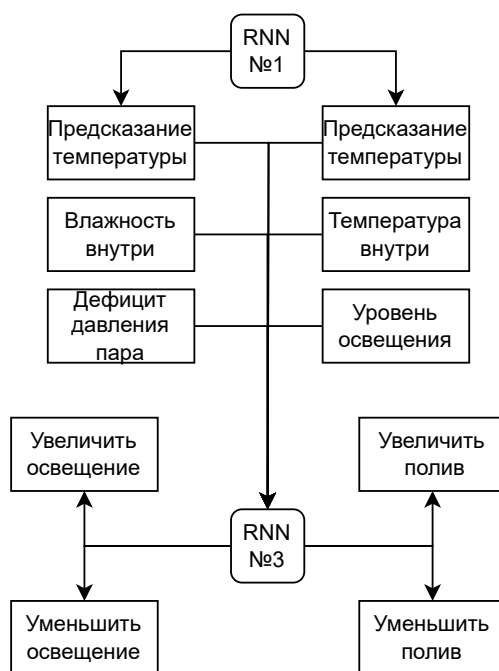


Рис. 4. Схема параметров нейронной сети № 3

После обучения модель будет способна выдавать рекомендации об изменении интенсивности работы средств регуляции среды: уровня освещения и полива в большую или меньшую сторону. Эти данные предполагается использовать для модели более высокого уровня с целью помощи в принятии решений.

Рассмотрим параметры, участвующие в моделировании нейронной сети № 4:

- дата и время;
- результаты работы нейронной сети № 1;
- результаты работы нейронной сети № 2;
- результаты работы нейронной сети № 3;
- температура воздуха внутри тепличной системы;
- влажность воздуха внутри тепличной системы;
- дефицит давления пара;
- мощность, подаваемая на устройство освещения.

После обучения нейронная сеть показывает свою оценку для решений преды-

дущих нейронных сетей на основе данных о состоянии микроклимата. Эта оценка отображается в пользовательском интерфейсе и работает как часть системы принятия решений. Поскольку на вход нейронной сети № 4 подаются те же данные что и на входы предыдущих, это приводит к двойной оценке состояния микроклимата. После проверки работоспособности концепции следует рассмотреть возможность убрать из входных данных нейронной сети № 4 данные о текущем состоянии микроклимата, и в качестве входных данных использовать только выходные данные предыдущих нейронных сетей.

Схема входных и выходных параметров нейронной сети № 4 представлена на рис. 5.

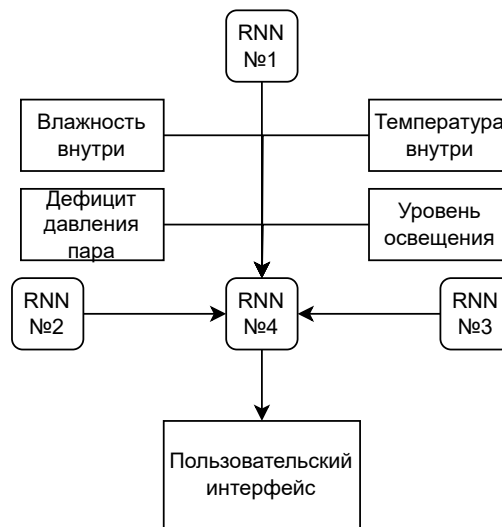


Рис. 5. Схема параметров нейронной сети № 4

3. Обработка данных внутри системы принятия решений

Данные о состоянии среды, предсказанные с помощью нейронных сетей, не стоит принимать как абсолютно истинные. Для оценки эффективности моделей стоит использовать метрики, измеряющие различия между предсказанными и фактическими значениям, такие как Mean Absolute Error (средняя абсолютная ошибка, MAE) и Root Mean Squared Error (среднеквадратичная ошибка, RMSE).

3.1. MAE

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|.$$

Результат вычисления просто интерпретируем и представляет собой среднее абсолютное отклонение между предсказанными и реальными данными. Алгоритм MAE не учитывает абсолютную величину ошибок, все ошибки имеют одинаковый вес в итоговой формуле. За счёт этого алгоритм устойчив к выбросам и случайным данным. Такие данные могут появиться в системе из-за ошибок чтения датчиков

влажности или температуры. Для значений используемых в системе, MAE должен не превышать следующих значений:

- Температура:
 - краткосрочные предсказания (1 измерение, 10 минут): 0.3–1.0 °C.
 - долгосрочные предсказания (среднее за сутки): 1.5–3.0 °C.
- Влажность:
 - краткосрочные предсказания (1 измерение, 10 минут): 1–3 %.
 - долгосрочные предсказания (среднее за сутки): 3–5 %.

Величина MAE может быть снижена с помощью увеличения объёма обучающей выборки, предобработки данных (нормализация, устранение шумов), увеличения точности датчиков.

3.2. RMSE

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}.$$

Результат вычисления представляет собой среднее квадратичное отклонение ошибок предсказаний. Из-за того, что учитывается не только величина отклонений, но и их распределение, метрика может быть менее интуитивной, чем MAE. Алгоритм RMSE более чувствителен к выбросам и случайным значениям за счёт использования квадратичного штрафа. Это делает его более строгим по сравнению с MAE. Из-за специфики вычислений факторы, позволяющие увеличить точность предсказаний и соответственно уменьшить величину метрики RMSE, совпадают с таковыми у MAE.

Из-за квадратичной природы вычислений значения метрики будут на 10–30 % выше, чем у MAE:

- Температура:
 - краткосрочные предсказания (1 измерение, 10 минут): 0.5–1.5 °C.
 - долгосрочные предсказания (среднее за сутки): 2.0–3.5 °C.
- Влажность:
 - краткосрочные предсказания (1 измерение, 10 минут): 1.5–4 %.
 - долгосрочные предсказания (среднее за сутки): 4–6%.

4. О месте системы принятия решений в функциональной схеме системы

В результате декомпозиции абстрактной тепличной системы её функциональные подсистемы были выделены следующим образом:

- Система Пользовательского Интерфейса – отвечает за предоставление пользователю актуальных данных о состоянии системы, передачу введённых пользователем данных Системе Управления.
- Система Управления – отвечает за выработку управляющих сигналов для Системы Регуляции Среды, получение данных от Системы Сбора Данных, предоставление данных для Системы Пользовательского Интерфейса.

- Система Регуляции Среды – отвечает за переключение средств регуляции среды в соответствии с управляющим сигналом от Системы Управления.
- Система Сбора Данных – отвечает за опрос датчиков с использованием различных протоколов связи низкого уровня, преобразования полученных данных и их отправку Системе Управления с использованием протоколов связи высокого уровня [6].

Декомпозицию тепличной системы, рассмотренную ранее, на первый взгляд, можно дополнить Системой Принятия Решений (см. рис. 6).

Однако при внимательном рассмотрении можно заметить, что Система Принятия Решений имеет те же связи и роль, что и Система Управления: обе системы отвечают за анализ данных среды, возможное переключение средств регуляции среды и взаимодействие с пользовательским интерфейсом. Таким образом, более корректно будет не выделять Систему Принятия Решений как отдельный функциональный компонент, а считать её частью Системы Управления (см. рис. 7).

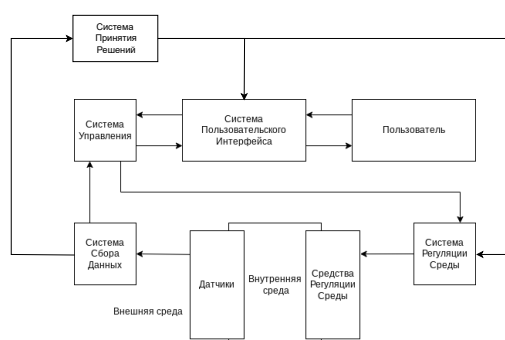


Рис. 6. Функциональная схема тепличной системы с выделением Системы Принятия Решений

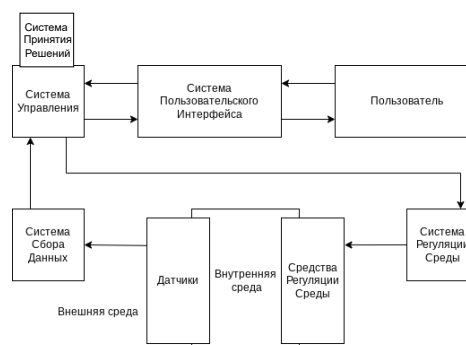


Рис. 7. Функциональная схема тепличной системы с Системой Принятия Решений как частью Системы Управления

Внедрение Системы Принятия Решений на основе RNN в тепличные хозяйства открывает новые возможности для повышения их эффективности и устойчивости. Использование каскадной архитектуры, где выходы одной нейронной сети являются входами другой, усложняет алгоритм и делает его менее прозрачным, однако позволяет добиться большей степени автоматизации. Применение Системы Принятия Решений как вспомогательной для пользователя позволяет упростить внедрение в уже существующие тепличные системы и снизить влияние человеческого фактора. Увеличение автоматизации, снижение человеческого фактора и адаптивность к изменениям условий делают такие системы перспективными. Будущие исследования и разработки в этой области могут еще больше улучшить технологии и методы, используемые в тепличных хозяйствах, способствуя их дальнейшему развитию и интеграции в глобальную систему продовольственного обеспечения.

Литература

1. Бекбергенева Д.Е., Баранник В.А. Продовольственная безопасность Российской Федерации // Экономика и бизнес: теория и прак-

- тика. 2022. № 4-1. URL: <https://cyberleninka.ru/article/n/prodovolstvennaya-bezopasnost-rossiyskoy-federatsii> (дата обращения: 24.10.2023).
2. Коммерческий тепличный рынок – Статистика размеров и отрасли. URL: <https://www.mordorintelligence.com/ru/industry-reports/commercial-greenhouse-marke> (дата обращения: 01.12.2024).
 3. Domingo Sancho-Knapik, Óscar Mendoza-Herrer, David Alonso-Forn, Miguel Ángel Saz, Rubén Martín-Sánchez, José Víctor dos Santos Silva, Jerome Ogee, José Javier Peguero-Pina, Eustaquio Gil-Pelegrín, Juan Pedro Ferrio. Vapor pressure deficit constrains transpiration and photosynthesis in holm oak: A comparison of three methods during summer drought. *Agricultural and Forest Meteorology*. 2022. Vol. 327. URL: <https://doi.org/10.1016/j.agrformet.2022.109218> (дата обращения: 14.11.2024).
 4. de Brito M.S. et al. A service orchestration architecture for Fog-enabled infrastructures // Second International Conference on Fog and Mobile Edge Computing (FMEC). Valencia, Spain, 2017. P. 127–132. URL: https://www.researchgate.net/publication/317639602_A_service_orchestration_architecture_for_Fog-enabled_infrastructures (дата обращения: 14.11.2024).
 5. Губарева Е.А. Нейронные сети в анализе временных рядов // Инновации и инвестиции. 2020. № 10. URL: <https://cyberleninka.ru/article/n/neyronnye-seti-v-analize-vremennyh-ryadov> (дата обращения: 09.11.2024).
 6. Костеннов Т.В. Применение методов системного анализа к построению структурно-функциональной схемы тепличной системы // Математические структуры и моделирование. 2022. № 3 (63). С. 69–78.

USING NEURAL NETWORKS TO CREATE A DECISION-MAKING SYSTEM FOR MANAGING THE MICROCLIMATE OF A GREENHOUSE SYSTEM

T.V. Kostennov

Ph.D. Student, e-mail: timofey.kostennov@gmail.com

Dostoevsky Omsk State University, Omsk, Russia

Abstract. This article describes the mechanism of using neural networks to predict microclimate parameters in the decision-making system of a greenhouse system. The use of several recurrent neural networks within the decision-making system is considered. Attention is paid to the correct determination of the place of the decision-making system in the functional diagram of the greenhouse system.

Keywords: greenhouse system, system analysis, forecasting, microclimate, neural network, decision-making system.

Дата поступления в редакцию: 24.11.2024

СИСТЕМАТИЗАЦИЯ ИНСТРУМЕНТОВ И МЕТОДОВ ДЛЯ ЭФФЕКТИВНОЙ РЕПЛИКАЦИИ БАЗЫ ДАННЫХ

Т.М. Опарина

старший преподаватель, e-mail: oparina2007@yandex.ru

Омский государственный университет им. Ф.М. Достоевского, Омск, Россия

Аннотация. Рассматриваются основные проблемы, с которыми сталкиваются преподаватели и студенты, изучая тему репликации, а также предлагается возможный путь их решения.

Ключевые слова: преподавание, репликация, безопасность баз данных.

Введение

Изучение основ построения защищённых баз данных является важной частью подготовки специалистов в области безопасности информационных технологий. Одной из ключевых тем в рамках дисциплины «Основы построения защищённых баз данных» является изучение механизмов обеспечения высокой доступности баз данных. В системах, где требуется высокая надёжность, доступность данных и минимизация времени простоя, важна репликация базы данных. Анализ учебных пособий, лабораторных практикумов в области баз данных и безопасности баз данных показал, что в большинстве из них отсутствует практическое освещение данной темы, и лишь в редких случаях она рассматривается с теоретической точки зрения [1–3]. Это создаёт пробел в образовательном процессе, что подчёркивает необходимость более глубокого изучения и внедрения практических аспектов настройки репликации баз данных в учебные программы. Однако обучение студентов и формирование у них практических навыков в этой области сталкивается с рядом проблем.

В данной статье будут рассмотрены основные трудности, с которыми сталкиваются студенты, изучая тему репликации, а также предложены возможные пути их решения.

1. Методологические проблемы при реализации репликации базы данных

Репликация базы данных – это процесс создания и поддержания копий базы данных на нескольких серверах для повышения доступности, надёжности и производительности базы данных. Наиболее часто используется организация репликации, при которой изменения могут выполняться только на одном главном (master) сервере. Изменения, выполненные на этом сервере, переносятся на все копии, которые могут использоваться только для чтения данных, и называются запасными (standby), или репликами (replica) [4].

При реализации репликации в образовательной среде студенты и преподаватели могут столкнуться с рядом проблем:

1. Длительность и сложность настройки при ограниченном времени занятия:

1.1. Многообразие технологий.

Существует множество технологий и подходов к репликации (например, синхронная и асинхронная репликация, мастер-мастер и т. д.). Преподаватели могут испытывать трудности с выбором подходящей технологии и её настройкой на лабораторных занятиях. Разные операционные системы, системы управления базами данных (СУБД) имеют свои уникальные решения и процедуры в области репликации, что может привести порой к сложности реализации и нехватки времени в рамках одного занятия для полноценного разбора.

1.2. Настройка сетевого окружения.

Для успешной реализации репликации необходимо правильно настроить сетевую инфраструктуру. Для настройки репликации требуется как минимум два сервера (мастер и реплика), что в учебных условиях сложно организовать. В этом случае студенты могут столкнуться с проблемами настройки сетевых адресов и маршрутизации, что может затруднить процесс обучения.

2. Проблемы с синхронизацией данных.

При использовании асинхронной репликации могут возникать конфликты данных, когда изменения в одной базе данных не успевают быть отражены в другой. Это может создать ложное представление о целостности данных у студентов и затруднить их понимание процессов, связанных с репликацией.

3. Управление отказами:

3.1. Обнаружение сбоев и восстановление работоспособности системы.

Студенты должны научиться выявлять сбои в системе репликации и восстанавливать нормальную работу СУБД. Это включает в себя диагностику проблем и применение методов восстановления данных.

3.2. Потеря данных.

В случае сбоя системы репликации существует риск потери данных. Необходимо проводить репликацию учебной тестовой базы данных. Обучающиеся должны быть осведомлены о возможных последствиях и о том, как минимизировать риски потери информации.

4. Безопасность данных. Уязвимости при передаче данных в момент репликации базы данных.

Репликация данных между серверами может создавать уязвимости, связанные с безопасностью. Если данные передаются по незащищённым каналам (от мастера к реплике), то злоумышленники могут перехватить информацию, включая учётные данные и личные данные пользователей. Студенты должны осознавать важность шифрования данных при передаче и применять меры

безопасности для защиты информации. Таким образом, тема «Методы шифрования в базе данных» должна стоять перед изучением репликации.

5. Ограниченные ресурсы.

В образовательном учреждении часто отсутствуют необходимые ресурсы (память, процессор, дисковое пространство) для полноценного обучения репликации баз данных. Студенты могут не иметь доступа к реальным техническим инструментам или программам, что ограничивает их практический опыт. Вместо работы с реальными данными студенты могут использовать симуляционные модели, которые требуют меньше ресурсов. Это позволит им понять основные принципы репликации данных без необходимости в больших объёмах информации.

6. Теоретические знания без практики.

Обучение может быть сосредоточено на теоретических аспектах без достаточного количества практических занятий. Это может привести к недостаточному пониманию студентами материала, связанного с репликацией баз данных.

2. Формирование компетенций в области репликации баз данных

Целью занятия по настройке репликации является изучение процессов и методов, позволяющих создавать копии данных с одного сервера на другой, обеспечивая тем самым высокую доступность, отказоустойчивость и балансировку нагрузки в СУБД. В ходе занятия студенты должны ознакомиться с различными типами синхронной репликации, а также получить практические навыки по настройке репликации в конкретной СУБД.

В настоящее время для изучения дисциплины «Основы построения защищённых баз данных» на факультете цифровых технологий и кибербезопасности Омского государственного университета им. Ф.М. Достоевского используется PostgreSQL.

Практическая часть занятия по репликации данных предполагает решение двух задач в PostgreSQL с использованием двух виртуальных машин:

1. Настройка синхронной репликации в PostgreSQL, когда одна виртуальная машина является мастером, а вторая – репликой, состоит из следующих этапов:
 - Создание двух виртуальных машин VMware (VM1 и VM2) с установленной операционной системой (например, Ubuntu Linux). Для реализации репликации используются две виртуальные машины под Ubuntu. Одна является мастером, а вторая – репликой.
 - Установка PostgreSQL на обе виртуальные машины.
 - Настройка сети виртуальной машины VM1 с IP-адресом и виртуальной машины VM2 с IP-адресом.
 - Настройка основной (мастер) машины.
 - Настройка резервной (реплика) машины.

- Создание серверов базы данных и проверка репликации.
2. Настройка синхронной репликации в PostgreSQL, когда одна виртуальная машина является мастером, а вторая – репликой, но при этом если сбой происходит на машине, которая является мастером, то вторая машина становится мастером на время сбоя первой. Работа с базой данных осуществляется с той машины, которая является мастером в кластере (под кластером здесь понимается серверный процесс, а также набор баз данных под его управлением). Только мастер может вносить изменения в базу данных, реплика же находится в режиме чтения, и с неё вносить изменения в базу данных нельзя.

Решение этой задачи связано с дополнительными техническими сложностями при настройке отказоустойчивого кластера PostgreSQL, с использованием Patroni и Etcd, но является более актуальной. Patroni – Framework, разработанный для построения кластера PostgreSQL высокой доступности на основе потоковой репликации [5]. Если мастер-узел становится недоступным, Patroni автоматически выбирает новую реплику и переводит её в режим мастера. Etcd – это распределённая система хранения ключей-значений, используемая для хранения конфигурации и метаданных. Patroni использует Etcd для синхронизации состояния узлов кластера, чтобы обеспечить отказоустойчивость. Etcd обеспечивает непрерывность работы кластера в случае сбоя отдельных узлов.

Это современное решение подходит для сред, где требуется минимизировать время простоя и обеспечить бесперебойную работу баз данных.

Настройка отказоустойчивого кластера PostgreSQL с использованием Patroni и Etcd предполагает:

- Создание двух виртуальных машин VMware (VM1 и VM2) с установленной операционной системой (например, Ubuntu Linux). Для реализации репликации используются две виртуальные машины под Ubuntu. Одна является мастером, а вторая – репликой.
- Установку PostgreSQL на обе виртуальные машины.
- Настройку сети виртуальной машины VM1 с IP-адресом и виртуальной машины VM2 с IP-адресом.
- Установку Etcd.
- Настройку Etcd (создание конфигурационного файла для каждой виртуальной машины). Студентам предлагается шаблон конфигурационного файла (см. рис. 1) с описанием каждого параметра.
- Настройку автоматизации запуска Etcd как сервиса в системе.
- Проверку корректной работы Etcd.
- Установку и настройку Patroni (создание конфигурационного файла Patroni для каждой виртуальной машины). Студентам также предлагается шаблон конфигурационного файла (см. рис. 2);
- Проверку корректной работы Patroni.
- Создание серверов базы данных и тестирование работы репликации.

```

1 name: "node1"
2 data-dir: "node1.etcd"
3 initial-advertise-peer-urls: http://192.168.1.36:2380
4 listen-peer-urls: http://192.168.1.36:2380
5 listen-client-urls: http://192.168.1.36:2379,http://127.0.0.1:2379
6 advertise-client-urls: http://192.168.1.36:2379
7 initial-cluster: "node1=http://192.168.1.36:2380,node2=http://192.168.1.37:2380"
8 initial-cluster-state: "new"
9 initial-cluster-token: "etcd-cluster"

```

Рис. 1. Пример конфигурационного файла Etcd для виртуальной машины VM1

```

1 scope: postgres
2 namespace: /db/
3 name: fatal1
4
5 restapi:
6 listen: 192.168.1.36:8008
7 connect_address: 192.168.1.36:8008
8
9 etcd3:
10 hosts: 192.168.1.36:2379,192.168.1.37:2379
11
12 bootstrap:
13 dcs:
14 ttl: 30
15 loop_wait: 10
16 retry_timeout: 10
17 maximum_lag_on_failover: 1048576
18 postgresql:
19 use_pg_rewind: true
20 initdb:
21 - encoding: UTF8
22 - data-checksums
23 pg_hba:
24 - host replication replicator 127.0.0.1/32 md5
25 - host replication replicator 192.168.1.36/0 md5
26 - host replication replicator 192.168.1.37/0 md5
27 - host all all 0.0.0.0/0 md5
28 users:
29 admin:
30 password: admin
31 options:
32 - createrole
33 - createdb
34
35 postgresql:
36 listen: 192.168.1.36:5432
37 connect_address: 192.168.1.36:5432
38 data_dir: /mnt/patroni
39 bin_dir: /usr/lib/postgresql/16/bin
40 pgpass: /tmp/pgpass
41 authentication:
42 replication:
43 username: replicator
44 password: '1234'
45 superuser:
46 username: postgres
47 password: '1234'
48 parameters:
49 unix_socket_directories: '.'
50 wal_level: replica
51 max_wal_senders: 10
52 max_replication_slots: 10
53 synchronous_commit: local
54
55 tags:
56 nofailover: false
57 noloadbalance: false
58 clonefrom: false
59 nosync: false

```

Рис. 2. Пример конфигурационного файла Patroni для виртуальной машины VM1

Заключение

Формирование практических навыков студентов при изучении темы «Настройка репликации базы данных» является важной задачей в рамках дисциплины «Основы построения защищённых баз данных». Несмотря на существующие проблемы, их решение возможно через использование современных инструментов, таких как виртуальные машины, системы автоматизации процесса репликации, упрощение настройки окружения, разработка подробных практических заданий. Это позволит студентам не только освоить теорию, но и приобрести необходимые навыки для успешной профессиональной деятельности в области информационных технологий.

Литература

1. Агафонов А.А., Юмаганов А.С. Безопасность систем баз данных: учеб. пособие. Самара: Изд-во Сам. ун-та, 2023. 272 с.
2. Дудник Е.А. Информационная безопасность баз данных: метод. указ. для самостоят. работы студентам направления подготовки «Информатика и вычислительная техника». Рубцовск: РИИ, 2021. 32 с.
3. Пригонюк Н.Д., Петров В.И. Основы построения защищённых баз данных: учеб. пособие. Воронеж: ООО «МИР», 2019. 76 с.
4. Новиков Б.А. Основы технологий баз данных: учеб. пособие. М.: ДМК Пресс, 2020. 582 с.
5. Отказоустойчивый кластер СУБД Postgres Pro на основе Patroni. URL: <https://postgrespro.ru/clusters/patroni> (дата обращения: 24.11.2024).

SYSTEMATIZATION OF TOOLS AND METHODS FOR EFFECTIVE DATABASE REPLICATION

T.M. Oparina

Assistant Professor, e-mail: oparina2007@yandex.ru

Dostoevsky Omsk State University, Omsk, Russia

Abstract. This article examines the main problems that teachers and students encounter why studying the topic of replication, and also suggests a possible way to solve them encounters.

Keywords: teaching, replication, database security.

Дата поступления в редакцию: 04.02.2025

ОБУЧЕНИЕ ГРАНУЛЯРНЫХ КОГНИТИВНЫХ КАРТ МЕТОДОМ APSO

С.Н. Чуканов¹

д.т.н., профессор, e-mail: ch_sn@mail.ru

И.С. Чуканов²

магистрант, e-mail: chukanov022@gmail.com

С.В. Лейхтер³

старший преподаватель, e-mail: leykhter@mail.ru

¹Институт математики им. С.Л. Соболева СО РАН, Омский филиал, Омск, Россия

²Уральский федеральный университет имени первого Президента России Б.Н. Ельцина,
Екатеринбург, Россия

³Омский государственный университет им. Ф.М. Достоевского, Омск, Россия

Аннотация. Рассматривается метод ускоренной оптимизации роя частиц для оценивания весов и обучения гранулярных когнитивных карт (GCM). Основная задача процедуры обучения – найти настройку весов GCM, которая приведёт GCM к требуемому устойчивому состоянию. Это достигается посредством минимизации правильно определённой целевой функции. Установленные алгоритмы в значительной степени зависят от начального приближения матрицы весов, которое предоставляется экспертами. В работе для обучения весов нечётких когнитивных карт используется метод ускоренной оптимизации роя частиц (APSO). Один из методов обучения GCM основан на методе ускоренной оптимизации роя частиц. APSO используется для определения соответствующих весовых матриц для системы путём минимизации заданной целевой функции.

Ключевые слова: гранулированные когнитивные карты, метод PSO, ускоренная оптимизация роя частиц.

Введение

В 1986 г. Б. Коско представил нечёткие когнитивные карты (Fuzzy Cognitive Maps, FCM) для причинно-следственной обработки и вычислительной обработки оценок причинно-следственных связей, показанных в виде ориентированного графа. Создание FCM требует знаний экспертов о рассматриваемой системе. FCM состоит из концепций $C_i, i = 1, \dots, N$. Каждый узел (концепция) содержит значение $A_i \in [0, 1], i = 1, \dots, N$. Концепции соединены взвешенными дугами, которые показывают отношения. Связь между двумя концепциями имеет вес w_{ij} . Направление указывает, является ли концепция причиной или следствием. Существует три типа весов:

$$\begin{cases} w_{ij} > 0, & \text{положительная причинность,} \\ w_{ij} < 0, & \text{отрицательная причинность,} \\ w_{ij} = 0, & \text{отсутствие причинности.} \end{cases} \quad (1)$$

На каждом шаге значение концепции A_i находится под влиянием значений других концепций:

$$A_i(k+1) = f \left(A_i(k) + \sum_{\substack{j=1 \\ j \neq i}}^n W_{ji} A_j(k) \right), \quad (2)$$

где w_{ji} – вес дуги, соединяющей концепцию – сигмоидальная функция фаззификации [1]:

$$f(x) = (1 + \exp(-\lambda(x - \sigma)))^{-1}, \quad (3)$$

где λ – значение крутизны сигмоидальной функции; σ – значение, при котором $f(\sigma) = 0,5$. Проектирование FCM опирается на мнения экспертов. Эксперты описывают влияние одной концепции на другую как «негативное», «положительное» или «отсутствие». Метод дефаззификации используется для преобразования лингвистического веса в числовое значение [2]. Преимущество этого подхода в том, что экспертам не нужно напрямую присваивать значения для создания корреляций и эффектов, но им необходимо сравнивать значения между понятиями. Экспертами задаётся начальная матрица весов $W^{initial} = [w_{ij}]; i, j = 1, \dots, N$, при $w_{ii} = 0; i = 1, \dots, N$. Процедуры обучения представляют собой средства повышения эффективности и надёжности FCM путём обновления матрицы весов. На сегодняшний день существует всего несколько алгоритмов обучения FCM, основанных на опыте обучения нейронных сетей [3]. Такие алгоритмы начинают с начального состояния и начальной матрицы весов $W^{initial}$ FCM и адаптируют веса, чтобы вычислить матрицу весов, которая приводит FCM к желаемому устойчивому состоянию.

GCM, о которых идёт речь в данной статье, являются разновидностью FCM и алгоритмы их обучения основываются на концепции FCM в целом.

1. Алгоритм ускоренного обучения PSO

Процесс обучения имеет решающее значение для проектирования когнитивной карты и её дальнейшей работы. Доступные для обучения данные поступают в виде последовательности векторов $A_c(k), k = 1, 2, \dots, N$. Эти векторы преобразуются в уровни активации $A_c(k)$, которые используются в процедуре обучения. Функция критерия, по экстремуму (минимуму) которой определяются требуемые значения вектора значений весов w_{ij} , находится как сумма расстояний между прогнозируемыми векторами активации узлов $A(k+1)$, которые вычисляются на основе $A(k)$, доступных в k -й момент времени, и векторами $\tilde{A}_i(k+1)$ известных уровней акти-

вазии узлов в $(k + 1)$ -й момент времени:

$$V = \sum_{k=1}^N \sum_{i=1}^c \left\| \tilde{A}_i(k+1) - f_i \left(\sum_{j=1}^c w_{ij} A_j(k) \right) \right\|^2. \quad (4)$$

Алгоритм обучения PSO (particle swarm optimization) [4,5] был предложен в качестве метода роевого интеллекта. В PSO, в отличие от традиционных методов оптимизации, реализуется поиск на основе популяции для поиска оптимальных решений. Возможные решения называются частицами. Во время полёта скорость и положение обновляются с точки зрения знаний о его собственных и всех частицах. По сути, PSO обновляет скорость и положение по уравнениям (5) и (6) соответственно:

$$V_i^{t+1} = V_i^t + \alpha r_1 (X_i^* - X_i^t) + \beta r (g^* - X_i^t), \quad (5)$$

$$X_i^{t+1} = X_i^t + V_i^{t+1} \Delta t, \quad (6)$$

где X_i^t, V_i^t – положение и скорость частицы i соответственно; а X_i^* и g^* – текущее глобальное лучшее решение и индивидуальное лучшее решение частицы i в истории поиска, здесь α, β – параметры обучения. В уравнении (2) временной интервал Δt можно установить $\Delta t = 1$. Базовый PSO получает полезные знания поиска как от каждой частицы, так и от всей популяции частиц. X_i^* используется для повышения разнообразия. Базовый PSO можно упростить, чтобы ускорить его сходимость, используя только g^* , что называется ускоренным PSO (APSO) [5,6]. В APSO скорость обновляется по уравнению (7):

$$V_i^{t+1} = V_i^t + \alpha r + \beta r (g^* - X_i^t). \quad (7)$$

Здесь случайная величина r вычисляется по распределению $N(0, 1)$ для замены второго члена с целью ускорения сходимости скорости. По сути, это обновление то же самое, что и в уравнении (6). Скорость бесполезна в APSO. Кроме того, уравнение (7) можно выразить другим способом:

$$X_t^{t+1} = (1 - \beta) X_t^t + \beta g^* + \delta^t r. \quad (8)$$

Обратите внимание, что скорость бесполезна в уравнении (8), и поэтому нам не нужно инициализировать её. Здесь APSO намного проще и легче реализуется. Здесь τ – член рандомизации. В большинстве случаев $\alpha = (0,1 \dots 0,5) L$, где L – число переменных, в то время как $\beta = 0.2 \dots 0.7$ подходит для большинства практических инженерных задач. Параметр α можно задать как

$$\alpha = \delta^t; 0 < \delta < 1. \quad (9)$$

Рассмотрим пример оценивания весов взаимного влияния концепций экологии города [7, 8] методом APSO (см. рис. 1).

В табл. 1 заданы временные ряды концепций. Определим веса взаимных влияний концепций методом APSO по информации табл. 1.

Экспертная оценка интервалов принадлежности значений весов взаимного влияния соответствует значениям: $w_{12} \in [0,0 \dots 0,2]$; $w_{16} \in [-0,4 \dots -0,2]$; $w_{23} \in$

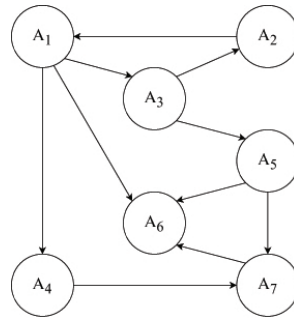


Рис. 1. Граф взаимных влияний концепций: A_1 – население города; A_2 – миграция в город; A_3 – модернизация; A_4 – мусор на единицу площади; A_5 – санитарные условия; A_6 – количество заболеваний; A_7 – количество бактерий на единицу площади

Таблица 1. Исходный временной ряд

t	A_1	A_2	A_3	A_4	A_5	A_6	A_7
0	0,30	0,35	0,40	0,45	0,50	0,55	0,60
1	0,04	0,79	0,69	0,90	0,97	0,69	0,63
2	0,00	1,00	0,89	0,99	1,00	0,14	0,66
3	0,01	1,00	0,98	0,99	1,00	0,00	0,82
4	0,02	1,00	0,99	0,99	1,00	0,00	0,96
5	0,02	1,00	0,99	0,99	1,00	0,00	0,99

$[0,6 \dots 0,8]$; $w_{31} \in [0,5 \dots 0,7]$; $w_{41} \in [0,8 \dots 1,0]$; $w_{53} \in [0,8 \dots 1,0]$; $w_{65} \in [-1,0 \dots -0,8]$; $w_{67} \in [0,7 \dots 0,9]$; $w_{74} \in [0,8 \dots 1,0]$; $w_{75} \in [-1,0 \dots -0,8]$. В результате оценки весов w_{ij} методом APSO получим значения, приведённые в табл. 2. В табл. 3 приведены временные ряды концепций $A_1 - A_7$ после оценки весов w_{ij} .

Таблица 2. Результаты оценки весов w_{ij}

	A_1	A_2	A_3	A_4	A_5	A_6	A_7
A_1		$w_{12} = 0,08$				$w_{16} = -0,3$	
A_2			$w_{23} = 0,73$				
A_3	$w_{31} = 0,6$						
A_4	$w_{41} = 0,9$						
A_5			$w_{53} = 0,9$				
A_6					$w_{65} = -1,0$		$w_{67} = 0,7$
A_7				$w_{74} = 0,9$	$w_{75} = -0,9$		

Таблица 3. Полученный после оценки весов w_{ij} временной ряд

t	A_1	A_2	A_3	A_4	A_5	A_6	A_7
0	0,3	0,35	0,4	0,45	0,5	0,55	0,6
1	0,03	0,81	0,69	0,90	0,97	0,43	0,63
2	0,01	0,95	0,89	0,99	1,00	0,32	0,66
3	0,01	0,99	0,98	0,99	1,00	0,14	0,82
4	0,01	0,99	0,99	0,99	1,00	0,03	0,96
5	0,01	0,99	0,99	0,99	1,00	0,01	0,99

2. Обучение GCM

Рассмотрим процесс грануляции данных на основе алгоритма кластеризации нечётких С-средних [9]. Задача кластеризации сводится к минимизации выбранной целевой функции Q , являющейся функцией параметров кластеров. Задача оптимизации формализуется как $\min_{U; v_1, v_2, \dots, v_c} Q(U; v_1, v_2, \dots, v_c)$, где минимизация выполняется относительно матрицы разбиения $U \in U$ и прототипов $v_1, v_2, \dots, v_c \in \mathbb{R}^n$. Для множества n -мерных данных $\{x_k\}$, $k = 1, 2, \dots, N$, целевая функция Q выражается как сумма квадратов расстояний с суммированием по данным и кластерам $Q = \sum_{i=1}^c \sum_{k=1}^N u_{ik}^m \sum_{j=1}^n (x_{kj} - v_{ij})^2$, где u_{ik} – степень принадлежности данных x_k к кластеру v_i . Минимизация Q осуществляется относительно $u_{ik} \in U$ и множества прототипов кластеров, $V = \{v_1, v_2, \dots, v_c\}$. Более явно мы запишем это следующим образом $\min_{u_{ik} \in U, v_1, v_2, \dots, v_c \in \mathbb{R}^n} Q$. В результате минимизации Q получим выражение для u_{ik} :

$$u_{ik} = \left(\sum_{j=1}^c \left(\frac{\|x_k - v_i\|^2}{\|x_k - v_j\|^2} \right)^{\frac{1}{m-1}} \right)^{-1}.$$

В результате минимизации по v_i получим следующее выражение для прототипов кластеров: $v_i = \left(\sum_{k=1}^N u_{ik}^m \right)^{-1} \sum_{k=1}^N u_{ik}^m x_k$. Рассмотрим пример оценивания весов взаимного влияния концепций экологии города [7, 8]. В табл. 4 приведены временные ряды значений концепций. В табл. 5 приведены гранулярные значения концепций. Определим веса w_{ij} взаимных влияний концепций методом APSO по информации таблицы 4. В результате оценки весов w_{ij} методом APSO получим значения, приведённые в табл. 6. В табл. 7 приведены гранулярные временные ряды концепций A_1 – A_7 после оценки весов w_{ij} . Сравнение значений табл. 5 и 7 указывает на удовлетворительные результаты оценивания гранулярных временных рядов. Целевой функционал $Q = \sum_{i=1}^c \sum_{k=1}^N u_{ik}^m \sum_{j=1}^n (x_{kj} - v_{ij})^2$ принимает значение 0,0407.

Таблица 4. Временные ряды значений концепций

t	A_1	A_2	A_3	A_4	A_5	A_6	A_7
0	0,3000	0,3500	0,4000	0,4500	0,5000	0,5500	0,6000
1	0,3407	0,5646	0,5399	0,6083	0,6726	0,5399	0,5275
2	0,3706	0,7079	0,6198	0,6963	0,7887	0,4288	0,4848
3	0,4075	0,7830	0,6647	0,7426	0,8446	0,3129	0,4510
4	0,4462	0,8171	0,6939	0,7718	0,8683	0,2363	0,4301
5	0,4785	0,8328	0,7157	0,7936	0,8793	0,1975	0,4223
6	0,5013	0,8413	0,7322	0,8098	0,8856	0,1806	0,4232
7	0,5156	0,8465	0,7438	0,8208	0,8898	0,1743	0,4280
8	0,5240	0,8500	0,7514	0,8277	0,8927	0,1724	0,4333
9	0,5286	0,8522	0,7561	0,8318	0,8945	0,1724	0,4378
10	0,5310	0,8536	0,7589	0,8341	0,8957	0,1729	0,4409
11	0,5322	0,8544	0,7604	0,8354	0,8964	0,1735	0,4430
12	0,5327	0,8549	0,7612	0,8360	0,8967	0,1740	0,4443

Таблица 5. Гранулярные значения концепций

v_i	A_1	A_2	A_3	A_4	A_5	A_6	A_7
v_1	0,40	0,68	0,59	0,67	0,76	0,46	0,51
v_2	0,50	0,80	0,69	0,73	0,83	0,23	0,45
v_3	0,55	0,82	0,72	0,81	0,86	0,20	0,46
v_4	0,82	0,82	0,73	0,81	0,87	0,20	0,47

Таблица 6. Результаты оценки весов w_{ij}

	A_1	A_2	A_3	A_4	A_5	A_6	A_7
A_1		0,30				-0,10	
A_2			0,77				
A_3	0,56						
A_4	0,83						
A_5			0,84				
A_6					-0,97		0,60
A_7				1,00	-0,96		

Таблица 7. Полученный после оценки весов w_{ij} временной ряд гранулярных значений

v_i	A_1	A_2	A_3	A_4	A_5	A_6	A_7
v_1	0,40	0,68	0,59	0,67	0,76	0,46	0,51
v_2	0,53	0,78	0,65	0,73	0,82	0,28	0,47
v_3	0,62	0,83	0,71	0,79	0,85	0,19	0,46
v_4	0,67	0,85	0,75	0,83	0,87	0,15	0,47

Заключение

Метод GCM успешно используется для проектирования сложных систем. В работе предложен алгоритм обучения для определения матриц с субоптимальными весами для GCM. Метод основан на минимизации целевой функции с использованием алгоритма APSO.

Благодарности

Работа выполнена по Госзаданию ИМ СО РАН, проект № FWNF-2022-0016.

Литература

1. Haykin S. Neural networks and machine learning (the third edition of the original book). 2011.
2. Van Leekwijck W., Kerre E. E. Defuzzification: criteria and classification // Fuzzy sets and systems. 1999. Vol. 108, No. 2. P. 159–178.
3. Papageorgiou E., Stylios C. D., Groumpos P. P. Activation hebbian learning rule for fuzzy cognitive maps // IFAC Proceedings Volumes. 2002. Vol. 35, No. 1. P. 319–324.
4. Eberhart R. C., Shi Y., Kennedy J. Swarm Intelligence (Morgan Kaufmann series in evolutionary computation). Morgan Kaufmann Publishers, 2001.
5. Yang X. S. Nature-inspired metaheuristic algorithms. Luniver press, 2010.
6. Wang G.G., Bai D., Gong W., Ren T., Liu X., Yan X. Particle-swarm krill herd algorithm // 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). IEEE, 2018. P. 1073–1080.
7. Lu W., Yang J., Liu X., Pedrycz W. The modeling and prediction of time series based on synergy of high-order fuzzy cognitive map and fuzzy c-means clustering // Knowledge-Based Systems. 2014. Vol. 70. P. 242–255.
8. Hagiwara M. Extended fuzzy cognitive maps // IEEEJ Transactions on Electronics, Information and Systems. 1994. Vol. 114, No. 3. P. 367–372.
9. Pedrycz W. An introduction to computing with fuzzy sets // IEEE ASSP Mag. 2021. Vol. 190.

LEARNING GRANULAR COGNITIVE MAPS USING APSO**S.N. Chukanov¹**

Dr.Sc. (Techn.), Professor, e-mail: a@a.ru

I.S. Chukanov²

Master's Degree Student, e-mail: chukanov022@gmail.com

S.V. Leykhter³

Assistant Professor, e-mail: leykhter@mail.ru

¹Sobolev Institute of Mathematics, Omsk branch, Omsk, Russia²Ural Federal University named after the first President of Russia B.N. Yeltsin, Ekaterinburg, Russia³Dostoevsky Omsk State University, Omsk, Russia

Abstract. The paper considers the accelerated particle swarm optimization method for weight estimation and training of granular cognitive maps (GCM). The main task of the training procedure is to find the GCM weight setting that will lead the GCM to the desired stable state. This is achieved by minimizing a well-defined objective function. The established algorithms largely depend on the initial approximation of the weight matrix, which is provided by experts. In this paper, the accelerated particle swarm optimization (APSO) method is used to train the weights of fuzzy cognitive maps. One of the GCM training methods is based on the accelerated particle swarm optimization method. APSO is used to determine the appropriate weight matrices for the system by minimizing a given objective function.

Keywords: granular cognitive maps, PSO method, accelerated particle swarm optimization.

Дата поступления в редакцию: 04.12.2024

ТЕОРИЯ ИГР В КИБЕРПРОСТРАНСТВЕ: РЕФЛЕКСИВНЫЙ ПОДХОД К МОДЕЛИРОВАНИЮ ВЗАИМОДЕЙСТВИЯ АДМИНИСТРАТОРА И ЗЛОУМЫШЛЕННИКА

Т.В. Вахний

к.ф.-м.н., доцент, e-mail: vahniytv@mail.ru

С.В. Вахний

студентка, e-mail: vakhniysv@mail.ru

Омский государственный университет им. Ф.М. Достоевского, Омск, Россия

Аннотация. Предложена расширенная модель биматричной игры, учитывающая различные уровни стратегической рефлексии администратора безопасности и злоумышленника.

Ключевые слова: компьютерная система, кибербезопасность, биматричная игра, рефлексия игроков, оптимальная стратегия.

Введение

В эпоху цифровых технологий кибербезопасность и защита данных приобретают критическое значение для успеха бизнеса. Последствия успешной кибератаки могут выйти далеко за рамки утечки данных и финансовых потерь, затронув операционную устойчивость, репутацию и даже само существование организации. Для эффективной борьбы с киберугрозами необходимо понимать возможные причины, лежащие в основе изменений в тактиках злоумышленника. В этой связи теоретико-игровые модели предлагают удобный инструмент для анализа взаимодействия между администратором безопасности и злоумышленником, позволяя оптимизировать выбор методов и средств защиты для построения надёжной системы безопасности, минимизирующей экономические риски компании [1–5].

В данной статье для моделирования стратегического взаимодействия в киберпространстве предлагается построить биматричную игру администратора безопасности со злоумышленником и решать её с введением рангов рефлексии для каждого игрока. Сопоставление полученных наилучших гарантированных результатов игроков позволит проанализировать, как учёт стратегической рефлексии влияет на выбор оптимальных стратегий, а также численно проверить, есть ли возможность увеличения выигрыша игрока при повышении его ранга рефлексии и может ли игнорирование этой способности привести к неоптимальным решениям. Предложенный подход будет полезен как для анализа применяемых, так и для поиска более эффективных стратегий защиты компьютерных систем от хакерских атак.

1. Постановка задачи и игровой подход

Один из подходов, моделирующий игру администратора безопасности и атакующего злоумышленника, основан на проведении биматричной игры, в которой интересы игроков не совпадают, а выигрыши задаются разными платёжными матрицами [2–4]. В этих матрицах строки соответствуют стратегиям одного игрока (программное средство или набор из программных средств), столбцы – стратегиям другого игрока, на их пересечении в первой матрице A стоит цена игры для администратора безопасности, а во второй матрице B – цена игры для злоумышленника. Проведение биматричной игры позволяет определить наиболее выигрышные стратегии для каждого игрока.

Если администратор для обеспечения безопасности системы может выбирать из S средств защиты, и при этом их можно использовать одновременно, то у него будет $N = 2^S - 1$ вариантов стратегий. Аналогично, если злоумышленник имеет L способов атаки, то у него будет $M = 2^L - 1$ вариантов вредоносных стратегий.

Ходом администратора безопасности является использование одной из N стратегий защиты компьютерной системы x_i ($i = 1, 2, \dots, N$), а ходом злоумышленника – применение одной из M стратегий атаки y_j ($j = 1, 2, \dots, M$) на компьютерную систему. Последовательно перебирая все стратегии игроков, можно заполнить две таблицы, в одной из них указывая ущерб администратора a_{ij} (табл. 1), а во второй – прибыль b_{ij} злоумышленника (табл. 2) соответственно при выборе стратегии защиты x_i ($i = 1, 2, \dots, N$) и способа атаки y_j ($j = 1, 2, \dots, M$).

Таблица 1. Платёжная матрица A

	y_1	y_2	...	y_M
x_1	a_{11}	a_{12}	...	a_{1M}
x_2	a_{21}	a_{22}	...	a_{2M}
...	...	...	...	...
x_N	a_{N1}	a_{N2}	...	a_{NM}

Таблица 2. Платёжная матрица B

	y_1	y_2	...	y_M
x_1	b_{11}	b_{12}	...	b_{1M}
x_2	b_{21}	b_{22}	...	b_{2M}
...	...	...	...	...
x_N	b_{N1}	b_{N2}	...	b_{NM}

Из табл. 1 и 2 можно выписать платёжные матрицы A и B , содержащие N строк и M столбцов с элементами a_{ij} и b_{ij} соответственно:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1M} \\ a_{21} & a_{22} & \dots & a_{2M} \\ \dots & \dots & \dots & \dots \\ a_{N1} & a_{N2} & \dots & a_{NM} \end{pmatrix}; \quad B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1M} \\ b_{21} & b_{22} & \dots & b_{2M} \\ \dots & \dots & \dots & \dots \\ b_{N1} & b_{N2} & \dots & b_{NM} \end{pmatrix}.$$

Здесь элементы a_{ij} платёжной матрицы администратора безопасности A вычисляются следующим образом:

$$a_{ij} = R(x_i, y_j) + G_i,$$

где G_i – затраты администратора безопасности на приобретение и использование средств защиты, необходимых для реализации i -й стратегии x_i ; $R(x_i, y_j)$ – величина ущерба от атаки y_j при использовании стратегии защиты x_i .

Аналогично элементы b_{ij} платёжной матрицы злоумышленника B вычисляются по формуле:

$$b_{ij} = P(x_i, y_j) - F_j,$$

где F_j – затраты злоумышленника на использование атаки y_j ; $P(x_i, y_j)$ – величина прибыли от атаки y_j при использовании администратором стратегии защиты x_i .

В биматричной игре игроки выбирают свои стратегии однократно, одновременно и независимо друг от друга. Процесс игры состоит в том, что администратор выбирает стратегию защиты x_i , злоумышленник выбирает стратегию атаки y_j , после чего вычисляется исход игры, заключающийся в том, что администратор терпит ущерб, равный a_{ij} , а злоумышленник получает прибыль b_{ij} . Решение биматричной игры сводится к поиску равновесных (оптимальных) стратегий игроков [2–4].

2. Критерии выбора оптимальных стратегий администратора безопасности и злоумышленника

В биматричной игре администратор безопасности стремится выбрать такую стратегию, которая позволит ему свести к минимуму наносимый компьютерной системе ущерб от реализации атак злоумышленника при минимизации финансовых затрат на покупку средств защиты. Поставим в соответствие каждой стратегии администратора x_i число $W_i(A)$, вычисляемое с помощью его платёжной матрицы A . Критерий выбора оптимальной стратегии x_{i_0} для администратора состоит в том, чтобы взять

$$W_{i_0}(A) = \min_i \max_j a_{ij}. \quad (1)$$

В свою очередь, злоумышленник стремится выбрать такую стратегию, которая позволит ему максимизировать свой возможный выигрыш от реализации тех или иных угроз. Если при этом он ориентируется на самые неблагоприятные условия, то ему нужно максимизировать свой возможный минимальный выигрыш. В таком случае критерий выбора оптимальной стратегии y_{j_0} для злоумышленника состоит в том, чтобы взять

$$W_{j_0}(A) = \max_j \min_i b_{ij}. \quad (2)$$

Таким образом, $W_{i_0}(A)$ и $W_{j_0}(B)$ – это наилучшие гарантированные результаты администратора безопасности и злоумышленника соответственно.

Минимаксная (1) и максиминная (2) стратегии уместны в тех случаях, когда администратор и злоумышленник не столько хотят выиграть, сколько не хотят проиграть [2]. Хотя они могут выбрать для себя и другие критерии для подбора наиболее оптимальных для них наборов программных средств [4–6].

3. Учёт стратегической рефлексии игроков

Взаимодействие администратора безопасности и злоумышленника в киберпространстве похоже на динамичную игру, в которой каждая сторона адаптирует свои стратегии, пытаясь предугадать действия оппонента. Классические биматричные игры являются полезным инструментом для моделирования таких сценариев, но они пренебрегают важным аспектом – стратегической рефлексией, т. е. способностью игроков думать о мыслях друг друга.

Дополним математическую модель биматричной игры введением рангов рефлексии для каждого игрока. Будем считать, что игрок с нулевым рангом рефлексии знает только собственную платёжную матрицу и действует, исходя из своей гарантированной наилучшей стратегии, не учитывая возможные действия оппонента, а игрок с ненулевым рангом рефлексии, равным k , предполагает, что его противник обладает чуть более низким рангом рефлексии, равным $k - 1$. Это позволяет ему выбирать оптимальную стратегию, основываясь не только на своей платёжной матрице, но и на анализе наиболее вероятных действий второго игрока [7]. При этом расчёт ожидаемых выигрышей и выбор стратегий производится им рекурсивно на основе уже рассчитанных стратегий оппонента на уровне $k - 1$.

В предложенной модели биматричной игры администратор безопасности и злоумышленник могут иметь разные уровни стратегической рефлексии, что позволяет моделировать сценарии, где один из них более искусен в анализе поведения оппонента.

4. Решение биматричной игры при задании у администратора безопасности и злоумышленника различных рангов рефлексии

Решение биматричной игры без учёта рефлексии игроков

Рассмотрим решение биматричной игры администратора безопасности и злоумышленника с платёжными матрицами небольшого размера A и B (табл. 3 и 4) для случая, когда игроки не пытаются предугадать наилучшие стратегии друг друга. Будем считать, что они играют осторожно, поэтому их оптимальные стратегии будем находить из условий минимакса (1) и максимина (2) соответственно.

Таблица 3. Платёжная матрица A

	y_1	y_2	y_3	y_4
x_1	3	5	7	3
x_2	6	4	2	3
x_3	8	2	1	4
x_4	6	7	6	2

Таблица 4. Платёжная матрица B

	y_1	y_2	y_3	y_4
x_1	7	7	5	6
x_2	4	7	3	4
x_3	2	3	8	5
x_4	4	6	8	5

В каждой строке платёжной матрицы администратора безопасности A выделим ячейки с наибольшими значениями ущерба, а в платёжной матрице злоумышлен-

ника B в каждом столбце выделим ячейки с наименьшими значениями выигрыша от атаки. Согласно формулам (1) и (2) получается, что наилучшие гарантированные результаты игроков равны $W_{i_0=2}(A) = \min(7, 6, 8, 8) = 6$ и $W_{j_0=4}(B) = \max(2, 3, 3, 4) = 4$. Поэтому для администратора безопасности и злоумышленника лучшими стратегиями будут x_2 и y_4 соответственно. При выборе игроками этих стратегий ущерб администратора безопасности будет составлять 3 у. е., а выигрыш злоумышленника – 4 у. е. (табл. 3 и 4).

Решение биматричной игры с ненулевым рангом рефлексии у злоумышленника

Рассмотрим, как изменится решение биматричной игры с теми же платёжными матрицами A и B (табл. 3 и 4), когда у злоумышленника будет ненулевой уровень рефлексии $k_B = 1$. В таком случае злоумышленник может догадаться, что администратор безопасности выберет стратегию защиты x_2 . Тогда $W_{j_0=2}(B) = \max(4, 7, 3, 4) = 7$, поэтому для злоумышленника лучшей стратегией будет y_2 . При выборе игроками этих стратегий ущерб администратора безопасности будет 4 у. е., а выигрыш злоумышленника – 7 у. е. Таким образом, повышение уровня рефлексии злоумышленника позволило ему увеличить выигрыш на 3 у. е. по сравнению с предыдущей ситуацией, когда его уровень рефлексии был нулевым. При этом ущерб администратора безопасности увеличился на 1 у. е.

Решение биматричной игры с ненулевым рангом рефлексии у администратора безопасности

Рассмотрим решение биматричной игры с теми же платёжными матрицами A и B (табл. 3 и 4), когда у администратора безопасности ненулевой уровень рефлексии $k_A = 1$. В таком случае администратор может догадаться, что злоумышленник выберет стратегию y_4 . Тогда $W_{i_0=4}(A) = \min(3, 3, 4, 2) = 2$, поэтому для администратора лучшей стратегией будет x_4 . При выборе игроками этих стратегий ущерб администратора безопасности будет 2 у. е., а выигрыш злоумышленника – 5 у. е. Таким образом, повышение уровня рефлексии администратора безопасности позволило ему уменьшить ущерб на 1 у. е. по сравнению с ситуацией, когда уровни рефлексии у обоих игроков были нулевыми.

Решение биматричной игры с учётом рефлексии обоих игроков

Рассмотрим решение биматричной игры с теми же платёжными матрицами A и B (табл. 3 и 4), но когда у администратора безопасности и злоумышленника ненулевые уровни рефлексии, равные $k_A = 2$ и $k_B = 1$ соответственно. В таком случае администратор может догадаться, что злоумышленник догадается, что администратор выберет стратегию x_4 , на основании чего он выберет стратегию y_3 . Тогда $W_{i_0=3}(A) = \min(7, 2, 1, 6) = 1$, поэтому для администратора лучшей стратегией будет x_3 . При выборе игроками этих стратегий ущерб администратора безопасности составит 1 у. е., а выигрыш злоумышленника – 8 у. е. Таким образом, дальнейшее повышение уровня рефлексии администратора безопасности позволило ему уменьшить ущерб на 2 у. е. по сравнению с ситуацией, когда уровни рефлексии у обоих игроков были нулевыми.

Если же администратор безопасности будет считать, что у злоумышленника уровень рефлексии $k_B = 1$, а на самом деле он будет нулевым, то злоумышленник выберет стратегию y_4 , а администратор – x_3 . При выборе игроками этих стратегий ущерб

администратора безопасности составит 4 у. е., а выигрыш злоумышленника – 5 у. е. Таким образом, если ранг рефлексии игрока превышает ранг рефлексии противника больше, чем на 1, то повышение ранга рефлексии не гарантирует ему улучшения исхода игры.

Проведённый анализ решений биматричной игры, моделирующей стратегическое взаимодействие администратора безопасности и злоумышленника с учётом рефлексивности их мышления, показал, что повышение ранга рефлексии даёт игроку преимущество, если его ранг рефлексии будет на единицу выше, чем у противника.

Заключение

В статье продемонстрировано решение рефлексивной биматричной игры, моделирующей стратегическое взаимодействие администратора безопасности и злоумышленника. Анализ полученных решений показал, что более высокий ранг рефлексии обычно даёт игроку преимущество, но при этом оптимальный ранг рефлексии зависит от ранга рефлексии противника. В частности, игрок, чей ранг рефлексии на единицу выше, как правило, выигрывает, в то время как при большей разнице в рангах повышение ранга рефлексии не гарантирует игроку улучшения его исхода игры.

Перспективным направлением дальнейших исследований является разработка алгоритмов для автоматического определения уровней рефлексии игроков на основе анализа их поведения в реальных сценариях и дополнительное проведение игры рангов игроков [7].

Литература

1. Вахний Т.В., Константинов П.В. Создание сервиса для анализа защищённости компьютерных систем на основе теоретико-игрового подхода и базы знаний MITRE ATT&CK // Математические структуры и моделирование. 2024. № 2 (70). С. 80–86.
2. Гуц А.К., Вахний Т.В. Теория игр и защита компьютерных систем: учеб. пособие. Омск: Изд-во ОмГУ, 2013. 160 с.
3. Вахний Т.В., Вахний С.В. Итеративное решение биматричной игры для оптимизации защиты компьютерной системы // Математические структуры и моделирование. 2022. № 1 (61). С. 105–114.
4. Вахний Т.В., Вахний С.В. Оптимизация выбора наилучшей стратегии защиты от вредоносных атак на основе решения биматричной игры с учётом рисков // Математические структуры и моделирование. 2024. № 1 (69). С. 103–109.
5. Вахний Т.В., Гуц А.К., Новиков Н.Ю. Матрично-игровая программа с выбором критерия для определения оптимального набора средств защиты компьютерной системы // Математические структуры и моделирование. 2016. № 2 (38). С. 103–115.
6. Шевченко Д.В. Методы принятия управленческих решений: задания и методические указания для выполнения расчётно-графической работы. Казань: Познание, 2014. 69 с.
7. Новиков Д.А., Чхартишвили А.Г. Рефлексия и управление: математические модели. М.: Изд-во физ.-математ. лит., 2013. 412 с.

**GAME THEORY IN CYBERSPACE: A REFLEXIVE APPROACH
TO MODELING THE INTERACTION BETWEEN
AN ADMINISTRATOR AND AN ATTACKER**

T.V. Vakhniy

Ph.D. (Phys.-Math.), Associate Professor, e-mail: vahniytv@mail.ru

S.V. Vakhniy

Student, e-mail: vakhniysv@mail.ru

Dostoevsky Omsk State University, Omsk, Russia

Abstract. This article offers an extended bimatrix game model that takes into account different levels of strategic reflection between the security administrator and the attacker.

Keywords: Computer system, cybersecurity, bimatrix game, reflection of the players, optimal strategy.

Дата поступления в редакцию: 27.02.2025

О ПРЕДСТАВЛЕНИИ НЕКОТОРЫХ РОЛЕВЫХ МОДЕЛЕЙ РАЗГРАНИЧЕНИЯ ДОСТУПА ОБЪЕКТНО-ОРИЕНТИРОВАННОЙ МОДЕЛЬЮ HRU: НАСЛЕДОВАНИЕ ИЕРАРХИЧЕСКИ ОХВАТНЫМ ЛИСТОВЫМ ПОДХОДОМ

С.В. Усов

к.т.н., доцент, e-mail: raintower@mail.ru

Омский государственный университет им. Ф.М. Достоевского, Омск, Россия

Аннотация. Публикуется заключительная часть работы по представлению некоторых ролевых политик безопасности средствами объектно-ориентированной модели с матрицей доступов.

Ключевые слова: компьютерная безопасность, ролевые модели разграничения доступа, иерархия ролей, OOHU.

Введение

В данной статье, являющейся заключительной частью работы «О представлении некоторых ролевых моделей разграничения доступа объектно-ориентированной моделью HRU», предыдущие части которой были представлены в [1] и [2], опубликованных ранее, рассматривается моделирование ещё одного подхода к построению иерархии ролей средствами OOHU. С более подробной исторической справкой и подробным описанием рассматриваемых моделей можно ознакомиться в упомянутых статьях.

Предмет рассмотрения – OOHU – представляет собой объектно-ориентированное развитие [3] дискреционной модели разграничения доступа, предложенной М. Харрисоном, У. Руззо и Д. Ульманом [4]. Объектно-ориентированный подход позволил как расширить охват компьютерных систем, которые можно описать с помощью модели HRU, так и сделал исходную модель совместимой с рядом актуальных задач.

Компьютерная система в OOHU рассматривается в виде множества объектов O , разбитых по множеству классов K (все объекты одного класса имеют одинаковый набор полей и методов), обладающих открытыми полями $f \in F$ и скрытыми полями $p \in P$, а также методами обработки полей $s \in S$. Здесь $F = \bigcup_{k \in K} k.F$ – множество всевозможных открытых полей всех объектов и классов; $k.F$ – множество открытых полей класса k (каждый объект класса k обладает тем же набором $k.F$ открытых полей), аналогично определяются P и S . Причём если поле $k.f$ наследуется классом k у класса k' , то соответствующее поле класса k' мы будем для удобства обозначать именно $k'.f$, подчёркивая тем самым их взаимосвязь (таким

образом, $f \in k.F$ и $f \in k'.F$). Пусть $O^k \in O$ – множество объектов класса $k \in K$. В случае, если требуется уточнить класс объекта, поле f объекта $o^k \in O^k$ будем обозначать $o^k.f$, поле f класса k – $k.f$. Для скрытых полей класса будем использовать аналогичные обозначения.

Для построения модели дискреционного разделения доступов для каждого объекта и для каждого класса вводится дополнительное скрытое поле M , содержащее локальную матрицу доступов, и методы работы с матрицей доступов. Строки матрицы доступов объекта o соответствуют объектам и классам системы, столбцы – полям и методам объекта o , а в ячейке, находящейся на пересечении строки, соответствующей объекту o' , и столбца, соответствующего полю либо методу $x \in X = F \cup S$, находится подмножество множества R прав доступа, определённых в системе, которое обозначается $o.M[o', x]$. Модификация матриц доступа производится посредством выполнения команд системы безопасности, о которых будет сказано ниже.

Модель безопасности OOHU называется иерархической (или моделью с иерархией), если на множестве объектов O задан частичный порядок-иерархия, и в любой момент работы системы для любых двух объектов $o, o' \in O$, таких что $o' \leq o$, для любого поля или метода $x \in X$, общего для объектов o и o' , и для любого поля или метода $x' \in X$ объекта $o'' \in O$ верно следующее: $o''.M[o, x'] \subset o''.M[o', x']$ и $o'.M[o'', x] \subset o.M[o'', x]$. Здесь и далее знак « $\leq$ » – отношение частичного порядка, задающее иерархию.

Состояние системы в модели HRU изменяется под действием команд, которые состоят из условной части и последовательности элементарных операторов [4], которая выполняется, только если истинна условная часть. Список элементарных операторов в OOHU включает [3]:

1. $Create(o^k, k)$ – создаёт объект o^k класса $k \in K$, если $o^k \in O$.
2. $Destroy(o^k)$ – уничтожает объект $o^k \in O$.
3. $Enter(r, o^k, o'^{k'}.f)$ – вносит право доступа r в $o'^{k'}.M[o^k, o'^{k'}.f]$, где o^k – объект класса k ; $o'^{k'}$ – объект класса k' .
4. $Delete(r, o^k, o'^{k'}.f)$ – удаляет право доступа r из $o'^{k'}.M[o^k, o'^{k'}.f]$.
5. $Grant(r, o^k, o'^{k'}.s)$ – разрешает вызов объектом o^k метода $o'^{k'}.s$.
6. $Deprive(r, o^k, o'^{k'}.s)$ – запрещает вызов объектом o^k метода $o'^{k'}.s$.

Изменения, производимые операторами, отражаются в матрицах доступа объектов системы. Подробное описание модели OOHU можно найти в [3].

В рамках данной работы изучается взаимосвязь между OOHU и ролевыми политиками разграничения доступа [5], в частности, возможность моделирования ролевых политик с помощью инструментария объектно-ориентированных HRU. Отметим, что в работах Р. Санду [6] уже было показано, что технология управления доступом на основе ролей обладает достаточной гибкостью для моделирования как дискреционных, так и мандатных политик безопасности. Однако в данных работах моделируются лишь частные дискреционные политики, близкие по структуре к Take-Grant, но заметно отличающиеся от HRU.

В связи с необходимостью избежать конфликта одинаковых обозначений, традиционных для дискреционных моделей матриц доступа и ролевых моделей, в рамках данной работы используются нестандартные обозначения для элементов ролевой модели:

U – множество пользователей, $u \in U$ – пользователь.

Z – множество ролей, $u.z$ – обозначение роли $z \in Z$, назначенной пользователю u .

Y – множество полномочий, $z.y$ – обозначение элементарного полномочия $y \in Y$, относящегося к роли z (поскольку каждая роль задаётся именно набором полномочий).

C – множество сеансов работы пользователей с системой. Управление доступом осуществляется на основе изменения множества активных сеансов системы. Каждому сеансу $c \in C$ сопоставляется пользователь $u_c \in U$, подмножество доступных пользователю в рамках данного сеанса ролей $Z_c \subset Z$ и подмножество допустимых полномочий $Y_c \subset Y$. Считается, что система функционирует безопасно, если пользователь u_c может осуществлять действия только в рамках полномочий из множества $Y_c = \bigcup_{z \in Z_c} z.Y$ во время сеанса $c \in C$.

Множество ролей в реальных системах, как правило, связано иерархией, принципы построения которой могут отличаться в зависимости от организационно-управленческих связей, в соответствии с которыми функционирует система. Как правило, подчинённость ролей в иерархии включает наследование прав и полномочий, которое может быть направлено как «снизу», так и «сверху». При наследовании «сверху» подчинённый субъект наследует права родительских субъектов. Однако базовые ролевые модели эксплуатируют наследование «снизу». Здесь можно выделить три ключевых подхода к построению иерархии [7]:

1. Строгий таксономический листовой подход. Всё множество полномочий разбивается на непересекающиеся подмножества, каждое из подмножеств приписывается листу дерева иерархии ролей. Роль в нелистой вершине наделяется множеством полномочий, являющимся объединением множеств полномочий непосредственно подчинённых ролей (соответствие между ролями и полномочиями, таким образом, индуцируется полномочиями листовых вершин и устанавливается при движении по дереву ролей снизу вверх).

2. Нестрогий таксономический листовой подход. Единственное отличие от предыдущего подхода – отсутствие требования на запрет пересечения подмножеств полномочий, сопоставленных листовым вершинам.

3. Иерархически охватный подход. Граф такой иерархии ролей не обязан быть деревом (но, конечно, не может содержать сильных циклов). При таком подходе считается, что если пользователю сопоставлена некоторая роль $r \in R$, то ему также должны быть сопоставлены и все роли, подчинённые r . Что позволяет исключить из роли r полномочия, уже содержащиеся в подчинённых ей ролях.

Иерархическая модель OOHU [3] предполагает, что объект o , находящийся на более низком уровне иерархии, чем объект o' , обладает меньшим набором прав (как в отношении доступа к другим объектам, так и в отношении ограничения доступа других объектов по отношению к себе) по сравнению с объектом o' . Такая структура напоминает решётку иерархии ролей, что позволяет предположить структурную близость данных моделей.

Связь между субъектно-объектной ролевой моделью без иерархии на множестве ролей и OOHU, равно как и связь между субъектно-объектной ролевой моделью с иерархией с наследованием «сверху» на множестве ролей и OOHU, были изучены

в работе [1]. Возможность моделирования таксономического листового подхода была представлена в [2]. Здесь будет разобран случай ролевой модели с иерархически охватным подходом к организации иерархии ролей.

Случай иерархически охватного подхода к наследованию

Базовая ролевая политика разграничения доступа подразумевает статичность подсистемы безопасности в рамках одного сеанса, т. е. полномочия неизменны, к каждой роли относится фиксированный набор полномочий, и каждому пользователю назначен неизменный в течение сеанса набор ролей. Таким образом, перераспределения прав доступа в смысле дискреционной политики безопасности не происходит. Построим иерархическую объектно-ориентированную модель HRU $\Sigma' = (O(t), M(t), K, R)$, соответствующую данной субъектно-объектной ролевой модели $\Sigma = (U, Z, Y, C)$.

Для начала необходимо определить множество R прав доступа модели Σ' , вычленив права доступа r из элементарных полномочий, а также множество X объектов системы. В ролевой модели указанные множества явно не определяются, но фактически каждое элементарное полномочие представляет собой возможность осуществить доступ определённого вида к массиву данных, который может быть представлен как объект с точки зрения дискреционной политики безопасности. Таким образом, элементарное полномочие можно рассматривать как пару $y = (x, r)$, где x отвечает объекту системы (или группе сходных объектов) в рамках субъектно-объектной парадигмы, а r – праву доступа к этому объекту (или группе объектов). В качестве r может также выступать право вызова, если x представляет собой активную сущность системы. Сформированное таким образом множество всех объектов системы обозначим через X , а множество всех прав доступа – через R .

Далее определим классы модели ООHRU:

- класс пользователей k_U ;
- класс сеансов k_C – эти два класса необходимы для задания отображения между множеством пользователей и множеством сеансов в ролевой модели;
- классы объектов k_o ;
- классы ролей k_α .

В целом построение этих классов практически не отличается от случая ролевой модели, свободной от иерархии [1].

Зададим классы объектов, а также вспомогательные класс пользователей и класс сеансов: в реализующей ролевую политику с иерархией модели ООHRU они задаются так же, как и в случае реализации ролевой политики без иерархии на множестве ролей.

Так, для определения классов объектов будем работать со сформированным ранее множеством X , вообще говоря содержащим как пассивные, так и активные сущности, поскольку полномочие может выражаться как в изменении каких-либо данных, так и в осуществлении какого-либо потока данных (исполнении файла, алгоритма). В частности, пассивной сущности $x \in X$ субъектно-объектной модели будет соответствовать объект o , имеющий поле f , содержащее информацию объекта x , и приватное поле M – матрицу доступов. Активной сущности x' субъектно-

объектной модели соответствует объект o' , содержащий метод s , выполняющий активные функции субъекта x' , и приватное поле M – матрицу доступов. Объекты будут в дальнейшем разбиты по классам исходя из их структуры и возможностей доступа к ним [1].

Каждой роли $z \in Z$ также будет сопоставлен класс $k_{\alpha(z)} \in K$ в модели ООHRU. Для описания этого сопоставления конкретизируем интерпретацию иерархически охватного подхода, которую будем использовать в дальнейшем:

- во-первых, в иерархии ролей каждому узлу соответствует ровно одна роль, и ни одна роль не встречается дважды;
- во-вторых, ни в одной ветви дерева иерархии ролей (нисходящей простой цепи от узла до листа; если иерархия ролей не является деревом, то она всё равно не может содержать сильных циклов) ни одно полномочие не определено более чем у одной роли;
- в-третьих, если пользователю назначается какая-то роль, то вместе с ней ему назначаются и все нижестоящие по иерархии роли.

Но в естественной иерархии классов модели ООHRU должно происходить наследование прав доступа от родителя к ребёнку. Это значит, что при построении иерархии классов, отвечающих ролям, в ООHRU, во-первых, граф иерархии будет перевёрнут относительно соответствующего графа ролевой модели. При этом корнем будет являться «пустой» класс k_0 (соответствует пустому множеству ролей), а его непосредственными наследниками – классы вида k_z , где z – листовая роль в иерархии ролей. Во-вторых, с целью учёта всех полномочий, которые фактически окажутся у пользователя при назначении ему определённой роли z , соответствующий класс в ООHRU определим как $k_{\alpha(z)}$, где $\alpha(z)$ – множество ролей, содержащее как роль z , так и все нижестоящие роли в иерархии ролей.

Таким образом, если в графе ролей z' является непосредственным наследником роли z , то в иерархии соответствующих ролям классов в ООHRU класс $k_{\alpha(z)}$ будет являться непосредственным наследником класса $k_{\alpha(z')}$, при этом множество $\alpha(z)$ будет являться объединением всех ролей, содержащихся во множествах родителей (таких, как $\alpha(z')$) и самой роли z .

Заполнение матриц доступов объектов происходит следующим образом: если множество ролей $\alpha \in 2^Z$ содержит роль с элементарным полномочием $y = (x, r) = (o.f, r)$, где x – объект субъектно-объектной модели, объектно-ориентированной интерпретацией которого является объект o , то в матрицу доступов объекта o в ячейку на пересечении строки поля f и столбца класса k_α заносится право доступа r : $o.M[k_\alpha, f] := o.M[k_\alpha, f] \cup r$.

Если множество ролей α содержит роль с элементарным полномочием $y = (x', r) = (o'.s, r)$, где x' – субъект субъектно-объектной модели, объектно-ориентированной интерпретацией которого является объект o' , а r – право вызова процедуры, то в матрицу доступов объекта o в ячейку на пересечении строки метода f и столбца класса k_α заносится 1: $o.M[k_\alpha, f] := 1$.

Отсутствие перераспределения доступов в рамках одного сеанса влечёт отсутствие таких элементарных операторов, как *Enter*, *Delete*, *Grant*, *Deprive*, в командах модели Σ' . Старт пользователем u сеанса c , в котором ему назначается роль z , а точнее (в рамках иерархически охватного подхода) множество ролей $\alpha = \alpha(z)$,

будет выражен в Σ' следующей командой:

$CommandStartSession_ \alpha(o_u : k_\alpha)$
 $Create(o_u, k_\alpha).$

Команда завершения сеанса этим же пользователем:

$CommandEndSession_ \alpha(o_u : k_\alpha)$
 $Destroy(o_u).$

Для проверки возможности пользователю u назначить роль z будут использоваться класс сеансов k_C и класс пользователей k_U . Для каждого сеанса c в класс k_C помещается объект o_c , содержащий поле-идентификатор сеанса, метод $user$ инициализации пользователя и матрицу доступов. Для каждого пользователя u в класс k_U помещается объект $o*_u$, содержащий метод $user$ и матрицу доступов. Аналогичный метод $user$ будут содержать все классы и объекты естественной иерархии подмножеств ролей. Роль этого метода формальна, нам потребуется лишь соответствующая ему ячейка в матрице доступов содержащего его объекта (класса).

Положим, что сеанс c подразумевает назначение пользователю u роли z , а значит и всех ролей из множества $\alpha = \alpha(z)$. В этом случае при создании объекта o_c происходит модификация матриц доступа объекта $o*_u$ и класса k_α следующей командой:

$CommandCreateSession_ \alpha(o_c : k_C; o*_u : k_U; k_\alpha)$
 $Create(o_c, k_C),$
 $Grant(o_c, o*_u.user),$
 $Grant(o_c, k_\alpha.user).$

При выполнении этой команды матрицы доступов модифицируются следующим образом: $o*_u.M[o_c, user] = 1$, $k_\alpha.M[o_c, user] = 1$ (объект o_c получает право запускать метод $user$ объекта $o*_u$, относящегося к пользователю u , и аналогичный метод класса k_α). Команда может быть выполнена только если пользователь u зарегистрирован в системе, при регистрации был создан соответствующий ему объект $o*_u$.

Теперь в команду сеанса мы можем включить проверку возможности пользователю u назначить множество ролей α :

$CommandStartSession_ \alpha(o_c : k_C; o*_u : k_U; o_u : k_\alpha)$
 If
 $o*_u.M[o_c, user] = 1 \text{ and } k_\alpha.M[o_c, user] = 1$
 then
 $Create(o_u, k_\alpha).$

Для смены набора назначенных ролей α на набор β пользователю необходимо завершить текущий сеанс и начать новый, в котором ему будет назначен требуемый набор ролей (конечно, если такой сеанс существует). В ООHRU это осуществимо последовательностью команд

$CommandEndSession_ \alpha,$
 $CommandStartSession_ \beta.$

Также средства ООHRU позволяют модифицировать множество привилегий, назначенных роли z . Модификация множества привилегий, назначенных роли z , осуществима следующим образом. Допустим, роли $z_1 \dots z_t$ являются непосредственными родителями роли z , которой необходимо добавить новое элементарное полномочие $y = (o.f, r)$, т. е. право доступа r к полю f объекта o некоторого класса k_o . Для

этого используем команду

$$\begin{aligned} & \text{CommandAddPermission_z_r_f}(k_{\alpha(z)}, o : k_o) \\ & \text{Enter}(r, k_{\alpha(z)}, o.f). \end{aligned}$$

Данная команда, однако, может быть выполнена только если в матрицы всех классов непосредственных потомков (соответствующих родительским ролям в иерархии ролей) $k_{\alpha(z_1)} \dots k_{\alpha(z_t)}$ уже было внесено право доступа, соответствующее привилегии y . При необходимости этого можно добиться, предварив данную команду цепочкой аналогичных команд для всех классов-потомков $k_{\alpha(z)}$, начиная с нижнего уровня. Дело в том, что в ООHRU оператор *Enter* может быть выполнен только если соблюдены так называемые условия целостности [3]:

$$(r \in o.M[k_{z_1}, f]) \& \dots \& (r \in o.M[k_{z_s}, f]).$$

Если элементарное полномочие содержит право вызова метода s объекта o некоторого класса k_o , используем команду несколько иного вида:

$$\begin{aligned} & \text{CommandAddPermission_z_s}(k_{\alpha(z)}, o : k_o) \\ & \text{Grant}(k_{\alpha(z)}, o.s). \end{aligned}$$

Как и ранее, построенная нами модель ООHRU будет однородной в зоне классов естественной иерархии подмножеств ролей.

Сформулируем полученный результат в виде теоремы.

Теорема 1. *Для любой субъектно-объектной иерархической ролевой модели с иерархически охватным подходом к наследованию существует реализующая ее иерархическая модель ООHRU.*

Литература

1. Усов С.В. О представлении некоторых ролевых моделей разграничения доступа объектно-ориентированной моделью HRU // Математические структуры и моделирование. 2018. № 4 (48). С. 127–137.
2. Усов С.В. О представлении некоторых ролевых моделей разграничения доступа объектно-ориентированной моделью HRU: наследование таксономическим листовым методом // Математические структуры и моделирование. 2020. № 4 (56). С. 126–135.
3. Усов С.В. Об отношении между дискреционными моделями объектно-ориентированных и субъектно-объектных компьютерных систем. // Проблемы информационной безопасности. Компьютерные системы. 2013. Т. 3, С.18–26.
4. Harrison M.A., Ruzzo W.L., Ulman J.D. Protection in Operating Systems // Communications of the ACM. 1975. P. 14–25.
5. Ferraiolo D.F., Kuhn D.R. Role-Based Access Control // 15th National Computer Security Conference. 1992. P. 554–563.
6. Sandhu R., Munawar Q. How to do discretionary access control using roles // 3rd ACM Workshop on Role-Based Access Control. 1998. P. 47–54.
7. Гайдамакин Н.А. Разграничение доступа к информации в компьютерных системах. Екатеринбург: Изд-во Ур. ун-та. 2003. 328 с.

**ON THE REPRESENTATION OF ROLE-BASED ACCESS CONTROL MODELS
BY OBJECT-ORIENTED HRU MODEL: HIERARCHICAL COVERAGE LEAF
APPROACH**

S.V. Usov

Ph.D. (Techn.), Associate Professor, e-mail: raintower@mail.ru

Dostoevsky Omsk State University, Omsk, Russia

Abstract. In this paper the possibility of representing of some types of role-based access control models by object-oriented discretionary access control model is considered. The role-based security models with non-taxonomic hierarchy are considered. The permissions of the role-based access control model are represented as a set of pairs of object and access right. A hierarchy of classes of the object-oriented HRU model, based on the role-based access control policy, is constructed.

Keywords: computer security, role-based access control model, object-oriented discretionary access control model, role hierarchy.

Дата поступления в редакцию: 02.02.2025

Памяти Владимира Никаноровича Ремесленникова

26.09.1938 – 13.05.2025



Редакция журнала «Математические структуры и моделирование» с глубокой скорбью сообщает о кончине выдающегося учёного, доктора физико-математических наук, профессора Владимира Никаноровича Ремесленникова, чьи работы стали фундаментом для современных исследований в теории групп, алгоритмической алгебре и математическом моделировании.

Научный путь

В.Н. Ремесленников родился в Ишиме Тюменской области, окончил Пермский государственный университет (1961), а его дальнейшая карьера была связана с ведущими научными центрами страны: Институтом математики СО РАН (Новосибирск) и Омским государственным университетом. Под руководством академика М.И. Каргаполова он защитил кандидатскую диссертацию (1967), а в 1975 году – докторскую, заложив основы новых направлений в алгебре.

Его труды в области алгебраической геометрии для свободных групп, алгоритмических проблем теории групп и разработки системы «Magnus» (системы компьютерной поддержки исследований в комбинаторной теории групп) стали классикой. Лауреат премии им. И.М. Виноградова (2007), он сочетал глубину теоретических исследований с прикладными решениями, что нашло отражение в более чем сотне статей.

Учитель и наставник

Владимир Никанорович воспитал плеяду учёных: под его руководством защитились 30 кандидатов и 8 докторов наук. Его лекции, семинары и научные школы в Новосибирске и Омске стали стартовой точкой для многих исследователей, работающих сегодня на стыке математики и компьютерных наук.

Наследие

Работы В.Н. Ремесленникова по сей день цитируются в ключевых трудах по теории групп, а его научные результаты используются в международных проектах. Редакция журнала особо отмечает его вклад в развитие компьютерных наук, в частности, по алгоритмическим вопросам теории групп – теме, которой он посвятил ряд новаторских статей, задав вектор для междисциплинарных исследований.

Выражаем искренние соболезнования семье, коллегам, ученикам и всем, кому посчастливилось работать с Владимиром Никаноровичем. Его научная смелость, мудрость и преданность делу навсегда останутся в истории математики.

Светлая память.

Редакция журнала

Научный журнал

Математические структуры И моделирование

№ 2 (74)

Главный редактор
А.К. Гуц

Зам. гл. ред., выпускающий редактор
Д.Н. Лавров

Зам. гл. ред., технический редактор
Н.Ф. Богаченко

Корректор:
Ф.А. Абдурашитова

Перевод на английский язык в авторской редакции

Адрес научной редакции

644077, Омская обл., г. Омск, пр-т Мира, д. 55а,
Омский государственный университет

E-mail: aguts@mail.ru, dmtrl@ya.ru

Электронная версия журнала:

<http://msm.omsu.ru>

<http://msm.univer.omsk.su>



Подписано в печать 02.06.2025. Формат 60 × 84 1/8.
Усл. печ. л. 17,2. Тираж 25 экз. Заказ № 161.

Отпечатано на полиграфической базе издательства ОмГУ им. Ф.М. Достоевского
644077, Омская обл., г. Омск, пр-т Мира, д. 55а

ISSN 2222-8772



9 772222 877005



25074 >